Relational Contracts, Incentives and Information

by

Jonathan David Levin

A.B., B.S., Stanford University (1994)
M.Phil., Oxford University (1996)

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

June 1999

Signature of Author ....
Department of Economics
May 11, 1999

Certified by ....
Bengt Holmström
Paul A. Samuelson Professor of Economics
Thesis Supervisor

Accepted by ....
Peter Temin
Chairman, Department Committee on Graduate Students

Relational Contracts, Incentives and Information

by

Jonathan David Levin

Submitted to the Department of Economics
on May 11, 1999, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Chapter 1: I study the design and functioning of self-enforcing incentive contracts under imperfect observability, using a model of repeated agency that allows for both common and private performance monitoring. When performance measures are mutually observed, optimal relational agreements always keep the parties on the Pareto frontier. When performance measures are privately observed, self-enforcing agreements involve the possibility of separation on the equilibrium path, but optimal contracts still take a basic "termination" form. Using these results, one can view optimal long-term contracts as the solution to a static optimization problem. I use this static program to describe the shape of optimal contracts and the nature of second-best inefficiencies. Under standard conditions, optimal moral hazard contracts are "one-step" — a fixed discretionary transfer is made to the agent any time performance is above some cut-off. Hidden information contracts are also characterized and it is shown that optimal contracts call for effort distortion by all types.

Chapter 2: This chapter considers self-enforcing relational contracts between a firm and many agents. Even when contracting opportunities are technologically independent, firms will benefit from reaching multilateral contracts that link their transactional arrangements. Optimal multilateral contracts equalize the shadow cost of incentive constraints on each relationship, something bilateral contracts will generally fail to do. I derive some novel implications for asset ownership and *ex ante* investment, and consider ways in which firms might be able to use existing relationships as "leverage" in reaching new agreements. I also investigate conditions under which firms might want to refrain from multilateral contracting and conduct relationships separately — this may be the case if firm is concerned about a breakdown in one relationship acting as a catalyst that brings down others. The results are applied to discuss two-tier workforce arrangements, supplier associations and the prevalence of diversified business groups in developing countries.

Chapter 3: A seminal theorem due to Blackwell (1951) shows that every Bayesian decision-maker prefers an informative signal Y to another signal X if and only if Y is statistically sufficient for X. Sufficiency is an unduly strong requirement in most economic problems because it does not incorporate any structure the model might impose. This chapter develops a general theory of information that allows a characterization of the information preferences of decision-makers based on how their marginal returns to acting vary with the underlying (unknown) state of the world. The analysis focuses on "monotone decision problems," in which all decision-makers in the relevant class choose higher actions when higher values of the signal are realized. This restriction allows a characterization of information preferences in terms of stochastic dominance orders over distributions of posterior beliefs. Conditions are also given under which one decision-maker has a higher marginal value of information than another decision-maker, and thus will acquire more information. The results are applied to oligopoly models, labor markets with adverse selection, hiring problems, and a coordination game. (This chapter is co-authored with Susan Athey.)

Chapter 4: This chapter revisits Akerlof's classic adverse selection market and asks the following question: do greater information asymmetries reduce the gains from trade? Perhaps surprisingly, the answer is no. Greater asymmetries worsen the "buyer's curse," thus lowering the demand curve, but may shift the supply curve as well. Whether trade increases or decreases depends on where the information impacts the market. A characterization is given for the case of partition information and then for the general case using a definition of information formulated in the previous chapter.

Thesis Supervisor: Bengt Holmström
Title: Paul A. Samuelson Professor of Economics

Acknowledgements

I enjoyed wonderful guidance and support at M.I.T. It was a privilege to work with Bengt Holmström. He sets a standard as a thinker and an economist towards which to strive. I also benefited from the outstanding advising of Glenn Ellison and Jerry Hausman. They provided keen insights and generosity at every turn. Listening and talking to Bob Gibbons greatly influenced my research interests and thinking. Susan Athey co-authored one of the chapters in this thesis and collaborated on other research; I learned a lot from my interaction with her. I am indebted as well to Paul Joskow and Jean Tirole for their advice and encouragement.

My friends in the department helped make the last few years enjoyable. In particular, I am delighted to acknowledge Max Amarante, Scott Ashworth, Mitali Das, Botond Koszegi, Paras Mehta, Markus Möbius, Sharon Nowak, Alan Sorensen, and Annette Vissing-Jorgensen.

My family has supported me in every way. Their intellectual vibrance is an inspiration. It has been, in addition, not a small pleasure to follow my father into economics.

And finally, Amy — who made it through with me.

Contents

1 Relational Incentive Contracts 9

2 Multilateral Contracting and Organizational Design 53

3 The Value of Information in Monotone Decision Problems (with Susan Athey) 85

4 The Relationship between Information and Trade: A Look at the Lemons Market 123

Chapter 1

Relational Incentive Contracts

1.1 Introduction

Economists working in the tradition of Coase and Williamson have focused on the ability of parties to carry out transactions when the potential for writing court-verifiable contingent contracts is limited. When court-enforced contracts alone are ineffectual, parties may rely on *relational contracts* that include both formal court-enforced and informal self-enforced provisions. Self-enforcement can be modeled with an appeal to repeated games — the problem of contract design becomes one of selecting and supporting a particular equilibrium. My interest in this chapter is in exploring the design and functioning of relational contracts under *imperfect observability* — that is, in situations where parties may be privately informed about their costs, actions may be observed imperfectly, or parties may have different interpretations of a given outcome. I develop a model of repeated agency that allows for both hidden information and moral hazard (hidden action), as well as private performance monitoring, and obtain sharp results characterizing both the dynamic structure of optimal relational contracts and the specific design of second-best incentives.

The importance of informational problems in the context of relational contracting has been stressed at length in the descriptive literature (Macaulay, 1963; Macneil, 1974, 1978; Williamson, 1985), but has not received a detailed theoretical analysis. Previous theoretical work on relational contracting has concentrated on contract enforceability — the general assumption being that although verifiable performance measures are poor or unavailable, parties within the relationship perfectly observe each other's behavior (Baker, Gibbons and Murphy, 1994, 1997; Bull, 1987; Greif, 1993; Klein and Leffler, 1981; Kreps, 1990; MacLeod and Malcomson, 1989; Shapiro and Stiglitz, 1984).[1] This study generalizes and extends the relational

[1] Baker, Gibbons and Murphy (1994) allow for noisy performance measures, and Kreps (1990) discusses an agency situation where one side has private cost shocks; both these papers study specialized, not optimal, contracts.

contracting framework and establishes a link to complete contracting, in which problems of asymmetric information and performance measurement are central (Holmström, 1979; Hart and Holmström, 1987). It departs from complete contract theory in assuming that key contracting variables are observable but not verifiable. It is this gap between observability and verifiability that creates a role for self-enforcing contingent contracts.[2]

In the agency model I develop, relational contracts combine a fixed court-enforced payment for each period with an informal agreement that suggests actions and discretionary payments or penalties corresponding to each contingency.[3] Such contracts have a natural interpretation in many contexts — for example, in an intermediate goods market, the formal contract states *ex ante* a fixed price to be paid on delivery, while the informal agreement describes how the price, or future prices, will be "adjusted" *ex post* to reflect delivered quality. I focus attention on incentive provision by assuming that both the firm (the principal) and the agent are risk-neutral.[4] The analysis proceeds in two steps: first by showing that the problem of long-term relational contract design can be reduced to a static optimization problem, then by using this static program to derive optimal incentive plans. A summary of the main points follows.

When nonverifiable performance measures are observed by both parties (monitoring is *common*), I show that optimal relational contracts *always* keep parties on the constrained efficient frontier. Moreover, I show that optimal incentives can be provided using the same payment plan in every period on the equilibrium path — that is, *stationary* agreements are optimal. When performance measures are observed by the firm alone (monitoring is *private*), stationary agreements are no longer effective, but I show that another type of simple contract suffices for optimality — one that calls for the same incentives in every period, but probabilistic termination following bad outcomes. These *termination* contracts can be formally represented as stationary contracts that call for the parties to burn money on the equilibrium path.

Having established that optimal contracts have a stationary representation, I turn to characterizing second-best incentive plans. Optimal relational contracts solve a constrained optimization problem similar to the one in static complete contracting models — in which the parties seek to maximize their joint surplus subject to static incentive constraints on their action choices — only the relational contracting program involves an additional dynamic self-enforcement constraint. This constraint limits the range of transfers between the parties — promised payments must be small relative to the future gains from continuing the relationship. The self-enforcement constraint does not bind when the discount factor is near one and the

[2]The incomplete contracting literature initiated by Grossman and Hart (1986) emphasizes a different mechanism to provide incentives in the absence of verifiable performance measures: the allocation of ownership or control rights. That approach is entirely compatible with the argument that parties will use ongoing informal incentives (for an eloquent discussion, see Baker, Gibbons and Murphy, 1997).

[3]In particular, I adopt the contracting assumptions of MacLeod and Malcomson (1989).

[4]For repeated game analyses that address the trade-off between insurance and incentives, see Atkeson (1991), Pearce and Stacchetti (1998), Radner (1985) and Spear and Srinivasta (1987).

parties' outside options are unattractive (meaning it is *as if* court-enforced contracts could be written). On the other hand, when the discount factor is near zero and the parties' outside options are appealing, the opposite obtains: the requirement of self-enforcement renders all relational contracts infeasible. The interesting case is the broad middle, where the requirement of self-enforcement constrains, but does not eliminate, incentive provision.

Under these conditions, I find that optimal relational contracts involve second-best distortions that are conceptually and qualitatively distinct from those found in static complete contracting models (where anything mutually observed is assumed to be verifiable). I analyze three leading cases in detail: pure moral hazard with common monitoring, pure moral hazard with private monitoring, and pure hidden information.[5] I show that under standard conditions, optimal moral hazard contracts are "one-step": the maximum feasible reward is given when output is above some cut-off, and no reward is given when output is below the cut-off. When monitoring is private, the "one-step" contract has an additional twist — following bad outcomes, the parties terminate the relationship. These results are driven by the combination of risk-neutrality (which makes "extreme" rewards desirable) and the limit on transfer payments (which defines the extremes), and are related to Abreu, Pearce and Stacchetti's (1990) bang-bang necessity result. The simplicity of optimal relational contracts is especially appealing in light of the often-made criticisms of the standard risk-averse agent complete contracting model, which is said to be overly sensitive to small changes in the information structure and to predict unreasonably complex incentive plans (Hart and Holmström, 1987).

Optimal hidden information contracts likewise involve interesting differences from those found in complete contracting models. The restriction on transfer payments imposed by self-enforcement limits both the amount of screening achieved by optimal contracts and the overall *level* of incentive provision. As a consequence, optimal contracts distort the effort choices of *all* cost types, a result that contrasts to the usual finding of "no distortion at the top." Optimal relational contracts also tend to involve more "pooling" of types than occurs in standard complete contracting environments.

My characterization of the dynamic structure of optimal contracts relies on the dynamic programming tools developed by Abreu (1988), Abreu, Pearce and Stacchetti (1990) and Fudenberg, Levine and Maskin (1994) to study repeated games. A conceptual distinction between relational contracting and standard repeated game models is the ability of the parties to make explicit transfer payments. The use of transfer payments means that although sanctions will need to be carried out on the equilibrium path, parties can "settle up" at the end of each period, and (if they decide to continue) move forward optimally. In the work of Green and Porter (1984), Abreu, Pearce and Stacchetti (1990), Fudenberg, Levine and Maskin (1994) and others

[5]Analyzing the cases in turn clarifies the results and facilitates comparison to existing work on complete contracting and repeated games. I note briefly how the cases might be combined.

on repeated games with imperfect observability, incentives are provided through the use of continuation payoffs, and optimal equilibria generally involve variation in play on the equilibrium path. The formal connection between contracting (where transfers are made explicitly) and optimal repeated game equilibria (where "transfers" are made through continuation play) has been stressed by Fudenberg and Levine (1994) and Fudenberg, Levine and Maskin (1994) in the process of studying the equilibrium set of general repeated games for large discount factors. The connection in this chapter is, of course, quite explicit and allows me to obtain detailed characterizations of optimal contracts for all discount factors.

1.2 The Model

There is one firm and one agent. Both are risk-neutral, live forever and have discount factor δ. Time is discrete, indexed by $t = 1, 2, \ldots$. At each date, the firm and the agent simultaneously announce whether to continue their relationship or terminate it. Either party can call off the relationship and termination is irreversible.[6] If the agent is employed at date t, he observes a private cost parameter θ_t (reflecting task difficulty, work environment, etc.). The cost parameter θ_t is an i.i.d. random variable with finite support $\Theta = \{\theta_1, \ldots, \theta_m\}$ and pdf $p(\theta)$, where $\theta_1 < \ldots < \theta_m$. The agent then chooses an action e_t at private cost $c(e_t, \theta_t)$.[7] The agent's interests are not perfectly aligned with the firm's, so effort is costly: c is strictly increasing and convex in e on $[0, \bar{e}]$. Zero effort yields zero disutility, i.e. $c(0, \theta) \equiv 0$. Assume that c is increasing in θ, and that the standard sorting condition holds: c is supermodular in (e, θ), i.e. c_e is increasing in θ. Once the agent has acted, output y_t is realized. Output is stochastic, with finite support $Y = \{y_1, \ldots, y_n\}$, $y_1 < \ldots < y_n$, and pdf $f(y|e)$, where $f(y|e) > 0$ for all y, e. I capture the idea that output is positively related to effort by requiring f to have the monotone likelihood ratio property ($f_e(y|e)/f(y|e)$ is increasing in y). After output is realized, the firm makes a report to the agent, interpreted as a statement of output, $m_t \in Y$. This announcement will be relevant when the firm privately observes output.

Time t payments are made at the end of the period and are composed of a formal (court-enforced) payment w_t and an informal (voluntary) bonus b_t. Letting W_t be the total time t payment, the agent's payoff is $W_t - c(e_t, \theta_t)$, while the firm's payoff is $y_t - W_t$. The surplus generated in the period is $y_t - c(e_t, \theta_t)$. To keep the problem well-behaved, take $f(y|e)$ and $c(e, \theta)$ to be continuous in e, and assume that for every θ, $E[y|e] - c(e, \theta)$ is concave and has a unique interior maximizer $e^{FB}(\theta)$. Note that c being supermodular implies that $e^{FB}(\theta)$ is

[6] The assumption that a relationship cannot be re-started is convenient but not necessary.

[7] When both θ_t and e_t are private to the agent, the parties would like the agent to make an announcement of his costs at the time he chooses effort (in order to tailor the incentive plan). Since this introduces extra notation that I will not need for the cases I consider, I will refrain from introducing it. Allowing for this additional communication does not affect the stationarity results.

decreasing in θ.

If the relationship has been terminated, the firm receives a per-period payoff of $\overline{\pi} \geq 0$ and the agent a per-period payoff of $\overline{u} \geq 0$. I do not explicitly model the parties' outside options, which might depend on market conditions, the potential for spot contracting, or individual production capabilities. I do assume, however, that ending the relationship yields at least as much joint surplus as continuing with zero effort: $E[y - c|e = 0] \leq \overline{u} + \overline{\pi}$. Absent a contingent contract, this means that the static Nash equilibrium of the one-shot game is to separate.[8]

It may help to think about how a one-shot game would look if court-enforceable incentive contracts could be written. If e were verifiable, with θ privately observed by the agent, there is pure hidden information, and the parties can write standard screening contracts. Alternatively, if y were verifiable, but not the action choice, the situation is one of pure moral hazard with incentive contracts based on output. The key premise here is that *neither* e nor y are verifiable, meaning it is impossible to provide incentives in a one-shot interaction. Incentives can be provided only by self-enforcing agreements based on observable information.

1.2.1 Payoffs

Let T denote the final period of employment. The agent's lifetime expected utility from period t is:

$$U_t = E\left\{\sum_{\tau=t}^{T} \delta^{\tau-t}\left[W_\tau - c(e_\tau, \theta_\tau)\right] + \delta^{T-t+1}\left[W_{T+1} + \overline{U}\right]\right\}, \qquad t = 1, ..., T+1, \tag{1.1}$$

where $\overline{U} = \overline{u}/(1-\delta)$ is the agent's lifetime expected utility from leaving employment, and W_{T+1} is the discounted present value of any severance payment. Similarly, if the firm employs the agent at t, its lifetime profits are

$$\Pi_t = E\left\{\sum_{\tau=t}^{T} \delta^{\tau-t}\left[y(e_\tau) - W_\tau\right] + \delta^{T-t+1}\left[\overline{\Pi} - W_{T+1}\right]\right\}, \qquad t = 1, ..., T+1, \tag{1.2}$$

where $\overline{\Pi} = \overline{\pi}/(1-\delta)$. The joint surplus from time t on is

$$S_t = U_t + \Pi_t = E\left\{\sum_{\tau=t}^{T} \delta^{\tau-t}\left[y(e_\tau) - c(e_\tau, \theta_\tau)\right] + \delta^{T-t+1}\left[\overline{U} + \overline{\Pi}\right]\right\}, \qquad t = 1, ..., T+1. \tag{1.3}$$

Let $\overline{S} = \overline{U} + \overline{\Pi}$ be the total surplus if the parties separate.

[8] Clearly, neither party will make a voluntary *ex post* payment in a one-shot game. But this means the agent's payoff is independent of performance, and he will never voluntarily incur effort costs. Thus regardless of fixed payments, at least one party must weakly prefer to separate. "Separation" can be interpreted as spot transacting with zero effort.

1.2.2 Relational Contracts

A relational contract can be viewed as an *ex ante* agreement between the parties that consists of a court-enforced fixed-price contract and informal terms governing action choices and discretionary payments. Following MacLeod and Malcomson (1989), I assume that *formal* contracts enforced by the court can specify only fixed payments to be paid at the end of each period. If the agent is employed, I write the fixed payment as w_t; if the relationship has just terminated, I write the separation payment as d_t. These payments can be contingent on the history of past payments and on employment status.[9] I rule out the possibility that the parties can write contracts involving transfers to third parties.[10] To provide incentives, the parties can make *informal* bonus payments based upon anything they have mutually observed. Let b_t denote the net transfer from the firm to the agent. I allow $w, b, d \in (-\infty, \infty)$, and by definition, $W_t = w_t + b_t + d_t$.

The informal part of the contract specifies proposed strategies s and σ for the agent and the firm. Strategies are maps from histories to actions. To define histories, divide each period into four sub-periods, t^0 (employment choices, then θ_t observed), t^1 (e_t chosen then y_t observed), t^2 (message m_t sent) and t^3 (payments made). A time t outcome is $o_t = (q_t, r_t, \theta_t, e_t, y_t, m_t, w_t, b_t, d_t)$, where $q_t \in \{0,1\}$ is the agent's decision to quit ($q = 0$) or work ($q = 1$), and $r_t \in \{0,1\}$ is the firm's decision to fire ($r = 0$) or retain ($r = 1$). Let $h(t^k)$ be the full history up to time t^k, e.g. $h(t^0) = (o_1, ..., o_{t-1})$, and let $h(1^0) = \emptyset$. The firm and the agent will generally not observe the full history; denote the histories they observe as $h^f(t^k)$ and $h^a(t^k)$, and let $h^m(t^k)$ denote the history that has been mutually observed. Let $H^a(t^k)$, $H^f(t^k)$, $H^m(t^k)$, $H(t^k)$ be the sets of possible t^k histories. The verifiable history $h^v(t) \in H^v(t)$ is $(q_1r_1, w_1, b_1, d_1, ..., w_{t-1}, b_{t-1}, d_{t-1}, q_tr_t)$ where q_tr_t is the product of q_t and r_t (the court cannot distinguish between a quit and a firing). A strategy s for the agent is three maps $Q_{t^0} : H^a(t^0) \to \{0,1\}$, $E_{t^1} : H^a(t^1) \to [0, \bar{e}]$, $B^a_{t^3} : H^a(t^3) \to [0, \infty)$ specifying quit, effort and voluntary payment decisions contingent on his observed history. A strategy σ for the firm is three maps $R_{t^0} : H^f(t^0) \to \{0,1\}$, $M_{t^2} : H^f(t^2) \to Y$, $B^f_{t^3} : H^f(t^3) \to [0, \infty)$, specifying hiring, message and bonus decision rules. I will write B_{t^3} as a map from history to *net* transfer payments, with the understanding that $B_{t^3} = B^f_{t^3} - B^w_{t^2}$.[11] A legally enforceable formal contract, ς, is a map $W_t : H^v(t) \to (-\infty, \infty)$.

A *relational contract* is a triple (s, σ, ς), consisting of strategies for the agent and the firm, as

[9]Note that no issue arises with the parties using court-enforced message games (along the lines of Moore and Repullo, 1988) to make observable performance measures verifiable. First, there is no mechanism for "screening" (since the firm cannot be prevented from obtaining y_t once e_t is chosen); and second, I rule out third-party payments, which these mechanisms require.

[10]Such contracts would allow the parties to "bond" themselves, lowering their joint outside options. This would facilitate cooperation.

[11]This might seem problematic because $B^f_{t^3}$ and $B^a_{t^3}$ apparently have different domains. In fact, both will be functions of the mutual history.

well as a formal contract. Let Ω be the set of all contractual agreements. If the strategies (s, σ) form a sequential equilibrium given ς, the agreement is *self-enforcing*. In this case, neither the firm nor the agent can benefit by unilaterally deviating from the agreement — neither party abandons the relationship unless the agreement calls for termination, the agent takes the proposed action, and promised bonus payments are delivered. Let Ω^* be the set of self-enforcing contractual agreements.

1.2.3 Optimal Contracts

Because utility can be costlessly transferred in the form of money and both parties are risk-neutral with respect to money income, the surplus from a relationship depends only on the effort profile and not the sequence of wage payments. Moreover, any individually rational division of the surplus can be arranged with the correct wage schedule (Theorem 1.2, below). Thus it is natural to believe that the parties will reach an agreement that optimizes the total level of surplus.

Definition 1.1 *A self-enforcing agreement* $(s, \sigma, \varsigma) \in \Omega^*$ *with expected surplus* S *is constrained efficient if there does not exist* $(\hat{s}, \hat{\sigma}, \hat{\varsigma}) \in \Omega^*$ *with expected surplus* $\hat{S} > S$.

I will also refer to self-enforcing agreements that are constrained efficient as *optimal* contracts.

Theorem 1.1 *A constrained efficient contract exists (see Appendix for proof).*

Note that if no self-enforcing agreement generates $S > \overline{S}$, then immediate separation is constrained efficient. I say that *a relationship is feasible* if there exists some self-enforcing contract that generates $S > \overline{S}$.

Given that a relationship is feasible, I will be particularly interested in *stationary* agreements, which are agreements that implement the same effort in every period using the same wage and bonus plan. To state a concise definition, it is useful to let $H^e(t^k|(s, \sigma, \varsigma))$ be the set of t^k histories that could be generated if the parties follow the contract (s, σ, ς). And let φ be the subset of (θ, e, y, m) that is mutually observable.

Definition 1.2 *An agreement* $(s, \sigma, \varsigma) \in \Omega$ *is stationary if* $Q_{t^0}(h) = 1$, $R_{t^0}(h) = 1$, $E_{t^1}(h) = e(\theta_t)$, $M_{t^2}(h) = m(y_t)$, $B_{t^3}(h) = b(\varphi_t)$, $W_t(h) = w$ *for any* t^k, $h \in H^e(t^k|(s, \sigma, \varsigma))$.

Note that stationary agreements have a parsimonious structure: bonus payments are contingent only on φ_t, the mutually observed time t variables, while the agent's effort choice is contingent only on his present costs, θ_t.

I now formalize the idea that the division of the surplus is arbitrary, so long as both parties get at least their outside option. Specifically, if there exists a self-enforcing contract that generates surplus $S > \overline{S}$, it is possible to give the agent any share $\phi \in [0,1]$ of the amount $S-\overline{S}$. To see this, suppose the original contract gives the agent share $\hat{\phi}$, or payoff $\hat{U} = \overline{U} + \hat{\phi}(S - \overline{S})$. Increase the first period wage by $U - \hat{U}$ (which may be negative), where $U = \overline{U} + \phi(S - \overline{S})$. The first period wage affects only the parties' decisions about whether to initiate the contract, something that both still desire. Following the initiation decision, play follows the original contract. Moreover, if the original contract is stationary, any division of the surplus can be attained with another stationary contract. Simply augment the wage in every period by $(1 - \delta)(U - \hat{U})$ and increase the severance payment by $U - \hat{U}$. It is easy to verify that this new contract will still be self-enforcing. To summarize:

Theorem 1.2 *If there exists some (stationary)* $(s, \sigma, \varsigma) \in \Omega^*$ *that generates surplus* $S > \overline{S}$, *there exists some (stationary)* $(s', \sigma', \varsigma') \in \Omega^*$ *that gives the agent* $U = \overline{U} + \phi(S - \overline{S})$ *for any* $\phi \in [0,1]$.

1.3 Common Monitoring (Objective Measurement)

By common monitoring, I refer to situations where any performance measure observed by the firm is also observed by the agent. The settings I have in mind are those where performance criteria cannot be verified by an outsider, but are *objective* in the sense that both parties involved can immediately agree on the outcome.[12] In this section, I show that when performance measures are commonly observed, optimal agreements *must* keep the parties on the constrained efficient frontier, and that this can be done using a stationary contract. I also show that these contracts can be designed to satisfy a strong version of renegotiation-proofness. Once I have established this stationarity result, I will turn in the next three sections to obtaining detailed characterizations of optimal contracts under perfect observability, hidden information, and moral hazard.

When monitoring is common, the firm never has private information and can base its decisions only on mutually available information.[13] Because of this, the agent can condition his effort choice at time t solely on the mutual history and his present costs, as they are the only payoff-relevant variables.[14] Thus self-enforcement means that the strategies must form a *perfect public equilibrium* (Abreu, Pearce and Stacchetti, 1990; Fudenberg, Levine and Maskin, 1994) given the formal contract ς. Note that perfect public equilibrium requires that strategies be "public", i.e. based on mutual information, and that play be a Nash equilibrium following every

[12]The repeated game literature uses the term "public monitoring," which might be misleading in this context. Even when the agent observes the same performance measures as the firm, these measures are *not* observable to outsiders (i.e. the court).

[13]This means the firm's messages are irrelevant; I will ignore them.

[14]This is a standard dynamic programming result.

history.[15]

1.3.1 Stationary Contracts

The key feature in this setting is that in an optimal contract, all incentives on the equilibrium path can be provided using balanced transfers at the end of each period. This makes it possible to restrict attention to stationary agreements.

Theorem 1.3 *If monitoring is common, and a relationship is feasible, there exists some stationary* $(s^*, \sigma^*, \varsigma^*) \in \Omega^*$ *that is constrained efficient.*

The intuition behind this result is straightforward. To provide the correct incentives for the agent, the firm must offer rewards following "good" observations, and punishments following "bad" observations. Either bonuses or continuation payoffs can be used to provide incentives. Since both parties are risk-neutral with respect to money income, these two instruments are perfect substitutes and there is no loss to using only bonuses. The fact that monitoring is common means that when the time comes to make bonus payments, the specified payment is common knowledge. So it obvious to both parties if the correct transfer has been made. This makes it easy to provide incentives for the parties to actually carry out the promised transfers — simply threaten the worst possible punishment (e.g. the threat of exit) if they are not delivered and the best possible continuation (the constrained efficient surplus) if they are. Along the equilibrium path, transfer payments are always made, so play remains on the constrained efficient frontier in every period.

Proof. The proof is constructive: I consider some arbitrary self-enforcing contract that is constrained efficient and construct a stationary contract that generates at least as much surplus. Suppose $(s, \sigma, \varsigma) \in \Omega^*$ is constrained efficient and generates surplus S, with payoffs to firm and agent of Π and U. Suppose that (s, σ, ς) implements effort $e(\theta)$ in the first period. Denote by $B_1(\varphi)$ the prescribed net bonus payment following a first period realization $\varphi_1 = \varphi$, and by $U(\varphi)$, $\Pi(\varphi)$ the continuation payoffs from $t = 2$ on. Because of the stationary recursive structure of perfect public equilibria, $U(\varphi)$, $\Pi(\varphi)$ are themselves PPE payoffs, i.e. correspond to some self-enforcing agreement. Thus for all φ, $U(\varphi) + \Pi(\varphi) \leq S$. Then,

$$S = E_{\theta,y}\left[y - c(e,\theta) \mid e = e(\theta)\right] + \delta E_{\theta,y}\left[U(\varphi) + \Pi(\varphi) \mid e = e(\theta)\right]$$

implies that

$$E_{\theta,y}\left[y - c(e,\theta) \mid e = e(\theta)\right] \geq S/(1-\delta).$$

[15]To fit this definition exactly, one wants to think of the agent as choosing a *vector* of actions, one for each cost type, and later, of the firm as choosing a *vector* of reports, one for each output realization.

I will design a new stationary contract that implements $e(\theta)$ in every period. But first, note the implications of (s, σ, ς) being self-enforcing at date 1. For all φ, both the firm and the agent must prefer to make the specified bonus payment and continue the relationship rather than refuse the bonus and terminate the relationship. So,

$$\begin{aligned} -B_1(\varphi) + \delta\Pi(\varphi) &\geq \delta\overline{\Pi} - \delta d, \\ B_1(\varphi) + \delta U(\varphi) &\geq \delta\overline{U} + \delta d, \end{aligned}$$

where d is the severance payment following $W_1 = w_1$, and

$$e(\theta) \in \arg\max_e E_\varphi \left[B_1(\varphi) + \delta U(\varphi) \mid e\right] - c(e, \theta).$$

Consider a new stationary contract as follows. So long as there has been no observed deviation, let the agent choose $e(\theta)$ at each date and the firm pay bonus

$$\hat{b}(\varphi) = B_1(\varphi) + \delta U(\varphi) - \delta\overline{U} - \delta d,$$

(which is non-negative from the inequality above). Let

$$\hat{w} = -E_{\theta,y}\left[\hat{b}(\varphi) - c(e, \theta) \mid e = e(\theta)\right] + (1 - \delta)\overline{U},$$

and $\hat{d} = 0$. Following a deviation (a missed bonus payment or the realization of some φ to which the parties assign probability zero), no further bonuses are paid and the parties separate as soon as possible (giving the mutual worst PPE payoff). This contract generates payoffs $\hat{U} = \overline{U}$ and $\hat{\Pi} \geq S - \overline{U}$.

To see that this new contract is self-enforcing, note that $\hat{U} = \overline{U}$ and $\hat{\Pi} \geq \overline{\Pi}$ means that both the agent and the firm are willing to continue the relationship so long as there have been no deviations. Moreover, for all φ, $\hat{b}(\varphi) \geq 0$, so the agent accepts the bonus payment. And also

$$-\hat{b}(\varphi) + \delta\hat{\Pi} \geq -B_1(\varphi) - \delta U(\varphi) + \delta\overline{U} + \delta d + \delta(S - \overline{U}) \geq -B_1(\varphi) + \delta\Pi(\varphi) + \delta d \geq \delta\overline{\Pi},$$

so the firm prefers to pay the bonus rather than withhold it. Finally, since the bonuses are credible,

$$e(\theta) \in \arg\max_e E_\varphi \left[B_1(\varphi) + \delta U(\varphi) \mid e\right] - c(e, \theta) = \arg\max_e E_\varphi \left[\hat{b}(\varphi) \mid e\right] - c(e, \theta) + \delta\left(\overline{U} + d\right),$$

so the agent will choose $e(\theta)$. Thus the new stationary contract is self-enforcing and constrained efficient. *Q.E.D.*

Several points should be noted. First, any optimal contract *must* generate the same expected surplus in each period on the equilibrium path. To see why this follows from the proof, suppose that in some period along the equilibrium path play generates surplus less than $(1 - \delta)S$. Then in some other period on the equilibrium path, play must generate surplus strictly greater than $(1 - \delta)S$ and must be self-enforcing. From the construction in the proof, there must be a stationary contract that generates this higher surplus in every period, contradicting the fact that the original agreement was constrained efficient. On the other hand, while it is not necessary for wage and bonus payments to be stationary in an optimal contract, it is always possible to generate optimal incentives using a stationary plan. The upshot is that in searching for an optimal agreement, there is no loss in restricting attention to stationary payments.

A second point is that the stationary agreement constructed in the proof does not rely on court-enforced separation payments. This has practical significance because, at least in the U.S., large penalties for breach of contract are generally unenforceable. It also means that, as noted by MacLeod and Malcomson in the special case of perfect observability, the use of formal contracts may not improve the parties' contracting possibilities. If the parties can make the fixed payment before effort is chosen, they can avoid written contracts altogether.

Finally, the contract designed in the proof gives all the surplus to the firm. An alternative stationary agreement without bonding or severance payments could have been constructed to give all surplus to the agent. Just let $\hat{b}(\varphi) = B_1(\varphi) - \delta[\Pi(\varphi) - \overline{\Pi}] - \delta d_2$, and $\hat{w} = E[y|e(\theta)] - E[\hat{b}] - (1 - \delta)\overline{\Pi}$. This contract combines high wages with negative bonuses — the agent makes up for a poor outcome by paying a fine or rebate — and can easily be shown to be self-enforcing. Intermediate divisions of the surplus can be achieved by taking linear combinations of these two plans.

Before moving on, it is useful to return briefly to the difference between the present model and standard repeated games that do not allow for transfer payments (e.g. the model of Green and Porter, 1984). In standard repeated game models, all incentives must be provided through variation in continuation payoffs. Thus, unless deviations are perfectly detectable, stationary agreements will be ineffective. In the present model, parties can "settle up" immediately at the end of the period and deviations at this time *are* perfectly detectable. Moreover, under common monitoring, the moral hazard is entirely one-sided, so the process of settling up is exact — there is no need to sanction both parties by going to inefficient continuation play on the equilibrium path.[16]

[16] In Section 7, when there is private monitoring by the firm, incentives must be provided to both parties (the firm needs an incentive to truthfully reveal output), and "budget balance" will be a problem.

1.3.2 Renegotiation

Perhaps the most natural interpretation of a relational contract is that *ex ante* the parties reach an "understanding" of certain unwritten terms. Given this interpretation, it seems plausible that if the parties reached a point in the relationship where they were no longer behaving optimally, they would simply update the agreement. Of course, the danger posed by this sort of renegotiation is that parties may unilaterally deviate from the original agreement, anticipating that threatened sanctions will not be carried out.[17] It turns out that under common monitoring, the parties can construct constrained efficient relational contracts that are extremely robust to renegotiation. These agreements are actually constrained efficient along *every* path of play.

The proof of Theorem 1.3 constructed an agreement that was always constrained efficient on the equilibrium path but required separation following a missed bonus payment. The requirement of separation can be relaxed. Following a missed bonus payment by one side, the parties continue to implement $e(\theta)$ using either the wage-bonus pair from the proof of Theorem 1.3 or from the discussion following it. Using these payment plans, it is possible to hold either party, in particular the party that withheld a promised bonus, to their outside option. Because the agent or firm views going to a constrained efficient continuation agreement that pays their outside option with the same distaste as separating, the incentive to deliver on a promised bonus is unchanged.

Proposition 1.1 *Under common monitoring, there exist stationary self-enforcing agreements that generate the constrained efficient surplus along every path of play and achieve any surplus division.*

Proof. Let $(s, \sigma, \varsigma), (s', \sigma', \varsigma') \in \Omega^*$ be constrained efficient stationary agreements that implement $e(\theta)$ in every period using wage bonus pairs $(w, b(\varphi))$ and $(w', b'(\varphi))$ and no separation payments, and hold, respectively, the agent and firm to their outside options. For an arbitrary division of the constrained efficient surplus, there is some stationary $(\hat{s}, \hat{\sigma}, \hat{\varsigma}) \in \Omega^*$ that achieves it by implementing $e(\theta)$ with bonuses $\hat{b}(\varphi)$, a wage $\hat{w}$ and no separation payments. Consider the following agreement: choose effort $e(\theta)$, wage $\hat{w}$ and bonuses $\hat{b}(\varphi)$ and no separation payments. If the firm ever reneges on a bonus payment, switch to $(w', b'(\varphi))$; if the agent reneges, switch to $(w, b(\varphi))$. Consider the subgame following the agent reneging. Note that it *must* be the case that $b(\varphi) \geq 0$ for all φ (because the agent would never voluntarily pay a penalty when he could refuse and still get his outside option). If the firm delivers the bonus, retain the same contract for the next period. If the firm reneges on the bonus, switch to $(w', b'(\varphi))$. The subgame involving $(w', b'(\varphi))$ involves the agent paying all bonuses. If the agent reneges, switch to $(w, b(\varphi))$.[18]

[17] A starting point to the literature on renegotiation in infinitely repeated games is Farrell and Maskin (1989).
[18] Thus, the renegotiation-proof agreement takes the form of an Abreu (1988) simple penal code.

This agreement is always constrained efficient, and it is easy to check that it is self-enforcing. Equilibrium play follows $(\hat{w}, \hat{b}(\varphi))$ so it achieves the required division. *Q.E.D.*

1.4 Perfect Observability

This section lays out the benchmark case where all relevant contracting variables (i.e. θ_t, e_t, and y_t) are commonly observed before time t wages and bonuses are paid. MacLeod and Malcomson (1989) studied a deterministic version of this problem and showed, using a different argument, that optimal contracts would be stationary. An important insight of their paper (also due to Bull (1987)) is that the incentive constraints for the firm and the agent can be pooled to yield a single incentive condition: the discounted future gains from the relationship must be larger than the present costs. It is worthwhile reviewing the basic logic of this result since it will apply to the more complex environments studied below.

Suppose a stationary contract implements effort $e(\theta)$ using bonuses $b(e(\theta))$, and gives lifetime payoffs of U,Π. The agent's incentive condition for effort is

$$-c(e(\theta), \theta) + b(e(\theta)) \geq -c(\hat{e}, \theta) + b(\hat{e}), \qquad \forall \hat{e} \neq e(\theta). \tag{1.4}$$

For the firm and agent to make the promised bonus payments rather that renege, it must be that for all e

$$b(e) + \delta U \geq \delta \overline{U} + \delta d, \tag{1.5}$$
$$-b(e) + \delta \Pi \geq \delta \overline{\Pi} - \delta d. \tag{1.6}$$

where d is the separation payment that follows no bonus payment. Clearly a necessary condition for the contract to be self-enforcing is that the sum of these two inequalities holds:

$$\delta \left(S - \overline{S}\right) \geq \max_e b(e) - \min_e b(e), \tag{1.7}$$

where S is the joint surplus:

$$S = \frac{1}{1-\delta} E\left[y - c(e, \theta) \mid e = e(\theta)\right]. \tag{1.8}$$

So long as (1.7) does not bind, one can vary the fixed wage (or the level of the bonuses), and adjust U, Π (and perhaps $b, -b$) so that neither (1.5) nor (1.6) will bind. Moreover, under perfect observability, one can combine the agent's static effort incentive constraint (1.4) with the dynamic reneging constraint (1.7) — to see how this works, set $b = c$ if $e = e(\theta)$ and $b = 0$ otherwise.

Theorem 1.4 *(MacLeod-Malcomson, 1989) There exists a stationary contract* $(s, \sigma, \varsigma) \in \Omega^*$ *with effort level* $e(\theta)$ *if and only if*

$$\delta\left[S - \overline{S}\right] \geq \max_{\theta} c(e(\theta), \theta). \tag{IC}$$

An important implication of this result is that optimal contracts can be characterized by a relatively straightforward programming problem: $\max_{e(\cdot)} S$ subject to (IC). Two basic points are immediate. If δ is sufficiently large and the outside options $\overline{S}$ sufficiently unattractive, the first-best is attainable. On the other hand, if δ is small and the outside options are relatively appealing, there may be no self-enforcing contracts that specify anything other than immediate separation.[19] The optimization program applies when a relationship is feasible, but the parties are incentive-constrained. It is interesting to note how an optimal relational contract will adjust to commonly observed cost shocks. In general, the incentive constraint will not bind in a systematic way — that is, the first-best may be available for high or medium or low cost realizations, with second-best incentives for other cost shocks.

1.5 Hidden Information

In this section, I consider situations where the firm and the agent are asymmetrically informed about each period's costs or task difficulty, but the agent's input is relatively easy for the firm to observe or infer. To model this, assume that the agent privately observes the cost parameter θ_t while his action choice is perfectly observed by the firm. A stationary contract (which is optimal) specifies a fixed payment w, and a menu of effort-bonus pairs $(e(\theta_i), b(\theta_i))$. The agent chooses from this menu in a revelatory way (although, unlike in complete contracting models, he is free to choose an effort level not on the menu). I show that the relational contracting problem can be analyzed using well-known properties of incentive compatibility, and characterize optimal agreements. Interestingly, optimal contracts distort the effort choices of *all* cost types, a result that overturns one of the most robust findings of static mechanism design, the idea that there should be "no distortion at the top." Moreover, relational contracts tend to involve less sorting of cost types (more pooling) than occurs in complete contracting settings.

Consider momentarily what the parties would do were e verifiable. Since there is symmetric information at the time of contracting, the parties can write a court-enforced contract to implement $e^{FB}(\theta)$, using the fixed payment to extract the expected information rents from the agent *ex ante* and achieve an arbitrary redistribution of the surplus (at least in expectation). Such a first-best screening contract will be infeasible under relational contracting if it calls for contingent payments that are not self-enforcing. In a relational contract, the parties cannot

[19]This idea, that relational contracting may fail if spot contracting (the outside option) is reasonably efficient, has been emphasized by Baker, Gibbons and Murphy (1994).

commit to make contingent payments — and so incentives are constrained by the need to keep reneging temptations small relative to future gains from the relationship. Thus, while relational contracts will share obvious features with static complete contracts (e.g. $e(\theta)$ will be decreasing, information rents are needed to induce revelation), they involve a different set of distortions than are present in more familiar static contracting environments (where distortions generally arise from distributional conflict and asymmetric information at the time of contracting).

To proceed, observe that if a contract can be found to implement a stationary effort profile $e(\theta) = \{e(\theta_i)\}_{i=1}^{m}$, the joint surplus is

$$S = \frac{1}{1-\delta}\left[\sum_{i=1}^{m} \left[E[y|e(\theta_i)] - c\left(e\left(\theta_i\right), \theta_i\right)\right] p(\theta_i)\right]. \tag{1.9}$$

Theorem 1.5 *There exists a stationary contract $(s, \sigma, \varsigma) \in \Omega^*$ with effort levels $e(\theta)$ if and only if $e(\theta)$ is decreasing and*

$$\delta\left[S - \overline{S}\right] \geq \sum_{i=1}^{m-1} \left[c(e(\theta_i), \theta_i) - c(e(\theta_{i+1}), \theta_i)\right] + c(e(\theta_m), \theta_m). \tag{IC}$$

The proof of the theorem is in the appendix: self-enforcement means that the parties face an endogenous "limited transfer" constraint, $\delta(S - \overline{S}) \geq \max b - \min b$; the derivation of (IC) follows using mechanism design arguments.[20] Note that the first term on the right hand side of (IC) is the information rents needed to induce truthful revelation; the second term is the "base" transfer $c(e(\theta_m), \theta_m)$ needed to keep the "worst" cost type from choosing zero effort (which may not be on the menu). These terms are analogous to the two terms that appear in complete contracting models — the information rents and the surplus that accrues to the worst type.

Given Theorem 1.5, optimal relational contracts with hidden information can be characterized by solving a static optimization program. The contracting problem is $\max_{e_1,\ldots,e_m} S$ subject to (IC) and $e(\theta)$ monotone decreasing, i.e.

$$\max_{e_1,\ldots,e_m} S$$

$$\text{s.t.} \quad \delta(S - \overline{S}) - c(e_1, \theta_1) - \sum_{i=2}^{m} \left[c(e_i, \theta_i) - c(e_i, \theta_{i-1})\right] \geq 0 \tag{IC}$$

$$\overline{e} - e_1 \geq 0,\ e_1 - e_2 \geq 0, \ldots, e_{m-1} - e_m \geq 0,\ e_m \geq 0. \tag{$M_0 : M_m$}$$

An immediate observation is that $e_i^* \leq e_i^{FB}$ for all i. This follows because if some proposed

[20] Somewhat similar effects could be found in a static model by making the *ad hoc* assumption that $\underline{b} < b(\theta) < \overline{b}$, for some $\underline{b}, \overline{b}$. Note that this is not the same as requiring limited liability on the part of the agent $(0 < w + b(\theta))$.

solution has $e_i > e_i^{FB}$, lowering e_i will raise S and relax (IC). The only potential problem is if $e_i - e_{i+1}$ binds, in which case $e_i = e_{i+1} > e_i^{FB} \geq e_{i+1}^{FB}$. But then by a similar argument, we must have $e_i = e_{i+1} = ... = e_m$, and lowering all of these effort levels gives greater surplus.

I now use the optimal contracting program to prove an important result. When the first-best is not available, an optimal contract will generally involve a uniform departure from the first-best. This is because (IC) constrains the overall level of incentive provision as well as the total information rents. The result is in stark contrast to the static contracting notion that only the choices of the bad types should be distorted. To gain some intuition for the difference, consider inducing a higher level of effort for cost type θ_i. In a complete contracting environment where the principal is concerned about distribution, her cost of doing so is proportional to $\Pr(\theta < \theta_i)$, meaning that for the "best" cost type, there is no cost. Here, there *is* a shadow cost to raising e_1 since it tightens (IC) — thus lowering the incentives that can be provided to other types.

The formal proof is straightforward. Let $e^* = (e_1^*, ..., e_m^*)$ solve the optimal contracting program. *Constraint qualification* holds if the gradient matrix of the binding constraints has full rank at the optimum schedule e^*.[21] Assuming constraint qualification, an interior e^* satisfies the Kuhn-Tucker conditions,

$$\begin{aligned}
\frac{\partial S}{\partial e_1}(e^*) &= \frac{\lambda}{1+\lambda\delta}\left[\frac{\partial c}{\partial e}(e_1^*, \theta_1)\right] + \frac{\mu_2}{1+\lambda\delta}, \\
\frac{\partial S}{\partial e_i}(e^*) &= \frac{\lambda}{1+\lambda\delta}\left[\frac{\partial c}{\partial e}(e_i^*, \theta_i) - \frac{\partial c}{\partial e}(e_i^*, \theta_{i-1})\right] + \frac{\mu_{i+1} - \mu_i}{1+\lambda\delta} \qquad i = 2, ..., m,
\end{aligned} \tag{1.10}$$

where $\lambda \geq 0$ is the Lagrange multiplier on the constraint (IC), and $\mu_i \geq 0$ is the multiplier on the monotonicity constraint (M_i). If $\lambda = 0$, the incentive constraint does not bind and $e_i^* = e_i^{FB}$ for all i — the optimal contract implements first-best effort. If $\lambda > 0$, then for every θ_i, the right hand side of (1.10) will be strictly positive. Since S is concave in e_i for all i, it must be the case that $e_i^* < e_i^{FB}$ for all i. Of course, it is easy to see that when δ is large and the outside options not too attractive, the (IC) constraint will not bind and the first-best will be attainable.

Proposition 1.2 *If $e^*(\theta)$ solves the optimal contracting program and satisfies constraint qualification, then either $e^* = e^{FB}$, or $e^*(\theta) < e^{FB}(\theta)$ for all θ.*

The fact that private cost shocks lead to action distortions for all types contrasts not only with the static hidden information paradigm, but with how relational contracts adapt to commonly observed cost shocks. As noted above, when θ is commonly observed, the self-enforcement constraint can conceivably bind for any subset of cost levels. So it might be possible to have high and low cost types choosing first-best effort, but effort distortion for mid-range costs.

[21] A discussion of constraint qualification can be found in virtually any text covering optimization. Note that from above, the constraint M_0 can be ignored.

Solving for the second-best effort schedule may be fairly involved. A standard simplification in the static mechanism design literature is to ignore the monotonicity constraints and check that the solution to this "relaxed program" is monotone. When it is not, the optimal contract involves "bunching" of cost types (i.e. limited screening). The next proposition identifies conditions under which solving only the "relaxed program" is a valid approach for the relational contracting model.

Proposition 1.3 *Define* $c(e, \theta_0) \equiv 0$. *Suppose that* $c(e, \theta_i) - c(e, \theta_{i-1})$ *is convex in* e *and supermodular in* (e, θ_i), *that* $p(\theta)$ *is decreasing in* θ, *and that there is some decreasing* $e(\theta)$ *that strictly satisfies* (IC), $e(\theta) \geq 0$. *Then the monotonicity constraints can be ignored. That is,* $e^*(\theta)$ *solves* $\max_{e(\theta)} S$ *s.t.* (IC), $e(\theta) \geq 0$. *Moreover, the Kuhn-Tucker conditions are necessary and sufficient for a solution.*

The proof is in the appendix. Interestingly, the assumptions needed to ensure monotonicity in the relational contracting program are stronger than in complete contracting models.[22] The difference arises in the distributional requirements on cost types — log-concavity will ensure separating contracts in most static models, but here concavity is required — and is easy to explain. As above, consider inducing a higher level of effort for cost type θ_i. The benefit (direct and through the incentive constraint) of doing so is proportional to the probability that θ_i will be realized, $p(\theta_i)$, but the (shadow) cost — the information rent term in (IC) — is independent of $p(\cdot)$. This means that when $p(\theta_i)$ is low, the parties would like to bias $e(\theta_i)$ down. This effect is less extreme in complete contracting models, where the information cost is proportional to $\Pr(\theta < \theta_i)$.

1.6 Moral Hazard with Common Monitoring

I now turn to a second important setting: situations of pure moral hazard where the agent's action choice is private and relational contracts must base payments on output, which is a noisy performance measure. To focus on moral hazard, assume that the distribution of cost parameters is degenerate ($c(e, \theta) = c(e)$), and that e is chosen privately. This section maintains the assumption of common monitoring, so that output is mutually observed and stationary contracts are optimal. A stationary contract specifies a fixed wage w, and bonuses $b_1, ..., b_n$ to be paid contingent on the output being $y_1, ..., y_n$.[23]

[22] For the model presented here, standard mechanism design techniques (see e.g. Fudenberg and Tirole (1991)) would require $v_{\theta ee}, v_{\theta\theta e} \geq 0$ and that $P(\theta)$ be log-concave (p/P decreasing).

[23] If both moral hazard *and* hidden information were present, the parties would want the agent to announce his cost type (θ), and relational contracts would make the payment plan contingent on this announcement (e.g. $b_1, ..., b_n$ would vary with the announced θ).

If output y were verifiable, a complete contingent contract could achieve the first-best given the risk-neutrality of the agent. Optimal relational contracts are typically second-best because promised bonuses are limited by the requirement of self-enforcement. My main finding is that, under standard conditions, optimal second-best contracts take an appealing "one-step" form — the maximum enforceable bonus is paid after "good" outcomes while the minimum is paid after "bad" outcomes.

The future surplus from a stationary contract that implements e is:

$$S = \frac{1}{1-\delta}\left[\sum_{i=1}^{n} y_i f(y_i|e) - c(e)\right].$$

Theorem 1.6 *There exists a stationary contract $(s, \sigma, \varsigma) \in \Omega^*$ with effort e if and only if there exist constants $b_1, ..., b_n$ such that for all $\hat{e} \neq e$*

$$\sum_{i=1}^{n} b_i \left[f(y_i|e) - f(y_i|\hat{e})\right] \geq c(e) - c(\hat{e}), \tag{IC_1}$$

and

$$\delta\left[S - \overline{S}\right] \geq \max_i b_i - \min_i b_i. \tag{IC_2}$$

The first condition (IC_1) is the agent's static incentive compatibility condition for effort choice. The second condition (IC_2) is the dynamic self-enforcement constraint — it says the range of transfers must be small relative to the gains generated by the relationship above and beyond the parties' outside options.

Proof. ($\Rightarrow$) Let b_i be the bonus paid when the output realization is y_i, for $i = 1, ..., n$. Rather than pay b_i and continue the relationship, the firm can pay nothing and fire the agent. For this deviation to be unprofitable,

$$-b_i + \delta\Pi \geq \delta\overline{\Pi} - \delta d, \tag{1.11}$$

where d is the severance pay if no bonus is paid. Similarly, the agent could refuse the bonus and quit. For him to cooperate requires

$$b_i + \delta U \geq \delta\overline{U} + \delta d. \tag{1.12}$$

Since each of these constraints must hold for every i, together they imply

$$\delta(S - \overline{S}) \geq \max_i b_i - \min_i b_i.$$

Moreover, for the agent to choose e, it is clear that (IC_1) must hold.

($\Leftarrow$) Let $\hat{b}_i = b_i - \min_i b_i$, and $\hat{w} = (1-\delta)\overline{U} + c(e) - E[\hat{b}|e]$, where $E[\hat{b}|e] = \sum_i \hat{b}_i f(y_i|e)$;

and let $\hat{d} = 0$. Then $U = \overline{U}$, $\hat{b}_i \geq 0$, and by (IC_1) the agent will always choose e. By definition $\Pi - \overline{\Pi} = S - \overline{S}$, so using (IC_2) and $\hat{d} = 0$, one can see that $-\hat{b}_i + \delta\Pi \geq \delta\overline{\Pi}$ for every i, hence the firm will always pay the bonus and choose to re-hire. *Q.E.D.*

An immediate consequence of Theorem 1.6 is that when

$$\frac{\delta}{1-\delta}\left[E[y|e^{FB}] - c(e^{FB}) - (1-\delta)(\overline{u}+\overline{\pi})\right] \geq y_n - y_1, \tag{1.13}$$

there is a feasible contract that sets $b_i = y_i$ for every i and implements the first-best effort in every period. So clearly the first-best is available for large enough discount factors when the outside options are not too attractive. On the other hand, when δ is small and the outside options relatively appealing, *no* relational contract will be feasible.

More generally, one wants to know the structure of optimal incentive plans for arbitrary discount factors. A remarkably strong characterization can be achieved provided that the static incentive constraints on effort choice, (IC_1), can be replaced by a single first-order condition.[24] Just as in the static principal-agent problem (Grossman and Hart, 1983; Rogerson, 1985), such a switch is possible provided that marginal returns to effort are stochastically decreasing. Let $F(y_i|e) = f(y_1|e) + ... + f(y_i|e)$ be the cumulative distribution function of output given effort.

Definition 1.3 *F satisfies the Convexity of the Distribution Function Condition (CDFC) if for any e, e', e'' and $\lambda \in [0,1]$ such that $c(e) = \lambda c(e') + (1-\lambda)c(e'')$, we have $F(y|e) \leq \lambda F(y|e') + (1-\lambda)F(y|e'')$ for all $y \in Y$.*

The CDFC says that the marginal returns to effort measured in cost units are decreasing in the sense of stochastic dominance. That is, each increase in effort shifts the distribution of outcomes up by first order stochastic dominance, but the shift becomes smaller and smaller as the cost of effort increases.

When the agent's first-order condition for effort choice replaces (IC_1), second-best informal incentives are "one-step". Below a certain output cut-off, the bonus payment is "low", above the cut-off the payment is "high". Moreover, an optimal contract never distorts effort above the first-best. While proving this requires some work, the intuition is not hard. The basic idea is that it is always possible to provide lower-powered incentives without violating (IC_2) by "flattening" the bonus plan. This means that an optimal contract will never implement too much effort, and in a second-best situation, the parties always want to increase the incentives. Doing this means providing the maximum possible reward for "good" outcomes (meaning the

[24]Such a switch is not necessarily valid. The problem is that at the solution (b, e) to the "relaxed program" using the first order condition, the agent's optimization problem may not be concave. Without concavity, there is no reason to believe that (b, e) also solves the "unrelaxed program" that imposes the more restrictive constraint (IC_1). See, e.g., Grossman and Hart (1983). Concavity is assured under MLRP and CDFC.

likelihood ratio f_e/f is positive), and the minimum possible reward for "bad" outcomes (f_e/f is negative). The fact that bonuses are monotone — and hence one-step — follows from the earlier (MLRP) assumption that the likelihood ratios are ordered (f_e/f is increasing in y).

Proposition 1.4 *Assume CDFC holds. There is an optimal stationary contract that either implements first-best effort e^{FB} with monotone bonuses, or implements $e^* < e^{FB}$ with a "one-step" bonus plan: that is, there is some k such that $b_1 = b_2 = ... = b_k < b_{k+1} = ... = b_n$, where $\delta(S - \overline{S}) = b_n - b_1$.*

Proof. As a first step, observe that for a given monotone bonus plan $b_1 \leq b_2 \leq ... \leq b_n$, the agent's optimization problem

$$\max_e \sum_i b_i f(y_i|e) - c(e)$$

is continuous and strictly concave. Assuming an interior solution, the optimal policy can be characterized by its first-order conditions,

$$\sum_i b_i f_e(y_i|e^*) = c'(e^*). \qquad (FOC)$$

A bonus plan $b_1, ..., b_n$ feasibly implements e if it satisfies both (IC_1) and (IC_2). Suppose some monotone bonus plan feasibly implements $e > e^{FB}$. I claim that there is a monotone bonus plan that feasibly implements e^{FB}. To see this, note that continuously "flattening" the bonus schedule (raising any $b_i = b_1$ by ε, and lowering any $b_i = b_n$ by ε) continuously lowers the left hand side of FOC. So the resulting maximizer $e(\varepsilon)$ is continuously decreasing in ε. Since $e((b_n - b_1)/2) = 0$ (this bonus plan has b_i constant), there is some ε such that $e(\varepsilon) = e^{FB}$. And the bonus plan corresponding to this ε is both monotone and feasible (since $b_n - b_1$ has been decreased by 2ε and the left hand side of (IC_2) has increased).

Next, I show that a non-monotone bonus plan will never be strictly optimal. Suppose some stationary contract implements $e \neq e^{FB}$ with bonuses b_i that are not monotone. Then there is some j with $b_j > b_j + 1$. Define a new bonus plan $\hat{b}_i$, with $\hat{b}_i = b_i$ for all $i \neq j, j+1$, and $\hat{b}_j = b_{j+1}$, $\hat{b}_{j+1} = b_j$. The new bonus plan reverses the bonuses for $j, j+1$. Introduce the dummy parameter τ, and let $b_i(\tau)$ equal b_i for $\tau = 0$, and $\hat{b}_i$ for $\tau = 1$. Then the agent's problem

$$\sum_i b_i(\tau) f(y_i|e) - c(e)$$

is bivariate single crossing in τ, e (using MLRP), so $e(\tau = 1) \geq e(\tau = 0)$ by Milgrom and Shannon (1994). Iterating this process (that is, seeing if $b(\tau)$ is monotone and if it is not, switching two non-monotone bonuses to obtain $b(\tau + 1)$), generates $b_i(\tau)$ for $\tau = 0, 1, .., \bar{\tau}$, where $b_i(\bar{\tau})$ is increasing in i (since there are only a finite number of outcomes, the bonus plan eventually becomes monotone). Since $\max_i b_i(\tau) - \min_i b_i(\tau)$ is constant in τ, all plans are feasible. And by Milgrom and Shannon (1994), the highest selection of $e(\tau)$ is monotone nondecreasing in τ. Then either $e(\bar{\tau}) \geq e^{FB}$ in which case there is a feasible monotone plan that implements e^{FB}, or $e(\bar{\tau}) < e^{FB}$. But in this case, $e(\bar{\tau}) \geq e(0)$ and $S(e(\bar{\tau})) \geq S(e(0))$, so the monotone plan $b(\bar{\tau})$ does at least as well as the non-monotone plan.

Combining the fact that implementing $e > e^{FB}$ with a monotone plan cannot be optimal, and that a non-monotone plan can always be replaced with a monotone plan that generates at least as much surplus, one sees that in searching for a optimal plan, it is possible to restrict attention to monotone bonuses that implement $e \leq e^{FB}$. I now show that if $e^* < e^{FB}$, then "one step" bonuses are optimal. Take an optimal contract that implements e^* with monotone bonuses. Because the bonuses are monotone, e^* solves (FOC). By MLRP, f_e is single crossing in y. So $f_e(y_i|e) \leq 0$ for all $i \leq k$, and $f_e(y_i|e) > 0$ for all $i > k$, some k. If $b_i < b_n$ for some $i > k$, then by raising b_i by ε, the principal can induce slightly higher effort from the agent without upsetting (IC_2), which would contradict the optimality of e^*. So it must be that $b_i = b_n$ for all $i > k$. An analogous argument works for all i such that $f_e(y_i|e^*) < 0$. If $f_e(y_i|e^*) = 0$, we arbitrarily take $b_i = b_1$. So it is optimal to set $b_i = b_1$ for all $i \leq k$. *Q.E.D.*

Proposition 1.4 demonstrates an important difference between the design of relational incentive contracts and court-enforced incentive contracts. Even if the performance measure is very fine (e.g. n is large), an optimal relational contract will take a simple cut-off form. This contrasts with the familiar risk-averse agent complete contracting model, where the twin assumptions of CDFC and MLRP guarantee only that an optimal incentive plan will be monotonic, but will very likely be irregular and sensitive to small changes in the information structure.[25] While changes in the information structure will affect both the cut-off point and the size of rewards here, there is a sense in which optimal relational contracts are *robust* to informational changes in that incentives always *look* roughly the same. Moreover, under appropriate distributional assumptions, the one-step result generalizes to the case of multiple actions and multiple signals.[26]

As in the hidden information case above, similar incentive schemes could be obtained in a static model by *exogenously* limiting the range of contingent payments, e.g. assuming that $0 < b_i < \bar{b}$, for some $\bar{b}$. Given risk-neutrality and the above distributional assumptions, a one-step bonus plan will be optimal.[27] A related idea has been pursued by Innes (1990), who

[25] Hart and Holmstrom (1987) contains a good discussion of these issues.

[26] Further distributional assumptions are needed to use the first-order approach in multi-signal models (Sinclair-Desgagné, 1994). The one-step result also applies in some cases where a subset of signals are verifiable. For example, take $y \sim f(y|e)$ to be nonverifiable and $z \sim g(z|e, m)$ to be verifiable, and suppose y, z are independent conditional on e, m. Informal incentives will be based solely on y and will be one-step. Without conditional independence, informal incentives will be based on both y and z (but will generally take a high/low form).

[27] With *risk-aversion*, a limited transfer restriction gives rise to an optimal schedule that is zero for low outcomes, monotone increasing for intermediate outcomes and equal to $\bar{b}$ for high outcomes. A stationary schedule of this sort will arise in a relational contracting model if the firm faces a sequence of risk-averse agents.

imposes a limited liability constraint ($w + b_i > 0$) along with a restriction on the *slope* of the incentive scheme ($b'(y) \leq 1$) in order to obtain contracts that pay zero up to some critical level of output, and increase one-one with output above this level.

From a repeated game perspective, Proposition 1.4 states that optimal agreements must necessarily use only extreme rewards and punishments. Abreu, Pearce and Stacchetti (1990) have shown a general result to this end for games with finite action spaces and a continuous signal space. Their result does not require distributional assumptions along the lines of MLRP and CDFC. Here, with a continuous action space and finite signal space, the distributional assumptions are needed to establish the validity of a first-order approach.

A consequence of the one-step result is that when the first-best is not available, the optimal relational contract assuming CDFC can be found by solving a reasonably easy program:

$$\max_{b,e,k} \quad \sum y_i f(y_i|e) - c(e)$$
$$\text{s.t.} \quad bF_e(y_k|e) - c'(e) = 0, \qquad (IC_1/FOC)$$
$$\delta\left[S - \overline{S}\right] = b, \qquad (IC_2)$$

where y_k is the cut-off point between "bad" outcomes (which imply zero bonus) and "good" outcomes (which imply a maximal bonus b), and $F(y_k|e)$ is the probability of a "bad" outcome given effort e. This program does not apply when a first-best contract is feasible, which is the case when for some k,

$$\delta\left[S^{FB} - \overline{S}\right] F_e(y_k|e^{FB}) \geq c'(e^{FB}).$$

Without imposing CDFC, it is much harder to characterize optimal contracts.[28] It is possible to show, for example, that the bonus plan cannot be decreasing. One can also see that if e is implementable, then the set of bonus plans that satisfy (IC_1) and (IC_2) is compact and convex. So e can be implemented in extreme points. Nevertheless, without replacing (IC_1) with the first-order condition, this fact does not seem to be of much use, since it is difficult to describe the extreme points.

One straightforward result that does hold regardless of distributional assumptions is that contracting possibilities improve monotonically with information quality (this should not be surprising in light of existing results on complete contracting (e.g. Holmström, 1979) and repeated games (Kandori, 1992)).

Proposition 1.5 *Suppose e can be implemented when the output is given by $f(y|e)$, and that f is a Blackwell garbling of g (with signals $g(z|e)$). Then e can be implemented using g.*

[28] The problem is similar to that faced by Grossman and Hart (1983).

Proof. If $g \succ_{BW} f$, then there exists some Markov kernel $k(y|z)$ such that for all $e, i = 1, ..., n$:

$$f(y_i|e) = \sum_{j=1}^{J} k(y_i|z_j) g(z_j|e). \tag{1.14}$$

Suppose $b_1, ..., b_n$ implement e under f, i.e. satisfy $(IC_1), (IC_2)$, and let $B_j = \sum_i b_i k(y_i|z_j)$ be a proposed bonus plan to implement e under g. Then for all $\hat{e} \neq e$,

$$\begin{aligned}
\sum_i b_i \left[f(y_i|e) - f(y_i|\hat{e})\right] &= \sum_i b_i \sum_j k(y_i|z_j) \left[g(z_j|e) - g(z_j|\hat{e})\right] \\
&= \sum_j \left[\sum_i b_i k(y_i|z_j)\right] \left[g(z_j|e) - g(z_j|\hat{e})\right] \\
&= \sum_j B_j \left[g(z_j|e) - g(z_j|\hat{e})\right] \\
&\geq c(e) - c(\hat{e}).
\end{aligned}$$

Since every B_j is a convex combination of b_i's, we have $\max_j B_j \leq \max_i b_i$ and $\min_j B_j \geq \min_i b_i$, so $B_1, ..., B_J$ satisfy both (IC_1) and (IC_2). Note that if $k(\bullet|\bullet) > 0$, these inequalities are strict and the parties can obtain *strictly* higher surplus under g than under f. *Q.E.D.*

The implication of this result is that *(i)* parties can base contracts on sufficient statistics, and *(ii)* improved monitoring will improve contracting possibilities.

1.7 Private Monitoring (Subjective Measurement)

The previous sections have shown that when performance measures are commonly observed, equilibrium play is always constrained efficient. But in many ongoing relationships, *objective* performance measures are unavailable and the parties must rely instead on *subjective* assessments. Subjective performance monitoring poses an additional difficulty for contracting because transfers at the end of each period must be based on reports. Contracts therefore need to induce truthful reporting. This section characterizes the dynamic structure of optimal contracts.

Subjective monitoring is captured by assuming that the agent's action e is chosen privately and that output y is observed solely by the firm. It is convenient to assume that there is no variation in the cost parameters, i.e. $c(e, \theta) = c(e)$. In what follows, I make an important simplification by restricting attention to strategies that depend only on the mutual history and not on privately known history. Thus strategies are public and again self-enforcement requires strategies to be a perfect public equilibrium.[29]

[29] The approach is roughly analogous to Compte (1998) and Kandori and Matsushima (1998) who introduce a

Given the restriction to public strategies, I first show that while the stationary contracts studied above are completely ineffective, another simple form of contract is optimal. These "termination" contracts also exhibit a form of stationarity: in each period the relationship is active, the contract calls for the same level of effort and bonuses; however, the contract calls for probabilistic separation depending on the firm's report. The possibility of separation is needed to provide incentives for both the firm and the agent — balanced transfers are not sufficient. The second point of this section concerns renegotiation. If one takes seriously the possibility of renegotiation, inefficient separation may not be credible and cooperative agreements may be simply impossible.

1.7.1 Termination Contracts

A first key observation is that for the firm to truthfully reveal its performance assessment, its future profits (inclusive of today's transfers) must be the same following any outcome. But in order to provide incentives for the agent, there must be variation in the agent's future surplus (inclusive of today's transfers). It follows immediately that stationary contracts are completely ineffective.

Proposition 1.6 *Under private monitoring, no stationary contract can implement a positive level of effort e.*

Proof. Suppose the contrary holds. Let b_i be the bonus following announcement m_i (which corresponds to y_i). The firm must be willing to announce truthfully ($m_i = y_i$ for all i). So for any i, j,

$$-b_i + \delta\Pi \geq -b_j + \delta\Pi. \tag{1.15}$$

since this must hold also for j, i, it must hold with equality. But then $b_i = b_j$ for all i,j. For the agent to choose e rather than zero effort requires

$$\sum_i (b_i + \delta U)\left[f(y_i|e) - f(y_i|0)\right] = 0 \geq c(e),$$

which is impossible for any $e > 0$. *Q.E.D.*

round of communication (but no transfer payments) at the end of each stage game and then study perfect public equilibria (or variants thereof) in repeated games with private monitoring. Strictly speaking, "communication" is not needed in the present model because by simply having the firm make varying bonus payments the parties can create a sufficient mutual history. Note that limiting attention to public strategies (or 1-PPE) is restrictive, as discussed by the above authors (see also Abreu, Milgrom, and Pearce 1990). The parties can do better by making more use of their private information. For example, if the discount factor is sufficiently high, the parties will be able to approximate the first-best by having the firm announce performance only every T periods (where $T \to \infty$ as $\delta \to 1$) (Compte, 1998) — as compared to Proposition 10.

Keeping the firm indifferent between performance outcomes while providing incentives for the agent requires the use of a contract that can punish the agent without rewarding the firm (and likewise reward the agent without punishing the firm). While there are many conceivable contracts that could do this, I will show that it suffices to study *termination contracts*: contracts that call for the same effort, bonus plan and separation probabilities in every period the relationship is active, but mandate probabilistic separation depending on the previous period's outcome.

Definition 1.4 *An agreement $(s,\sigma,\varsigma) \in \Omega$ is a termination contract if $Q_{t^0}(h) = R_{t^0}(h) = 1$ with probability $\alpha(b_{t-1})$, and $Q_{t^0}(h) = R_{t^0}(h) = 0$ with probability $1-\alpha(b_{t-1})$, $E_{t^1}(h) = e$, $M_{t^2}(h) = y_t$, $B_{t^3}(h) = b(m_t)$, and $W_t(h) = w$ for any $h \in H^e(t^k|s,\sigma,c)$ in which the relationship has not yet ended.*

Termination contracts are particularly tractable because, at a formal level, they are equivalent to *stationary* contracts that have the parties *burn money*. To understand this point, consider a termination contract that generates surplus S, and calls for separation with probability $1-\alpha$ (giving surplus $\overline{S}$) following some outcome. From the point of view of incentive provision, there is no difference between this stipulation and one that calls for the parties to burn $(1-\alpha)(S-\overline{S})$ and then continue for sure generating S. Money-burning and termination can be seen as equivalent ways to dissipate rents. The proof of the next theorem exploits this fact.

Theorem 1.7 *If monitoring is private, and a relationship is feasible, there exists a termination contract $(s^*,\sigma^*,\varsigma^*) \in \Omega^*$ that is constrained efficient.*

Proof. The proof proceeds in two steps: first by showing that a stationary contract with money-burning can achieve the constrained efficient surplus, then by constructing an equivalent termination contract. In a stationary contract with money burning, the firm pays direct bonus $b(y)$ and publicly "burns" $\tau(y)$ following announced output y (where to keep the firm honest, $b(y)+\tau(y)$ is constant). Suppose $(s,\sigma,\varsigma) \in \Omega^*$ is constrained efficient and generates surplus S. If (s,σ,ς) implements e in the first period, then self-enforcement ensures

$$\begin{aligned}
e \in \arg\max_{\tilde{e}} \quad & \sum_i (b_i + \delta U_i) f(y_i|\tilde{e}) - c(\tilde{e}), \\
-b_i + \delta\Pi_i &= -b_j + \delta\Pi_j, \qquad \text{for all } i,j, \\
b_i + \delta U_i &\geq \delta\overline{U} + \delta d, \\
-b_i + \delta\Pi_i &\geq \delta\overline{\Pi} - \delta d.
\end{aligned}$$

Now define a new contract with informal payments

$$\hat{b}_i = b_i + \delta U_i - \delta\overline{U} - \delta d,$$

$$\hat{\tau}_i = \max_j [b_j + \delta U_j] - b_i - \delta U_i.$$

Also let $\hat{d} = 0$ and $\hat{w} = (1-\delta)\overline{U} - E[\hat{b}|e] + c(e)$. Consider a stationary agreement that calls for these payments and truthful reporting by the firm, and punishes any observable deviation with immediate separation. This agreement gives payoff $\overline{U}$ to the agent, and generates per-period surplus

$$\begin{aligned}
(1-\delta)\hat{S} &= E[y|e] - c(e) - E[\hat{\tau}_i|e] \\
&= E[y|e] - c(e) - b_k - \delta U_k + E[b_i + \delta U_i|e] \qquad \text{for some } k \\
&= E[y|e] - c(e) - \delta S_k + \delta E[S_i|e] \\
&= S - \delta S_k \geq (1-\delta)S.
\end{aligned}$$

The last inequality follows because S_k corresponds to some perfect public equilibrium and S is constrained efficient. Note that it holds with equality *only if* $S_k = S$, meaning that in *any* optimal agreement, there must be some positive probability of remaining on the Pareto frontier.

The proposed agreement is self-enforcing. The agent's effort choice is still e because his objective function is unchanged:

$$\sum_i \hat{b}_i f(y_i|\tilde{e}) - c(\tilde{e}) = \sum_i (b_i + \delta U_i) f(y_i|\tilde{e}) - c(\tilde{e}).$$

Moreover, $\hat{b}_i \geq 0$, and continuing employment pays $\overline{U}$, so the agent will accept the bonus and will not quit. The firm is willing to make a truthful announcement because

$$-\hat{b}_i - \hat{\tau}_i = \delta\overline{U} + \delta d - \max_j [b_j - \delta U_j]$$

does not depend on i. The firm will not withhold the bonus because

$$\begin{aligned}
-\hat{b}_i - \hat{\tau}_i + \delta(\hat{S} - \overline{U}) &= \delta\hat{S} - \max_j [b_j + \delta U_j] + \delta d \\
&= \delta\hat{S} - b_k - \delta U_k + \delta d \qquad \text{for some } k \\
&\geq -b_k + \delta\Pi_k + \delta d \geq \delta\overline{\Pi}.
\end{aligned}$$

This stationary money-burning contract is equivalent to the following termination contract. Define $b_i^* = \hat{b}_i$, $w^* = \hat{w}$, $d^* = \hat{d} = 0$, and let the probability of continuing the relationship after bonus b_i^* be

$$\alpha_i^* = 1 - \frac{\hat{\tau}_i}{\delta\left[\hat{S} - \overline{S}\right]}.$$

The contract calls for the agent to choose e in every period the relationship is active, for truthful reporting and for any observable deviation to be punished by immediate termination. Then,

$$S^* = E[y|e] - c(e) + \delta S^* - (1 - E[\alpha_i^*|e])\delta(S^* - \overline{S}).$$

It follows that

$$(1-\delta)(S^* - \hat{S}) = -E[\hat{\tau}_i|e]\frac{S^* - \overline{S}}{\hat{S} - \overline{S}} + E[\hat{\tau}_i|e] = E[\hat{\tau}_i|e]\frac{\hat{S} - S^*}{\hat{S} - \overline{S}}.$$

Since $E[\hat{\tau}_i|e]$ is strictly positive, this means that $S^* = \hat{S}$. Finally, the proposed termination contract is self-enforcing. Note that $U_i^* = \overline{U}$ in every eventuality, so the agent's optimization is unchanged as are his quit decisions. The firm will announce output truthfully and pay the required bonus because

$$-b_i^* + \alpha_i^*\delta\left[S - \overline{U}\right] + (1 - \alpha_i^*)\delta\overline{\Pi} = -\hat{b}_i - \hat{\tau}_i + \delta\left[S - \overline{U}\right]$$

does not depend on i, and is greater than $\overline{\Pi}$. *Q.E.D.*

1.7.2 Renegotiation

With common performance monitoring, there are optimal contracts that are extremely robust to renegotiation. But when monitoring is private, optimal contracts *must* involve inefficient outcomes along the equilibrium path. Imposing a renegotiation criterion causes cooperative agreements to unravel completely. The following renegotiation restriction seems appropriate to the relational contracting setting. Let $\Gamma \subseteq \Omega^*$. An agreement $(s, \sigma, \varsigma) \in \Gamma$ is renegotiation-proof if after any t period history, the continuation agreement is in Γ, and every agreement in Γ generates the same surplus.[30] See the Appendix for the next proof.

Proposition 1.7 *Under private monitoring, no self-enforcing agreement other than immediate separation is renegotiation-proof.*

This result raises several interesting points. First, it emphasizes the critical role of objective measures in sustaining relationships. Second, looking at the current model as a two-sided moral

[30]This requirement is motivated by, but more stringent than, Farrell and Maskin's (1989) weak renegotiation-proofness, which requires that no two elements of Γ be strictly Pareto ranked. WRP seems far too weak here because if there are two agreements in Γ that are weakly Pareto ranked ($U' \geq U$, $\Pi' \geq \Pi$, one inequality strict), a simple transfer could strictly rank them. In fact, the WRP restriction has no bite at all in this context. This is because any division of the constrained efficient surplus can be achieved in a WRP manner using ex ante transfers and a constrained effiicient continuation agreement that specifies $U = \overline{U}$ after any history. Instead, the definition is closer to the notion of "sequential efficiency" found in the complete contracting literature (e.g. Fudenberg, Holmström and Milgrom, 1990).

hazard problem, it points to the importance of having an inactive monitor to soak up residual surplus (as in the static team problem). Third, it offers a potential role for arbitration or mediation in relational contracting — to resolve problems of "disagreement" over outcome measures. These points, however, are put aside to concentrate on optimal self-enforcing agreements that are not renegotiation-proof.

1.8 Moral Hazard with Private Monitoring

This section characterizes optimal moral hazard contracts under private monitoring. Optimal contracts under private monitoring are similar to the incentive contracts discussed above under common monitoring — for instance, under CDFC optimal incentives are one-step — but there are some crucial differences. Most importantly, equilibrium play does not remain on the constrained efficient frontier, and even optimally planned relationships may fail after bad performance outcomes. Moreover, the maximum attainable surplus is bounded away from the first-best, even as δ goes to one.

Consider a termination contract that implements e in every period of the relationship, and continues with probability α_i following bonus b_i. The surplus generated, S, can be decomposed as follows:

$$
\begin{aligned}
S &= \sum_i y_i f(y_i|e) - c(e) + \delta\overline{S} + \delta\left[S - \overline{S}\right]\sum_i \alpha_i f(y_i|e) \\
&= \frac{E\left[y|e\right] - c(e) + \delta\left(1 - E\left[\alpha|e\right]\right)\overline{S}}{1 - \delta E\left[\alpha|e\right]} \qquad (1.16)
\end{aligned}
$$

Theorem 1.8 *An effort level e can be implemented with a termination contract if and only if there exist constants $\alpha_i \in [0,1]$, $i = 1, ..., n$ such that for all $\hat{e} \neq e$,*

$$
\delta\left[S - \overline{S}\right]\sum_{i=1}^{n}\alpha_i\left[f(y_i|e) - f(y_i|\hat{e})\right] \geq c(e) - c(\hat{e}) \qquad (IC)
$$

An optimal contract solves $\max_{e,\alpha_1,...,\alpha_n} S$ subject to $\alpha_i \in [0,1]$ and (IC). As above, under CDFC, the incentive constraints can be replaced with a local incentive condition:

$$
\delta\left[S - \overline{S}\right]\sum_{i=1}^{n}\alpha_i f_e(y_i|e) - c'(e) = 0 \qquad (1.17)
$$

In this case, "nearly" one-step contracts are optimal (nearly because the bonus may be intermediate at the cut-off point). Above some output cut-off y_k, the contract calls for a high bonus payment and continuation with probability one. Output below the cut-off leads to a low bonus payment and separation with probability one. The logic is similar to the common monitoring case — when the likelihood ratio f_e/f is sufficiently high (greater than some strictly positive threshold), the contract gives maximal rewards to the agent; a low likelihood ratio implies minimal rewards for the agent, and termination is used to keep the firm honest. The proof of the following Proposition, and those of the next two, are in the Appendix.

Proposition 1.8 *Under CDFC, the optimal termination contract implements effort $e^* \leq e^{FB}$ with a "nearly one-step" bonus plan: there is some k such that bonuses are $b_1 = ... = b_{k-1} < b_{k+1} = ... = b_n$ and continuation probabilities are $\alpha_1 = ... = \alpha_{k-1} = 0$ and $\alpha_{k+1} = ... = \alpha_n = 1$. When output is y_k, the bonus is $b_k \in [b_1, b_n]$ and the relationship continues with probability $\alpha_k \in [0,1]$.*

A second positive result is that, just as above, increasing the quality of monitoring expands the range of feasible contracts.

Proposition 1.9 *Suppose e can be implemented with a termination contract when the output is given by $f(y|e)$, and that f is a Blackwell garbling of g (with signals $g(z|e)$). Then e can be implemented with a termination contract using g.*

I emphasized above that termination contracts are equivalent to stationary contracts in which the parties burn money. An important consequence of this is that because providing incentives to the agent necessitates some variation in the bonuses, and the firm must pay out an identical amount each period (to induce truth-telling), an optimal contract must require some money to be burned with positive probability in every period. Thus the expected surplus is bounded away from the first-best in every period meaning that contracts remain second-best even as δ goes to one.[31]

Proposition 1.10 *Suppose $s(\delta)$ is the per-period constrained efficient level of surplus for $\delta \in (0,1)$, and $s^{FB} = \max_e E[y - c|e]$. Then $\sup_{\delta \in (0,1)} s(\delta) < s^{FB}$.*

As was briefly suggested above, the fact that optimal contracts always involve throwing away money suggests that the firm and the agent might recruit a third-party who would pay ex ante for the informal privilege of collecting the money that must be burned. Or, other agents might serve to "break the budget balance," perhaps through a relative performance system. As always, the concern with such a plan is the potential for "secret" side agreements. A further possibility is that by finding additional incentive "instruments," the parties could generate mutual information and improve their contracting ability. For example, if the firm must take some observable ex post decision, the returns to which depend on y_t, the parties can use this decision in their relational contract to screen the firm for its private information.

[31]Note that the following result relies on the restriction to public strategies (see note 29).

On the other hand, in many relationships "money-burning" may be an apt description of the "settling up" that actually occurs, for instance through the use of inefficient perquisites or in-kind transfers (rather than cash) as rewards. So long as these inefficient transfers can be fine-tuned, the theory suggests that they may in fact be optimal.

1.9 Two Solved Examples

In this section, I work through two examples: one with moral hazard and one with hidden information. The examples make use of and illustrate the results developed above, and highlight how the degree of private information affects optimal contract design.

1.9.1 A Moral Hazard Example

Consider the following two-action, two-outcome version of the model. Let $y_t \in \{0, y\}$ be the output realization and $e_t \in \{0, 1\}$ the effort input, where effort costs are $c(0) = 0$, $c(1) = c$; and let q_e be the probability that $y_t = y$ given effort e, with $0 < q_0 < q_1$. Assume that

$$(q_1 - q_0)y > c,$$

so that implementing e is the first-best outcome. Zero effort is always feasible and is taken to be the outside option:

$$\overline{S} = \frac{q_0 y}{1 - \delta}.$$

Perfect Observability. From Theorem 4, implementing $e_t = e$ is possible if $\delta \left[S - \overline{S}\right] \geq c$, where $S = [q_1 y - c] / (1 - \delta)$, i.e. if

$$\frac{\delta}{1 - \delta} \left[(q_1 - q_0)y - c\right] \geq c.$$

Common Monitoring. Let the firm pay bonus b after output $y_t = y$, and no bonus after output $y_t = 0$. It follows from Theorem 6 that implementing $e_t = e$ requires

$$\frac{\delta}{1 - \delta} \left[(q_1 - q_0)y - c\right] \geq \frac{c}{q_1 - q_0}. \tag{1.18}$$

Private Monitoring. To implement $e_t = e$, it is optimal to set $\alpha(y) = 1$, and $\alpha(0) = \alpha$. The incentive compatibility condition from Theorem 8 is

$$\delta \left[S - \overline{S}\right] (q_1 - q_0)(1 - \alpha) \geq c,$$

where

$$S = \frac{q_1 y - c}{1 - \delta\left[q_1 + (1-q_1)\alpha\right]} + \frac{\delta(1-q_1)(1-\alpha)}{1 - \delta\left[q_1 + (1-q_1)\alpha\right]}\overline{S}.$$

Combining these two conditions, one obtains

$$\frac{\delta(1-\alpha)}{1 - \delta\left[q_1 + (1-q_1)\alpha\right]}\left[(q_1 - q_0)y - c\right] \geq \frac{c}{q_1 - q_0}. \tag{1.19}$$

The left hand side of this equation is decreasing in α — that is, the strongest effort incentives are provided by the threat of automatic termination following a bad outcome. An optimal relational contract will choose the largest α that satisfies (1.19). For *any* such α to exist requires

$$\frac{\delta}{1 - \delta q_1}\left[(q_1 - q_0)y - c\right] \geq \frac{c}{q_1 - q_0}. \tag{1.20}$$

This condition is more stringent than the common monitoring constraint above. So it is harder to implement a single period of high effort under private monitoring than to achieve the first-best under common monitoring. Of course, even if δ satisfies (1.20), the optimal private monitoring contract will involve inefficient reversion to zero effort after some time. Solving for the optimal contract parameter:

$$\alpha(\delta) = \frac{(q_1 - q_0)\left[(q_1 - q_0)y - c\right] - c(\frac{1}{\delta} - q_1)}{(q_1 - q_0)\left[(q_1 - q_0)y - c\right] - c(1 - q_1)}.$$

For any $\delta < 1$, $\alpha(\delta) < 1$; the optimal contract is strictly worse than the first-best.

1.9.2 A Hidden Information Example

I now consider another variant of the model with cost shocks and observable effort. Let $c(e,\theta) = \frac{1}{2}\theta e^2$, where $\theta \in \Theta = \{\theta_1, ..., \theta_m\}$ is, as above, an i.i.d. cost parameter with pdf $p(\theta)$. The agent chooses e from $[0, \bar{e}]$, which leads to expected output of $E[y|e] = Re$. Then,

$$e^{FB}(\theta) = \arg\max_e Re - \frac{1}{2}\theta e^2 = \frac{R}{\theta},$$

$c(e^{FB}(\theta), \theta) = \frac{1}{2}\frac{R^2}{\theta}$, and $(1-\delta)S^{FB} = E_\theta\left[\frac{1}{2}\frac{R^2}{\theta}\right]$. Note that $c(e^{FB}(\theta), \theta)$ is decreasing in θ — higher cost parameters lead to lower absolute costs at the first-best level of effort.

Assume that the outside option is spot contracting with zero effort, so $\overline{S} = 0$.

Commonly Observed Cost Shocks. Suppose first that θ is commonly observed. From Theorem 4, the first-best will be available whenever

$$\frac{\delta}{1-\delta}E_\theta\left[\frac{1}{\theta}\right] \geq \frac{1}{\theta_1}. \tag{1.21}$$

When the first best is not available, the optimal schedule $e_1, ..., e_m$ solves

$$\max_{e_1,\ldots,e_m} \frac{\delta}{1-\delta}\sum_i \left[Re_i - \theta e_i^2/2\right] p(\theta_i)$$

$$\text{s.t.} \qquad \frac{\delta}{1-\delta}\sum_i \left[Re_i - \theta e_i^2/2\right] p(\theta_i) \geq c(e_j, \theta_j) \qquad \text{for all } j.$$

Letting λ_j be the Lagrange multiplier attached to the incentive constraint on e_j, we obtain

$$e_i = \frac{R}{\theta_i}\left[\frac{\frac{\delta}{1-\delta}(1+\sum_j \lambda_j)p(\theta_i)}{\frac{\delta}{1-\delta}(1+\sum_j \lambda_j)p(\theta_i) + \lambda_i}\right]. \tag{1.22}$$

The bracketed term can be thought of as the "distortion" away from the first-best (R/θ_i). Several points are immediate. First, if $e_i = e_i^{FB}$ for some cost type θ_i, then $e_k = e_k^{FB}$ for all $k \geq i$, i.e. if some cost type is at first-best effort, then all higher cost types must be at first-best effort. So in this sense, optimal contracts will involve relatively more distortion for "better" (i.e. low) cost types. Note also that if $e_i < e_i^{FB}$, then distortion will be lower the greater is the discount factor δ, and the greater the probability of cost type i being realized (because the returns to raising e_i are higher).

Privately Observed Cost Shocks. Now suppose θ_i is privately observed by the agent, so there is hidden information. From Theorem 5, the first-best will be available if

$$\frac{\delta}{1-\delta}E_\theta\left[\frac{1}{\theta}\right] \geq \frac{1}{\theta_1} + \sum_{i=2}^{m}\frac{1}{\theta_i}\left(1 - \frac{\theta_{i-1}}{\theta_i}\right). \tag{1.23}$$

Comparison with (1.21) reveals the additional difficulty of inducing first-best effort when cost information is private — not only must the range of transfers be large enough to compensate the agent for his maximal costs, it must be large enough to compensate for costs and provide information rents.

When the first-best is not available, one can use the contracting program in Section 5 to obtain the optimal second-best effort plan. Assume that $p(\theta)$ is decreasing, and that $\theta_i - \theta_{i-1}$ is increasing in i, where $\theta_0 \equiv 0$, so that the monotonicity constraints can be ignored. Then a second-best relational contract under hidden information sets

$$e_i = \frac{R}{\theta_i}\left[\frac{\frac{\delta}{1-\delta}(1+\lambda)p(\theta_i)}{\frac{\delta}{1-\delta}(1+\lambda)p(\theta_i) + \lambda(1-\frac{\theta_{i-1}}{\theta_i})}\right],$$

where λ is the Lagrange multiplier attached to the single incentive constraint. The bracketed term can again be thought of as the distortion from the first-best; and again, several points can be made. First, as was shown generally in Section 5, the effort choices of all types are distorted below the first-best. Second, observe that e_i will be less distorted the higher are $p(\theta_i)$ and δ, and the higher is θ_{i-1}/θ_i. So if cost type i has costs that are "close" to $i-1$, then e_i will be less distorted because fewer information rents are needed to separate the types. Finally, note that depending on the parameters, there can be relatively more effort distortion for either high or low cost types. If $p(\theta_i)$ and $\theta_i - \theta_{i-1}$ are constant in i, then the effort of low cost types is more distorted. But if $p(\theta_i)$ is decreasing and θ_{i-1}/θ_i is decreasing, then there is more distortion for high cost types. The intuition is that when high cost types are less likely, and it is relatively costly to provide incentives for them, it makes sense to distort their effort down.

1.10 Conclusion

A rough interpretation of optimal relational contracting is that at the end of each period, the parties "settle up" and if they decide not to separate, move forward optimally. The process of "settling up" is mutually understood *ex ante*, but not formally contracted upon. In fact, the parties have a great deal of flexibility in how they arrange the "settling up"; for convenience I have mostly studied bonus payments and penalties, but they could also settle up by adjusting their formal contract to shift future surplus back and forth. In all cases, the observed pattern is one of reciprocity.

This general process of long-term contracting seems to be a reasonable approximation of many relationships within and between firms. To that end, I should stress that the model admits many interpretations: a hand-in-glove supply relationship with specific investment or any supplier discretion, a firm contracting out for customized services, an internal division contracting with headquarters, a basic employment relationship, a board of directors contracting with top management, a regulated firm bribing a captured official. The general framework could also prove useful for extending existing models of long-term borrowing agreements and financial contracts.

There are several theoretical extensions that may be of interest for applications. First, the model I presented is stationary — shocks are i.i.d. and the technology and outside options are time-independent. In a non-stationary model, optimal contracts will still have the feature that full settling-up can take place at the end of each period, but the dynamic self-enforcement constraint will reflect shifting future surplus, meaning that the parties' ability to provide incentives will vary over time. A second extension is to allow for "multi-tasking" or multidimensional effort along with multiple performance measures. Again, the general principles are the same, but the static constraints must account for the larger choice spaces. A third extension is to allow for multiple parties, something that I explore in the next chapter.

Let me finish by mentioning one omission from the analysis. While the environments I have studied have parametric uncertainty, the model misses out on a fundamental aspect of contracting that concerns ongoing "governance," or how long-term transactional relationships evolve over time. This aspect of relational contracting is hard to over-emphasize (Macneil, 1974, 1978; Williamson, 1985; Kreps, 1990). The model fails to completely address it because despite the minimal role played by ex ante formal contracts, informal agreements are fully contingent. All bridges are crossed in advance, at the outset, before the relationship begins. This means that it is difficult to conceptualize how contractual arrangements will adjust to shocks that alter the environment in ways only vaguely anticipated. Making progress along these lines remains an important objective for future work.

1.11 Appendix

Proof of Theorem 1.1. I prove this for the public monitoring case; the private monitoring case is analogous. One needs to show that the equilibrium payoff set $\Sigma = \{S : \exists(s,\sigma,\varsigma) \in \Omega^*$ with surplus $S\}$ is compact. Note that Σ is bounded (by $\overline{S}$ and S^{FB}), and convex, and also contains its lower bound, $\overline{S} \in \Sigma$. So it suffices to show that if $\Sigma \neq \{\overline{S}\}$, it still contains a maximal element. To do this, I show that the set $\Sigma^s = \{S : \exists(s,\sigma,\varsigma) \in \Omega^*$ that is stationary and generates surplus $S\} \subseteq \Sigma$ is compact. This will imply that Σ^s has a largest element, and from the proof of Theorem 1.3, that there is no bigger element in Σ.

Consider a sequence $S^1, S^2, ...$ in Σ^s. Corresponding to each S^k is a stationary equilibrium $(s^k, \sigma^k, \varsigma^k)$. Each $(s^k, \sigma^k, \varsigma^k)$ specifies a wage w^k, a finite vector of effort levels e^k, and a vector of bonuses b^k, indexed by φ, where φ is the mutual information available to the firm and the agent at the end of the period. Note that if $e \notin \varphi$, then b^k is a finite vector. If $e \in \varphi$, then without loss of generality, we can reduce b^k to a finite vector by keeping $b^k(e^k(\theta), \cdot)$ constant, and letting $b^k(e,\cdot) = b^k(0,\cdot)$ for any $e \neq e(\theta)$ some θ. Now for each k,

$$e^k(\theta) \in \arg\max_e E[b^k - c|\theta, e], \qquad \forall \theta \in \Theta,$$

and

$$\begin{aligned} b^k(\varphi) + \delta U^k &\geq \delta \overline{U}, \\ -b^k(\varphi) + \delta \Pi^k &\geq \delta \overline{\Pi}, \end{aligned}$$

where the first inequality is the agent's effort incentive constraint, and the latter two are the agent's and the firm's incentive compatibility constraints for making bonus payments. I have assumed that severance payments are zero; the proof of Theorem 1.3 shows that this is without loss of generality. In the above agreements, $S^k \leq S^{FB}$, $U^k \geq \overline{U}$, and $\Pi^k \geq \overline{\Pi}$ after any history, so the incentive constraints imply that $|b^k(\varphi)| \leq \delta(S^{FB} - \overline{S})$ and $w \in [E[c(e^{FB}, \theta)] - \delta(S^{FB} - \overline{S}), E[y|e^{FB}] + \delta(S^{FB} - \overline{S})]$. And of course, $e \in [0, \overline{e}]$. Thus it possible to find a convergent subsequence w^l, e^l, b^l converging to w, e, b. I claim there is a stationary self-enforcing agreement corresponding to w, e, b. Punish any deviation at the bonus stage by immediate separation. By continuity,

$$e(\theta) \in \arg\max_e E[b - c|\theta, e] \qquad \forall \theta \in \Theta$$

and clearly the two bonus constraints hold for $b(\varphi)$, all φ. So the stationary agreement is self-enforcing and generates surplus $S = E_{y,\theta}[y - c|e(\theta)] = \lim_l E_{y,\theta}[y - c|e^l(\theta)] = \lim_l S^l$. Thus S^k has a convergent subsequence, and Σ^s is compact. *Q.E.D.*

Theorem 1.9 ***(Constraint Reduction)***: *Let $\{b(\theta_i) = b_i\}_{i=1}^m$ be given. For all $\theta_i, \hat{\theta} \in \Theta$,*

$$b_i - c(e(\theta_i), \theta_i) \geq b(\hat{\theta}) - c(e(\hat{\theta}), \theta_i), \tag{1.24}$$

if and only if

$$b_i - c(e(\theta_i), \theta_i) \geq b_{i+1} - c(e(\theta_{i+1}), \theta_i), \tag{1.25}$$

$$b_i - c(e(\theta_i), \theta_i) \geq b_{i-1} - c(e(\theta_{i-1}), \theta_i). \tag{1.26}$$

Moreover, if either condition holds, then $e(\theta_i)$ and $b(\theta_i)$ must be decreasing.

Proof. See, for example, Stole (1997, p. 70). *Q.E.D.*

Proof of Theorem 1.5. ($\Rightarrow$) Suppose $(s, \sigma, \varsigma) \in \Omega^*$ is stationary. For every effort level $e(\theta)$, the firm prefers to pay the specified bonus $b(\theta)$ rather than to pay nothing and fire the agent immediately. So for all θ,

$$-b(\theta) + \delta\Pi \geq \delta\overline{\Pi} - \delta d.$$

Similarly, the agent prefers to play the prescribed equilibrium rather than set effort to zero and quit immediately. So

$$-c(e(\theta), \theta) + b(\theta) + \delta U \geq \delta\overline{U} + \delta d.$$

Combining these two conditions, we see that

$$\delta(S - \overline{S}) \geq \max_\theta b(\theta) - \min_\theta \{b(\theta) - c(e(\theta), \theta)\}. \tag{1.27}$$

Let $\tilde{\theta} = \arg\min_\theta \{b(\theta) - c(e(\theta), \theta)\}$.

Furthermore, the agent prefers to play his prescribed effort levels rather than the effort level of some other type. That is, for every θ_i ,

$$b(\theta_i) - c(e(\theta_i), \theta_i) \geq b(\hat{\theta}) - c(e(\hat{\theta}), \theta_i). \tag{1.28}$$

By the constraint reduction theorem (above), this implies that $e(\theta)$ is decreasing, and also that $b(\theta)$ is decreasing. Thus $\max_\theta b(\theta) = b(\theta_1)$. And moreover,

$$b(\theta_i) - c(e(\theta_i), \theta_i) \geq b(\theta_m) - c(e(\theta_m), \theta_i) \geq b(\theta_m) - c(e(\theta_m), \theta_m),$$

so $\min_\theta \{b(\theta) - c(e(\theta), \theta)\} = b(\theta_m) - c(e(\theta_m), \theta_m)$. Combining (1.27) and (1.28),

$$\begin{aligned} \delta(S - \overline{S}) &\geq b(\theta_1) - b(\theta_m) + c(e(\theta_m), \theta_m) \\ &= \sum_{i=1}^{m-1} [b(\theta_i) - b(\theta_{i+1})] + c(e(\theta_m), \theta_m) \\ &\geq \sum_{i=1}^{m-1} [c(e(\theta_i), \theta_i) - c(e(\theta_{i+1}), \theta_i)] + c(e(\theta_m), \theta_m). \end{aligned}$$

($\Leftarrow$) Let the formal contract ς specify $d = 0$, $w = (1-\delta)\overline{U} - E_\theta [b(\theta) - c(e(\theta), \theta)]$, where $b(\theta)$ will be defined momentarily. Let the firm's candidate strategy be: if there has been no deviation, hire in every period and after observing effort $e(\theta_i)$, pay bonus $b(\theta)$,where $b(\theta_m) = c(e(\theta_m), \theta_m)$, and $b(\theta_i) = \sum_{j=i}^{m-1} [c(e(\theta_j), \theta_j) - c(e(\theta_{j+1}), \theta_j)] + c(e(\theta_m), \theta_m)$ for all $i < m$. Following any deviation (i.e. if the firm observes some $e \neq e(\theta)$ for some θ, or if the agent refuses a bonus payment), the firm pays no bonus and fires immediately. For the agent, the candidate strategy specifies: if there has been no deviation, work in every period and choose $e(\theta_i)$ after observing θ_i; following any deviation, set effort to zero and quit immediately. If play remains on the equilibrium path, then clearly this agreement is stationary and involves effort $e(\theta)$ in every period.

This agreement is self-enforcing. Following any deviation, it is clear that neither the firm nor the agent can benefit by deviating. Moreover,

$$b(\theta_i) - b(\theta_{i+1}) = c(e(\theta_i), \theta_i) - c(e(\theta_{i+1}), \theta_i),$$

so that

$$b(\theta_i) - c(e(\theta_i), \theta_i) \geq b(\theta_{i+1}) - c(e(\theta_{i+1}), \theta_i),$$

and

$$\begin{aligned} b(\theta_i) - c(e(\theta_i), \theta_i) &= b(\theta_{i-1}) - c(e(\theta_{i-1}), \theta_i) + \{c(e(\theta_i), \theta_{i-1}) \\ &\quad -c(e(\theta_{i-1}), \theta_{i-1}) + c(e(\theta_{i-1}), \theta_i) - c(e(\theta_i), \theta_i)\} \\ &\geq b(\theta_{i-1}) - c(e(\theta_{i-1}), \theta_i). \end{aligned}$$

By the constraint reduction theorem, this guarantees the agent will choose according to his true type. And since $b(\theta) \geq c(e(\theta), \theta)$ for all θ, the agent prefers to choose $e(\theta)$ than any other e, and also to accept the bonus. Finally, $U = \overline{U}$ so the agent does not quit.

Since $b(\theta) > 0$, the firm's most profitable deviation is to withhold the bonus payment when $e = e(\theta_1)$, which means immediate termination. But this cannot be desirable, since

$$-b(\theta_1) + \delta\Pi = -b(\theta_1) + \delta(S - U) = -b(\theta_1) + \delta(S - \overline{U}) \geq \delta\overline{\Pi},$$

the last inequality following from (IC). *Q.E.D.*

Proof of Proposition 1.3. The optimal effort schedule solves the unrelaxed program

$$\max_{e_1,\ldots,e_m} S$$

$$\text{s.t.} \quad \delta(S - \overline{S}) - c(e_1, \theta_1) - \sum_{i=2}^{m} [c(e_i, \theta_i) - c(e_i, \theta_{i-1})] \geq 0 \qquad (IC)$$

$$e_1 - e_2 \geq 0, \ldots, e_{m-1} - e_m \geq 0, \; e_m \geq 0 \qquad (M_1 : M_m)$$

The relaxed program eliminates the constraints $M_1, .., M_{m-1}$. Because $c(e_i, \theta_i) - c(e_i, \theta_{i-1})$ is convex in e, both are concave programs and the nonempty interior condition (or Slater constraint qualification) guarantees that both programs can be characterized using their Kuhn-Tucker conditions. In particular, there exists some $\lambda \geq 0$ such that the solution to the relaxed program solves

$$\max_{e_1,\ldots,e_m \in [0,\overline{e}]^m} (1+\lambda\frac{\delta}{1-\delta}) \sum_{i=1}^{m} [E[y|e_i] - c(e_i, \theta_i)]\, p(\theta_i) - \lambda \left[\sum_{i=2}^{m} [c(e_i, \theta_i) - c(e_i, \theta_{i-1})] + c(e_1, \theta_1) \right].$$

Considering the optimization pointwise, e_i solves

$$e_i = \arg\max_{e \in [0,\overline{e}]} (1 + \lambda\frac{\delta}{1-\delta}) [E[y|e] - c(e, \theta_i)] - \lambda \frac{1}{p(\theta_i)} [c(e, \theta_i) - c(e, \theta_{i-1})].$$

The assumptions that $p(\theta_i)$ is decreasing in i and that $c(e, \theta_i)$ and $c(e, \theta_i) - c(e, \theta_{i-1})$ are supermodular in (e, i) guarantee that this objective is submodular in (e, i). So Topkis' Theorem (see Milgrom and Shannon, 1994) ensures a decreasing solution to the relaxed program. *Q.E.D.*

Proof of Proposition 1.7. Suppose some agreement $(s, \sigma, \varsigma) \in \Omega^*$ implements $e \neq 0$ at some time t. Let $S_1, \ldots, S_n$ be the continuation surpluses following output $y_1, \ldots, y_n$. I claim they cannot all be the same. The incentive constraint,

$$e \in \arg\max_{\tilde{e}} E\left[b + \delta U - c|\tilde{e}\right],$$

implies that there is some variation in $b + \delta U$ across outcomes. So there exist outcomes y_i,y_j

such that

$$b_M + \delta U_M > b_m + \delta U_m.$$

The message constraint is

$$-b_M + \delta \Pi_M = -b_m + \delta \Pi_m.$$

Adding these two constraints implies that $S_M \geq S_m$. *Q.E.D.*

Proof of Theorem 1.8. ($\Rightarrow$) Let b_i, α_i be the bonus and probability of continuation following output realization y_i. If the firm observes y_i it cannot prefer to report y_j:

$$-b_i + \delta \left[\alpha_i \Pi + (1 - \alpha_i)(\overline{\Pi} - d_i)\right] \geq -b_j + \delta \left[\alpha_j \Pi + (1 - \alpha_j)(\overline{\Pi} - d_j)\right]. \tag{1.29}$$

Of course, the reverse inequality must hold as well. Thus, for $i = 1, ..., n$,

$$b_i = b_1 + \delta \left[(\alpha_i - \alpha_1)\left(\Pi - \overline{\Pi}\right) + (1 - \alpha_1)d_1 - (1 - \alpha_i)d_i\right]. \tag{1.30}$$

For the agent to choose e rather than some other effort level $\hat{e}$, it must be that for every $\hat{e} \neq e$,

$$\sum_i \left(b_i + \delta \left[\alpha_i U + (1 - \alpha_i)\left(\overline{U} + d_i\right)\right]\right) \left[f(y_i|e) - f(y_i|\hat{e})\right] \geq c(e) - c(\hat{e}) \tag{1.31}$$

Substituting for b_i in this inequality, and simplifying, implies that for every $\hat{e} \neq e$,

$$\delta \left[S - \overline{S}\right] \sum_{i=1}^{n} \alpha_i \left[f(y_i|e) - f(y_i|\hat{e})\right] \geq c(e) - c(\hat{e}). \tag{1.32}$$

($\Leftarrow$) Let $b_1 = 0$, and define $b_2, ..., b_n$ as above. Since the agent's choice of e is assured, as well as the firm's truthful reports, it remains to find some w, d such that $U \geq \overline{U} + \max_i d_i$, and $\Pi \geq \overline{\Pi} - \min_i d_i$. This is easily done; for instance, let $d_i = 0$ for all i, and $w = (1-\delta)\overline{U} + c(e) - E[b|e]$. Then $U = \overline{U}$, and by definition $\Pi - \overline{\Pi} = S - \overline{S} > 0$. *Q.E.D.*

Proof of Proposition 1.8. The direct proof used in the common monitoring case fails here. Instead, I use an argument similar to Rogerson (1985). I show that the agent's incentive constraints can be replaced by a first-order condition, and that under this substitution optimal contracts are one-step. It then follows that $e^* \leq e^{FB}$.

I work with the equivalent money burning contract and look for the most efficient contract to implement a given effort level $e > 0$. Such a contract solves:

$$\min_{b,\tau} \sum_i \tau_i f(y_i|e)$$

s.t. $\quad \tau_i + b_i$ constant in i,

$$e \in \arg\max_{\tilde{e}} \sum_i b_i f(y_i|\tilde{e}) - c(\tilde{e}),$$

$$\frac{\delta}{1-\delta}\left(\sum_i (y_i - \tau_i) f(y_i|e) - c(e) - (1-\delta)\overline{S}\right) \geq \tau_i + b_i - \min_j b_j.$$

It is optimal to take $\min_j \tau_j = 0$. So for all i, $\tau_i + b_i = \max_j b_j$. Letting $\overline{b} = \max_j b_j$, and without loss of generality setting $\min_j b_j = 0$, one can rewrite the problem in terms of $b_1, ..., b_n, \overline{b}$:

$$\max_{b_1,\ldots,b_n,\overline{b}} -\overline{b} + \sum_i b_i f(y_i|e)$$

$$\text{s.t.} \qquad e \in \arg\max_{\tilde{e}} \sum_i b_i f(y_i|\tilde{e}) - c(\tilde{e}), \qquad (IC_1)$$

$$\frac{\delta}{1-\delta}\left(-\frac{1}{\delta}\overline{b} + \sum_i (b_i + y_i) f(y_i|e) - c(e) - (1-\delta)\overline{S}\right) \geq 0, \qquad (IC_2)$$

$$b_1 \geq 0, ..., b_n \geq 0,$$

$$\overline{b} - b_1 \geq 0, ..., \overline{b} - b_n \geq 0.$$

The relaxed program replaces (IC_1) with the agent's first-order condition:

$$\sum_i b_i f_e(y_i|e) - c'(e) = 0 \qquad (FOC)$$

Suppose $(b^*, \overline{b}^*)$ solves the relaxed program. If the agent's objective function is concave given bonus plan b^*, then $(b^*, \overline{b}^*)$ is also a solution to the original program. Note that the relaxed program is *linear* in $b, \overline{b}$. After some manipulation, the partial derivative of its Lagrangian with respect to b_i is

$$\left[1 + \gamma\frac{\delta}{1-\delta}\right] + \lambda\frac{f_e(y_i|e)}{f(y_i|e)} + \mu_i - \nu_i,$$

where γ is the multiplier on (IC_2), λ is the multiplier on (IC_1), μ_i the multiplier on the nonnegativity constraint (divided by $f(y_i|e) > 0$), and ν_i the multiplier on the constraint that $b_i \leq \overline{b}$ (similarly divided).

Suppose $\lambda \leq 0$. Then from MLRP the bonuses that solve the program will be decreasing or constant with at most one bonus being interior: $b_1 = ... = b_{k-1} = \overline{b}$ and $b_{k+1} = ... = b_n = 0$ (MLRP implies that f_e/f is strictly increasing in i). But this means the agent must choose $e = 0$. This cannot be correct, so λ must be nonnegative, and it follows immediately (using MLRP) that $b_1 = ... = b_{k-1} = 0$ and $b_{k+1} = ... = b_n = \overline{b}$ (the bonus b_k may be interior). That is, any solution to the relaxed program uses nearly one-step bonuses to implement e. Moreover, because nearly one-step bonuses are monotone and we have assumed CDFC, the agent's objective is concave and the solution to the relaxed program solves the original program. So optimal bonuses are nearly one-step. Note also that if $f_e(y_i|e) > 0$ then $b_i = \overline{b}$, but if $f_e(y_i|e) < 0$, it may still be the case that $b_i = \overline{b}$. With common monitoring, $b_i = \overline{b}$ if and only if $f_e > 0$, so the cut-off is "lower" under private monitoring.

To see that $e^* \leq e^{FB}$, consider some nearly one-step bonus plan that implements $e > e^{FB}$. It cannot be optimal because by lowering $\overline{b}$ by ε, the parties can induce a slightly lower effort choice and still satisfy (IC_2). This reduces money burning and increases surplus. The stated Proposition uses the equivalence of money-burning and termination contracts. *Q.E.D.*

Proof of Proposition 1.9. Suppose a termination contract under f uses continuation probabilities α to implement effort e. I will show that e is implementable under g using continuation probabilities a. Let $a_j = \sum_i \alpha_i k(y_i|z_j)$, where $k(y|z)$ is a Markov kernel such that for all e

$$f(y_i|e) = \sum_j k(y_i|z_j)g(z_j|e) \tag{1.33}$$

Then

$$\begin{aligned} E[a|e] &= \sum_j a_j g(z_j|e) \\ &= \sum_j \sum_i \alpha_i k(y_i|z_j)g(z_j|e) \\ &= \sum_i \alpha_i f(y_i|e) = E[\alpha|e] \end{aligned}$$

so the surplus generated under (f, α) is the same as under (g, a). Denote this S. Also, for all $\hat{e} \neq e$,

$$\delta(S - \overline{S}) \sum_j a_j \left[g(z_j|e) - g(z_j|\hat{e})\right] = \delta(S - \overline{S}) \sum_i \alpha_i \left[f(y_i|e) - f(y_i|\hat{e})\right] \geq c(e) - c(\hat{e}).$$

Finally, since each a_j is a convex combination of α_i's, and for all i, $\alpha_i \in [0,1]$, we have $a_j \in [0,1]$ for all j. Again, if $k(\bullet|\bullet) > 0$, contracting possibilities are strictly better under g than under f. *Q.E.D.*

Proof of Proposition 1.10. Consider optimal money burning contracts. To implement any $e > 0$ requires

$$\sum b_i \left[f(y_i|e) - f(y_i|0)\right] \geq c(e) > 0.$$

Thus, there must be some i_e, j_e such that $b_{i_e} - b_{j_e} \geq c(e)$. So truthful revelation implies that the parties must burn at least $T(e) = c(e)f(y_{j_e}|e) > 0$ in expectation. Now, fix some $\underline{e} > 0$ such that $E[y|\underline{e}] - c(\underline{e}) < s^{FB}$. Then to implement any $e \in [\underline{e}, \overline{e}]$, the parties must burn at least $T = \min_{e \in [\underline{e}, \overline{e}]} T(e) > 0$ in expectation. So the largest per-period surplus that can be obtained by a relational contract is bounded above by $\max\{E[y|\underline{e}] - c(\underline{e}), s^{FB} - T\}$, which is strictly smaller than s^{FB}. *Q.E.D.*

1.12 References

Abreu, Dilip (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

Abreu, Dilip, Paul Milgrom and David Pearce (1991): "Information and Timing in Repeated Partnerships," *Econometrica*, 59, 1713–1733.

Abreu, Dilip, David Pearce and Ennio Stacchetti (1990): "Towards a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041–1063.

Atkeson, Andrew (1991): "International Lending with Moral Hazard and Risk of Repudiation," *Econometrica*, 59, 1069–1089.

Baker, George, Robert Gibbons and Kevin Murphy (1994): "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, 109, 1125–1156.

Baker, George, Robert Gibbons and Kevin Murphy (1997): "Relational Contracts and the Theory of the Firm," Mimeo, M.I.T.

Bull, Clive (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102, 147–159.

Compte, Olivier (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66, 597-626.

Farrell, Joseph and Eric Maskin (1989): "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1, 327-360.

Fudenberg, Drew, Bengt Holmström and Paul Milgrom (1990): "Short-Term Contracts and Long-Term Agency Relationships," *Journal of Economic Theory*, 51, 1–31.

Fudenberg, Drew and David Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62, 103–135.

Fudenberg, Drew, David Levine and Eric Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997–1039.

Fudenberg, Drew and Jean Tirole (1991): *Game Theory*, Cambridge: MIT Press.

Green, Edward and Robert Porter (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52, 87–100.

Greif, Avner (1993): "Contract Enforceability and Economic Institutions in Early Trade: the Maghribi Traders' Coalition," *American Economic Review*, 83, 525–548.

Grossman, Sanford and Oliver Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51, 7–45.

Grossman, Sanford and Oliver Hart (1986): "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration," *Journal of Political Economy, 94, 691–719.*

Hart, Oliver and Bengt Holmström (1987): "The Theory of Contracts," in *Advances in Economic Theory*, Truman Bewley ed., Cambridge: Cambridge University Press.

Holmström, Bengt (1979): "Moral Hazard and Observability," *Bell Journal of Economics*, 10, 74–91.

Innes, Robert (1990): "Limited Liability and Incentive Contracting with Ex-ante Action Choices," *Journal of Economic Theory*, 52, 45–67.

Kandori, Michihiro (1992): "The Use of Information in Repeated Games with Imperfect Monitoring," *Review of Economic Studies*, 59, 581–593.

Kandori, Michihiro and Matsushima, Hitoshi (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66, 627–652.

Klein, Benjamin and Keith Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 615–641.

Kreps, David (1990): "Corporate Culture and Economic Theory," in *Essays on Positive Political Economy*, J. Alt and K. Shepsle ed., Cambridge: Cambridge University Press.

Macaulay, Stewart (1963): "Non-Contractual Relations in Business: A Preliminary Study," *American Journal of Sociology*, 28, 55–67.

MacLeod, Bentley and James Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57, 447–480.

Macneil, Ian (1974): "The Many Futures of Contracts," *Southern California Law Review*, 47, 691–816.

Macneil, Ian (1978): "Contracts: Adjustment of Long-Term Relationships under Classical, Neoclassical, and Relational Contract Law," *Northwestern Law Review*, 72, 854–906.

Milgrom, Paul and Christina Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

Moore, John and Rafael Repullo (1988): "Subgame Perfect Implementation," *Econometrica*, 56, 1191–1220.

Pearce, David and Ennio Stacchetti (1998): "The Interaction of Implicit and Explicit Contracts in Repeated Agency," *Games and Economic Behavior,* 23, 75–96.

Radner, Roy (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica*, 53, 1173-1198.

Rogerson, William (1985): "The First-Order Approach to Principal-Agent Problems, *Econometrica*, 53, 1357–1367.

Shapiro, Carl and Joseph Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433-444.

Sinclair-Desgagné, Bernard (1994): "The First-Order Approach to Multi-Signal Principal-Agent Problems," *Econometrica*, 62, 459–465.

Spear, Stephen and Sanjay Srinivasta (1987): "On Repeated Moral Hazard with Discounting," *Review of Economic Studies*, 54, 599–617.

Stole, Lars (1997): *Lectures on the Theory of Contracts and Organizations*, unpublished monograph.

Williamson, Oliver (1985): *The Economic Institutions of Capitalism*, New York: Free Press.

Chapter 2

Multilateral Contracting and Organizational Design

2.1 Introduction

A growing literature on organizations takes as a starting point the premise that ongoing transactions within and between firms are governed by a combination of formal court-enforced contracts and informal goodwill or reciprocal agreements. By modeling these *relational contracts* as equilibrium points of repeated games, this work has shed light on aspects of economic organization from the employment relationship to medieval trade arrangements to consumer markets, the structure of supply chains, corporate culture, and the choice between debt and equity financing.[1]

All of this work has been motivated at least in part by two seminal papers by Alchian and Demsetz (1972) and Jensen and Meckling (1976), who were among the first to emphasize the contractual nature of the firm. The current literature, however, has largely failed to get at the critical assertion of these authors, namely that the firm must be viewed as a "central contractual agent" (to use Alchian and Demsetz's words) or as a "nexus of contracts" (to use Jensen and Meckling's). Instead, the great bulk of this recent work deals with a single transactional relationship viewed in isolation. For instance, efficiency wage theory studies "matches" between individual jobs and workers. Studies of customer markets (Klein and Leffler, 1981) and integration decisions (Baker, Gibbons and Murphy, 1997; Rotemberg, 1991) model one upstream and one downstream party. These last two papers build theories of the firm based on optimal asset ownership with regard to a *single* transaction. In this, they are broadly consistent with the incomplete contracting view of organizations (e.g. Hart, 1995), which has tended to focus

[1] See for instance, Bull (1987) and Shapiro and Stiglitz (1984) on employment, Greif (1993) and Greif, Milgrom and Weingast (1994) on medieval trade, Klein and Leffler (1981) on customer markets, Taylor and Wiggins (1997) on supply chains, Kreps (1990) on corporate culture, and Fluck (1998) on corporate finance.

on "marginal" transactions at the the firm boundary.[2]

In this chapter, I argue that when contracting is relational, attempts to model employment or firm structure with reference to a single transaction are likely to present a very partial picture of the general problem of organizational design. The reason for this is that even when transactions are technologically separate, contracting possibilities will be interlinked by cross-monitoring and reputation. If the firm is involved in transactions with several agents, its behavior vis a vis any one agent will impact on its other relationships. Mathematically, this idea takes the form of a simple and appealing result: when a firm contracts with a single agent, a self-enforcing agreement can be reached if and only if the relationship generates a significantly positive quasi-rent. When one takes into account the fact that the firm contracts with many parties, it need no longer be the case that each relationship independently generate a quasi-rent. Instead, the sum of all relationships within the firm must generate a large enough rent.[3] The advantage of the multi-contract firm is that it can correctly allocate these rents and globally optimize incentives.

To make this point, I develop a general agency model in which the firm conducts ongoing transactions with many parties. The transactions may of may not be directly technologically linked. One can envision these transactions taking place within the firm — with the agents interpreted as workers, or divisions within a multi-unit organization, or franchisees — or between firms — with the agents interpreted as subcontractors or upstream firms in the supply chain. The two main assumptions are that key contracting variables are not court-verifiable, and that all parties are risk-neutral. Following MacLeod and Malcomson (1989), I assume that the parties can write a court-enforced contract specifying only fixed payments, but may use an informal agreement to govern action choices, rewards and penalties. In an employment relationship, this corresponds to having a fixed wage contract, and an informal agreement about discretionary bonuses or future wages. In the supply chain or subcontracting setting, the formal contract specifies a fixed price for the product or service, with informal "price adjustments" that depend on nonverifiable outcome measures.

An important extant result is that self-enforement of a bilateral agreement requires that the discounted future gains from the relationship (the quasirents) outweigh the present costs (Bull, 1987; MacLeod and Malcomson, 1989). Thus, if the firm reaches separate bilateral agreements

[2] Incomplete contracting models that do include multiple parties, such as Hart and Moore (1990), Stole and Zwiebel (1996) and Rajan and Zingales (1998), emphasize the technological connections between transactions. In these papers, surplus is divided according to the Shapley value. So technological interdependence leads to strategic effects. Holmström and Milgrom (1994) also emphasize the interdependence of firm decisions, and the importance of viewing incentives within firms as a whole.

[3] Coleman has made a related point in *Foundations of Social Theory*, although his argument does not directly suggest a formal approach. Coleman suggests that a crucial characteristic of the modern firm is *global viability*, the fact that only the sum total of firm transactions must generate a positive balance. He contrasts this with the position taken by Barnard (1938) that each relationship the firm undertakes must be *individually viable*.

with n agents whose production is technologically independent, the incentives are limited by n independent constraints. I show that when the agents can *cross-monitor*, i.e. observe each other's behavior, and have the freedom to write side contracts, all the bilateral constraints can be pooled into a single requirement that the sum of present costs is less than the total discounted gains from *all* the relationships. Optimal contract design reduces to an optimization problem with a single constraint. So even when relationships are technologically independent, the firm can shift quasirents between relationships and "cross-subsidize" relationships that are particularly constrained. That such subsidization will be optimal can be naturally interpreted in terms of shadow prices. In general, there is a shadow price attached to the incentive constraint on each relationship. Under technological independence, when the firm reaches bilateral agreements, these shadow prices are independent. But under cross-monitoring, an optimal agreement *equalizes* the shadow prices on each relationship.

These basic insights can be extended and applied in a number of directions. For instance, I show that if the parties can make decisions ex ante about observability (and thus about whether to contract bilaterally or multilaterally), the gain from multilateral contracting will be greater when the agent's inputs are substitutes than when they are complements. I also show that in a market equilibrium setting, even if multilateral contracting does not allow the firm to provide a higher level of incentives, the ability to write a multilateral contract may improve the firm's bargaining power and create a competitive advantage. The idea is that when a firm has important existing relationships, it has more credibility in making promises, and can use a wider variety of contracts than firms that must contract bilaterally. The model also has implications for ex ante and ongoing investment and for asset ownership that are distinct from the insights derived from the static property rights models of Grossman and Hart (1986) and Hart and Moore (1990)). Under multilateral contracting, investments that deepen relationships (increase the level of dependency) are particularly valuable because they relax the single firm constraint. The fact that such investments have *no strategic value* in the standard property rights model has been a main point of criticism (Holmström, 1996). Multilateral contracting also means that *coordination* of investments has particular importance — because an investment in any one relationship has implications for all other relationships.

In the initial model I develop, multilateral contracting can only improve on bilateral contracting. This raises the question of whether cross-monitoring might have a downside. Intuitively, one might suspect that if relationships are designed to be highly contingent on other relationships, then a breakdown in any one might threaten to bring down the whole complicated arrangement. To answer this question, I extend the model to allow for the resolution of uncertainty over, so that agents are learning about their respective futures with the firm. In this case, the firm's reputation can be so impaired by a breakdown in one relationship that its other relationships also fail, and bilateral contracting may be optimal.

In addition to the relational contracting literature noted above, my results are related to work on multi-market contact by Bernheim and Whinston (see also Telser (1981)). In their 1990 paper, they show that if two firms attempt to simultaneously collude in two markets, their collusive arrangement need only satisfy a single incentive constraint for each firm. In a sense, I couple this insight with the result of Bull (1987) and MacLeod and Malcomson (1989) dicussed above, which says that in a bilateral relational contracting situation, the parties can pool their two incentive constraints into a single constraint. Bull (1987) and Kreps (1990) have also emphasized the idea that the firm can sustain a reputation precisely because it conducts many transactions. In these papers, the firm is long-lived, and individual agents short-lived. The work on medieval institutions and trade done by Greif (1993) and Greif, Milgrom and Weingast (1994) is also related. These authors also emphasize the benefits of multilateral reputation mechanisms as a motive for institutional design. Bendor and Mookherjee (1990) have studied the benefits of multilateral enforcement in a prisoner's dilemma game.

2.2 The Model

2.2.1 Ongoing Transactions

There is one firm and I agents, where I denotes both the number and the set of agents. All parties are risk-neutral, live forever, and have discount factor δ. Time is discrete, indexed by $t = 1, 2, \ldots$. The firm's output (or gross benefit) at time t is $y_t(e_t)$, where $e_t = (e_t^1, \ldots, e_t^I)$ is the profile of time t inputs, e_t^i being the input or effort supplied by agent i. Agent i's cost is $c_t^i(e_t)$ which may also depend on the entire time t effort profile.

The firm and the agents engage in repeated transactions. Each period, $t \geq 1$, each agent currently involved with the firm simultaneously chooses whether or not to sever its attachment, and the firm has the opportunity to release any subset of its agents. I assume for convenience that separation is irreversible. If agent i's relationship ends at t, I write $e_\tau^i = \emptyset$ for $\tau \geq t$. Each agent i active at t chooses an effort level $e_t^i = [0, \overline{e}]$. Once actions are taken and y_t is realized, the firm and all active agents observe (or perfectly infer) e_t and y_t. Payments are made at the end of the period, where W_t^i denotes the net transfer from the firm to agent i. If I_t is the set of agents engaged with the firm at time t, each has payoff $W_t^i - c_t^i(e_t)$ and the firm's payoff is $y_t(e_t) - \sum_{i \in I_t} W_t^i$.

Effort is costly: in particular, for any $\tilde{I} \subseteq I$ with $i \in \tilde{I}$, $\sum_{j \in \tilde{I}} c_t^j(e)$ is continuous, strictly increasing and strictly convex in e^i, for $e^i \in [0, \overline{e}]$. Effort is positively related to output: $y_t(e)$ is continuous, increasing and concave in e^i. Finally, zero effort is costless $c^i(e^i = 0, e^{-i}) = 0$, but unproductive: $y_t(0, e^{-i}) - \sum_{j \in I} c^j(0, e^{-i}) \leq y_t(\emptyset, e^{-i}) - \sum_{j \in I} c^j(\emptyset, e^{-i})$.

Ignoring the presence of the courts, it is clear the only equilibrium of the one-shot game would be immediate separation. To see this, note that no one will be willing to make voluntary payments ex post, which means that there is no incentive to provide any effort. Since separation is more efficient than production with zero effort, no relationships will be initiated. Thus the only hope for cooperative agreements rests on either court-enforcement or long-term self-enforcement.

2.2.2 Long-Term Payoffs

Agents always have the option of separating from the firm and taking up their "outside option" (this might mean continuing to trade with the firm, but through a spot market). If agent i has separated from the firm at t, his *lifetime* utility from that point is $\overline{U}_t^i \geq 0$. If agent i is employed up to time T^i, his lifetime utililty from time t is

$$U_t^i = W_t^i - c^i(e_\tau) + \delta U_{t+1}^i \qquad t = 1, ..., T^i$$

where $U_{T^i+1}^i = W_{T^i+1}^i + \overline{U}_{T^i+1}$, with $W_{T^i+1}^i$ representing any court-enforced separation payments. The firm's lifetime payoff if all relationships have been terminated is $\overline{\Pi}_t - \sum_{i \in I_{t-1}} W_t^i$, where $\overline{\Pi}_t \geq 0$, and its payoff from time t is

$$\Pi_t = y_t(e_t) - \sum_{i \in I_{t-1}} W_t^i + \delta \Pi_{t+1}, \qquad t = 1, ..., T,$$

where its wage payments include separation payments to departed workers. Also $T = \max_i T^i$ and $\Pi_{T+1} = \overline{\Pi}_{T+1} - \sum_{i \in I_T} W_T^i$. The total surplus generated from time t by the firm and its time $t-1$ agents is

$$S_t = \Pi_t + \sum_{i \in I_{t-1}} U_t^i = y_t(e_t) - \sum_{i \in I_t} c^i(e_t) + \sum_{i \in I_{t-1} - I_t} \overline{U}_t^i + \delta S_{t+1}.$$

Also, let $\overline{S}_t = \sum_{i \in I_{t-1}} \overline{U}_t^i + \overline{\Pi}_t$.

2.2.3 Relational Contracts

A relational contract combines a formal court-enforced contract with an informal agreement governing behavior. Court-enforcement is limited. Following MacLeod and Malcomson (1989), I assume that legal contracts between the firm and its agents can specify only a fixed payment w_t^i to be made if the agent produces for the firm at t, and a separation payment d_t^i to be made if relationship i ends at t (the court cannot determine which party caused a separation). These payments may be contingent on the relationship status and wage histories of *all* of the agents, and can be positive or negative. I rule out the possibility that the parties can design court-enforced message games designed to elicit information that is observable but not directly verifiable.[4] On the other hand, the parties are free to make discretionary payments or contract adjustments based on any mutual information; I let b_t^i denote the net voluntary transfer from the firm to agent i. I allow $w^i, d^i, t^i \in (-\infty, \infty)$. By definition, the total payment to agent i is $W_t^i = w_t^i + d_t^i + b_t^i$.

The informal part of the contract is a set of proposed strategies. A strategy for agent i specifies termination, effort and payment decisions as a function of observed history. A strategy for the firm specifies separation and payment decisions as a function of *observed* history. A formal contract is a map from *verifiable* history to wage payments. A *relational contract* is a set of proposed strategies for the firm and each agent, and a formal contract. It is *self-enforcing* if, given the formal contract, the strategies form an equilibrium.

I consider two notions of equilibrium contracting. A first pass at equilibrium contracting is to require subgame perfection: no party has a profitable unilateral deviation after any history. This turns out to be a very weak requirement, leading me to consider a stronger version of equilibrium which allows the agents to forge agreements among themselves. I say that agents can side-contract if they can (1) coordinate their actions and make informal side payments, and (2) write legally enforceable contracts specifying side payments based on verifiable history. When agents can side contract, they can make multilateral deviations from the original contract. In such a world, a relational contract will be self-enforcing if and only if after no history is there a set of agents that can profitably deviate.

2.3 Self-Enforcing Contracts

Because all parties are risk-neutral and there are no limits on transfer payments, the total surplus generated by an agreement depends only on the action profiles at each time, $(e_t)_{t=1}^{\infty}$ (note that specifying $(e_t)_{t=1}^{\infty}$ also specifies the set of active active agents at each time t). Moreover, because the parties are free to make ex ante payments, *any* division of the surplus is possible provided that it ensures each party at least their outside option. Thus, it is reasonable to believe that the parties will reach a self-enforcing agreement which maximizes their joint surplus. That is, they will choose $(e_t)_{t=1}^{\infty}$ to maximize S_1 subject to the constraint that $(e_t)_{t=1}^{\infty}$ can be enforced with a relational contract. Call any profile that solves this problem *constrained efficient.* To characterize constrained efficient contracts, it is necessary to ask which action profiles $(e_t)_{t=1}^{\infty}$ are enforceable.

I give two results to this end. First, I show that when the agents cannot side contract, *any* sequence of action profiles is enforceable so long as it generates surplus $S \geq \overline{S}_1$ and involves

[4] Given the assumption that e_t is observed *after* the realization of costs and benefits, this assumption is not restrictive. On the other hand, if e_t was observed *before* the firm reaped the benefit $y_t(e_t)$, a suitable message game could implement the first best.

multiple active agents. The idea is that the parties can use the court to commit the firm to paying a fixed total level of wages — by carefully structuring payments, a bonus plan can be found that induces effort from all agents, and is even robust to one-time informal coordinated deviations by the agents. Thus, the presence of multiple agents may eliminate moral hazard problems caused by non-verifiability even though the space of available contracts is very small. This is a powerful result, but somewhat unreasonable — the agreements needed to achieve these outcomes are susceptible to collusion by the agents. Adopting a stronger notion of equilibrium contracting, I give an appealing characterization result: the parties can enforce an agreement if and only if the total discounted future gains from all relationships are larger than the total present costs.

2.3.1 Full Implementation without Side Contracts

The power of even limited formal contracts when there are many parties can be demonstrated with a simple example. Suppose the firm wants to induce two agents to exert effort $e = (e^1, e^2)$ for just one period. Define bonuses $b^i = c^i(e)$. If both agents choose the correct effort, the firm pays b^i to each agent i; if only one agent chooses the correct effort, the firm pays the whole bonus $b^1 + b^2$ to that agent and nothing to the other; if neither exerts the correct effort, the firm pays $b^1 + b^2$ to agent 1 and nothing to agent 2. To support this agreement, the parties write a court-enforced contract specifying fixed wages w^1, w^2, with a clause stating that if the actual total wage paid $W^1 + W^2$ falls short of $\sum_i (w^i + b^i)$, the firm must pay the difference to agent 1 as severance. So long as $y(e) - c^1(e) - c^2(e)$ is greater than the sum of the outside options, the parties can find w^1 and w^2 to make this self-enforcing.[5] The equilibrium will even be robust to informal agreements between the two agents: agent 1 could in principle compensate agent 2 if both deviated, but he cannot credibly promise to do so, and ex post, he would not.

The point is that if the court will compel a certain total level of payments, the firm can credibly threaten to give agent i's bonus to agent j. This general logic extends to arbitrary effort profiles so long as there are always two agents and the parties are willing *ex ante* to initiate the contract.

Proposition 2.1 *Suppose workers may not side contract. The agents can reach a self-enforcing agreement to implement effort* $(e_1, e_{2,\ldots})$ *if and only if* $S_1 \geq \overline{S}_1$*, and, if there exists* $\hat{T} < T$ *such that* $I_{\hat{T}} = \{i\}$*, then for all* $t \geq \hat{T}$

$$\delta \left[U^i_{t+1} + \Pi_{t+1} - \overline{U}^i_{t+1} - \overline{\Pi}_{t+1} \right] \geq c^i(e_t) \qquad (IC^1)$$

[5] A closely related point is made by Malcomson (1986), who shows in a one-period model that by committing to a fixed bonus pool and using a rank-order tournament, the principal can induce first-best effort.

The proof is somewhat complicated by the need to check various cases. In addition to using the sort of bonus scheme discussed above, the parties must employ a bonding scheme to prevent unplanned separations (which are a concern since potentially $S_t < \overline{S}_t$). Again, the scheme relies on being able to punish both the firm and agent i for i's unplanned termination by giving a large reward to some agent j. These sorts of schemes do not work when there is only a single agent because there is no third party to soak up promised bonus payments after a deviation. The last requirement, (IC^1), reflects this "budget balance" constraint. It says that if after some time only one agent remains, the total gain from continuing must outweigh present costs. This is exactly the condition that appears in MacLeod and Malcomson's (1989) analysis of bilateral contracting.

2.3.2 A Characterization Theorem

Schemes that reward j for i's malfeasance seem unlikely in that they give i and j strong incentives to collude. Account for this possibility leads to a more intuitive characterization of long-term contracting possibilities.

Consider some self-enforcing agreement that implements $(e_t)_{t=1}^{\infty}$. At time t, the agents can always set their efforts jointly to zero, refuse to make any bonus transfers, and then quit, using a side contract to re-divide the spoils. For this to be unprofitable,

$$\sum_{i\in I_t}\left[-c^i(e_t)+b_t^i+\delta U_{t+1}^i\right] \geq \sum_{i\in I_t}\left[b_t^i(0)+\delta\overline{U}_{t+1}^i+\delta d_{t+1}^i(0)\right], \tag{2.1}$$

where $b_t^i(0)$ are the bonuses the firm pays after observing zero effort, and $d_{t+1}^i(0)$ are the court-enforced severance payments that follow these bonuses.[6] At the same time, the firm can always behave as if the agents had shirked — since the contract is preferred,

$$-\sum_{i\in I_t} b_t^i + \delta\Pi_{t+1} \geq -\sum_{i\in I_t}\left[b_t^i(0)+\delta d_{t+1}^i(0)\right]+\delta\overline{\Pi}_{t+1}. \tag{2.2}$$

Since both these conditions must be satisfied, clearly their sum must be satisfied:

$$\delta\left[\sum_{i\in I_{t+1}} U_{t+1}^i + \Pi_{t+1} - \sum_{i\in I_{t+1}}\overline{U}_{t+1}^i - \overline{\Pi}_{t+1}\right] \geq \sum_{i\in I_t} c^i(e_t). \tag{2.3}$$

In other words, the total future gains from all relationships must outweigh the total present costs. In fact, (2.3) is a sufficient condition for self-enforcement as well as a necessary one. If some effort profile $\{e_t\}_{t=1}^{\infty}$ satisfies (2.3), a payment plan can be found to implement it. Thus

[6] Because the realization of these bonuses is verifiable, if the inequality fails, the agents can write a contingent contract so that each agent prefers to shirk and separate.

despite the fact that the firm has many ongoing relationships, they are all governed by a single incentive constraint.

Proposition 2.2 *Suppose agents can side-contract. There exists a self-enforcing agreement that implements* $\{e_t\}_{t=1}^{\infty}$ *if and only if* $S_1 \geq \overline{S}_1$ *and for all* t,

$$\delta\left[\sum_{i\in I_{t+1}} U_{t+1}^i + \Pi_{t+1} - \sum_{i\in I_{t+1}} \overline{U}_{t+1}^i - \overline{\Pi}_{t+1}\right] \geq \sum_{i\in I_t} c^i(e_t). \qquad (IC)$$

Proof. ($\Rightarrow$) Has been shown. ($\Leftarrow$) Let W_t be some bounded sequence of total wages that satisfies $U_1^i \geq \overline{U}_1^i$ for all i, and $\Pi_1 \geq \overline{\Pi}_1$ — such a sequence exists because $S_1 \geq \overline{S}_1$. Define a formal contracts as follows: let $w_t^i = W_t^i - c^i(e_t)$. Let $d_t^i = U_t^i - \overline{U}_t^i$ for all $i \in I_t$, and $d_t^i = 0$ for all $i \in I_{t-1} - I_t$, where I_t is known from $\{e_t\}_{t=1}^{\infty}$. Informally, agents choose effort e_t and the firm pays bonuses $b_t^i = c^i(e_t)$. Following any deviation, all agents exert zero effort, the firm pays no bonuses and all parties separate immediately. It is clear that following a deviation, the strategies are unimprovable by any individual or group of agents. Moreoever, no agent or group of agents benefits by shirking or quitting. The firm's best deviation is to withhold all bonuses and immediately terminate all the agents, which is not profitable because

$$-\sum_{i\in I_t} c^i(e_t) + \delta\Pi_{t+1} \geq \delta\overline{\Pi}_{t+1} - \delta\sum_{i\in I_{t+1}}\left[U_{t+1}^i - \overline{U}_{t+1}^i\right] = \delta\overline{\Pi}_{t+1} - \delta\sum_{i\in I_t} d_t^i$$

Note that if $W_t = \overline{U}_t^i - \delta\overline{U}_{t+1}^i + c^i(e_t)$, then $U_t^i = \overline{U}_t^i$ for all i, t and the parties can implement the effort sequence without bonding or severance payments. *Q.E.D.*

The rest of the chapter will take Proposition 2.2 as the description of feasible multilateral contracting. In the context of inter-firm transactions, it may seem odd to explicitly allow collusion between subcontractors or suppliers. But at the same time, the contracts used in Proposition 2.1 also seem somewhat unreasonable. In particular, the conclusion of Proposition 2.2 obtains without side contracting under the alternate assumption that court-enforced payments to agent i can depend only on the verifiable history of relationship i. Moreover, the conclusion of Proposition 2.2 also obtains when there is *no court-enforcement* regardless of whether the agents can collude.[7] Thus it may be an appropriate description of internal firm arrangements such as internal capital allocation and budgeting that tend to be beyond the domain of the court.

A few observations follow immediately from the (IC) condition. When the future is very important (δ near 1) and the outside options not too attractive, the first best will always be at-

[7]The court-enforced payment w_i can be made before e is chosen (see MacLeod and Malcomson (1989) on this point).

tainable. On the other hand, when the future is unimportant (δ near 0) and the outside options relatively appealing, there may be no feasible self-enforcing contracts. The case of interest is intermediate, when self-enforcing contracts are possible but the parties are still incentive-constrained. Also, the proof of sufficiency constructs an equilibrium which does not involve side payments, formal or informal, between the agents. Thus, *if* $(e_t)_{t=1}^{\infty}$ *can be achieved with some self-enforcing contract, it can be achieved without side agreements.* The firm can always assume the position of central contractual agent — simultaneously balancing the accounts of every relationship.

A final point of importance concerns the feasibility of court-enforced severance payments. Under U.S. law, the courts will not enforce large penalties for breach.[8] Nor, as many authors have pointed out, will courts enforce worker bonding. Thus, there is a strong case for not allowing severance payments in the model — and in some later sections I rule them out. Fortunately, even if eliminating severance payments, Proposition 2.2 remains valid.

2.4 Benefits of Cross-Monitoring

This section considers the problem of optimal contract design when the firm contracts bilaterally with each agent and when it fashions a multilateral agreement with all the agents at once. The possibility of multilateral contracting hinges on *observability* across relationships — if the agents cannot observe each others' inputs (or outputs), bilateral contracts are the only option. So the comparison can also be seen as capturing the difference that cross-monitoring provides. In the present setting a multilateral agreement can only improve on bilateral agreements, since any set of bilateral self-enforcing contracts will remain self-enforcing under cross-observability. In a world where uncertainty is resolved over time, this will not necessarily be the case.

2.4.1 Stationary Environments and Technologial Independence

To obtain sharp results, it will be interesting to "stack the deck" against cross-monitoring and assume both that the environment is stationary and that relationships are technologically independent. The environment is stationary if y_t, c_t, $\overline{\Pi}_t$, $\overline{U}_t$ do not depend on t. For this case, I generalize a result of MacLeod and Malcomson (1989) and show that it possible to dispense with bonding and severance payments, and more importantly, that one can always use stationary agreements to achieve constrained efficient outcomes. A stationary agreements is one which implements the same effort profile in each period using the same bonuses and formal payment plan.

[8] Restatement (Second) of Contracts, §356(1): "A term fixing unreasonably large liquidated damages is unenforceable on grounds of public policy as a penalty."

Proposition 2.3 *In a stationary environment, (i) any constrained efficient agreement implements the same unique effort profile e^* in each period; (ii) if such a profile exists, it can be implemented with a stationary agreement, and any division of the surplus can be attained without separation payments or nonstationary payments.*

I now assume that there are no *direct technological interdependencies* between relationships. Thus $y(e) = \sum_i y^i(e^i)$, and $c^i(e) = c(e^i)$ for all i.[9] If production is separable, then there are no complementarities in production. Similarly, individualistic effort means that there are no morale effects or cost externalities. When cost and production technologies satisfy these assumptions, the *only* gain from bringing workers together comes from the potential for cross-monitoring and rent-shifting.

2.4.2 The Problem of Second-Best Design

Ex ante, regardless of the division of the surplus, the parties reach an agreement that maximizes their joint surplus subject to the relevant incentive constraints. As a reference point, define the first best outcome:

$$V^{FB} = \max_e \sum_{i \in I} y^i(e^i) - c(e^i).$$

Let e^{FB} be the vector of effort levels that solves this program, and assume it is optimal for all agents to be active.

If the firm reaches separate relational contracts with each agent, or if there is no cross-monitoring, the best attainable outcome is

$$V^B = \max_e \sum_{i \in I} \left[y^i(e^i) - c(e^i) \right]$$

$$s.t.\ \forall i \in I \qquad y^i(e^i) - c(c^i) - \overline{\pi}^i - \overline{u}^i \geq \frac{1-\delta}{\delta} c(e^i), \qquad (IC^i)$$

where $\overline{\pi}^i$ is the best per period payoff the firm could get in the absence of i. A convenient assumption for the purposes of this section and the next is that for all i, there exists some $e^i \in [0, \overline{e}]$ which satisfies (IC^i).[10] Under this assumption, the unique effort profile solves the Kuhn-Tucker conditions:

$$\frac{y^{i\prime}(e^{i,B}) - c'(e^{i,B})}{c'(e^{i,B}) - \delta y^{i\prime}(e^{i,B})} = \lambda^{i,B}, \qquad (KT^B)$$

[9] Given the separability assumptions, there is no loss of generality in assuming that each agent has the same cost function.

[10] This implies that in any optimal bilateral or multilateral contracting arrangement, all agents will be active. In general, not all agents will be active — but any agent active under bilateral contracting will remain active under multilateral contracting.

where $\lambda^i \geq 0$ is the Lagrange multiplier attached to relationship i. The multiplier λ^i has a natural interpretation as a shadow price: it gives the per-period value of relaxing the incentive constraint on the firm's relationship with worker i. If $\lambda^i = 0$, then agent i will be choosing first best effort levels. Note that the left hand side of (KT^B) is decreasing in $e^{i,B}$, so the shadow price on relationship i falls as agent i's effort increases toward first best.

If the firm reaches a multilateral agreement with its agents, the parties achieve

$$V^M = \max_e \sum_{i \in I} \left[y^i(e^i) - c(e^i)\right]$$

$$\text{s.t.} \qquad \sum_{i \in I} \left[y^i(e^i) - c(e^i) - \overline{\pi}^i - \overline{u}^i\right] \;\geq\; \frac{1-\delta}{\delta} \sum_{i \in I} c(e^i). \qquad (IC^M)$$

The solution e^M uniquely solves the Kuhn-Tucker conditions:

$$\frac{y^{i\prime}(e^{i,M}) - c'(e^{i,M})}{c'(e^{i,M}) - \delta y^{i\prime}(e^{i,M})} = \lambda^M, \qquad (KT^M)$$

where now the same multiplier, $\lambda \geq 0$, attaches to each relationship. At the optimum, the shadow prices on each relationship are equalized. From an efficiency standpoint, relaxing the constraint on each relationship has equal value. Of course, when the firm write separate bilateral contracts, there is no reason to think that these shadow prices will be equal.

Proposition 2.4 *Under the above assumptions,*

(i) $V^{FB} \geq V^M \geq V^B$.

(ii) For all i, $e^{i,B}, e^{i,M} \leq e^{i,FB}$.

(iii) Either $e^M = e^{FB}$, *or* $V^{FB} > V^M$ *and for all* i, $e^{i,M} < e^{i,FB}$.

(iv) Either $V^M > V^B$ *or* $\lambda^{i,B}$ *is constant.*

The Proposition follows directly from the Kuhn-Tucker conditions. Part *(ii)* says that when the first-best is not attainable, incentives will be too weak. Part *(iii)* is an application of the general theory of the second best. It says that under cross-monitoring, the second-best optimum will distort the effort choices of *all* agents. In particular, suppose the firm could reach a bilateral contract with agent i that induced first-best effort (so $\lambda^{i,B} = 0$). Then so long as some relationship j is incentive constrained, the optimal multilateral contract will generally weaken agent i's incentives! That is, the quasirents from relationship i will generally be transferred to relationship j as the firm attempts to balance incentive provision.

The final statement of the Proposition says that the only time bilateral agreements perform as well as a multilateral contract is when the shadow prices on each separate relationship are already equalized. When this is the case, there are no gains from shifting quasirents across relationships — their marginal value is already equalized — and thus no gains from cross-monitoring. One special case when this occurs is the following:

Corollary 2.1 *Suppose production is separable and effort individualistic. Suppose relationships are symmetric, i.e.* $y^i(e) = y^j(e)$, $\overline{u}^i + \overline{\pi}^i = \overline{u}^j + \overline{\pi}^j$ *for all* $i, j \in I$. *Then* $V^B = V^M$.

Bernheim and Whinston (1990) and Bendor and Mookherjee (1990) have obtained a similar conclusion in somewhat different contexts.

2.4.3 Specific Investments and Dependence

The results above have implications for investment in specific assets and for how the level of mutual dependence influences a relationship. Consider an investment in dedicated assets for relationship i, k^i, at cost $C(k^i)$. At least two kinds of specific investments are of interest: the first are those that "deepen" a relationship, but do not affect the marginal returns to acting, i.e. $y^i(e, k^i)$ is increasing in k^i but $y_{ek} = 0$ for positive effort choices (e.g. $y^i(e) + k^i \mathbf{1}_{\{e>0\}}$). The second are specific investments that both deepen the relationship and increase i's marginal product, i.e. $y^i(e, k^i)$ is increasing in k and supermodular (y_{ek} positive). Call such an investment *complementary.*

Property rights models of organizations emphasize the importance of structural changes (i.e. the allocation of residual decision rights) that are equivalent to the second type of investment (Hart, 1995). But incentives within a relationship are *independent of the level of specificity* (see, for instance, the discussion in Holmström (1996)). Under relational contracting, investments which deepen relationships have incentive value because they increase the level of dependency and relax incentive constraints. Moreover, the value of investments depends on whether contracting is bilateral or multilateral.

Proposition 2.5 *Retain the assumptions of Proposition 2.4. Under bilateral contracting, any specific investment in relationship* i *improves incentives for* i *and has no effect on any other relationship* j. *Under multilateral contracting, a deepening of relationship* i *improves incentives for all agents; a complementary investment in relationship* i *increases* e^i *but has ambiguous effect on incentives for others.*

Proof. The result under bilateral investment is immediate. Under multilateral contracting, a deepening of relationship i relaxes (IC^M). Unless both are already first-best, either e^i must increase or some e^j must increase. Any KT^M) implies that at the optimum, for any $j \neq i$,

$$\frac{y_e^i(e^i, k) - c'(e^i)}{c'(e^i) - \delta y_e^i(e^i, k)} = \frac{y_e^j(e^j) - c'(e^j)}{c'(e^j) - \delta y_e^j(e^j)}. \tag{2.4}$$

The left hand side of (2.4) is constant in k, and decreasing in e^i and the right hand side is decreasing in e^j, so e^i must increase and so must e^j for all $j \neq i$. A complementary investment in relationship i must increase either e^i or e^j some $j \neq i$. Suppose e^j increases, which decreases the right hand side of (2.4). Since the left hand side of (2.4) is now increasing in k, e^i must also increase. So e^i must increase for sure. But the effect on e^{-i} is ambiguous. *Q.E.D.*

Relationship i can also deepen if one party's outside option, say $\overline{u}^i$, decreases. Reducing agent i's outside option creates a lock-in effect that enhances incentives. An interesting point about multilateral contracting is that deepening any relationship has exactly the same value to the group as a whole. This is immediate from the shadow price logic of above — the price attached to relaxing each constraint is identical. Moreover, even if relationship i could already sustain first-best incentives with a bilateral contract, deepening investments will generally have value so long as it is tied with other relationships.

2.4.4 Cross-Monitoring and Complementarities

When the relationships are not technologically independent, the scope for contracting with agent i depends on the firm's ability to provide incentives for agent j even when contracting is bilateral. In particular, if the agents' inputs are complements, inducing more effort from agent j will typically increase the set of feasible contracts with agent i regardless of whether contracting is bilateral or multilateral. If inputs are substitutes, however, inducing more effort from agent j can *decrease* the level of effort attainable in relationship i when contracting is bilateral. An improved ability to contract with agent j decreases the marginal gains from relationship i and improves the firm's fall-back position if it holds up i. This problem does not arise under multilateral contracting, suggesting that the gains from cross-monitoring are greatest when inputs are substitutes.

To illustrate these possibilities, consider a simplified model with two symmetric agents, a stationary technology, and no court-enforced severance payments. I assume that the firm is the ex post residual claimant on the profit stream (so $U = \overline{U}$), and moreover that $\overline{\pi} = y(\emptyset, \emptyset)$.[11] Optimal contracts might take the form of a fixed payment $w^i = \overline{u}^i$ and a discretionary payment $b^i = c(e^i)$.

Let S^i denote the surplus generated if the firm contracts optimally *only* with agent i, and S denote the surplus generated if the firm contracts bilaterally with *both* agents. Bilateral

[11]Symmetry and stationarity are assumed for convenience. Ruling out severance payments and making the firm the ex post claimant on profits are restrictive assumptions, though not hard to justify (imagine there are many firms, so a worker can always transfer to an identical situation). If the workers were ex post claimants, bilateral contracts would be as good as multilateral. The last assumption, that $\overline{\pi} = y(\emptyset, \emptyset)$ is only needed at one point, which will be pointed out.

contracts can enforce a stationary effort profile (e^1, e^2) generating S if and only if[12]

$$\delta(S - S^2 - \overline{U}^1) \geq c(e^1), \qquad (IC^1)$$

$$\delta(S - S^1 - \overline{U}^2) \geq c(e^2), \qquad (IC^2)$$

$$\delta(S - \overline{\Pi} - \overline{U}^1 - \overline{U}^2) \geq c(e^1) + c(e^2). \qquad (IC^M)$$

If the first (second) condition fails, the firm has an incentive to renege on its agreement with agent one (two). If the last condition fails, the firm has an incentive to renege on both agreements. Of course, if contracting is multilateral, Proposition 2 says that only (IC^M) must be satisfied. So multilateral contracting will be superior to bilateral contracting exactly when either (IC^1) or (IC^2) binds.

Proposition 2.6 *(i) If agents' actions are complements (y is supermodular), then bilateral contracting always does as well as multilateral contracting. And the presence of agent j improves incentives for agent i. If agents' actions are substitutes (y is submodular), bilateral contracting does strictly worse unless it achieves the first best. And the presence of agent j reduces incentives for agent i.*

A second point of interest concerns the incentive effects of a deepening of one of the relationships. Retain the assumption of symmetry in production (y symmetric) and identical costs. Now consider the effect of decreases in the parties' outside options which increase the level of dependency.

Proposition 2.7 *(i) Under multilateral contracting, a decrease in $\overline{\pi}$, $\overline{u}^i$ or $\overline{u}^j$ improves incentives for both i and j. (ii) Under bilateral contracing, a decrease in $\overline{u}^i$ improves incentives for i and improves (reduces) incentives for j if efforts are complements (substitutes). A decrease in $\overline{\pi}$ increases (decreases) incentives for both i and j if efforts are complements (substitutes).*

2.4.5 Stochastic Production and Incentive Smoothing

Any asymmetry between relationships, or competitive input supply, can cause multilateral contracting to improve on bilateral contracts. It is clear that multilateral contracting can also improve incentive provision arises if production opportunities are staggered — for instance agents 1 and 2 produce in alternate periods. In this case, only one agent at a time incurs "costs," but the future gains from both relationships are factored into (IC^M). A similar situation arises, in a less obvious way, when the production technology is stationary but subject to stochastic shocks. For instance, suppose costs are $c(e, \theta)$ where $\theta \in \Theta$ is a stochastic productivity shock —

[12] See the Appendix for a derivation.

so c is decreasing in θ (assume for simplicity that c_e is constant in θ). If agents' productivities are perfectly correlated, then multilateral contracting offers no benefits over bilateral contracting — the same incentive constraints apply for each realization of θ. But if productivities differ over time, multilateral contracting can allow a form of "incentive smoothing." The fact that some agents can have high productivity while others have low introduces an asymmetry so incentives in these periods are better under multilateral than under bilateral contracting. Moreover, this extra surplus translates into slack incentive constraints in the symmetric periods, improving incentives. The same effect can arise if the firm does not observe effort and output is stochastic. The self-enforcement constraint means that *total compensation* is limited. By using relative performance evaluation, the firm can provide incentives with a low total bonus.

2.5 Implications of Multilateral Contracting

In this section I explore two implications of the above results. First, I explore the possibility that multilateral contracting can provide a rationale for concentrating decision rights in the hands of large actors. I then look at the use of prior relationships as leverage in designing new relationships.

Both examples focus on the structure of incentive provision. If effort profile e satisfies (IC^M), there are many payment schemes that can implement e. For instance, a pre-specified payment of $(1-\delta)\overline{U}^i$ coupled with a bonus $c^i(e)$ gives the firm an opportunity to "hold up" the worker ex post by withholding the bonus. On the other hand, a high fixed payment with a built in premium for effort $(1-\delta)(\overline{U}+c^i(e)/\delta)$ with termination or a penalty of $c^i(e)$ if the agent holds up the firm also implements e, but by giving the agent residual claimancy status and the opportunity to hold up the firm. The point is that the party with the larger claim on residual profits from the relationship will be less likely to renege and therefore should be entrusted with hold-up opportunities. Of course, separation penalties for breach can shift these hold up incentives — I rule these out a priori.

2.5.1 Firm as Residual Claimant

Because the firm is central to all contracts, it is natural to conjecture that it will have the "most to lose" from a breakdown. From the above argument, the party with the most to lose has the greatest incentives to behave reciprocally and can be granted hold-up opportunities without endangering cooperation. Relational contracting might thus provide a rationale for an empirical regularity that is hard to reconcile with existing property rights models of organizations — namely the fact that assets are extremely clustered rather than spread out among those with important specific investment decisions.

This idea can be explored within the context of the general model. Suppose there are two agents and that production opportunities are staggered. By Proposition 2.2, to implement effort profile (e^1, e^2), I require that for every period t in which agent 1 acts,

$$\delta(S_{t+1} - \overline{S}) \geq c(e^1).$$

Disaggregating this condition:[13]

$$\begin{aligned} \delta(U^1_{t+1} - \overline{U}^1) &\geq c(e^1) - b^{1f} \\ \delta(U^2_{t+1} - \overline{U}^2) &\geq 0 \\ \delta(\Pi_{t+1} - \overline{\Pi}) &\geq b^{1f} \end{aligned}$$

An optimal contract must equalize the shadow prices on each constraint — so all these inequalities must be satisfied with equality. Thus it will be optimal to set $U^2_{t+1} = \overline{U}^2$. But since $U^i_t \geq \overline{U}^i$ must hold in every period, this is possible only if $U^2_t = \overline{U}^2$ for every t. A similar argument applies for agent 1, implying that $U^1_t = \overline{U}^1$. Thus the firm will optimally be residual claimant on the surplus after any *ex ante* redistributive payments are made. The idea is straightforward: since all of the firm's relationships are contingent on total compliance with the agreement, and the firm's reneging opportunities are spread out, the firm has strong incentives to behave responsibly in any given decision. The model is still a step away from asset concentration, but only a small one. If one believes that residual control (property rights) over assets should generally be associated with uncontracted on claims on their associated revenue streams, then the firm, as residual claimant on the profit stream, should also hold any specific assets.

2.5.2 Prior Relationships as Leverage

The results above imply that rent-shifting between relationships can be used to expand the set of attainable input levels. Tying contracts together improves monitoring, allows for more dramatic punishments and hence relaxes incentive constraints. It turns out that even if cross-monitoring does not allow the firm to provide *greater* incentives, it may change the *type* of available incentives. This in turn, may allow the firm to extract a large share of the surplus. I illustrate this in the context of a general equilibrium efficiency wage model.

In this example, there are L type B workers and one type A worker (L is large). At cost k a manager can install a production technology that produces a single perishable good $y = pe^B$, where $e^B \in \{0, 1\}$ is effort supplied by a type B worker. Exerting effort costs c^B. Suppose that each firm becomes obsolete from one period to the next with exogenous probability $1 - \gamma$. Ignoring the type A worker, an efficiency wage equilibrium of this economy will have a fixed

[13] I disregard the possibility that the agents will make transfer payments.

wage w (I again rule out severance payments). To sustain effort, there must be unemployment in equilibrium. Let J be the total number of jobs in the economy, and assume that vacant firms can immediately find an unemployed worker. Equilibrium can be characterized by the following equations:

$$\begin{aligned} U - \overline{U} &= c^B/\delta \\ \overline{U} &= \frac{J - \gamma J}{L - \gamma J} U + \frac{L - J}{L - \gamma J} \delta U \\ U &= \frac{w - c^B}{1 - \delta\gamma} + \frac{\delta(1-\gamma)}{1 - \delta\gamma} \overline{U} \\ \Pi &= \overline{\Pi} = k = \frac{p^B - w}{1 - \delta\gamma} \end{aligned}$$

where the first condition is the incentive compatibility condition for workers, the second and third are the recursive equations for workers' utilities, out of and in employment, and the last is the zero-profit entry equation. Equilibrium values can be easily worked out. The point to note is that when a firm and a worker meet, the worker takes *all* the surplus from the match. Since there is equilibrium unemployment, the firm has no incentive to ever make a bonus payment. So the most efficient relational contract for any pair involves giving all the surplus to the worker in the form of an efficiency wage. Free-entry ensures that only the most efficient arrangement survives in equilibrium.[14]

Now suppose that exactly one firm has a second technology $y^A = p^A e^A$, where $e^A \in \{0, 1\}$ is effort by a type A worker, which costs c^A. Suppose that $p^A - c^A$ is large, so that this firm, by virtue of its idiosyncratic technology, is able to successfully contract for effort with the one type A worker. Since there is only one type A worker and one firm with this technology, the quasi-rent from the relationship is

$$(p^A - c^A)/(1 - \delta\gamma) >> \delta\gamma c^A. \tag{2.5}$$

I now claim that this firm can leverage its relationship with worker A to extract rents from a type B worker it meets in the market. Since efficient effort was feasible before, having worker A around doesn't change the total surplus generated from a worker B but it can change the allocation of this surplus. In particular, suppose A threatens to quit if the firm ever misses a bonus payment to worker B. Since the potential loss from reneging is now very large, this firm can now offer its type B worker performance pay rather than an efficiency wage. This is something that *no other firm can do*. Offered a performance pay contract that specifies a bonus c^B, B will be willing to accept it, and exert effort if $\hat{U} = \overline{U}$, where $\hat{U}$ is the lifetime utility from

[14] MacLeod and Malcomson (1998) is a good reference for these sorts of efficiency wage models.

being in employment with this firm.

$$\hat{U} = \overline{U} = \frac{\hat{W} - c^B}{1 - \delta} \tag{2.6}$$

where $\hat{W} = \hat{w} + \hat{b} = \hat{w} + c^B$ is the overall wage rate. But then, in equilibrium

$$\hat{W} - W = (1 - \delta\gamma)c^B/\delta \tag{2.7}$$

Having worker A around to enforce bonus payments translates into substantially lower wages (in particular, no efficiency wage). To compute the whole equilibrium, one needs to account for the general equilibrium effect of having this one firm use a different incentive package (it will immediately lower $\overline{U}$, but then entry will drive will drive it back up). Profits are also greater, since more of the surplus can be extracted from the relationship:

$$\hat{\Pi} = k + c^B/\delta > k = \Pi. \tag{2.8}$$

2.6 Costs of Cross-Monitoring

In the model considered thus far, there is no downside to tying relational contracts together. Cross-monitoring can only expand the set of available outcomes. Nevertheless, it is reasonable to suppose that there are strong reasons for keeping certain relationships separate, even when cross-monitoring might help with enforcement. Here, I discuss one reason why this might be so — namely, that the release of information from one relationship could impair the workings of other relationships. By keeping the relationships separate, the firm gives up the gains from cross-monitoring, but may also avoid the chance of complete breakdown.

To formalize this, I extend the model to allow for the possibility of an adverse shock which causes one relationshp to collapse. The shock also reveals information about the future productivity of the firm's other relationships — if this information becomes common knowledge between the firm and its other agents (as it will if the relationships are tied), cooperation may break down. By separating the relationships, the firm can limit the effect of bad news. In the model I present below, there are also gains from cross-monitoring — full cooperation can be sustained initially only with cross-monitoring. So the firm must trade off between the gains from cross-monitoring and the costs of information linkage.

The idea that observability across relationships brings with it the potential for costly information spillovers has been explored by Fudenberg and Kreps (1987). They show using a chain store model that informational linkage between two markets can help or hurt an incumbent who is trying to maintain a reputation for "toughness." While the model I give below is somewhat different, the effects are similar.

2.6.1 Costly Informational Externalities

To make this as clear as possible, I again work with a simplified version of the general model. Consider a firm with two agents, 1 and 2. Agent i chooses effort $e^i \in \{0,1\}$, at cost $c^i e^i$, and produces $y^i(e^i) = p^i e^i$. Suppose that $p^i > c^i$, so it is optimal for both workers to exert effort, and normalize the outside options to zero. I allow for information externalities by adding a stochastic element to the model. With poisson probability $1-\gamma$, a technological shock occurs which changes the state of the world and immediately lowers the marginal productivity of the second agent; in particular $p^2 = 0$ following the shock. With probability q, this shock has no implications for the productivity of the first relationship and p^1 remains unchanged forever. With complementary probability $1-q$, the productivity of relationship 1 becomes endangered and with poisson probability $1-\phi$ the firm and agent 1 become unproductive. The consequence of the initial shock for relationship 1 is learned only by those who observe the inner workings of relationship 2.

I make the following Assumptions, discussed below:

$$\frac{1}{\delta\gamma}c^1 < p^1 < \frac{1}{\delta\phi}c^1 \quad (A1)$$

$$c^2 < p^2 < \frac{1}{\delta\gamma}c^2 \quad (A2)$$

Cross-Monitoring Contract

Consider an optimal contracting arrangement under cross-observability. Following a shock which ends relationship 2, Assumption $(A1)$ implies that relationship 1 will be able to continue if and only if the shock does not threaten relationship 1. So following a shock which ends relationship 2, an optimal relational contract will yield surplus $(p^1-c^1)/(1-\delta)$. Using Proposition 2, this implies that effort can be induced from both agents from the start if and only if

$$\delta\left[\gamma S(1,1) + (1-\gamma)q\frac{p^1-c^1}{1-\delta}\right] \geq c^1 + c^2 \quad (2.9)$$

where

$$S(1,1) = \frac{p^1-c^1}{1-\delta\gamma} + \frac{p^2-c^2}{1-\delta\gamma} + \frac{\delta(1-\gamma)}{1-\delta\gamma}q\frac{p^1-c^1}{1-\delta}. \quad (2.10)$$

So full cooperation is possible whenever

$$p^1 - \frac{1}{\delta\gamma}c^1 + \frac{1-\gamma}{\gamma}q\frac{p^1-c^1}{1-\delta} + p^2 - \frac{1}{\delta\gamma}c^2 \geq 0 \quad (2.11)$$

Supposing (2.11) holds, equilibrium play runs as follows. Both agents start off exerting effort. Eventually, a shock arrives which affects relationship 2, and the relationship terminates (or equivalently, continues with low effort). If the shock threatens relationship 1, then that relationship terminates as well. Note that this termination is suboptimal in the sense that relationship might well remain productive for some time. If the shock does not affect agent 1, he continues exerting effort forever. The total surplus is $S^M = S(1,1)$.

Bilateral Contracting with Symmetric Learning

Suppose instead that the firm reaches separate relational contracts with each agent and cross-observability is impossible. By Assumption ($A2$) and Proposition 2.2, the firm *cannot* induce effort from agent 2. Without cross-monitoring from agent 1, the incentives to renege are too high — despite effort being efficient, no self-enforcing agreement can be found to make it feasible.

Nevertheless, this arrangement may offer better contracting opportunities for the firm and agent 1. To see how this might occur, suppose that neither the firm nor agent 1 learns when the original shock occurs (this might happen either because the firm never hires agent 1, or because the information is only revealed if agent 1 is choosing high effort). In this case, at any given time, the probability that p^1 will remain unchanged is at least $q + (1-q)\gamma$. Therefore cooperation can be maintained from the beginning if[15]

$$p^1 \geq \frac{1}{\delta\left[\gamma + (1-\gamma)(q + (1-q)\phi)\right]} c^1 \tag{2.12}$$

If this condition holds, then

$$S^{1,B} = \frac{p^1 - c^1}{1-\delta\gamma} + \frac{\delta(1-\gamma)}{1-\delta\gamma}\left[q\frac{p^1 - c^1}{1-\delta} + (1-q)\frac{p^1 - c^1}{1-\delta\phi}\right].$$

A sufficient condition for the firm to prefer bilateral contracting is that $S^{1,B} \geq S^M$, i.e.

$$(1-q)\frac{\delta(1-\gamma)}{1-\delta\gamma}\left(p^1 - c^1\right) \geq p^2 - c^2 \tag{2.13}$$

It is easy to find parameter values that satisfy this condition, without violating ($A1$), ($A2$), (2.11), or (2.12). Notice that (2.13) is more likely to obtain if the relationships are more asymmetric, i.e. if $p^1 - c^1$ is much greater than $p^2 - c^2$. This is because by opting for a single relationship with agent 1, the firm avoids any informational spillover, but also gives up cooperation with agent 2.

[15] More generally, the optimal relational contract would involve a period of buildup, where $e^1 = 0$ at the beginning. If some time passed, and the relationship still appeared productive (i.e. p^1 was unchanged), then $e^1 = 1$ could be maintained (since the firm and worker 1 would become more optimistic that the shock had no adverse consequences for their relationship.

If (2.13) holds, the firm does better to pass up a relationship with agent 2, and concentrate on worker 1. However, I assumed that *neither* agent 1 *nor* the firm could track p^2. This seems plausible if the firm foregoes entirely a relationship with agent 2, but less likely if the firm conducts both relationships separately. Then, following a shock to agent 2's productivity, only the firm would realize the ramifications for agent 1. In this case, following a shock that affects p^2, there is asymmetric information: only the firm knows the true probability that relationship 1 will remain productive. It turns out that by keeping the agent uninformed, cooperation may be possible.

To demonstrate this possibility, consider the following efficiency wage contract. The firm pays agent 1 $w = p^1$ in every period the relationship is productive. Agent 1 is expected to exert effort. Following any deviation or attempt to renegotiate, the parties immediate punish each other by separating (i.e. playing the worst possible continuation equilibrium). Agent 1 will exert effort in every period, even following an observed breakup of relationship 2 if

$$\delta\left[q\frac{1}{1-\delta}+(1-q)\frac{1}{1-\delta\phi}\right](p^1-c^1)\geq c^1, \tag{2.14}$$

which can be re-written as

$$p^1\geq\frac{1+\delta\left[q\frac{1}{1-\delta}+(1-q)\frac{1}{1-\delta\phi}\right]}{\delta\left[q\frac{1}{1-\delta}+(1-q)\frac{1}{1-\delta\phi}\right]}c^1. \tag{2.15}$$

Assuming this obtains, surplus will again be $S^{1,B}$, so the firm will again prefer bilateral contracting if (2.13) holds. And again, it is easy to find parameters that satisfy $(A1), (A2)$, (2.11), and (2.15).

To summarize the findings from this section:

Proposition 2.8 *In the model with learning, multilateral contracting can be strictly worse than conducting bilateral contracting despite there being gains from cross-monitoring.*

This extension of the basic model suggests that linking relationships may have costs as well as benefits. The story is perhaps most natural if the model is given a reputational interpretation. Then multilateral contracting allows for cross-monitoring, but it also exposes the firm to reputational damage if one of the relationships breaks down. By contracting bilaterally, the firm can secure its important relationships.

2.7 Applications

My emphasis on multilateral contracting arrangements would be misguided were cross-monitoring relevant and important empirically. In fact, the formal results are consistent with a number of organizational features that have been stressed in descriptive studies. In this section, I briefly suggest three — the employment contract and the recent increase in two-tier work arrangements, supplier associations in Japanese subcontracting systems, and the prevalence of diversified business groups in developing countries.

2.7.1 The Employment Contract and Two-Tier Workforces

An oft-cited point in descriptive studies of internal labor markets is that employers find it extremely difficult to selectively cut wages or fire workers (Doeringer and Piore, 1971; Bewley, 1997). These authors usually cite "morale concerns" or "workplace norms" as the source of this difficulty — firing a subset of workers will alienate the others, which will translate to a loss of motivation. However, most models of employment take the match between a worker and an individual job as the primitive unit of study. So there is no room for formal analyses of this type of cross-monitoring. As the model illustrates, optimal relational contracting will typically call for the firm's implicit agreements with different workers to be interlinked. The existence of general workplace norms, and the observed difficulty of selective wage cuts or layoffs, are consistent with this interweaving of employment agreements.

The extension of the model in the previous section also suggests a story for one of the most widely-observed trends in U.S. employment — the move toward "temporary work arrangements." If workforce adjustment is seen as a violation of the relational contract between the firm and its long-term workers, explicitly establishing a second tier of workers — with no expectation of long tenure or job security, but perhaps with high current wage compensation — can be a sensible strategy. Why the current trend? The model suggests that greater asymmetry, while it raises the returns to cross-monitoring, can also lead to separation of relationships. If core workers are more important now that before, or if skill-biased technological change has magnified asymmetries in schooling or training, firms may want to "protect" their relationships with core or high-skilled workers, at the expense of relationships with peripheral or low-skilled labor.

2.7.2 Supply Associations in Japan

The relational aspects of Japanese supply systems have been emphasized by many authors. These arrangements have important aspects of cross-monitoring and group enforcement. An example is the widely studied and highly regarded supply system run by Toyota. Toyota's ties with some of its suppliers go back sixty years to the company's inception. As Williamson

(1985) has pointed out, an important feature of Toyota's supply system is its organized supplier associations. Williamson (1985, p. 121) describes the supplier associations as a "safeguard against opportunism" — a formal mechanism to promote the swift and accurate cross-reporting of reputationally-relevant information. Dore (1983) finds similar relational contracting arrangements taking place in the large Japanese enterprise groups, or *keiretsu.* He notes that "the only thing which formally defines the identity of the group is the lunch on the last Friday of the month when the Presidents of every company in the group get together" (p. 467).

2.7.3 Diversified Business Groups

A remarkable share of economic activity in developing countries is accounted for by widely diversified, and typically family-owned, "business groups." For instance, Ghemawat and Khanna (1998) report that in the late 1980s, more than two thirds of industrial assets in India were owned by the largest 20 business groups, of which more than two thirds were family-owned. Two frequent explanations for this phenomenon are the lack of available external financing and the possibilitiy that there might be a sort of "returns to scale" in bribing corrupt government officials (Ghemawat and Khanna). The model laid out here provides another (complementary) explanation because it identifies a broad economy of scope in contracting when court-enforcement is lacking. The minimal court enforcement of contracts in developing countries, and some of the resulting problems, is described in McMillan (1995). The results here suggest that given these circumstances, there may be gains from bringing together even technologically unrelated businesses under a single umbrella. Moreover, since the possibilities for multilateral contract enforcement depend on agent's j willingness to punish the firm for holding up i, there should be distinct advantages to explicity linking businesses within a group to a common (family) name. Common ownership, if not observed, seems far less likely to allow for effective cross-monitoring. The model also suggests that the returns to extremely broad diversification will shrink as contract law takes firmer root.

2.8 Conclusion

I began this chapter by arguing that a general theory of organizational design needs to explicitly account for the fact that firms firms typically are engaged in multiple ongoing transactional relationships. Focusing on marginal transactions at the firm boundary, or studying individual relationships in isolation cannot provide a complete picture of the firm, because even if the various relationships are technologically independent, they are likely to be linked by cross-monitoring and reputational concerns. As a step in this direction, I have investigated how a firm might optimally design relational contracts when it cares about maintaining cooperative outcomes with many agents. I have argued that when the agents can cross-monitor, the firm will be able to shift quasirents across relationships and improve incentives relative to a situation where relationships are conducted separately using bilateral agreements. The structure of optimal multilateral contracts has implications for investments in dedicated assets, for the coordination of decisions, for residual claimancy status. I have also suggested that multilateral contracting may have costs as well as benefits. Namely, tying relationships together may increase information flows and lead to situations where the breakdown of a single relationship can endanger many others. Bilateral contracting can avoid this fragility, at the cost of giving up the incentive gains from cross-monitoring.

2.9 Appendix

Proof of Proposition 2.1. ($\Rightarrow$) Let $(e_1, e_2, ...)$ be given, and suppose there is some $(s, \sigma, c) \in \Omega^{*0}$ that implements it. Then all parties must agree to initiate the contract, so for all $i \in I_1$, $U_1^i \geq \overline{U}_1$ and $\Pi_1 \geq \overline{\Pi}$, and thus $S_1 \geq \overline{S}_1$. Moreoever, suppose from time $\hat{T}$ there is a single worker i. Then this worker must prefer to exert effort e_t^i at t rather than set effort to zero and quit immediately, and the firm must prefer to pay any promised bonus rather than reneging and firing the worker:

$$\begin{aligned} -c^i(e_t) + b_t + \delta U_t^i &\geq \delta \overline{U}_t^i + \delta d_{t+1}^i \\ -b_t + \delta \Pi_{t+1} &\geq \delta \overline{\Pi}_{t+1} - \delta d_{t+1}^i. \end{aligned}$$

Summing these two constraints yields (IC^1).

($\Leftarrow$) Let $(e_1, e_2, ...)$ be given satisfying $S_1 \geq \overline{S}_1$ and (IC^1). Note that this gives $I_1, I_2, ...$ Let W_1 be date one wages, and suppose for any $i \in I_1$, relationship i begins (this will be verified momentarily). I construct an agreement recursively. Suppose that for all $\tau \leq t$, $\hat{I}_\tau = I_\tau$, and $\hat{W}_\tau^i = W_\tau^i$ where hatted values are those the court has observed, and unhatted values are specified in the contract. Let $w_t^i = \overline{U}_t^i - \delta \overline{U}_{t+1}^i$ for all $i \in I_t$. I now proceed by cases. If $|I_t| = \{i\}$, let $d_{t+1}^i = 0$, informally the parties agree to effort e_t a bonus $b_t^i = c^i(e_t)$ if the agent chooses e_t^i and zero otherwise, and to separate at $t+1$ if the worker shirks of the firm withholds the bonus. Then $W_t^i = w_t^i + c^i(e_t)$. If $|I_t| > 1$, but $I_{t+1} = \emptyset$, the informal agreement specifies effort e_t, and bonuses $b_t^i = c^i(e_t)$, where if i shirks, then b^i is paid to the agent in I_t with lowest index, if that agent shirks, his bonus goes to the agent with second-lowest index, and if all agents shirk, all bonuses go to agent 1. The formal contract specifies $d_{t+1}^i = 0$ if $\sum \hat{W}_t \geq \sum W_t$ and if $\sum W_t - \sum \hat{W}_t > 0$, this is paid in severance to the agent with lowest index, the others get zero. If $|I_t| > 1$ and $I_{t+1} \neq \emptyset$, the informal agreement specifies effort e_t, with the same bonus payments. Severance payments at $t+1$ are specified as follows: if $\hat{I}_{t+1} = I_{t+1}$, then $d_{t+1}^i = 0$; if $\hat{I}_{t+1} = \emptyset$, then all agents who separated get $d_{t+1} = -k$, where k is "large," except the agent in I_t with lowest index, who gets $d_{t+1} = K > |I|k$. And if $\hat{I}_{t+1} = I_{t+1} - \tilde{I}$, some strict subset of I_{t+1}, then all agents who separated get $d_{t+1} = -k$, and those who remain get $w_{t+1} = 0$, and $d_{t+2} = K/\delta$. Finally, if $\hat{I}_{t+1} = I_{t+1}$, but $\sum W_t - \sum \hat{W}_t \neq 0$, then arrangements are identical only the agent in I_{t+1} with lowest index has his fixed salary augmented by $\sum W_t - \sum \hat{W}_t$.

Supposing that the parties initiate this contract, this contract enforces $(e_1, e_2, ...)$. To see this, note that if $|I_t| = 1$, the agent is happy to exert effort and remain in employment, and the firm is happy to pay the bonus and retain the agent because its payoff from doing so is larger than from withholding the bonus:

$$-b_t^i + \delta\Pi_{t+1} = -c^i(e_t) + \delta(S_{t+1} - \overline{U}_{t+1}^i) \geq \delta\overline{\Pi}_{t+1}$$

by IC^1. And of course, separation is a credible threat because it is the static nash equilibrium. If $|I_t| > 1$, the agents are again willing to exert effort, and the firm willing to pay the specified bonuses. Moreover, at time $t+1$, the parties are willing to make the correct separation decisions provided the "off-path" severance payments are chosen large enough. So I only need to verify that there is some time 1 fixed wage which will induce the parties to initiate the contract. Since the contract will generate surplus $S_1 \geq \overline{S}_1$, such wages must certainly exist. One can verify as well that this contract is robust to one-time group deviations by subsets of agents.[16] *Q.E.D.*

Proof of Proposition 2.3. *(i)* For stationary effort e, the incentive compatibility condition from above can be re-written as

$$y(e) - \frac{1}{\delta}\sum_{i \in I_1} c^i(e) - \sum_{i \notin I_1} \overline{u}^i - \overline{\pi} \geq 0$$

where $\overline{u}^i = (1-\delta)\overline{U}^i$, $\overline{\pi} = (1-\delta)\overline{\Pi}$. The problem of finding an optimal stationary agreement is to maximize $y(e) - \sum_{i \in I} c^i(e)$ subject to (IC). From the above assumptions, the objective is strictly concave and the constraint set convex; assuming there is some $e \neq (\emptyset, .., \emptyset)$ which satisfies (IC), the program has unique solution e^*.

(ii) To implement a stationary profile e without bonding or severance payments, consider punishing any deviation by complete breakdown, i.e. no working, no bonuses, immediate quitting and firing. Under this scenario, the firm will be willing to pay any bonus which satisfies:

$$\sum_{i \in I} \hat{b}^i \leq \frac{\delta}{1-\delta}\left[y(\hat{e}) - \sum_{i \in I} \hat{W}^i - \overline{\pi}\right], \tag{2.16}$$

For workers not to deviate from the proposed equilibrium, no subgroup $\tilde{I} \in \{i : e_i \neq \emptyset\} = I_1$

[16]The agreement is not robust to long-term cooperative agreements by the agents. If the agents can reach informal side-agreements, but cannot write legally enforceable side-contracts, I believe any effort profile can be enforced so long as $S_t \geq \overline{S}_t$ for all t.

must want to deviate by shirking:

$$\sum_{i\in\bar{I}}\hat{b}^i \geq \frac{\delta}{1-\delta}\sum_{i\in\bar{I}}\left[\frac{1}{\delta}c^i(\hat{e}) - \hat{W}^i + \overline{u}^i\right]. \tag{2.17}$$

Note that a necessary and sufficient condition for (2.17) is in fact that no individual worker i want to shirk. Combining the incentive conditions and re-arranging terms, one sees that a sufficient condition for e to be implemented without bonding is that

$$y(\hat{e}) - \sum_{i\in I}\frac{1}{\delta}c^i(\hat{e}) - \overline{\pi} - \sum_{i:e_i\neq\emptyset}\overline{u}^i \geq 0. \tag{2.18}$$

But this is exactly condition (IC) from Proposition 2.2, that is, the necessary and sufficient condition for $\hat{x}$ to be implementable. *Q.E.D.*

Derivation from Section 4.4. For agent i not to renege, it must be the case that

$$-c(e^i) + b^i + \delta U^i \geq \delta\overline{U}^i + \delta d^i$$

For the firm not to renege on worker i alone requires

$$-b^i + \delta\Pi \geq \delta\Pi^j - \delta d^i$$

and for the firm not to renege on both workers:

$$-b^i - b^j + \delta\Pi \geq \delta\overline{\Pi} - \delta d^i - \delta d^j$$

Assuming that $U^i = \overline{U}^i$, this shows the conditions in the text are necessary. To show sufficiency, consider fixed wage payments $w^i = \overline{u}^i$, no severance payments, and bonuses $b^i = c(e^i)$, and punishing any deviation with no bonus payments and immediate termination of that relationship. Following the termination of relationship i, the firm and j recontract optimally. It is easy to show that if e satisfies the conditions in the text, this agreement will be self-enforcing. Note that the derivation of the conditions in the text relies on hte firm being the residual claimant. If the agents are the *ex post* residual claimants, so $\Pi \equiv \overline{\Pi}$, the above conditions reduce to the single multilateral contract condition that $\delta(U^i + U^j - \overline{S}) = \delta(S - \overline{S}) \geq c(e^1) + c(e^2)$. *Q.E.D.*

Proof of Proposition 2.6. Let $\hat{e}^{FB}$ denote the first best if the firm contracts with only one agent, and (e^{FB}, e^{FB}) denote the first best if if contracts with both. Note that if efforts are complements, the $e^{FB} \geq \hat{e}^{FB}$, and the opposite holds if efforts are substitutes (by Topkis' Theorem). Suppose that if the firm contracts with agent i alone, it can at best implement effort

$\hat{e} < \hat{e}^{FB}$. To implement (e^1, e^2) bilaterally requires

$$y(e^1,e^2) - \frac{1}{\delta}c(e^1) - c(e^2) - \overline{u}^1 - y(\emptyset,\hat{e}) + c(\hat{e}) \;\geq\; 0 \qquad (IC^1)$$

$$y(e^1,e^2) - c(e^1) - \frac{1}{\delta}c(e^2) - \overline{u}^2 - y(\hat{e},\emptyset) + c(\hat{e}) \;\geq\; 0 \qquad (IC^2)$$

$$y(e^1,e^2) - \frac{1}{\delta}c(e^1) - \frac{1}{\delta}c(e^2) - \overline{u}^1 - \overline{u}^2 - \overline{\pi} \;\geq\; 0 \qquad (IC^M)$$

The last condition can be re-written:

$$\left\{y(e^1,e^2) - \frac{1}{\delta}c(e^1) - \overline{u}^1 - y(\emptyset,\hat{e})\right\} + \left\{y(\emptyset,\hat{e}) - \frac{1}{\delta}c(e^2) - \overline{u}^2 - \overline{\pi}\right\} \geq 0 \qquad (IC^{M*})$$

Letting $e^2 = \hat{e}$, the second term will equal zero. If y is supermodular, the first term is positive if $e^1 \leq \hat{e}$. So if efforts are complements, then the firm can implement a symmetric contract with (e,e), $e > \hat{e}$. Moreover, for a contract to implement (e^1, e^2), with $e^i > \hat{e}$, it need only satisfy (IC^M), since subtracting the left hand side of (IC^i) from the left hand side of (IC^M) gives

$$y(\emptyset,\hat{e}) - c(\hat{e}) - \overline{u} - \overline{\pi} - \frac{1-\delta}{\delta}c(e^i) < 0.$$

So if y is supermodular, the only relevant constraint is (IC^M) and bilateral contracting is equivalent to multilateral contracting. It is easy to check that the optimal contract is symmetric and implements $e^i > e$, so the presence of agent j always improves incentive provision for agent i.

If y is submodular, then (IC^{M*}) fails to hold for (e^1, e^2), $e^i \geq \hat{e}$. So the best symmetric contract must implement $e < \hat{e}$ under both bilateral and multilateral contracting. Moreover, subtracting the left hand side of (IC^i) implies

$$y(\emptyset,\hat{e}) - c(\hat{e}) - \overline{u} - \overline{\pi} - \frac{1-\delta}{\delta}c(e^i) > 0.$$

So if y is submodular and the firm implements a symmetric contract under bilateral contracting, the relevant (binding) constraints are (IC^1) and (IC^2) and (IC^M) is slack. So if symmetric contracts are chosen, then $e^B < e^M < \hat{e}$.

I still need to verify that symmetric contracts are optimal. It is easy to verify that the optimal multilateral contract is symmetric so long as the interiority condition is satisfied — in this case, e^i must satisfy the Kuhn-Tucker condition

$$y_i(e^i, e^j) - c'(e^i) = \frac{\lambda}{1+\lambda}\frac{1-\delta}{\delta}c'(e^i)$$

and so $e^i = e^j$ regardless of whether y is supermodular or submodular. The optimal bilateral contract under complementarity is characterized by the same condition, so it must be symmetric. When contracting is bilateral and actions are substitutes, the Kuhn-Tucker conditions are

$$y_i(e^i, e^j) - c'(e^i) = \frac{\lambda^i}{1+\lambda^i+\lambda^j}\frac{1-\delta}{\delta}c'(e^i)$$

so long as $e^i > 0$. If $e^i > e^j$ then (IC^j) cannot bind, so $\lambda^j = 0$. But then clearly $e^j = 0$. So the optimal bilateral contract is either symmetric or involves contracting with only a single agent, i.e. $(\hat{e}, \emptyset)$. *Q.E.D.*

Proof of Proposition 2.7. *(i)* The optimal multilateral contract will be symmetric regardless of changes in $\overline{u}^i + \overline{u}^j + \overline{\pi}$, so a decrease which relaxes (IC^M) will increase $e^i = e^j$ regardless of the sign of y_{ek}. *(ii)* Now consider bilateral contracts. Let $\hat{e}^i, \hat{e}^j$ denote the optimal effort levels when there is a single agent. First suppose efforts are complements. I show that only (IC^M) will bind. Either $e^i > \hat{e}^i$ or $e^j > \hat{e}^j$ (otherwise (IC^i) and (IC^j) both bind for $e^i \leq \hat{e}^i$ and $e^j \leq \hat{e}^j$ which is impossible with y supermodular). Suppose $e^i > \hat{e}^i$. Then if (e^i, e^j) satisfy (IC^M), (IC^j) must be strictly satisfied, so (IC^j) cannot bind. On the other hand, if $e^j < \hat{e}^j$ then (IC^M) must be slack if (IC^i) binds. But it cannot be optimal to have $e^j < \hat{e}^j$ and only (IC^i) binding, since increasing e^j will relax (IC^i) and increase the surplus. Thus $e^i \geq \hat{e}^i$, $e^j \geq \hat{e}^j$, strict for either i or j and only (IC^M) is binding. The result follows immediately. Now suppose that efforts are substitutes. An analagous argument shows that (IC^M) will be slack. First, it cannot be the case that $e^i \geq \hat{e}^i$ and $e^j \geq \hat{e}^j$, since then (IC^M) implies

$$\begin{aligned}
& y(e^1, e^2) - \frac{1}{\delta}c(e^1) - \frac{1}{\delta}c(e^2) - \overline{u}^1 - \overline{u}^2 - \overline{\pi} \\
= \; & \{y(e^1, e^2) - y(e^1, \emptyset) - y(\emptyset, e^2) + y(\emptyset, \emptyset)\} + \{\overline{\pi} - y(\emptyset, \emptyset)\} \\
& + \left\{y(e^1, \emptyset) - \frac{1}{\delta}c(e^1) - \overline{u}^1 - \overline{\pi}\right\} + \left\{y(\emptyset, e^2) - \frac{1}{\delta}c(e^2) - \overline{u}^2 - \overline{\pi}\right\} \geq 0,
\end{aligned}$$

an impossibility since each of the four terms in the second expression must be negative (here is where the $\overline{\pi} = y(\emptyset, \emptyset)$ assumption comes in). Suppose that $e^i < \hat{e}^i$ but that $e^j > \hat{e}^j$. Then if (IC^M) is satisfied, (IC^i) will be slack. And if (IC^j) is satisfied, (IC^M) is slack, so the only relevant constraint is (IC^j), an impossibility since increasing e^i relaxes that constraint. So $e^i \leq \hat{e}^i$, $e^j \leq \hat{e}^j$ with one inequality strict, and (IC^M) is slack. Now a decrease in $\overline{\pi}$ increases both $\hat{e}^i$ and $\hat{e}^j$, tightening (IC^i) and (IC^j), raising λ^i, λ^j and lowering e^i, e^j. Finally, letting $\overline{s}^i$ be the surplus from a bilateral relationship with i alone, then for any value of $\overline{u}^i$, at an optimum $\overline{u}^i + \overline{s}^j + \frac{1-\delta}{\delta}c(e^i) = \overline{u}^j + \overline{s}^i + \frac{1-\delta}{\delta}c(e^j)$. A decrease in $\overline{u}^i$ (and corresponding increase in $\overline{s}^i$ must either increase e^i or decrease e^j or both. But if e^j decreases, it is clear that e^i must rise, and vice-versa. *Q.E.D.*

2.10 References

Abreu, Dilip (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

Alchian, Armen and Harold Demsetz (1972): "Production, Information Costs, and Economic Organization," *American Economic Review*, 62, 777–795.

Baker, George, Robert Gibbons and Kevin J. Murphy (1997): "Relational Contracts and the Theory of the Firm," Mimeo, M.I.T.

Barnard, Chester (1938): *The Functions of the Executive,* Cambridge:.Harvard University Press

Bendor, Jonathan and Dilip Mookherjee (1990): "Norms, Third-Party Sanctions, and Cooperations," *Journal of Law, Economics, and Organization," 1, 33-63.*

Bernheim, B. D. and Michael Whinston (1990): "Multimarket Contact and Collusion," *Rand Journal of Economics*, 2, 1–26.

Bewley, Truman (1999): *Why Wages Don't Fall During a Recession*, forthcoming, Cambridge: Harvard University Press.

Bull, Clive (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102, 147–159.

Coleman, James (1990): *Foundations of Social Theory*, Cambridge: Harvard University Press.

Doeringer, Peter B. and Michael J. Piore (1971): *Internal Labor Markets and Manpower Analysis*, Lexington, Massachusetts: Heath.

Dore, Ronald (1983): "Goodwill and the Spirit of Market Capitalism," *British Journal of Sociology*, 34, 459–482.

Eccles, Robert G. (1981): "The Quasifirm in the Construction Industry," *Journal of Economic Behavior and Organization*, 2, 335–357.

Fluck, Zsuzanna (1998):"Optimal Financial Contracting: Debt versus Outside Equity," *Review of Financial Studies*, 11, 383–418.

Fudenberg, Drew and David Kreps (1987): "Maintaining a Reputation with Simultaneous Opponents," *Review of Economic Studies*, 54, 541–568.

Ghemawat, Pankaj and Tarun Khanna (1998): "The Nature of Diversified Business Groups: A Research Design and Two Case Studies," *Journal of Industrial Economics*, 46, 35–61.

Greif, Avner (1993): "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *American Economic Review*, 83, 525–548.

Greif, Avner, Paul Milgrom and Barry R. Weingast (1994): "Coordination, Commitment and Enforcement: The Case of the Merchant Guild," *Journal of Political Economy*, 102, 745–776.

Grossman, Sanford and Oliver Hart (1986): "The Costs and Benefits of Ownership: A Theory of Lateral and Vertical Integration," *Journal of Political Economy, 94, 691–719.*

Hart, Oliver (1995): *Firms, Contracts and Financial Structure.* Oxford: Oxford University Press.

Hart, Oliver and John Moore (1990): "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98, 1119–58.

Holmström, Bengt (1996): "The Firm as a Subeconomy," Mimeo, M.I.T.

Holmström, Bengt and Paul Milgrom (1994): "The Firm as an Incentive System," *American Economic Review, 84, 972–991.*

Jensen, Michael and William Meckling (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3, 305–360.

Klein, Benjamin and Keith Leffler (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89, 615–641.

Kreps, David (1990): "Corporate Culture and Economic Theory," in *Essays on Positive Political Economy*, Ed. J. Alt and K. Shepsle, Cambridge: Cambridge University Press.

Macauley, Stewart (1963): "Non-Contractual Relations in Business: A Preliminary Study," *American Journal of Sociology*, 28, 55–67.

MacLeod, Bentley and James Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57, 447–480.

MacLeod, Bentley and James Malcomson (1998): "Motivation and Markets," *American Economic Review*, 88, 388–411.

Macneil, Ian (1974): "The Many Futures of Contracts," *Southern California Law Review*, 47, 691–816.

Malcomson, James (1986): "Rank-Order Contracts for a Principal with Many Agents," *Review of Economic Studies*, 53, 807–817.

McMillan, John (1995): "Reorganizing Vertical Supply Relationships," in *Trends in Business Organization: Do Participation and Cooperation Increase Competitiveness?*, Ed. H. Siebert, Tubingen: Mohr.

Rajan, Raghuram and Luigi Zingales (1998): "Power in a Theory of the Firm," *Quarterly Journal of Economics*, 113, 387–432.

Rotemberg, Julio (1991): "A Theory of Inefficient Intrafirm Transactions," *American Economic Review*, 81, 191–209.

Shapiro, Carl and Joseph Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74, 433-444.

Simon, Herbert (1951): "A Formal Theory of the Employment Relationship," *Econometrica*, 19, 293–305.

Stole, Lars and Jeffrey Zwiebel (1996): "Organizational Design and Technology Choice under Intrafirm Bargaining," *American Economic Review*, 85, 195–222.

Taylor, Curtis R. and Steven N. Wiggans (1997): "Competition or Compensation: Supplier Incentives under the American and Japanese Subcontracting Systems," *American Economic Review*, 87, 598–618.

Telser, Lester (1981): "A Theory of Self-Enforcing Agreements," *Journal of Business*, 22, 27–44.

Williamson, Oliver (1985): *The Economic Institutions of Capitalism*, New York: Free Press.

Chapter 3

The Value of Information in Monotone Decision Problems (with Susan Athey)

3.1 Introduction

In a Bayesian decision problem, an agent who is uncertain about the true state of the world must choose an action after observing an imperfectly informative signal. Blackwell (1951, 1953) proved the seminal result that *every* agent faced with such a decision problem will prefer (ex ante) an informative signal $\tilde{y}$ to another signal $\tilde{x}$ if and only if $\tilde{y}$ is statistically sufficient for $\tilde{x}$. This notion of "better information" is useful and intuitive, but it is also quite demanding (as noted by Blackwell himself).[1] For economic modeling, one might expect that sufficiency is far stronger than needed to compare information structures— in sharp contrast to most economic models, Blackwell's theorem places no restrictions on the decision-maker's payoff function.

This chapter shows that by exploiting properties of the decision-maker's preferences, it is possible to relax sufficiency and derive comparisons among a richer set of information structures. For many different classes of decision-makers, conditions are derived under which a signal $\tilde{y}$ is preferred to another signal $\tilde{x}$ by all members of the class. In most cases, the informativeness order can be represented as a stochastic dominance ordering over the distribution of posteriors that might arise from each information source. Conditions are also provided under which one decision-maker will have a higher marginal value of information than another decision-maker,

[1] See Blackwell and Girshik (1954). For $\tilde{y}$ to be sufficient for $\tilde{x}$, all of the posteriors, no matter how unlikely, generated by $\tilde{x}$ must be in the convex hull of the set of posteriors generated by $\tilde{y}$. There are many unsatisfying examples of distributions which cannot be ranked according to sufficiency. Unless a signal $\tilde{x}$ is normally distributed, $\tilde{x}$ cannot be more informative than a normally distributed signal $\tilde{y}$ in the Blackwell order. See Lehmann (1988) for further examples.

and thus will acquire more information. The results are applied to oligopoly models, labor markets with adverse selection, hiring problems, and a coordination game.

The starting point of the analysis is to focus attention on *monotone decision problems*, each consisting of a prior belief on the state of the world $H(\omega)$, a class of payoff functions U, and a set of admissable signals $\{\tilde{x}\}$. The class of payoff functions is defined by how the decision-maker's incremental returns to taking an action (a) vary with the state of the world (ω). For example, in many economic problems, it is assumed that the incremental returns to an action are nondecreasing in the state of the world (i.e., the payoff function $u(\omega, a)$ is supermodular). Given such a class, it is possible to (partially) order posterior beliefs about the state of the world so that "higher" beliefs induce higher actions for all decision-makers in the class. For a signal to be admissable to a monotone decision problem, it is required that all of the posteriors that might be generated from different realizations of the signal can be ordered in this way. To illustrate, for the class of decision-makers with supermodular payoffs, higher posteriors in the sense of First Order Stochastic Dominance (FOSD) induce higher actions, and the signal $\tilde{x}$ is admissable only if the corresponding set of posteriors can be totally ordered by FOSD. In general, different classes of decision-makers will induce different "stochastic dominance" orderings over posteriors (for instance, if the marginal returns to acting are concave in the state of the world, the relevant order is Second Order Stochastic Dominance (SOSD)).

A natural information ordering over signals is available for monotone decision problems. A signal $\tilde{y}$ is preferred to $\tilde{x}$ for a class of decision-makers (and thus $\tilde{y}$ is "more informative"), if the "high" posteriors induced by $\tilde{y}$ are (on average) higher than the "high" posteriors induced by $\tilde{x}$, and the "low" posteriors induced by $\tilde{y}$ are (on average) lower than the "low" posteriors induced by $\tilde{x}$. The terms "high" and "low" refer to the stochastic dominance order induced by the restriction on the decision-maker's payoff function.

To fix ideas, return again to the example of supermodular payoff functions. Recalling that the posteriors of the signal $\tilde{x}$ are totally ordered, index them with a parameter α_x, so that higher values of the index correspond to higher ordered (by FOSD) posteriors. So that this index will be comparable across signals, let it be normalized to have a uniform distribution ex ante: with probability α_x, a posterior lower (in the FOSD order) than $F(\omega|\alpha_x)$ is realized after observing $\tilde{x}$. Similarly, index the posteriors $G(\omega|\alpha_y)$ generated by $\tilde{y}$. Then a signal $\tilde{y}$ more informative than $\tilde{x}$ for all supermodular payoff functions if, for all $\alpha \in [0, 1]$,

$$G(\omega|\alpha_y \geq \alpha) \succeq_{FOSD} F(\omega|\alpha_x \geq \alpha).$$

The expression $G(\omega|\alpha_y \geq \alpha)$ represents the average over the highest $1 - \alpha$ fraction of the posteriors generated by $\tilde{y}$. Thus, a signal $\tilde{y}$ is more informative than $\tilde{x}$ if high realizations of ω are more likely when highly-ranked posteriors are realized. For other classes of payoff functions, other stochastic dominance orders (such as SOSD) will be relevant for the comparison of average posteriors.

The conditions derived in this chapter are *sufficient* for all decision-makers in a given class to prefer one signal to another; they are *necessary* when comparing small (differential) changes in the signal structure. The theory can be generalized by considering orders based on single crossing properties rather than stochastic dominance.

The chapter's second objective is to derive conditions under which one decision-maker will have a higher marginal value for information than another, and thus will acquire more information when information is costly. If u and v are in the same class of payoff functions, and the two decision-makers consider purchasing signals ranked according to our criteria, then decision-maker u buys more information than v if u's preferences over the distribution of posteriors when using an optimal decision rule are "more sensitive" than v's. The meaning of "sensitive" is determined by the class of payoff functions under consideration. Although the conditions depend on the optimal policies chosen by the agents, and thus are not primitive, they can be verified in many applications.

Our results are related to work in statistics by Lehmann (1988). He considered one specific class of monotone decision problems — those where the decision-maker's payoffs satisfy a single crossing property, and where the signals satisfy the monotone likelihood ratio property. For such problems, Lehmann (1988) derived a new information ordering that relaxes Blackwell's sufficiency criterion. Lehmann's *effectiveness* ordering has already entered economics in the context of auctions (Persico, 1997), principal-agent problems (Jewitt, 1997), and implicit incentive models (Dewatripont, Jewitt and Tirole, 1999). While the methods used here are quite different - Lehmann takes an approach based on statistical hypothesis testing - we obtain the effectiveness ordering as a special case. Closely related to Lehmann's work is that of Persico (1996), who studied the same specific class of decision problems. His paper develops an approach to ranking decision-makers in terms of their incentives to acquire information.[2] Our approach to information acquisition builds directly on his, and provides a significant generalization to other classes of monotone decision problems.

The chapter develops as follows. The next section describes the model and briefly reviews the standard approach to information and some preliminary results on stochastic orderings. In Section 3.3, we introduce the idea of monotone decision problems (MDPs), and discuss some important classes of MDPs. Section 3.4 includes the main theorems on ordering information structures, and characterizes the monotone information order for several classes of MDPs. Section 3.5 presents results on ordering payoff functions in terms of their marginal value for information. Section 3.6 gives some economic applications — to information gathering by firms, adverse selection in labor markets, a coordination game under uncertainty, and a hiring

[2] Persico (1997) further established that similar techniques can be used to rank the revenue to the auctioneer under different auction formats.

problem. The last section discusses some possible extensions and concludes.

3.2 The Model

3.2.1 The Bayesian Decision Problem

A decision-maker (DM) who is uncertain about the true state of the world must take an action after observing an informative signal. The state of the world is denoted $\omega \in \Omega$, where $\Omega \subseteq \mathbb{R}$ is an interval. Let $\mathcal{P}$ denote the set of all probability distributions on Ω. The DM must choose an action $a \in A \subset \mathbb{R}$. Her payoff $u(\omega, a)$ depends on both her action and the true state; we assume u is a bounded measurable function taking $\Omega \times \mathbb{R} \to \mathbb{R}$. Throughout, we maintain the following assumption about the set of available actions.

(A) Either A is finite, or A is a compact interval of $\mathbb{R}$ and $u(\omega, a)$ is continuous in a.

The DM has prior distribution $H(\omega)$. Before acting, the DM observes some informative random variable $\tilde{x}$, with support $\mathcal{X} \subseteq \mathbb{R}$, and forms a posterior distribution $F(\omega|x)$. The joint distribution of (ω, x) is then written $F(\omega, x)$, while the marginal distributions are denoted $F(x)$ and $F(\omega) \equiv H(\omega)$.[3] We refer to F as an "information structure."

Observe that many different information structures can be equivalent from the perspective of decision-making, since only the posterior generated by a signal realization affects behavior (the value of the signal does not). In particular, it does not matter if the DM observes $\tilde{x}$ or some $T(\tilde{x})$, so long as $F(\omega|T(\tilde{x}) = T(x)) = F(\omega|\tilde{x} = x)$ for all x. The payoff-relevant features of an information structure, F, can be *uniquely* characterized in terms of the probability measure induced on the space of posteriors $\mathcal{P}$, which we write as μ_F.[4] We begin by working with this abstract characterization of an information structure before mapping back to the first formulation in terms of the joint distribution $F(\omega, x)$.

The decision-maker's problem, given a posterior distribution $P \in \mathcal{P}$, is to solve

$$\max_{a \in A} \int_{\Omega} u(\omega, a) dP(\omega) \tag{3.1}$$

[3]Note that $F(\omega) = H(\omega)$ is needed to ensure that the expectation of the posterior is the prior. Taking the joint distribution of (ω, x) as primitive is a departure from the analyses of Blackwell (1951) and Lehmann (1988), which take as primitive the distribution of $x|\omega$. This means that their information orders are the same *for all prior distributions H*. Our information rankings will be given *relative to a fixed prior H*. Thus, we may sensibly consider averages over subsets of the posterior distributions.

[4]This construction is introduced in Blackwell's original article (Blackwell, 1951). Endow $\mathcal{P}$ with the weak topology. Then $F(\omega|x)$ is measurable with respect to the Borel σ-algebra and $\mu_F(E) = \int_{x:F(\omega|x)\in E} dF(x)$.

to obtain an optimal action $a^*(P)$, and a realized payoff $u(\omega, a^*(P))$. We define the (ex ante) *value* of the decision problem as

$$V^*(F,u) = \int_{\mathcal{P}} \int_{\Omega} u(\omega, a^*(P)) dP(\omega) d\mu_F(P). \tag{3.2}$$

3.2.2 The Classical Approach to Information

The classical approach to information, due to Blackwell (1951, 1953), begins by writing $V^*(F,u)$ as

$$V^*(F,u) = \int_{\mathcal{P}} \left\{ \max_{a(P)} \int_{\Omega} u(\omega, a(P)) dP(\omega) \right\} d\mu_F(P) \equiv \int_{\mathcal{P}} u^*(P) d\mu_F(P). \tag{3.3}$$

Using revealed preference, it is straightforward to show that $u^*(P)$ will be convex in P. Simply observe that for $\lambda \in [0,1]$, $P^1, P^2 \in \mathcal{P}$,

$$\max_a \left\{ \lambda \int_{\Omega} u dP^1 + (1-\lambda) \int_{\Omega} u dP^2 \right\} \leq \max_a \lambda \int_{\Omega} u dP^1 + \max_a (1-\lambda) \int_{\Omega} u dP^2. \tag{3.4}$$

Now suppose G is an alternative information structure which induces a measure μ_G over posteriors distributions, and that μ_G *stochastically dominates* μ_F *for convex functions*, i.e. $\int_{\mathcal{P}} \varphi d\mu_G \geq \int_{\mathcal{P}} \varphi d\mu_F$ for any convex function $\varphi : \mathcal{P} \to \mathbb{R}$. Clearly, if μ_G stochastically dominates μ_F for convex functions, then $V^*(G,u) \geq V^*(F,u)$ for any payoff function u and action set A, i.e. every decision-maker will find G more valuable than F; Blackwell (1951) showed the converse via a separation argument. Thus, Blackwell's order can be thought of as a multivariate generalization of the (perhaps more familiar) mean-preserving spread of Rothschild and Stiglitz (1970). The geometric representation of a mean-preserving spread in multiple dimensions can be formalized using what is known as a "dilation," so that μ_G is a dilation of μ_F.[5] It can further be shown that for a signal $\tilde{y}$ to be sufficient for $\tilde{x}$, the set of posteriors generated by $\tilde{x}$ must lie in the convex hull of those generated by $\tilde{y}$.

Sufficiency can be understood in the context of the following three state, two signal example. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, where $\omega_1 < \omega_2 < \omega_3$, and suppose the prior on ω is uniform $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Consider a signal $\tilde{y}$ that is equally likely to induce posterior beliefs $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$ and $(0, \frac{1}{2}, \frac{1}{2})$, and another signal $\tilde{x}$ that is equally likely to induce posterior beliefs $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ and $(\frac{1}{6}, \frac{5}{12}, \frac{5}{12})$. As illustrated in Figure 1, the posteriors generated by $\tilde{y}$ are a mean-preserving spread of the posteriors generated by $\tilde{x}$. So $\tilde{y}$ will be sufficient for $\tilde{x}$. But this is clearly very special, since $\tilde{y}$ can be sufficiency ranked relative to another signal $\tilde{z}$ only if the posteriors generated by $\tilde{z}$ lie on the line between $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$ and $(0, \frac{1}{2}, \frac{1}{2})$. The ranking will be upset by even the slightest

[5] A dilation $D : P \to D_P$ is a mapping from $\mathcal{P}$ into the set of probability measures on $\mathcal{P}$ such that $\int_{\mathcal{P}} Q dD_P(Q) = P$. That is, D associates with each $P \in \mathcal{P}$ a non-degenerate probability measure on $\mathcal{P}$ with mean P. For our problem, $\mu_G = \int_{\mathcal{P}} D_P d\mu_F(P)$. Strassen (1965) has shown that other stochastic dominance relations also have this sort of representation.

perturbation of $\tilde{z}$ leading to a potential posterior belief off of the line. This lack of robustness is unappealing.[6] Figure 3.1 also highlights the fact that *no* two-point information structure can be sufficient for $\tilde{y}$ (the prior is fixed, and so no further spreading is possible). In spite of this, $\tilde{y}$ is clearly not uniquely suited to making all types of inferences. For instance, $\tilde{y}$ provides no information at all as to the relative likelihood of ω_2 versus ω_3. And while $\tilde{y}$ provides significant information as to whether ω is low or high, many other two-point information structures might be more informative on this question. Since it seems reasonable to suppose that particular classes of decision-makers — especially the classes of decision-makers that appear in economic models — care only about making certain types of inferences, the example suggests that it should be possible to derive other informativeness criteria.

Unfortunately, the classical approach is ill-suited to exploiting any structure one might impose on the decision-maker's preferences. Most important from our perspective, and from that of many economic models, are assumptions about how the returns to taking a higher action, $r(\omega) = u(\omega, a^H) - u(\omega, a^L)$, depend on the state — for instance $r(\omega)$ might be increasing, or concave, or polynomial in the state of the world, or positive if and only if the state variable is greater than some critical level ω_0. Using the above approach it is hard to see how to incorporate such a restriction. For example, suppose one restricted attention to the class of supermodular payoff functions ($r(\omega)$ increasing in ω). First, it is not clear that this implies anything about $u^*(\omega, P) = u(\omega, a^*(P))$; and second, even if it did, we would still need to connect this property with a meaningful characterization of "better" information. In the next few sections, we show that by placing a total order over the posteriors (such as a stochastic dominance order), it is possible to overcome these difficulties.

3.3 Monotone Decision Problems and Stochastic Orders

3.3.1 Preliminaries: Stochastic Dominance and Single Crossing

Since we will make extensive use of stochastic orderings, we first review a few definitions and known facts.[7] Assume that U is some set of measurable functions taking $\mathbb{R}^n \to \mathbb{R}$, and let P^H, P^L be two probability distributions on $\mathbb{R}^n$. We say that P^H *stochastically dominates* P^L with respect to U, written $P^H \succ_{SD-U} P^L$ if

$$\int u(\mathbf{z})dP^H(\mathbf{z}) \geq \int u(\mathbf{z})dP^L(\mathbf{z}) \quad \text{for all } u \in U. \qquad (SD)$$

[6] Indeed, it has motivated work in statistics on "approximate sufficiency"; see Le Cam (1964).

[7] For further treatments of this material, see *inter alia*, Karlin and Studden (1966), Karlin (1968), Jewitt (1986), Border (1991) and Athey (1998a,1998b).

If U is the set of nondecreasing univariate functions, then $\succ_{SD-U}$ is the standard First Order Stochastic Dominance (FOSD) relationship. If U is the set of concave univariate functions, then $\succ_{SD-U}$ corresponds to Second Order Stochastic Dominance (SOSD).

Any set of functions U induces a stochastic dominance order. Some critical features of stochastic dominance orders for our purposes can be understood using the notion of a closed convex cone, as follows.

Definition 3.1 *A set U is a closed convex cone (ccc) if (a) $u, v \in U$ implies that $\alpha u + \beta v \in U$ for any $\alpha, \beta > 0$, and (b) U is closed under the weak topology.*

If $\int u dP^H \geq \int u dP^L$ for all $u \in U$, then the same inequality will hold for any function v in the closed convex cone generated by U (denoted $ccc(U)$). Further, since P^H and P^L are probability distributions, then (SD) holds for the functions $u(\mathbf{z}) \equiv 1$ and $u(\mathbf{z}) \equiv -1$ (denoted $\{\mathbf{1}, -\mathbf{1}\}$). Thus, the set U generates the same stochastic dominance order as $ccc(U \cup \{\mathbf{1}, -\mathbf{1}\})$.[8]

A weaker notion than stochastic dominance is that of *stochastic single crossing.*[9] We write $P^H \succ_{SC-U} P^L$ if

$$\int u(\mathbf{z}) dP^L(\mathbf{z}) \geq 0 \quad \Rightarrow \quad \int u(\mathbf{z}) dP^H(\mathbf{z}) \geq 0 \quad \text{for all } u \in U. \tag{SC}$$

If $\{\mathbf{1}, -\mathbf{1}\} \in U$, then one can show (Athey, 1998a) that $\succ_{SC-U}$ is equivalent to $\succ_{SD-U}$.[10] If U does not contain constant functions, however, $\succ_{SC-U}$ is weaker than $\succ_{SD-U}$. The distinction between stochastic dominance and stochastic single crossing will become relevant in our analysis of information orders, since some sets of payoff functions we wish to consider (most notably, single crossing payoff functions, which arise in auction games and portfolio problems) do not contain the constant functions.

The stochastic single crossing order is used in our analysis because it implies a comparative statics prediction about monotonicity of the optimal policy.[11]

[8] In fact, just this insight allows us to characterize stochastic dominance orderings, since we can also consider a the order induced by a much smaller set E_U (for the case of nondecreasing functions, a set of indicator functions), loosely referred to as extreme points, so long as $U = ccc(E_U \cup \{\mathbf{1}, -\mathbf{1}\})$. See Border (1991) or Athey (1998a) for more discussion.

[9] Various definitions of single crossing properties (using different combinations of weak and strict inequalities) have been proposed by different authors in economics (Milgrom and Shannon, 1994; Shannon, 1995) and statistics (Karlin, 1968). The definition in (SC) is not the most natural variation for comparative statics theorems. However, as it involves only weak inequalities, (SC) is especially appropriate for working with closed convex cones and stochastic orders.

[10] This result hinges on the assumption that P^H and P^L are probability distributions.

[11] To simplify the notation and the statement of the result we consider only sufficiency here, but with minor qualifications, the single crossing condition is in fact necessary for comparative statics as well. See Milgrom and Shannon (1994) and Shannon (1995).

Lemma 3.1 *Let U_1 be some set of functions taking $\not\leq \to \mathbb{R}$. Suppose that $u(\omega, a) : \Omega \times \mathbb{R} \to \mathbb{R}$ and that for all $a^H > a^L$, $u(\omega, a^H) - u(\omega, a^L) \in U_1$. Let $a^*(P) = \arg\max_{a \in A} \int_\Omega u(\omega, a) dP(\omega)$. Then $P^H \succ_{SC-U_1} P^L$ implies that there exists a selection $\hat{a}^*(P)$ from $a^*(P)$ such that $\hat{a}^*(P^H) \geq \hat{a}^*(P^L)$.*

Proof. Define a function $v(\tau, a) : \{0,1\} \times \mathbb{R} \to \mathbb{R}$, with $v(1, a) = \int_\Omega u(\omega, a) dP^H(\omega)$ and $v(0, a) = \int_\Omega u(\omega, a) dP^L(\omega)$. Then $v(\tau, a)$ is bivariate weak single crossing in (τ, a) and by Shannon (1995), there exists a selection from $a^*(\tau) = \arg\max_{a \in A} v(\tau, a)$ that is nondecreasing. *Q.E.D.*

3.3.2 Monotone Decision Problems

Our approach is to restrict attention to particular classes of decision problems. Each class is characterized by a restriction on the DM's payoff function and a corresponding restriction on the set of admissable information structures.

Definition 3.2 *The pair $(U_2, \mathcal{F})$ constitutes a class of monotone decision problems if there exists a prior $H(\omega)$ and some set U_1 of bounded measurable functions taking $\not\leq$ into $\mathbb{R}$, such that:*

(MDP-U) *For all $u \in U_2$, if $a^H, a^L \in \mathbb{R}$ and $a^H \geq a^L$, then $u(\omega, a^H) - u(\omega, a^L) \in U_1$. And moreover, if $u : \Omega \times \mathbb{R} \to \mathbb{R}$ is some bounded measurable function with incremental returns in U_1, then $u \in U_2$.*

(MDP-$\mathcal{F}$) *For all $F \in \mathcal{F}$, $F(\omega) \equiv H(\omega)$, and further if x^H, $x^L \in$support(F) and $x^H \geq x^L$, then $F(\omega|x^H) \succ_{SC-U_1} F(\omega|x^L)$. And moreover, if F is an information structure with prior $H(\omega)$ and posteriors completely ordered in x by $\succ_{SC-U_1}$, then $F \in \mathcal{F}$.*

If $(U_2, \mathcal{F})$ satisfy this definition, we refer to them as an *MDP pair.* The first condition, *(MDP-U)*, implies that every DM under consideration has incremental returns to acting that rely on the state in some pre-specified way (for example, they might be nondecreasing or concave); the second, *(MDP-$\mathcal{F}$)*, implies that for every admissable information structure, the induced posterior beliefs are completely ordered in the sense of stochastic single crossing.

What do these restrictions buy us? Let $(U_2, \mathcal{F})$ be an *MDP* pair. By *(MDP-U)* and Lemma 3.1, each DM $u \in U_2$ has an optimal policy $a^*(P)$ that is monotone in P when posteriors are ordered by $\succ_{U_1-SC}$. And *(MDP-$\mathcal{F}$)* implies that if $F \in \mathcal{F}$, then *every* posterior associated with F can be ordered in this way and indexed by x. Thus, each DM $u \in U_2$ takes higher actions when she receives higher realizations of the signal.

In this way, the *MDP* restrictions impose an *ordinal* structure, so that one can heuristically think of the posteriors generated by $\tilde{x}$ going from low to high along a single dimension. However, the goal in this paper is to compare different information structures. While *(MDP-$\mathcal{F}$)* implies that the posteriors for a given signal can be totally ordered, it provides no guidance as to how to compare the posteriors arising from different signals. Thus, we introduce a *cardinal* index for the posteriors generated by each signal, $\alpha \in [0,1]$, so that we may "match" for comparison posteriors from different signals. This is accomplished using a strictly increasing function $T_F : \mathcal{X} \to [0,1]$, so that $T_F(x) = \alpha$.

For an important set of cases, we will show that it is appropriate to match posteriors according to their percentile in the ex ante signal distribution. Formally, for an information structure $F(\omega, x)$ with marginal distribution $F(x)$, we let $T_F(x) = F(x)$. Using this index, we can let $\tilde{F}(\omega, \alpha) = F(\omega, F^{-1}(\alpha))$ on $\Omega \times [0,1]$,[12] so that $\tilde{F}(\omega|\alpha)$ is the "α-percentile posterior." The probability of observing a posterior below $\tilde{F}(\omega|\alpha)$ in the $\succ_{U_1-SC}$ order is given by α, so that $\tilde{F}(\alpha)$ is the uniform distribution on $[0,1]$. As illustrated in Figure 3.2, we can use such an index to compare the realizations of two signals, $\tilde{x}$ and $\tilde{y}$, according their α-percentile.

Using this construction, we can represent a decision-maker's policy function $a : [0,1] \to A$ in terms of the action it prescribes for an α-indexed posterior. Policy functions $a(\alpha)$ can then be analyzed without reference to a particular information structure. Using this representation, for any probability distribution $\tilde{F}$ on $\Omega \times [0,1]$, we define the ex ante expected value from using the policy a by

$$V(\tilde{F}, u, a) = \int_{[0,1]} \int_{\Omega} u(\omega, a(\alpha)) d\tilde{F}(\omega, \alpha). \tag{3.5}$$

An important consequence of *(MDP-U)* follows: for any nondecreasing $a(\alpha)$,

$$u(\omega, a(\alpha^H)) - u(\omega, a(\alpha^L)) \in U_1 \text{ for } \alpha^H > \alpha^L. \tag{3.6}$$

This in turn implies that $u(\omega, a(\alpha))$ viewed as a function of (ω, α) can be extended to be in U_2. In words, whenever the DM uses a monotone policy, the incremental returns to having a higher α-indexed posterior retain the properties we assumed for the incremental returns of the primitive payoff function, and thus $u(\omega, a(\alpha))$ inherits the properties specified by U_2.

The discussion of this subsection can be formally summarized as follows.

Theorem 3.1 *Suppose $(U_2, \mathcal{F})$ are an MDP pair and that (A) holds. Consider any $(u, F) \in (U_2, \mathcal{F})$. Then for any $T_F : \mathcal{X} \to [0,1]$ continuous and strictly increasing, $\tilde{F}(\omega, \alpha) \equiv F(\omega, T_F^{-1}(\alpha)) \in \mathcal{F}$. Further, there exists some nondecreasing $a^F : [0,1] \to A$ such that:*

[12] We can, without loss of generality, take $F(x)$ to be continuous and strictly increasing. Lehmann (1988, p. 527) provided a construction which shows that one can always take $F(x)$ to be continuous. If $F(x)$ is constant over some interval $[x_0, x_1]$, notice that the DM would experience no payoff loss from observing $x^* = x$ on $x < x_0$, $x^* = x_0$ on $[x_0, x_1]$ and $x^* = x - (x_1 - x_0)$. And $F^*(x^*)$ will be strictly increasing (one needs a more involved construction if $\mathcal{X}$ is not compact). Then $F(x) : \mathcal{X} \to [0,1]$ is a bijection (F is continuous and strictly monotone), so $\tilde{F}$ is well defined.

(i) *For all* $\alpha \in [0,1]$, $a^F(\alpha) \in \arg\max_{a \in A} \int_\Omega u(\omega, a) d\tilde{F}(\omega|\alpha)$.

(ii) $V(\tilde{F}, u, a^F(\cdot)) = V^*(F, u)$.

(iii) *For any* $\alpha^H > \alpha^L$, $u(\omega, a^F(\alpha^H)) - u(\omega, a^F(\alpha^L)) \in U_1$.

(iv) *There exists* $u^F : \Omega \times \mathbb{R} \to \mathbb{R}$, $u^F \in U_2$, *with* $u^F(\omega, \alpha) = u(\omega, a^F(\alpha))$.

Proof. *(i)* Because A is compact, for all $P \in \mathcal{P}$, $\int_\Omega u(\omega, a) dP(\omega)$ attains a maximum on A. For any $\alpha^H \geq \alpha^L$, monotonicity of T and *(MDP-$\mathcal{F}$)* implies that $\tilde{F}(\omega|\alpha^H) \succ_{SC-U_1} \tilde{F}(\omega|\alpha^L)$. So from Lemma 3.1, there will exist a nondecreasing selection from the set of optimizers, denoted $a^F(\alpha)$. *(ii)* By (i), $\int_\Omega u(\omega, a^F(\alpha)) d\tilde{F}(\omega|\alpha) = \max_{a \in A} \int_\Omega u(\omega, a) dF(\omega|T_F^{-1}(\alpha))$. The result then follows immediately from the definitions of the distributions. *(iii)* Let $a^H = a^F(\alpha^H)$, $a^L = a^F(\alpha^L)$ and note that $a^H \geq a^L$. But then $u(\omega, a^F(\alpha^H)) - u(\omega, a^F(\alpha^L)) = u(\omega, a^H) - u(\omega, a^L) \in U_1$ by *(MDP-U)*. *(iv)* The function $u^F(\omega, \alpha)$ can be defined to equal $u(\omega, a^F(\alpha))$ for $\alpha \in [0,1]$, to equal $u(\omega, a^F(0))$ for $\alpha < 0$, and to equal $u(\omega, a^F(1))$ for $\alpha > 1$. It follows from *(MDP-U)* that $u^F \in U_2$. *Q.E.D.*

3.3.3 Examples

Our framework covers many problems of economic interest. In this section, we discuss four common classes of problems and then show briefly how the theory can be applied to other cases.

Example.1. *Supermodular (Incremental returns increasing in ω).* Suppose U_1 is the set of nondecreasing functions. Then U_2 is the set of functions $u(\omega, a)$ that are supermodular in (ω, a). In words, the incremental return to a higher action is nondecreasing in the state of the world. Numerous economic applications take this form (see Milgrom and Roberts (1990)); for example, a might represent the level of investment for a firm, where the marginal returns are indexed by ω. Since U_1 contains constant functions, we have $P^H \succ_{SC-U_1} P^L$ if and only if P^H is higher than P^L according to FOSD, denoted $P^H \succ_{FOSD} P^L$. So (*MDP-$\mathcal{F}$*) implies that if $F \in \mathcal{F}$, the induced posteriors must be ranked by FOSD, which requires that $F(\omega|x)$ is decreasing in x: a higher signal corresponds to a higher probability that the state of the world is high. Figure 3.3 illustrates the FOSD ordering over posteriors for a 3-state example.

Example 2. *Concave returns (Incremental returns concave in ω).* Suppose U_1 is the set of concave functions. Then U_2 is the set of functions $u(\omega, a)$ that have concave incremental returns. Such a payoff function can arise naturally in binary decision problems ($a \in \{0, 1\}$) where the DM must decide whether or not to undertake some risky venture with concave payoff $r(\omega)$ (in Section 3.5 we present a hiring problem of this type). More generally, if $u \in U_2$, then the DM becomes strongly more risk averse as she raises her action.[13] For this case $P^H \succ_{SC-U_1} P^L$ if and only if P^H dominates P^L according to SOSD ($P^H \succ_{SOSD} P^L$): the DM takes a higher action when the posterior about ω is "less risky." So if $F \in \mathcal{F}$, then $F(\omega|x^H) \succ_{SOSD} F(\omega|x^L)$. This requires that $E[\omega|x]$ is constant in x and for any ω, $\int_{-\infty}^{\omega} F(\tilde{\omega}|x)d\tilde{\omega}$ is nonincreasing in x. Alternatively, for any $x^L < x^H$, $F(\omega|x^L)$ can be attained from $F(\omega|x^H)$ by a sequence of mean-preserving spreads (Rothschild and Stiglitz, 1970).

Example 3. *WSC(ω_0) (Incremental returns weak single crossing at ω_0).* Suppose U_1 is the set of functions $r(\omega)$ such that $r(\omega) \leq 0$ for $\omega < \omega_0$ and $r(\omega) \geq 0$ for $\omega > \omega_0$, i.e. the functions that cross zero from below at ω_0. We say such a function satisfies WSC(ω_0). Payoff functions in this class arise in the context of investment under uncertainty problems, where ω might represent the return on a risky asset, ω_0 is the return on a risk-free asset, a is the portfolio weight on the risky asset, and investor payoffs are given by $v(a\omega + (1-a)\omega_0)$. In this case, if P^H, P^L have densities p^H, p^L with respect to Lebesgue measure, then $P^H \succ_{SC-U_1} P^L$ if and only if $p^H(\omega) - \frac{p^H(\omega_0)}{p^L(\omega_0)} p^L(\omega)$ satisfies $WSC(\omega_0)$ in ω (Athey, 1998b). Thus if $F \in \mathcal{F}$, this reduces to $\frac{f(\omega|x^H)}{f(\omega|x^L)} \leq (\geq) \frac{f(\omega_0|x^H)}{f(\omega_0|x^L)}$ as $\omega < (>) \omega_0$. A high signal means that it is more likely that the true state ω is greater than ω_0. Since the class of $WSC(\omega_0)$ functions does not contain constant functions, a stronger condition is required to order posteriors by stochastic dominance. For convenience, define the set of functions $\overline{WSC}(\omega_0)$ as the set of functions $r(\omega)$ satisfying $WSC(\omega_0)$ and $r(\omega_0) = 0$. We have $P^H \succ_{SD-U_1} P^L$ if and only if $p^H(\omega) - p^L(\omega)$ is $\overline{WSC}(\omega_0)$. In particular, the densities must cross at ω_0, a restriction not imposed by $\succ_{SC-U_1}$.

Example 4. *Weak Single Crossing (Incremental returns single crossing in ω).* Suppose U_1 is the set of functions $r(\omega)$ that cross zero from below at some point ω_0. In other words, U_1 is the union of all the WSC(ω_0) sets. Payoff functions $u(\omega, a)$ with incremental returns that are single crossing also arise throughout economics — for instance in bidding and pricing problems.[14] For the class of single crossing functions, we have $P^H \succ_{SC-WSC} P^L$ if and only if $P^H \succ_{SC-WSC(\omega_0)} P^L$ for all ω_0, which implies that (where the densities exist) $p^H(\omega)/p^L(\omega)$ is nondecreasing in ω. This order has received a great deal of attention in economics and statistics, and it is known as the Monotone Likelihood Ratio (MLR) order ($P^H \succ_{MLR} P^L$). Then if $F \in \mathcal{F}$, $F(\omega|x)$ must be ordered by MLR, and if a density exists, $f(\omega|x^H)/f(\omega|x^L)$ is nondecreasing in ω.[15]

[13] Recall that $r^H(\omega)$ is strongly more risk averse than $r^L(\omega)$ if there is some $\lambda \geq 0$ such that $r^H(\omega) - \lambda r^L(\omega)$ is concave. See Ross (1981) or Jewitt (1986).

[14] See Milgrom and Shannon (1994) for an extensive discussion of the single crossing property, and Athey (1998b) for applications in stochastic problems.

[15] Athey (1998b) also shows that the set of log-supermodular payoff functions (where u is positive and $u(\omega, a^H)/u(\omega, a^L)$ is nondecreasing in ω) induce the same stochastic single crossing order, $\succ_{SC-U_1}$, despite the fact that the set of log-supermodular functions is smaller than the set of payoff functions with incremental

Note that the class of payoff functions with single crossing incremental returns includes the class of supermodular payoff functions and the class of payoff functions with incremental returns that are WSC(ω_0). But expanding the set of payoff functions comes at a cost — we have to limit the set of information structures that we can attempt to compare. For instance, requiring the posteriors to be ordered by MLR is significantly stronger than requiring the posteriors to be ordered by FOSD.[16] This is illustrated in Figure 3.3.

More generally, our theory applies if U_1 is some arbitrary closed convex cone of functions, such as linear functions or quadratic functions or functions with positive n^{th} derivatives. For any such case, there is a rich theory of stochastic dominance that allows one to characterize the relation $\succ_{SD-U_1}$, and hence the relevant restriction on posterior beliefs, in terms of the "extreme points" of the cone U_1.[17] We leave this pursuit to the reader.

3.4 Ordering Information Structures

This section contains our main results on ordering information structures. We begin by identifying a *sufficient* condition for all DMs with payoffs $u \in U_2$ to prefer an information structure $G \in \mathcal{F}$ to another information structure $F \in \mathcal{F}$. We call this condition the "monotone information order" *(MIO)* for $(U_2, \mathcal{F})$; it can be characterized using the stochastic dominance order induced by U_1. We show that *(MIO)* is not just sufficient, but also *necessary*, for small (differential) changes in the information structure when U_1 is a closed convex cone that contains the constant functions. Section 3.4.3 provides a general analysis of informativeness for classes of MDPs based on stochastic single crossing. Recall from Section 3.3.1 that when U_1 does not contain the constant functions, $\succ_{SC-U_1}$ is weaker than $\succ_{SD-U_1}$. Thus, the ordering based on stochastic single crossing can be used to compare a greater variety of information structures when U_1 does not contain constant functions. We derive Lehmann's (1988) effectiveness order as a corollary, and relate our work to his. Examples are provided.

3.4.1 Monotone Information Orders Using Stochastic Dominance

We begin by stating a sufficient condition for G to be more informative than F to a given class of decision-makers.

---

returns that are single crossing. However, the set of log-supermodular payoffs does not satisfy *(MDP-U)*, and thus our analysis below does not apply directly to this class of payoffs. To see an example where log-supermodular payoffs arise, suppose that the action a is price, firms maximize $(a-c)D(a,\omega)$, and higher states of the world correspond to more inelastic demand.

[16]In particular, assuming F is indexed by the MLR is equivalent to assuming the $F(\omega|x)$ is ordered by FOSD *for all prior distributions* $H(\omega)$ (Milgrom, 1981).

[17]That is, we check (SD) for some set of functions E_U such that $ccc(E_U \cup \{\mathbf{1}, -\mathbf{1}\}) = U$. See Karlin and Studden, 1966; Border, 1991; Jewitt, 1986; Athey 1998a.

Theorem 3.2 *Let $(U_2, \mathcal{F})$ be and MDP pair and suppose (A) holds. Consider any $F, G \in \mathcal{F}$. Then $V^*(G,u) \geq V^*(F,u)$ for all $u \in U_2$ if for all $\alpha \in (0,1)$*

$$G(\omega|G(y) \geq \alpha) \succ_{SD-U_1} F(\omega|F(x) \geq \alpha). \qquad \textit{(MIO)}$$

If the assumptions of Theorem 3.2 hold and *(MIO)* obtains, we write $G \succ_{MIO-U_1} F$, i.e. G is greater than F in the monotone information order induced by payoff functions with incremental returns in U_1. The condition *(MIO)* is easy to interpret. It says that the high posteriors induced by G (where "high" means $G(\tilde{y}) \geq \alpha$) are on average higher (according to stochastic dominance) than the corresponding high posteriors induced by F.[18] This condition is also equivalent to saying that the low posteriors induced by G are lower: observe that

$$\alpha F(\omega|F(x) \leq \alpha) + (1-\alpha)F(\omega|F(x) > \alpha) = H(\omega), \tag{3.7}$$

where $H(\omega)$ is the prior. So an equivalent expression to *(MIO)* is that for all $\alpha \in [0,1]$,

$$F(\omega|F(y) \leq \alpha) \succ_{SD-U_1} G(\omega|G(x) \leq \alpha). \tag{3.8}$$

Notice that in this condition, signals are implicitly indexed by their *ex ante* percentile. This motivates us to use the index $\alpha_x = F(x)$, as described in Section 3.3.2 and illustrated in Figure 3.2. Thus, throughout this subsection, we let $\tilde{F}(\omega,\alpha) = F(\omega, F^{-1}(\alpha))$ and let $\tilde{G}(\omega,\alpha) = G(\omega, G^{-1}(\alpha))$.

Proof of Theorem 3.2. We show that (MIO) actually implies a stronger result, namely that under *(MIO)*, for any $u \in U_2$, $V(\tilde{G},u,a) \geq V(\tilde{F},u,a)$ for *any* $a(\alpha)$ nondecreasing. This will imply that G is more informative than F under the conditions of Theorem 3.2. To see this, observe that $a^F(\alpha) \in \arg\max_a \int_\Omega u(\omega,a)dF(\omega|F(x){=}\alpha)$ is nondecreasing, and thus by revealed preference

$$V^*(G,u) \geq V(\tilde{G},u,a^F) \geq V(\tilde{F},u,a^F) \equiv V^*(F,u).$$

Suppose that $A = \{a_1, ..., a_n\}$ is finite, with $a_{i+1} > a_i$. Define

$$r_i(\omega) \equiv u(\omega, a_{i+1}) - u(\omega, a_i).$$

[18] Analogous to Blackwell, one can interpret *(MIO)* as saying that the G posteriors are "more spread out" than the F posteriors. To see this, note that *(MIO)* is equivalent to saying that for all $\alpha \in [0,1]$,

$$\int_0^\alpha G(\omega|G(y) = \tilde{\alpha})d\tilde{\alpha} \prec_{SD-U_1} \int_0^\alpha F(\omega|F(x) = \tilde{\alpha})d\tilde{\alpha}$$

which can be interpreted in a manner similar to the second order stochastic dominance condition for comparing two distributions, F^H and F^L, on $[0,1]$, which requires that $\int_0^z F^H(\tilde{z})d\tilde{z} \leq \int_0^z F^L(\tilde{z})d\tilde{z}$ with equality when $z=1$.

Consider some arbitrary monotone increasing policy $a : [0,1] \to A$. We can find $\alpha_0 = 0 \leq \alpha_1 \leq \dots \leq \alpha_{n-1} \leq \alpha_n = 1$ such that $a(\alpha) = a_i$ on $[\alpha_{i-1}, \alpha_i]$. Then we have

$$\begin{aligned}
V(\tilde{F}, u, a) &\equiv \int_\Omega \int_{[0,1]} u(\omega, a(\alpha)) d\tilde{F}(\omega, \alpha) \\
&= \int_\Omega \sum_{i=1}^{n} u(\omega, a_i) \left[\tilde{F}(\alpha_i|\omega) - \tilde{F}(\alpha_{i-1}|\omega)\right] d\tilde{F}(\omega) \\
&= E[u(\omega, a_1)] + \int_\Omega \left\{ \sum_{i=1}^{n-1} [u(\omega, a_{i+1}) - u(\omega, a_i)] \left[1 - \tilde{F}(\alpha_i|\omega)\right] \right\} d\tilde{F}(\omega) \\
&= E[u(\omega, a_1)] + \sum_{i=1}^{n-1} (1 - \alpha_i) \int_\Omega r_i(\omega) d\tilde{F}(\omega | \tilde{\alpha}_x \geq \alpha_i) \\
&\leq E[u(\omega, a_1)] + \sum_{i=1}^{n-1} (1 - \alpha_i) \int_\Omega r_i(\omega) d\tilde{G}(\omega | \tilde{\alpha}_y \geq \alpha_i) \quad = V(\tilde{G}, u, a).
\end{aligned}$$

The equalities follow by algebraic manipulation and Bayes' rule. The inequality follows directly from (MIO) since, by *(MDP-U)*, we know that for each i, $r_i(\omega) \in U_1$. The case of A compact follows via a limiting argument and is deferred until the next section. *Q.E.D.*

The idea in the proof is very simple. Starting from action a_1, every new increment to the choice of action entails a new (interim) gamble on ω, described by $r_i(\omega) \equiv u(\omega, a_{i+1}) - u(\omega, a_i) \in U_1$. Formally, payoffs for a given action can be written as the sum over these incremental gambles:

$$u(\omega, a_i) = u(\omega, a_1) + \sum_{j=1}^{i-1} r_j(\omega). \tag{3.9}$$

Jumping up to a higher action (say, from a_i to a_{i+1}) at some posterior indexed by $F(x) = \alpha_i$ means taking on the gamble for all higher-ranked posteriors as well, since the DM uses a monotone policy. The effect on *ex ante* payoffs is then given by $E[r_i(\omega)|F(\tilde{x}) \geq \alpha_i]$. Preferences over the gambles $r_i(\omega)$ are described by $\succ_{SD-U_1}$, and *(MIO)* requires that *every* such gamble is more favorable under G than under F. As this argument applies when the DM uses the optimal policy for signal $\tilde{x}$, there is no possible way for a decision-maker in the class to do better, from an *ex ante* standpoint, using F instead of G.

In the proof, we showed that *(MIO)* implies $V(\tilde{G}, u, a) \geq V(\tilde{F}, u, a)$ for *any* $u \in U_2$ and $a(\alpha)$ nondecreasing. Since *(MIO)* has such powerful consequences, it might seem to be "too strong" as an informativeness order. At a minimum, it seems reasonable to compare $V(\tilde{G}, u, a^F) \geq V(\tilde{F}, u, a^F)$ only for policies a^F which are optimal for information structure F. However, for important classes of monotone decision problems, we show that restricting the policy function in this way does not permit a weakening of *(MIO)*.

To see why this might be true, consider the simplest case of $A = \{0,1\}$ and some $u \in U_2$. The optimal policy for u (denoted $a^{F,u}$) entails jumping from $a = 0$ to $a = 1$ at a posterior indexed by $F(\tilde{x}) = \alpha^{F,u}$. Now consider utility functions of the form $w(\omega, a) = u(\omega, a) + aK$, where $K \neq 0$ is an arbitrary constant. There is no particular reason for the policy of jumping to $a = 1$ at $\alpha^{F,u}$ to be optimal for w under F; however, a DM with payoff w who does use $a^{F,u}$ (suboptimally), will prefer G to F if and only if the DM with payoff u does as well:

$$\begin{aligned} & V(\tilde{G}, u, a^{F,u}) - V(\tilde{F}, u, a^{F,u}) - [V(\tilde{G}, w, a^{F,u}) - V(\tilde{F}, w, a^{F,u})] && (3.10) \\ = \; & (1-\alpha) \int_\Omega K dG(\omega | G(\tilde{y}) \geq \alpha) - (1-\alpha) \int_\Omega K dF(\omega | F(\tilde{x}) \geq \alpha) = 0. && (3.11) \end{aligned}$$

Similarly, a DM with payoff u who uses $a^{F,w}$ (suboptimally) prefers G to F if and only if $V(\tilde{G}, w, a^{F,w}) - V(\tilde{F}, w, a^{F,w})$.[19] To extend this idea, suppose that for every K, $w_K = u + aK \in U_2$. Then by varying K, we potentially can recover *any* policy of the form $a = 0$ for $F(\tilde{x}) < \alpha$, $a = 1$ for $F(\tilde{x}) \geq \alpha$ as the optimal policy of some w_K. And thus for *all* payoff functions $\{w_K \in U_2\}$ to prefer G to F using their optimal response to F, it will be *necessary* that $V(\tilde{G}, u, a) \geq V(\tilde{F}, u, a)$ for all monotone policies $a(\alpha)$.

This heuristic argument can be made exact under the following condition:

(U_1-C) $U_1 = ccc(U_1 \cup \{\mathbf{1}, -\mathbf{1}\})$.

That is, U_1 is a closed convex cone, and it contains the constant functions. Under this condition, $r \in U_1$ implies that $r + K \in U_1$ for any K. Thus – crucially for our purposes – if $u \in U_2$, then $u + aK \in U_2$.

The following theorem establishes that, when (U_1-C) holds, then for small (differential) changes in the information structure, *(MIO)* is both *necessary and sufficient* for a ranking of information structures.

Theorem 3.3 *Let $(U_2, \mathcal{F})$ be an MDP pair such that (U_1-C) and (A) holds. Let $F^\theta(\omega, x)$ be smoothly parametrized by θ, with $F^\theta \in \mathcal{F}$ for all θ. Then $\frac{d}{d\theta} V^*(F^\theta, u) \geq 0$ for all $u \in U_2$, if and only if $F^{\theta + d\theta} \succ_{MIO-U_1} F^\theta$.*

Proof. Sufficiency follows from above. We prove a stronger result, that *(MIO)* is necessary for $V(\tilde{G}, u, a^F) \geq V(\tilde{F}, u, a^F)$, where a^F is the optimal policy for information structure F. We

[19]Note that this argument turns critically on our decision (implicit in (MIO)) to compare G to F using (fixed) policy functions that depend on a posterior's rank in the ex ante distribution, $\alpha_y = G(y)$. Had we used policy functions $a : [0,1] \to A$ based on some other index (not the marginal), the policy of increasing the action at $\alpha_y = \alpha$ for information structure G would no longer correspond to choosing a higher action when $G(\tilde{y}) \geq \alpha_i$ and (3.11) would fail.

proceed by assuming that *(MIO)* fails and constructing a contradiction. Suppose $G \not\succ_{MIO-U_1} F$. Then there is some $\bar{\alpha}$, and $\bar{r} \in U_1$ such that

$$\int_\Omega \bar{r}(\omega) dG(\omega | G(\tilde{y}) \leq \bar{\alpha}) > \int_\Omega \bar{r}(\omega) dF(\omega | F(\tilde{x}) \leq \bar{\alpha}). \tag{3.12}$$

Let $A = \{0, 1\}$. Then compute $K_{\bar{\alpha}} \equiv E[\bar{r}(\omega) | F(\tilde{x}) = \bar{\alpha}]$, and define

$$\bar{u}(\omega, a) = a\left(\bar{r}(\omega) - K_{\bar{\alpha}}\right).$$

By (U_1-C), since $\bar{r} \in U_1$, then $\bar{u} \in U_2$. An optimal policy for this DM is to set $a^F(\alpha) = 1$ if $\alpha > \bar{\alpha}$ and zero otherwise. Then $V(\tilde{G}, u, a^F) \geq V(\tilde{F}, u, a^F)$, if and only if

$$\int_{[0,1]} \int_\Omega \bar{u}(\omega, a^F(\alpha)) dG(\omega, G^{-1}(\alpha)) \geq \int_{[0,1]} \int_\Omega \bar{u}(\omega, a^F(\alpha)) dF(\omega, F^{-1}(\alpha)) \tag{3.13}$$

$$\iff \quad \int_\Omega [\bar{r}(\omega) - K_{\bar{\alpha}}] dG(\omega | G(\tilde{y}) \leq \bar{\alpha}) \leq \int_\Omega [\bar{r}(\omega) - K_{\bar{\alpha}}] dF(\omega | F(\tilde{x}) \geq \bar{\alpha}) \tag{3.14}$$

But, the expected value of a constant function is simply that constant. So (3.13) and (3.14) contradict (3.12).

For the case of small changes, consider F^θ, a smoothly parameterized family, and let $a^\theta(\alpha)$ be the optimal policy. Then, by the envelope theorem, $\frac{d}{d\theta} V(F^\theta, u, a^\theta)|_{\theta=\bar{\theta}} = \frac{\partial}{\partial\theta} V(F^\theta, u, a^{\bar{\theta}})\Big|_{\theta=\bar{\theta}}$. Then apply the above argument to compare $F^{\bar{\theta}+d\theta}$ and $F^{\bar{\theta}}$ at the policy $a^{\bar{\theta}}$. *Q.E.D.*

It is also possible to view the necessity result through the lens of statistical hypothesis testing. Consider testing the null hypothesis $H_0 : r(\omega) \geq 0$ against the alternative $H_1 : r(\omega) < 0$, for some $r(\omega) \in U_1$. However, we impose a constraint on the decision-maker: the test has an *average size* of $1 - \alpha$. That is, the *ex ante* probability (over all possible signals) of accepting the null hypothesis is $1 - \alpha$. A decision-maker facing this problem would then solve:

$$\begin{aligned} \max_{a(x) \in \{0,1\}} \quad & \int_\mathcal{X} \int_\Omega a(x) r(\omega) dF(\omega, x) \\ \text{s.t.} \quad & \int_\mathcal{X} a(x) dF(x) = 1 - \alpha \end{aligned}$$

Assuming the conditions of Theorem 3.3 apply, an optimal policy for this testing problem will set $a(x) = 1$ if and only if $F(\tilde{x}) \geq \alpha$. The *ex ante* expected payoff will be $(1-\alpha) \cdot E[r(\omega) | F(\tilde{x}) \geq \alpha]$, which, by *(MIO)*, is larger than the payoff from solving the inference problem associated with G, $(1-\alpha) \cdot E[r(\omega) | G(\tilde{y}) \geq \alpha]$. Thus, *(MIO)* can be interpreted as requiring that better information allows (on average) better inference about the returns to taking a higher action.[20]

3.4.2 Examples

We now revisit two of our examples from above.

Example 1. *Supermodular (Incremental returns increasing in ω).* Since U_1 is the set of nondecreasing functions, the relation $\succ_{SD-U_1}$ corresponds to FOSD. Thus G is higher than F in the supermodular monotone information order *(MIO-SPM)* if and only if for all $\alpha \in (0,1)$, $\tilde{G}(\omega|\tilde{\alpha}_y \geq \alpha) \succ_{FOSD} \tilde{F}(\omega|\tilde{\alpha}_x \geq \alpha)$. That is, high signals from G lead on average to higher posterior beliefs than high signals from F. Recall from above that if $F, G \in \mathcal{F}$, then $\tilde{F}(\omega|\alpha)$, $\tilde{G}(\omega|\alpha)$ are increasing in the sense of FOSD as α increases. So a higher signal is "good news" about the state of the world. If $G \succ_{MIO-U_1} F$ then high signals are (on average) "better news" under G than under F. Since high signals lead to high actions, one can think intuitively about G bringing about a better match between actions and the true state of the world.

It is interesting to relate *(MIO-SPM)* to sufficiency. To do this, we return to our earlier three-state, two-signal example. Recall the DM has uniform prior over the three states $\omega_1 < \omega_2 < \omega_3$. As illustrated in Figure 3.4, let $\tilde{y}$ be a signal that put equal likelihood on posteriors $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$ and $(0, \frac{1}{2}, \frac{1}{2})$. Note that $\tilde{y}$ is admissable for supermodular MDPs since $(0, \frac{1}{2}, \frac{1}{2}) \succ_{FOSD} (\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$. Consider an information structure $\tilde{z}$ that puts equal likelihood on posteriors $(\frac{2}{3}, \frac{1}{3}, 0)$ and $(0, \frac{1}{3}, \frac{2}{3})$. Clearly $\tilde{z}$ is also admissable for supermodular MDPs. Further, from the discussion in Section 3.3.3, $\tilde{y}$ and $\tilde{z}$ cannot be compared by Blackwell's sufficiency criteria. However, all supermodular decision-makers prefer $\tilde{z}$ to $\tilde{y}$, i.e. $\tilde{z} \succ_{MIO-SPM} \tilde{y}$. The reason is simple: $\tilde{z}$'s posteriors are more spread out according to FOSD. When the DM sees a low signal realization from $\tilde{z}$, her beliefs about the state are truly pessimistic, while the converse holds for high signal realizations.

Figure 3.5 illustrates this shift in a different way. For each signal, we label the lower (by FOSD) posterior L, and the higher posterior H. We see that signals that are higher according to *(MIO-SPM)* place more probability weight on (ω_2, L), and (ω_3, H). In fact, using ideas from bivariate stochastic dominance, it is possible to show that *(MIO-SPM)* can in general be characterized this way. Results from Meyer (1991) can be used to show that $G \succ_{MIO-SPM} F$ if and only if G is obtained from F using a "marginal-preserving spread" of the form illustrated in Figure 3.3.[21]

[20] To see why this alternative approach is essentially equivalent to our necessity proof, let $\bar{\lambda}$ be the Lagrange multiplier on the constrained optimization problem for $\bar{r}$, $\bar{\alpha}$ (defined as in the proof). Then $\bar{u}$ in the proof equals $a(\bar{r}(\omega) + \bar{\lambda})$.

[21] See also Levy and Paroush (1974) and Shaked and Shantikumar (1997) for more discussion of the stochastic dominance order associated with supermodular functions.

As a further comparison, consider the signal structure $\tilde{x}$ that puts equal likelihood on posteriors $(\frac{2}{3}, 0, \frac{1}{3})$ and $(0, \frac{2}{3}, \frac{1}{3})$. It follows easily that $\tilde{x}$ is admissable for supermodular decision problems and that $\tilde{y}, \tilde{z} \succ_{MIO-SPM} \tilde{x}$. Yet no two-point information structure is sufficient for $\tilde{x}$. Moreover, if we perturb $\tilde{x}$ so that it generates posteriors $(\frac{2}{3}-\varepsilon, \varepsilon+\eta, \frac{1}{3}-\eta)$ and $(\varepsilon, \frac{2}{3}-\varepsilon-\eta, \frac{1}{3}+\eta)$ with equal probability (where $\varepsilon > 0, \eta > 0$, and $\varepsilon + 2\eta < \frac{1}{3}$), then we still have $\tilde{y} \succ_{MIO-SPM} \tilde{x}$, even though the two signals are not Blackwell comparable (illustrated in Figure 3.4). So there is a sense in which our information order is *robust.*

Our examples highlight the fact that when there are three or more states of the world, there is a significant restriction in Blackwell's requirement that the posteriors of the "bad" signal must lie in the convex hull of the "good" signal posteriors. If there are only two states of the world, $|\Omega| = 2$ (the state of the world is a "dichotomy"), then this restriction is no longer severe, since all posteriors can be summarized in a single dimension. It can be shown that on dichotomies *(MIO-SPM)* is equivalent to sufficiency.

Example 2. *Concave (Incremental returns concave in ω).* When U_1 is the set of concave functions, the relation $\succ_{SD-U_1}$ corresponds to SOSD. Recall from above that (*MDP-$\mathcal{F}$*) implies that $\tilde{F}(\omega|\alpha)$ becomes "less risky" (by SOSD) as α increases. So $G \succ_{MIO-CV} F$ if, for any α, $\tilde{G}(\omega|\tilde{\alpha}_y \geq \alpha) \succ_{SOSD} \tilde{F}(\omega|\tilde{\alpha}_x \geq \alpha)$: high signals lead on average to less risky posteriors under G than under F. Since high signals lead to high actions and high actions lead to interim preferences that are more risk-averse, one can see immediately that having a better match between high signals and low risk will be valuable to the DM. We give an example in Section 3.6 where such preferences arise. Interestingly *(MIO-CV)* implies sufficiency when there are only two *or* three possible states of the world.

Just as we could find an *MDP* pair $(U_2, \mathcal{F})$ corresponding to *any* set of incremental return functions U_1 that forms a closed convex cone, we can describe the relevant monotone information order condition. Once again, $\succ_{MIO-U_1}$ says that if the average posteriors under G stochastically dominate the average posteriors under F (with respect to functions in U_1), then G is more informative than F for the class of problems in $(U_2, \mathcal{F})$. And we know that any stochastic dominance relationship induced by a closed convex cone can be described in terms of the extreme points of the cone.[22]

[22] It is interesting to observe in passing that if the DM's marginal returns are *both* nondecreasing and concave, we can concatenate our conditions — as is often done for straightforward preferences over gambles. For example, such a DM would be better off after a sequence of two changes to $\tilde{F}(\omega|\tilde{\alpha}_x \geq \alpha)$: one that makes higher states more likely (according to FOSD) and one that reduces risk (according to SOSD).

3.4.3 Monotone Information Orders Using Stochastic Single Crossing

In the previous section we indexed posteriors by their *ex ante* percentile and then compared information structures, finding that informativeness rankings could be derived using familiar notions of stochastic dominance. For this approach to be "tight," we required U_1 to contain the positive and negative constant functions. When U_1 does not contain the constants, but does contain some other function $\hat{r}$ and its additive inverse $-\hat{r}$, we can extend our results using stochastic single crossing in place of stochastic dominance.

We introduce an additional pair of restrictions (jointly denoted (R)) in order for $(U_2, \mathcal{F})$ to be admissable:

(R-U) U_1 is a closed convex cone, and there some $\hat{r}(\omega) \geq 0$ such that $\hat{r}, -\hat{r} \in U_1$.

(R-$\mathcal{F}$) For an $\hat{r}$ satisfying (R-U), for all $F \in \mathcal{F}$ with associated signal $\tilde{x}$, there exists $b > 0$ such that $E[\hat{r}(\omega)|x] \geq b$ for all x in support$(\tilde{x})$.

If U_1 contains the constant functions, then *(R)* holds with $\hat{r}(\omega) = 1$. If U_1 is the class of WSC(ω_0) functions, then *(R)* holds with $\hat{r} = 1_{\{\omega \approx \omega_0\}}$ (that is, $\hat{r}$ is an indicator function for a "small" interval containing ω_0).

We begin with a Lemma that generalizes a key step in the proof of sufficiency above.

Lemma 3.2 *Let U_2 satisfy (MDP-U) and assume (A) holds. Consider two information structures F and G, where $F(\omega) = G(\omega)$, and continuous strictly increasing functions $T_F : \mathcal{X} \to [0,1]$, $T_G : \mathcal{Y} \to [0,1]$. Define $m(\omega, \alpha) = G(\omega, T^{G-1}(\alpha)) - F(\omega, T_F^{-1}(\alpha))$. Then*

$$\int_\Omega \int_{[0,1]} u(\omega, a(\alpha)) dm(\omega, \alpha) \geq 0 \tag{3.15}$$

for all $u \in U_2$ and all $a : [0,1] \to A$ nondecreasing, if and only if for all $r \in U_1$ and all $\alpha \in (0,1)$,

$$\int_\Omega r(\omega) d_\omega m(\omega, \alpha) \leq 0. \tag{3.16}$$

An important thing to note is that $m(\cdot, \alpha)$ is *not necessarily* a probability distribution on ω; checking (3.16) when $r(\omega) \equiv 1$ entails checking that $\Pr(T_G(\tilde{y}) \leq \alpha) \leq \Pr(T_F(\tilde{x}) \leq \alpha)$. Thus, our choice of index in the last subsection can be immediately understood: if $\{\mathbf{1}, -\mathbf{1}\} \in U_1$, (3.16) holds for all $r \in U_1$ only if $\Pr(T_G(\tilde{y}) \leq \alpha) = \Pr(T_F(\tilde{x}) \leq \alpha)$, which is always true when $T_F(x) = F(x)$ and $T_G(y) = G(y)$.

When $\hat{r}, -\hat{r} \in U_1$, then (3.16) requires that

$$\int_\Omega \hat{r} d_\omega m(\omega, \alpha) = 0. \tag{3.17}$$

This motivates our new choice of indexing functions, T_F, T_G:

$$T_F(x) = \frac{\int_\Omega \hat{r}(\omega) d_\omega F(\omega, x)}{\int_\Omega \hat{r}(\omega) dF(\omega)} = F(x) \frac{E[\hat{r}|\tilde{x} \leq x]}{E[\hat{r}]}. \tag{3.18}$$

This choice of T_F, T_G guarantees that (3.17) will hold, meaning that it remains only to identify a condition under which (3.16) will hold for all $r \in U_1$, $r \neq \hat{r}$. If $\hat{r} = 1$, then $T_F(x) = F(x)$ as above. If U_1 is the class of WSC(ω_0) functions, then $T_F(x) = F(x|\omega_0)$. In general, (R-$\mathcal{F}$) implies that $T_F : \mathcal{X} \to [0, 1]$ is strictly increasing.[23]

Theorem 3.4 *Let $(U_2, \mathcal{F})$ be an MDP pair, suppose (A) and (R) hold. Let $F, G \in \mathcal{F}$. Then $V^*(G, u) \geq V^*(F, u)$ for all $u \in U_2$ if for all $\alpha \in (0, 1)$,*

$$G(\omega|T_G(\tilde{y}) \geq \alpha) \succ_{SC-U_1} F(\omega|T_F(\tilde{x}) \geq \alpha) \tag{MIO$'$}$$

where T_F and T_G are defined by (3.18).

(MIO') generalizes (MIO) by requiring that the average posteriors be ordered using stochastic single crossing rather than the (stronger) stochastic dominance order. To interpret this, the stochastic single crossing requirement can be restated as follows: for each α, and for all $r \in U_1$,

$$E[r(\omega)|T_G(\tilde{y}) \geq \alpha] \geq 0 \implies E[r(\omega)|T_F(\tilde{x}) \geq \alpha] \geq 0. \tag{3.19}$$

This contrasts with stochastic dominance, which requires $E[r(\omega)|T_G(\tilde{y}) \geq \alpha] \geq E[r(\omega)|T_F(\tilde{x}) \geq \alpha]$. If U_1 contains the constant functions, then stochastic dominance and stochastic single crossing coincide. For this reason, we maintain the notation $G \succ_{MIO-U_1} F$ even when applying (MIO').

To apply Theorem 3.4, it is useful to have an alternative condition for (MIO') that requires checking only a single inequality: for all $\alpha \in (0, 1)$ and $r \in U_1$,

$$E[r(\omega)|T_G(\tilde{y}) \geq \alpha] \cdot \Pr(T_G(\tilde{y}) \geq \alpha) \geq E[r(\omega)|T_F(\tilde{x}) \geq \alpha] \cdot \Pr(T_F(\tilde{x}) \geq \alpha). \tag{MIO$''$}$$

The proof of Theorem 3.4 proceeds by first establishing the equivalence of (MIO') and (MIO''), and then applying Lemma 3.2 to show that the result is implied by (MIO'').

Proof. It can be shown[24] that when U_1 is a closed convex cone, then for each α, (3.19) holds

[23] And continuity can be ensured (without changing the information content) using Lehmann's (1988) construction.

[24] See Jewitt (1986), Gollier and Kimball (1995), or Athey (1998a) for alternative proofs.

if and only if there exists a $\lambda(\alpha) \geq 0$ such that

$$E[r(\omega)|T_G(\tilde{y}) \geq \alpha] - \lambda(\alpha)E[r(\omega)|T_F(\tilde{x}) \geq \alpha] \geq 0. \tag{3.20}$$

for all $r \in U_1$. ($\lambda(\alpha)$ can be interpreted as the Lagrange multiplier in the problem of minimizing $E[r(\omega)|T_G(\tilde{y}) \geq \alpha]$ subject to the constraint that $E[r(\omega)|T_F(\tilde{x}) \geq \alpha] \geq 0$, and the left-hand side of (3.20) is the minimized value of the objective). But, checking (3.20) for $\hat{r}$ and $-\hat{r}$, both in U_1 by $(R\text{-}U)$, implies that $\lambda(\alpha) = E[\hat{r}(\omega)|T_G(\tilde{y}) \geq \alpha]/E[\hat{r}(\omega)|T_F(\tilde{x}) \geq \alpha]$. But, the latter ratio is in turn equal to $\Pr(T_G(\tilde{y}) \geq \alpha)/\Pr(T_F(\tilde{x}) \geq \alpha)$ (simply substitute in from the definitions of T_F and T_G in (3.18)). So, (3.20) is equivalent to (MIO''). To complete the proof, let $\tilde{F}(\omega,\alpha) = F(\omega, T_F^{-1}(\alpha))$ and likewise for G. By Lemma 3.2, (MIO'') is equivalent to the following: for all $u \in U_2$ and all monotone decision policies $a : [0,1] \to A$, $V(\tilde{G}, u, a(\cdot)) \geq V(\tilde{F}, u, a(\cdot))$. So for all $u \in U_2$, if a^F is an optimal (monotone) policy for u under F, $V^*(G,u) \geq V(\tilde{G}, u, a^F) \geq V^*(F,u)$. *Q.E.D.*

We now show that *(MIO')* is tight for a larger class of decision problems than that identified in Theorem 3.3 above.

Theorem 3.5 *Let $(U_2, \mathcal{F})$ be an MDP pair and suppose (A) and (R) hold. Let $F^\theta(\omega, x)$ be smoothly parametrized by θ, with $F^\theta \in \mathcal{F}$ for all θ. Then $\frac{d}{d\theta}V(F^\theta, u) \geq 0$ for all $u \in U_2$, if and only if $F^{\theta+d\theta} \succ_{MIO-U_1} F^\theta$.*

Proof. As in Theorem 3.3, we prove a stronger result, that *(MIO)* is necessary for $V(\tilde{G}, u, a^F) \geq V(\tilde{F}, u, a^F)$, where a^F is the optimal policy for information structure F. Suppose $G \not\succ_{MIO-U_1} F$. Using the representation (MIO''), this implies that there is some $\bar{\alpha}$, and $\bar{r} \in U_1$ such that

$$\int_\Omega \bar{r}(\omega)dG(\omega, T_G^{-1}(\bar{\alpha})) > \int_\Omega \bar{r}(\omega)dF(\omega, T_F^{-1}(\bar{\alpha})). \tag{3.21}$$

Let $A = \{0,1\}$. Then define

$$K_{\bar{\alpha}} = E[r(\omega)|T_F(\tilde{x}) = \bar{\alpha}]/E[\hat{r}(\omega)|T_F(\tilde{x}) = \bar{\alpha}].$$

The denominator is non-negative by $(R\text{-}\mathcal{F})$. Then define

$$\bar{u}(\omega, a) = a\left(\bar{r}(\omega) - K_{\bar{\alpha}}\hat{r}(\omega)\right).$$

By $(R\text{-}U)$, $\bar{u} \in U_2$. By *(MDP-$\mathcal{F}$)*, $E[\bar{r}(\omega) - K_{\bar{\alpha}}\hat{r}(\omega)|T_F(\tilde{x}) = \alpha]$ is single crossing in α, so that an optimal policy for this DM is to set $a^F(\alpha) = 1$ if $\alpha > \bar{\alpha}$ and zero otherwise. Then

$V(\tilde{G},u,a^F) \geq V(F,u,a^F)$, if and only if

$$\int_{[0,1]}\int_{\Omega} \bar{u}(\omega,a^F(\alpha))dG(\omega,T_G^{-1}(\alpha)) \geq \int_{[0,1]}\int_{\Omega} \bar{u}(\omega,a^F(\alpha))dF(\omega,T_F^{-1}(\alpha)) \qquad (3.22)$$

$$\iff \qquad \int_{\Omega}[\bar{r}(\omega)-K_{\bar{\alpha}}\hat{r}(\omega)]dG(\omega,T_G^{-1}(\bar{\alpha})) \leq \int_{\Omega}[\bar{r}(\omega)-K_{\bar{\alpha}}\hat{r}(\omega)]dF(\omega,T_F^{-1}(\bar{\alpha})) \qquad (3.23)$$

But, we have defined T_F and T_G so that $\int_{\Omega}\hat{r}(\omega)d_{\omega}G(\omega,T_G^{-1}(\alpha)) = \int_{\Omega}\hat{r}(\omega)d_{\omega}F(\omega,T_F^{-1}(\alpha))$ for all α. So (3.23) contradicts (3.21). As in Theorem 3.3, necessity for the case of a smoothly parameterized distribution follows from an application of the Envelope Theorem. *Q.E.D.*

If $(U_2,\mathcal{F})$ are an MDP pair corresponding to U_1, and $F,G \in \mathcal{F}$, the posteriors induced by each information structure are not required by *(MDP-$\mathcal{F}$)* to be totally ordered by stochastic dominance unless U_1 contains the constant functions. If, however, they do happen to be ordered by stochastic dominance, we can show that *(MIO)* and *(MIO′)* are equivalent.

(MDP-SD) *For all $F \in \mathcal{F}$, if x^H, x^L ∈support(F) and $x^H \geq x^L$, then $F(\omega|x^H) \succ_{SD-U_1} F(\omega|x^L)$.*

Proposition 3.1 *Let $(U_2,\mathcal{F})$ be an MDP pair and suppose (A), (R), and (MDP-SD) hold. Let $F,G \in \mathcal{F}$. Then (MIO′) is equivalent to (MIO).*

Proof. Note that $\int \hat{r}dF(\omega|x)$ and $-\int \hat{r}dF(\omega|x)$ are both increasing in x, so it must be the case that $\int \hat{r}dF(\omega|x) = \int \hat{r}dF(\omega)$ is constant in x. And likewise for G. Simplifying the expressions from (3.18), $T_F(x) = F(x)$, $T_G(x) = G(x)$. It follows immediately that *(MIO″)* (which we know is equivalent to *(MIO′)*) is equivalent to *(MIO)*. *Q.E.D.*

The result obtains because when posteriors are ordered by stochastic dominance and *(R-U)* holds, $E[\hat{r}|x]$ must be constant in x. It follows that $T_F(x)$ reduces immediately to $F(x)$, our earlier index. In contrast, when the posteriors are ordered by $\succ_{SC-U_1}$, $E[\hat{r}|x]$ can vary with x.

3.4.4 Examples

We now interpret the monotone information orders corresponding to Examples 3 and 4 from above. In the process we relate our information orders to Lehmann's efficiency criteria.

Example 3. *WSC(ω_0) (Incremental returns weak single crossing at ω_0).* Recall that if r is WSC(ω_0), then $r(\omega) \leq (\geq) 0$ as $\omega < (>) 0$. So $\hat{r} = 1_{\{\omega\approx\omega_0\}}$, and $T_F(x) = F(x|\omega_0)$ using Bayes' Rule. Then (MIO′) reduces to

$$\Pr(G(\tilde{y}|\omega_0)>\alpha\,|\,\omega) < \Pr(F(\tilde{x}|\omega_0)>\alpha\,|\,\omega) \qquad \text{for } \omega < \omega_0$$

$$\Pr(G(\tilde{y}|\omega_0)>\alpha\,|\,\omega) > \Pr(F(\tilde{x}|\omega_0)>\alpha\,|\,\omega) \qquad \text{for } \omega > \omega_0$$

Alternatively, $F(F^{-1}(\alpha|\omega_0)|\omega) - G(G^{-1}(\alpha|\omega_0)|\omega)$ is *WSC*(ω_0). Recall that the DM is essentially interested in knowing whether ω is above or below ω_0 — and *(MDP-$\mathcal{F}$)* ensures that high signals are "good news" about ω being above ω_0. The information condition implies that if the true state of the world $\omega > \omega_0$, then the probability of receiving "good news" under G (defined as $\tilde{y} > G^{-1}(\alpha|\omega_0)$) is higher than under F. If we impose the additional assumption (MDP-SD), it follows that for all $\tilde{x}$, $\tilde{y}$, $f(\omega_0|x) = g(\omega_0|y) = h(\omega_0)$ (the prior). Then, conditions (MIO) and (MIO′) are equivalent, and we can simply check that $g(\omega|G(\tilde{y})>\alpha) - f(\omega|F(\tilde{x})>\alpha)$ is *WSC*(ω_0) for all α. That is, conditional on a high posteriors, low states are less likely and high states are more likely under G as opposed to F, while neither signal is informative about the state ω_0.

The monotone information order for *WSC*(ω_0) can also be interpreted in terms of hypothesis testing. Consider testing the null hypothesis $H_0 : \omega = \omega_0$ against the alternative $H_1 : \omega < \omega_0$. However, the test is constrained to have a *size* of $1-\alpha$: conditional on H_0, the probability of rejecting the null is $1-\alpha$. Thus, we reject when $F(x|\omega_0) > \alpha$, and likewise for G. Since *MIO-WSC*(ω_0) is equivalent to requiring that $F(F^{-1}(\alpha|\omega_0)|\omega) - G(G^{-1}(\alpha|\omega_0)|\omega)$ if $\omega \geq (\leq)\omega_0$, we have the following interpretation. The probability of rejecting the null when the alternative is true ($\omega \leq \omega_0$) is greater for G than F; and the probability of rejecting the null when the alternative is false ($\omega > \omega_0$) is smaller for G than F. Thus, the hypothesis test using G is uniformly more powerful for a given size. This is exactly the interpretation given by Lehmann (1988) in his analysis, which will be discussed in more detail below.

To see a very simple example where this order can be applied, return to the portfolio allocation problem discussed in Section 3.3.3, where payoffs are given by $u(\omega, a) = v(a\omega + (1-a)\omega_0)$. For this problem, F is more informative than G if it provides a more powerful inference about whether or not investment in the risky asset is worthwhile.

Example 4. *Single Crossing (Incremental returns single crossing in ω).* The monotone information order for the case of single crossing payoff functions and signal distributions with monotone likelihood has been derived by Lehmann (1988). He showed that G is more informative than F for this class of problems ($G \succ_L F$) if for all $x \in \mathcal{X}$, $G^{-1}(F(x|\omega)|\omega)$ is nondecreasing in ω.[25] Jewitt (1997) has given further equivalent characterizations. We now show that Lehmann's order can be obtained as a special case of our result.

Recall that we noted above that the set of all functions that cross zero once from below at some ω_0 was the union of all the WSC(ω_0) sets. By quantifying over all crossing points ω_0, we can recover the information preferences of DMs with singe crossing payoff functions.

[25] Lehmann (1988) also considered necessity, although his theorem formally answers a slightly different question. He finds that the efficiency order is necessary and sufficient for one signal to return higher payoffs than another in every possible state (instead of on average, given the prior).The discussion following Example 3 above provides an outline of Lehmann's approach.

Theorem 3.6 *The following are equivalent: (i) $G \succ_L F$; (ii) $G \succ_{MIO-WSC(\omega_0)} F$ for all $\omega_0 \in \Omega$; (iii) $G \succ_{MIO-SPM} F$ for all prior distributions $H(\omega)$.*

Proof. First note from above that if $F(\omega|x)$ is increasing in $\succ_{SC-WSC(\omega_0)}$ for all ω_0 as x increases, then $F(\omega|x)$ is ordered by MLR.. And similarly, if $F(\omega|x)$ is increasing by FOSD as x increases for all priors H, then F is ordered by MLR.. We first show that (i) and (ii) are equivalent. Suppose $G \succ_{MIO-WSC(\omega_0)} F$. This means that $F(F^{-1}(\alpha|\omega_0)|\omega) - G(G^{-1}(\alpha|\omega_0)|\omega)$ is $WSC(\omega_0)$. For $WSC(\omega_0)$ to hold for every ω_0, the expression must be increasing in ω. Letting $x = F^{-1}(\alpha|\omega_0)$, and taking $G^{-1}(\cdot|\omega)$ of each term, we have $G^{-1}(F(x|\omega)|\omega) - G^{-1}(F(x|\omega_0)|\omega_0)$ is $WSC(\omega_0)$. This will hold for every ω_0 if and only if $G^{-1}(F(x|\omega)|\omega)$ is increasing in ω, which is exactly Lehmann's condition. Now consider the equivalence of (ii) and (iii). Suppose $G \succ_{MIO-SPM} F$. Then using the definition of FOSD, we have $F(\omega|F(x) \leq \alpha) \leq G(\omega|G(y) \leq \alpha)$. Applying Bayes' Rule, this is equivalent to $F(F^{-1}(\alpha)|\tilde{\omega} \leq \omega) \leq G(G^{-1}(\alpha)|\tilde{\omega} \leq \omega))$. Letting $x = F^{-1}(\alpha)$ and taking $G^{-1}(\cdot|\tilde{\omega} \leq \omega)$ of both sides, we see this is equivalent to $G^{-1}(F(x|\tilde{\omega} \leq \omega)|\tilde{\omega} \leq \omega) \leq G^{-1}(F(x))$. This is implied by $G \succ_L F$, and quantifying over all two-point priors, it implies $\succ_L$. *Q.E.D.*

To see how Lehmann's order can depart from the supermodular monotone information order in practice, return to Figure 3.4. In that example, recall that $\tilde{z} \succ_{MIO-SPM} \tilde{y} \succ_{MIO-SPM} \tilde{x}$. Using Lehmann's order, $\tilde{z} \succ_L \tilde{y}$, but $\tilde{x}$ does not satisfy MLR, so Lehmann's order does not apply to it. Of course, *none* of the three are ordered by sufficiency.

The analysis can be extended to other payoff classes. For instance, it follows from results in Athey (1998b) that when U_2 is the set of log-supermodular functions the appropriate information order is again $\succ_L$. Or in another example, consider the set of payoffs with incremental returns that positive only in some intermediate range between ω_0 and ω_0', where $\omega_0 < \omega_0'$. This case works out similarly to single crossing at a point.

3.5 Ordering Payoff Functions

We now provide conditions under which one decision-maker has a higher incremental return to improving her information than another decision-maker. Persico (1997) has investigated this question for the case of single crossing payoff functions and has shown that if $\frac{\partial}{\partial \alpha}u(\omega, a^F(\alpha)) - \frac{\partial}{\partial \alpha}v(\omega, a^F(\alpha))$ is single-crossing in ω, the first decision-maker (u) benefits more from a small increase in information than the second (v), according to Lehmann's information order for single crossing payoff functions. We generalize this result. To do this, we first limit our attention to marginal changes in the information structure. If two signal distributions are linked by a smoothly parameterized family of distributions which is information-ranked all along the way from F to G, we can then compare the incremental return to increasing the signal strength from F to G.

We need a small amount of notation: let Θ be a convex subset of $\mathbb{R}$, and suppose $\{F^\theta(\omega, x) : \theta \in \Theta\}$ is a family of information structures smoothly parametrized by θ. Let $a^{\theta,u}(\alpha)$ be a nondecreasing selection from $\int_\Omega u(\omega, a) d\tilde{F}^\theta(\omega|\alpha)$, and let $u^\theta(\omega, \alpha) = u(\omega, a^{\theta,u}(\alpha))$.

Theorem 3.7 *Suppose $(U_2, \mathcal{F})$ is an MDP pair and that the family $\{F^\theta(\omega, x) : \theta \in \Theta\}$ is smoothly parametrized on Θ, with $F^\theta \in \mathcal{F}$ for all θ. Let u, v be bounded measurable payoff functions. If, for some $\overline{\theta}$, $u^{\overline{\theta}}(\omega, \alpha) - v^{\overline{\theta}}(\omega, \alpha) \in U_2$, then $\frac{\partial}{\partial\theta} V(F^\theta, u) \geq \frac{\partial}{\partial\theta} V(F^\theta, v)$ at $\overline{\theta}$.*

Proof. By the envelope theorem, we have

$$\frac{\partial}{\partial\theta} V(F^\theta, u) = \int_\Omega \int_{[0,1]} u(\omega, a^{\theta,u}(\alpha)) d\left[\frac{\partial}{\partial\theta} \tilde{F}^\theta(\omega, \alpha)\right]. \tag{3.24}$$

So letting $w(\omega, \alpha) = u(\omega, a^{\theta,u}(\alpha)) - v(\omega, a^{\theta,v}(\alpha))$ and $m(\omega, \alpha) = \frac{\partial}{\partial\theta} \tilde{F}^\theta(\omega, \alpha)$, we have

$$\frac{\partial}{\partial\theta} V(u; F^\theta) - \frac{\partial}{\partial\theta} V(v; F^\theta) = \int_\Omega \int_{[0,1]} w(\omega, \alpha) d_\omega m(\omega, \alpha). \tag{3.25}$$

The assumption that $\tilde{F}^\theta$ is increasing in $\succ_{MIO-U_1}$ as θ increases implies that for any $r(\omega) \in U_1$, $\alpha' \in (0, 1)$

$$\int_\Omega r(\omega) d_\omega m(\omega, \alpha) \geq 0. \tag{3.26}$$

Lemma 3.2 then implies that (3.25) evaluated at $\overline{\theta}$ is nonnegative. *Q.E.D.*

Theorem 3.7 says that if $u^{\overline{\theta}} - v^{\overline{\theta}}$ is in U_2, i.e. if $u^{\overline{\theta}}$ is "more U_2", than $v^{\overline{\theta}}$, then the decision-maker with payoff function u has a higher marginal value for information than the DM with payoff function v. The theorem *does not* require that $u, v \in U_2$, or that the marginal value of information is nonnegative for each agent, although it is somewhat hard to imagine applying the Theorem when this is not the case (since policies might not be monotone).

If u is more sensitive to information than v in response to every signal in the family, we can compare changes in information that are not marginal. From this, we can derive comparative statics on the amount of amount of information acquired.

Theorem 3.8 *Suppose the conditions of Theorem 3.7 are satisfied, and that $u^{F^{\overline{\theta}}}(\omega, \alpha) - v^{F^{\overline{\theta}}}(\omega, \alpha) \in U_2$ for every $\theta \in \Theta$, where Θ is a closed interval. Let $C : \Theta \to \mathbb{R}$ be the cost of information; and let $\theta^*(u) = \arg\max_{\theta \in \Theta} V(\theta; u) - C(\theta)$. Then $\theta^*(u) \geq \theta^*(v)$ (in the strong set order).*

Proof. Let $V(\gamma^H; \theta) = V(u; F^\theta)$, and $V(\gamma^L; \theta) = V(v; F^\theta)$. Applying Theorem 3.7, $V(\theta; \gamma)$ is supermodular in (θ, γ), so $V(\theta, \gamma) - C(\theta)$ is supermodular in (θ, γ). By Topkis' (1978) Monotonicity Theorem, $\theta^*(\gamma)$ is nondecreasing in the strong set order. *Q.E.D.*

A few words are in order about the results of this section. A main drawback to Theorems 3.7 and 3.8 is that the conditions on u and v are not primitive, since they depend on the properties of the objective function evaluated at its optimum. In general, verifying the necessary condition may require knowing something about the shape of the optimizer, or about the curvature of the payoff functions. To take an apparently simple example, suppose $u(\omega, a) = v(\omega, a) + g(a)$, where v is supermodular. One can verify that *sufficient* conditions for u to acquire more information than v, where information is ranked by the supermodular information order, are that $a^F(\alpha; u) \geq a^F(\alpha; v)$, that $a^F_\alpha(\alpha; u) \geq a^F_\alpha(\alpha; v)$, and $v_{\omega aa} \geq 0$, (i.e. *marginal returns* are supermodular). The first condition will hold if $g(a)$ is increasing,[26] and the last is (reasonably) transparent, but even if g is linear, the second condition requires significant assumptions on the curvature of v. What conclusion should we draw from this? While we think that these results hold promise for applied modeling, it may be necessary to place a fair amount of structure on the model to make them operational.

On the other hand, the existing literature (prior to Persico (1997)) provides minimal guidance in this area. For example, an increase in the Blackwell information order increases the ex ante expected value of any function which is convex in the posterior beliefs. Thus, agent u will buy more Blackwell-ordered information than v if and only if $u^*(P) - v^*(P) = \int_\Omega [u(\omega, a^{*,u}(P)) - v(\omega, a^{*,v}(P))]\, dP(\omega)$ is convex in the posterior $P(\omega)$. While convexity of $u^*(P)$ and $v^*(P)$ follows as a simple consequence of optimality, convexity of the difference is not at all transparent. Checking the condition would almost certainly involve non-primitive assumptions similar to the ones described in this section.

3.6 Applications

We now present several applications of the results developed above. The first two applications are standard decision problems; the third examines adverse selection in a labor market equilibrium, while the fourth studies a coordination game. The examples demonstrate both strengths and weaknesses of the approach. In particular, while comparisons of information structures may be obtained immediately in a broad range of cases, comparative statics on information acquisition often require additional assumptions.

3.6.1 Application: Cost Uncertainty for Producers

A growing literature in Industrial Organization considers the value of information to firms in oligopoly models (see for instance Mirman, Samuelson, and Schlee (1994) and references

[26] To see this, let $u(\omega, a, \tau) = v(\omega, a)$ when $\tau = 0$, and $v(\omega, a) + g(a)$ when $\tau = 1$. Then u is supermodular in (a, τ) for every ω, so $a^*(\alpha, \tau)$ will be increasing in (α, τ). That is, $a^*(\alpha; u) \geq a^*(\alpha; v)$ for every α.

therein). The methods outlined above allow for some new results in this general class of problems.

Consider the problem faced by a producer who must choose quantity, q, to maximize some objective. We will compare the importance of information to firms under different market structures. However, we will restrict attention to covert information gathering, whereby we hold the strategies of other agents fixed when we analyze the effects of gathering information.

Write the firm's gross surplus as a general function, $R(q;\beta)$, where β parameterizes the market structure. In particular, we will consider $\beta \in \{M, D, S\}$, representing monopoly, duopoly, and the social planner's total surplus. The firm faces uncertainty about its cost: the total cost of producing q is given by $C(q;\omega)$. Thus, the firm's profits are given by $\pi(q,\omega;\beta) = R(q;\beta) - C(q;\omega)$. We consider a smoothly parameterized family of information structures, $\{F^\theta(\omega,\alpha) : \theta \in \Theta\}$, where $F^\theta(\alpha) = \alpha$ for each θ and all $\alpha \in [0,1]$. Expected profits given an α-percentile posterior are denoted $\overline{\pi}(q,\alpha;\beta) = R(q;\beta) - E[C(q,\omega)|\alpha]$. Assume that each of these functions is differentiable in q. Under appropriate assumptions, the optimal choice of quantity is then determined by $MR(q;\beta) = \frac{\partial}{\partial q}R(q;\beta) = \frac{\partial}{\partial q}E[C(q,\omega)|\alpha] = E[MC(q,\omega)|\alpha]$.

Suppose that θ indexes $F^\theta(\omega,\alpha)$ according to *(MIO-SPM)*, and that α indexes $F^\theta(\omega|\alpha)$ according to FOSD for all θ. We first apply our analysis from above to characterize increasing information for this problem. If we assume that $MC(q;\omega)$ is decreasing in ω, then $\pi(q,\omega;\beta)$ is supermodular in (q,ω). Since *(MDP-$\mathcal{F}$)* is satisfied, the optimal choice, $q^\beta(\alpha;\theta)$, will be nondecreasing in α. Better information (according to *(MIO-SPM)*) leads to higher ex ante expected payoffs.

We now turn to consider how the marginal value of information changes under different market conditions. We begin with the following result.

Proposition 3.2 *Consider the model described above, where for each θ, α^θ indexes $F^\theta(\omega|\alpha^\theta)$ according to FOSD. Assume that the smoothly parameterized family $\{F^\theta(\omega,\alpha)\}$ is ordered by MIO-SPM. In addition, assume that $MC(q;\omega)$ is decreasing in ω, increasing in q, and submodular in (ω,q). Finally, assume that $q^H(\alpha) > q^L(\alpha)$ and $q^{H\prime}(\alpha) > q^{L\prime}(\alpha)$. Then the marginal value of θ is higher for firm β^H than firm β^L.*

Proof. By Theorem 3.7, it suffices to show that $\pi_q(q^H(\alpha),\omega;\beta^H)q^{H\prime}(\alpha) - \pi_q(q^L(\alpha),\omega;\beta^L)q^{L\prime}(\alpha)$ is nondecreasing in ω. If $q^{H\prime}(\alpha) > q^{L\prime}(\alpha)$, then our result obtains if $\pi_q(q^H(\alpha),\omega;\beta^H) - \pi_q(q^L(\alpha),\omega;\beta^L)$ is nondecreasing in ω (since each term is separately nondecreasing). Recall that $\pi_q(q^\beta(\alpha),\omega;\beta) = MR(q^\beta(\alpha);\beta) - MC(q^\beta(\alpha),\omega)$ for each firm. Since the first term does not vary with ω, we can consider whether $MC(q^L(\alpha),\omega) - MC(q^H(\alpha),\omega)$ is nondecreasing in ω, which follows since $MC(q,\omega)$ is submodular. *Q.E.D.*

To interpret the conditions on the cost function, note that they are satisfied if $C(q,\omega) = q^2/\omega$. To interpret the conditions on the quantities chosen, which of course are not primitive conditions, it is useful to consider some examples. First consider comparing a monopolist's problem with that of the social planner. Let $R(q;\beta^M) = P(q) \cdot q$ and $R(q;\beta^S) = \int_0^q P(t)dt$.

As usual, it follows directly that $q^S(\alpha) > q^M(\alpha)$, that is, the monopolist underprovides quantity. Now consider the terms $q^{S\prime}(\alpha)$ and $q^{M\prime}(\alpha)$. The implicit function theorem yields

$$q^{\beta\prime}(\alpha) = \frac{\frac{\partial}{\partial\alpha} E[MC(q^\beta(\alpha),\omega)|\alpha]}{\frac{\partial}{\partial q} MR(q^\beta(\alpha);\beta) - \frac{\partial}{\partial q} E[MC(q^\beta(\alpha),\omega)|\alpha]}.$$

Thus, $q^{S\prime}(\alpha) > q^{M\prime}(\alpha)$ if $MC(q,\omega)$ is submodular in (q,ω), if $MC(q,\omega)$ is convex in q, and if $P'(q^S(\alpha)) > 2P'(q^M(\alpha)) + q^M(\alpha)P''(q^M(\alpha))$. If P is linear or convex, a sufficient condition for the latter inequality is the commonly assumed requirement that the marginal revenue curve is steeper than the demand curve.

We can also compare the value of information for a monopolist and a duopolist. We suppose that both duopolists have the same initial signal structure as the monopolist, but only one duopolist. has the opportunity to (covertly) purchase additional information (it is also possible to study the information acquisition game, but we will not consider that case here). With these symmetry assumptions, the result $q^M(\alpha) > q^D(\alpha)$ obtains so long as marginal revenue is downward sloping. Further, $q^{M\prime}(\alpha) > q^{D\prime}(\alpha)$ if marginal cost is submodular, and if

$$2P'(q^M(\alpha)) + q^M(\alpha)P''(q^M(\alpha)) > 2P'(2q^D(\alpha)) + q^D(\alpha)P''(2q^D(\alpha)).$$

This condition is satisfied for the constant elasticity demand curve.

Summarizing, we see that under several reasonable conditions, the social planner has a larger incentive to acquire information than the monopolist, but the monopolist may have a larger incentive to acquire information than the duopolist.

3.6.2 Application: Screening for a "Target" Hire

An entrepreneur or manager needs to hire a worker or subcontractor to perform a very specific task. She interviews an interested party, who appears to have roughly the right characteristics. The manager now has to make a single take-it-or-leave it monetary offer based on information gleaned from the interview. What sort of signal from the interview will cause the manager to make a large offer? And what would it mean for the screening process to be more effective?

To model such a situation, we assume that there is some optimal task outcome ω^*. Hiring the agent would result in outcome ω, and a payoff of $r(\omega)$, which achieves a maximum at ω^*, and decreases as ω moves away in both directions. So $r(\omega)$ is concave (but not necessarily symmetric—it's not just the distance from perfection that matters, but the direction). A wage offer a will be accepted with probability $p(a)$, which is increasing. The manager has a signal x about ω (arising from an information structure $F(\omega, x)$), and chooses a to maximize

$p(a)E[r(\omega) - a|x]$. We can write the payoff function $u(\omega, a)$ as $u(\omega, a) = p(a)[r(\omega) - a]$. In order for signals to be ordered in such a way that the manager will offer more money following a "good" signal, regardless of the exact shape of p and r, it must be the case that x orders the posteriors by SOSD (see Example 2 in Section 3.3.3 for interpretations). If the potential signals from the interview satisfy this condition, $a^F(x)$ will be monotone regardless of the exact shape of p and r.

Now for the question of what constitutes better information. Applying *(MIO-CV)*, G will be a more informative interview that F if $\forall \alpha \in [0, 1]$,

$$G(\omega|G(x) > \alpha) \succeq_{SOSD} F(\omega|F(x) > \alpha).$$

The intuition is simple: after the interview the expected outcome if the agent is hired is always ω^*, but there is some residual uncertainty. The manager does not like this residual risk—in particular, the less risk she believes there is, the more aggressively she will pursue the candidate. A more informative screening process is one where high signals (which are more likely to lead to hires) imply less residual risk.

3.6.3 Application: Adverse Selection and Labor Markets

Our next example suggests how our information orders can be applied to adverse selection markets. Consider the following stylized situation. Workers live for two periods and spend the first period training (in school). Each worker's productivity is unknown with prior $H(\omega)$, and schooling generates a noisy signal $\tilde{x}$ about underlying ability. Suppose that the joint distribution of (ω, x) is $F(\omega, x)$. This signal is observed by the school, but not by the general labor market. Instead, only the top $1 - \alpha$ fraction of the class "graduates," while the rest do not — the labor market observes only if a given worker graduated. Each firm has production function $J(\omega)$, increasing in ω, and the labor market is competitive so that workers receive a wage $E[J(\omega)|\mathfrak{I}]$, where $\mathfrak{I}$ is the information available to the market about productivity.

The market wage for graduates will be $E[J(\omega)|F(\tilde{x}) \geq \alpha]$, and for non-graduates $E[J(\omega)|F(\tilde{x}) < \alpha]$. Now suppose that the schooling process becomes more informative about the worker's ability, in the sense that F increases to G in the supermodular monotone information order *(MIO-SPM)*. It follows immediately that the wage for graduates will increase since

$$E[J(\omega)|G(\tilde{y}) \geq \alpha] \geq E[J(\omega)|F(\tilde{x}) \geq \alpha],$$

while the wage decreases for those who do not graduate. The point is that the labor market interprets a failure to graduate as bad news about ability, and as worse news when the schooling process is more revealing. Note that the average wage (and average production) for the whole economy is constant at $E[J(\omega)]$, but inequality increases with information.

To extend this, suppose now that a fraction $1-\beta$ of workers, $\beta < \alpha$ can be hired into jobs that are "skill-sensitive" — that is, have production function $S(\omega)$, with $S'(\omega) > J'(\omega)$ for all ω. And suppose that $E[S(\omega)] = E[J(\omega)]$ so that "on average" productivity is the same at the two jobs. Assuming that each worker gets her expected marginal product as a wage, the equilibrium in this two-job labor market has a fraction β of workers, all of whom graduated, going to "skill-sensitive" jobs at a wage $E[S(\omega)|F(\tilde{x}) \geq \alpha]$, the remaining graduates taking less skill-sensitive jobs at a lower wage (they are rationed), and non-graduates receiving $E[J(\omega)|F(\tilde{x}) < \alpha]$ as before.

Consider an increase in the information generated by school screening in the sense of *(MIO-SPM)*. This will increase the wage for graduates and decrease the wage for non-graduates. But it will also increase the total production of the economy by leading to better matching between high-skill workers and skill-sensitive jobs. Moreover, as the fraction of skill-sensitive jobs, β, increases toward α, the social returns to better screening by the schools increase (assuming the planner cares only about gross production and not inequality). Similarly, if $S(\omega,\tau)$ is supermodular in τ, then the social returns to better screening will be increasing in τ.

3.6.4 Application: Coordination Under Uncertainty

This section considers a game where both players can choose to acquire information. The game involves two players with symmetric payoffs $u(\omega, a_i, a_j)$ where a_i is player i's action. Player i receives signal $\tilde{\alpha}_i$, with joint distribution $F^{\theta_i}(\omega, \alpha_i)$, where $F^{\theta_i}(\alpha_i) = \alpha_i$ and θ_i reflects signal quality. Assume the $\tilde{\alpha}_i$'s are independent conditional on ω. We let $u(\omega, a_1, a_2) = a_1[\omega + \gamma a_2 + K] - \frac{1}{2}\tau a_1^2$, and restrict $\tau > \gamma > 0$. In this supermodular game, players make unobserved choices of signal quality (θ_1, θ_2) and then choose their strategies (a_1, a_2). We look for symmetric Nash equilibria. Conditional on the information structure, Player 1 maximizes

$$\int_{\Omega} \left[a_1[\omega + \gamma E[a_2|\omega; \theta] + K] - \frac{1}{2}\tau a_1^2 \right] dF^{\theta_1}(\omega|\alpha_1)$$

to find an optimal action $a_1(\alpha_1) = \frac{1}{\tau}\left[E[\omega|\alpha_1;\theta] + \gamma E[a_2|\alpha_1;\theta] + K\right]$. In the unique (essentially) symmetric equilibrium, $a_i(x) = \frac{1}{\tau-\gamma}\left[E[\omega|\alpha_i;\theta_i] + K\right]$.

The first question to ask is when the optimal policy will be monotone in α_i. Clearly requiring that the posteriors are ordered by FOSD is sufficient; actually all that is needed is that the posteriors be ordered according to their first moments. Since any two distributions can be ranked according to their means, assuming *(MDP-$\mathcal{F}$)* is simply notational. The linearity of payoffs arises since when j plays her equilibrium strategy, $E[a_j|\omega] = \frac{1}{\tau-\gamma}[\omega + K]$, and so $u_i(a_i, \omega) = a\frac{\tau}{\tau-\gamma}[\omega + K]$. Thus marginal returns are of the form $u_a(\omega) = c + d\omega$, where $d > 0$. The set of marginal returns satisfying these restrictions forms a convex cone, with corresponding *(MDP)* and *(MIO)* conditions. The condition for monotonicity of the policy is that $E[\omega|\alpha_i]$

be increasing in α_i. The condition under which G is more informative than F for all decision-makers with linear marginal returns can be derived immediately from Theorem 3.2: $\forall \alpha \in [0,1]$,

$$E[\omega|F(x) < \alpha] \geq E[\omega|G(y) < \alpha]. \tag{MIO-LIN}$$

This is the linear monotone information order.

We now consider the information acquisition choice. Assume that θ_i indexes F_i according to (*MIO-LIN*). Expected payoffs to player i are $\Pi(\theta_i, \theta_j) = E[u(\omega, a_i(\alpha_i), a_j(\alpha_j))|\theta_1, \theta_2]$, so the symmetric equilibrium must also satisfy $\Pi_i(\theta_i, \theta_i) - C'(\theta_i) = 0$, the first order condition for information acquisition. We are also interested in how changes in the parameters alter the amount of information gathered in equilibrium. Straightforward algebra yields

$$u^{\theta_i}(\omega, \alpha_i) = \frac{\tau}{(\tau-\gamma)^2}\left[E[\omega|\alpha_i; \theta_1] + K\right]\left[(\omega + K) - \frac{1}{2}(E[\omega|\alpha_i;\theta_i] + K)\right],$$

$$\frac{\partial}{\partial \alpha_i} u^{\theta_i}(\omega, \alpha_i) = \frac{\tau}{(\tau-\gamma)^2}\left[\omega - E\left[\omega|\alpha_i; \theta_i\right]\right]\frac{\partial}{\partial \alpha_i} E[\omega|\alpha_i;\theta_i].$$

Comparative statics follow immediately from an application of Theorem 3.7. A change in K, the known marginal benefit to acting, has no effect on the amount of information gathered. An increase in the quadratic cost of action, τ, decreases the amount of information gathered. And as is intuitive, an increase in γ, the returns to coordination, increases information gathering. Similarly, agent's actions are increasing γ and K and decreasing in τ. Endogenizing the information structure in this game *reinforces* known complementarities.

3.7 Conclusion

This chapter has obtained a new set of results for the standard Bayesian decision problem. We have provided a general analysis of when one signal is more valuable than another signal to a given class of decision-makers, and also provided conditions under which decision-makers within a given class will differ systematically in their marginal value for information. We have applied this general framework to several payoff classes of particular economic relevance, and described their corresponding monotone information orders. The results can be applied to a variety of economic settings, allowing characterizations of what sort of information is valuable in a given environment, and which sorts of agents care most about acquiring it.

An important, but potentially difficult extension of this research, is to analyze the value of information in strategic settings. In such a setting, information may have both a direct and a strategic value. In standard incentive problems, information has only strategic value, while in Persico's (1997) study of auctions, information is acquired secretly so there is only a direct value. In the industrial organization literature discussed in Section 3.6.1, results combining both direct and strategic benefits are obtained, but they typically rely on restrictive functional form assumptions. The examples from this literature suggest that unambiguous results about the value of information may not be available in many settings. It remains to be seen whether general characterization results can be obtained in game-theoretic models.

3.8 Appendix

Proof of Lemma 3.2. Assume first that A is finite, $A = \{a_1, ..., a_n\}$ with $a_{i+1} > a_i$. Then for any monotone nondecreasing $a(\alpha)$, there exist $\alpha_0 = 0 \leq \alpha_1 \leq ... \leq \alpha_{n-1} \leq \alpha_n = 1$ such that $a(\alpha) = a_i$ on $[\alpha_{i-1}, \alpha_i]$. Then

$$
\begin{aligned}
\int_\Omega \int_{[0,1]} u(\omega, a(\alpha)) dm(\omega, \alpha) &= \int_\Omega \sum_{i=1}^{n} u(\omega, a_i) \left[m(\alpha_i|\omega) - m(\alpha_{i-1}|\omega) \right] dm(\omega) \\
&= \int_\Omega \left\{ \begin{array}{c} u(\omega, a_n) m(1|\omega) - u(\omega, a_1) m(0|\omega) \\ - \sum_{i=1}^{n-1} \left[u(\omega, a_{i+1}) - u(\omega, a_i) \right] m(\alpha_i|\omega) \end{array} \right\} dm(\omega) \\
&= -\int_\Omega \left\{ \sum_{i=1}^{n-1} \left[u(\omega, a_{i+1}) - u(\omega, a_i) \right] m(\alpha_i|\omega) \right\} dm(\omega) \\
&= -\sum_{i=1}^{n-1} \int_\Omega r_i(\omega) dm_\omega(\omega, \alpha_i)
\end{aligned}
$$

for some $r_1(\omega)$,...,$r_{n-1}(\omega)$ with $r_i(\omega) \in U_1$. The third equality uses the fact that $m(1|\omega)$, $m(0|\omega)$ are zero a.e. with respect to $m(\omega)$. The last step follows from *(MDP-U)*. Then clearly (3.16) implies (3.15). Moreover, suppose (3.16) fails for some $\hat{\alpha} \in (0, 1)$, $\hat{r} \in U_1$. Then define $\hat{u}(\omega, a) = a\hat{r}(\omega)$. And let $\hat{a}(\alpha) = a_1$ on $[0, \hat{\alpha})$ and $\hat{a}(\alpha) = a_n$ on $[\hat{\alpha}, 1]$. By *(MDP-U)*, $\hat{u} \in U_2$, and $\hat{a}$ is clearly nondecreasing. Substituting into the derivation above,

$$
\int_\Omega \int_{[0,1]} \hat{u}(\omega, \hat{a}(\alpha)) dm(\omega, \alpha) = -(a_n - a_1) \int_\Omega \hat{r}(\omega) d_\omega m(\omega, \hat{\alpha}) < 0.
$$

So (3.15) must imply (3.16).

Now suppose A is compact. We know from above that (3.16) is equivalent to (3.15) holding for all $a(\alpha)$ step functions. We show that (3.15) holds for all nondecreasing functions $a(\alpha)$ if and only if it holds for all step functions. Let $a(\alpha)$ be some nondecreasing function, and let $a^1(\alpha), a^2(\alpha), ...$ be a sequence of step functions converging to $a(\alpha)$. Then since u is continuous in a, then $u(\omega, a^k(\alpha))$ will be converging to $u(\omega, a(\alpha))$. And since $u(\omega, a)$ is bounded, we can apply the Lebesgue Convergence Theorem to show

$$\int_{\Omega\times[0,1]} u(\omega, a^k(\alpha))dm(\omega,\alpha) \longrightarrow \int_{\Omega\times[0,1]} u(\omega, a(\alpha))dm(\omega,\alpha).$$

Then clearly, (3.15) will hold for all a nondecreasing if and only if it holds for all step functions. *Q.E.D.*

3.9 References

Athey, Susan (1998a): "Characterizing Properties of Stochastic Objective Functions," M.I.T. Working Paper 96-1R.

Athey, Susan (1998b): "Comparative Statics under Uncertainty: Single Crossing Properties and Log-Supermodularity," M.I.T. Working Paper 97-11R.

Blackwell, David (1951): "Comparisons of Experiments," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics*, 93–102.

Blackwell, David (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Blackwell, David and M. Girshik (1954): *Theory of Games and Statistical Decisions*, John Wiley & Sons.

Border, Kim (1991): "Functional Analytic Tools for Expected Utility Theory," in *Positive Operators, Riesz Spaces, and Economics*, ed. C.D. Aliprantis, K. Border and W. Luxemburg. Springer Verlag.

Dewatripont, M., I. Jewitt, and J. Tirole (1999): "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies, 66, 183-200.*

Gollier, C. and M. Kimball (1995): "Toward a Systematic Approach to the Economic Effects of Uncertainty I: Comparing Risks," Mimeo, University of Toulouse.

Jewitt, Ian (1986): "A Note on Comparative Statics and Stochastic Dominance," *Journal of Mathematical Economics*, 15, 249–254.

Jewitt, Ian (1997): "Information and Principal Agent Problems," Mimeo, University of Bristol.

Karlin, Samuel (1968): *Total Positivity*: *Volume I.* Stanford University Press.

Karlin, Samuel and William Studden (1966): *Tchebycheff Systems with Applications in Analysis and Statistics*, New York: Interscience.

Le Cam, Luc (1964): "Sufficiency and Approximate Sufficiency," *Annals of Mathematical Statis-*

tics, 35, 1419–1455.

Lehmann, E.L. (1988): "Comparing Location Experiments," *Annals of Statistics*, 16, 521-533.

Levy, Haim and Jacob Paroush, "Toward Multivariate Efficiency Criteria," *Journal of Economic Theory*,7, 129-142.

Meyer, Margaret (1990): "Interdependence of Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality under Uncertainty," Nuffield College Discussion Paper, no. 49.

Milgrom, Paul (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

Milgrom, Paul and John Roberts (1990): "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–1277.

Milgrom, Paul and Christina Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

Mirman, L., L. Samuelson, and E. Schlee (1994): "Strategic Manipulation of Information in Duopolies." *Journal of Economic Theory*, 36, 364-384.

Persico, Nicola (1996): "Information Acquisition in Affiliated Decision Problems," Mimeo, UCLA.

Persico, Nicola (1997): "Information Acquisition in Auctions," Mimeo, UCLA.

Ross, Stephen (1981): "Some Stronger Measures of Risk Aversion in the Small and the Large with Applications," *Econometrica* 49, 621-638.

Rothschild, Michael, and Joseph Stiglitz (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory* 2, 225-243.

Shaked, Moshe and George Shantikumar (1997): "Supermodular Stochastic Order and Positive Dependence of Random Variables," *Journal of Multivariate Analysis*, 61, 86–101.

Shannon, Chris (1995), "Weak and Strong Monotone Comparative Statics," *Economic Theory*, 5, 209-227.

Strassen, V. (1965): "The Existence of Probability Measures with Given Marginals," *Annals of Mathematical Statistics*, 36, 423–439.

Topkis, Donald (1978): "Minimizing a Submodular Function on a Lattice," *Operations Research*, 26, 305–321.

Figure 1: Sufficiency with 3 states and 2-point signal distributions.

States: $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Assume $\omega_1 < \omega_2 < \omega_3$. Prior: $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Two signals, $\tilde{x}$ and $\tilde{y}$, where $\tilde{y}$ is sufficient for $\tilde{x}$.

Each signal has two (equally likely) possible realizations:

$\mathcal{X} = \{x^L, x^H\}$, $\mathcal{Y} = \{y^L, y^H\}$, $\Pr(\tilde{x} = x^L) = \Pr(\tilde{y} = y^L) = \frac{1}{2}$.

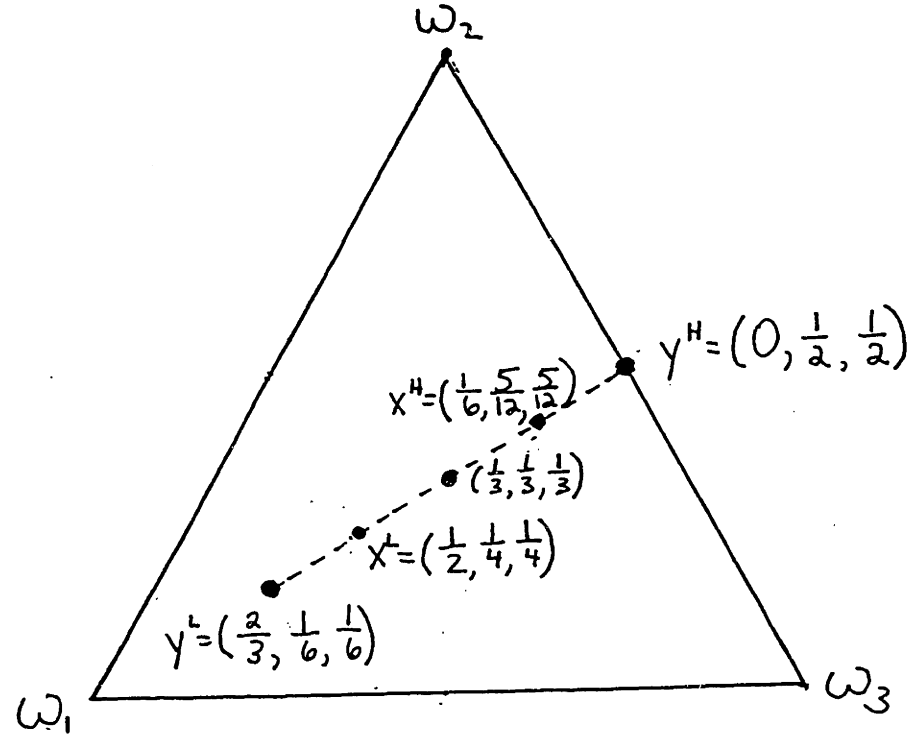

Figure 2: Indexing Posteriors by Ex Ante Percentile

$\alpha = F(x_\alpha)$, $\alpha = G(y_\alpha)$.

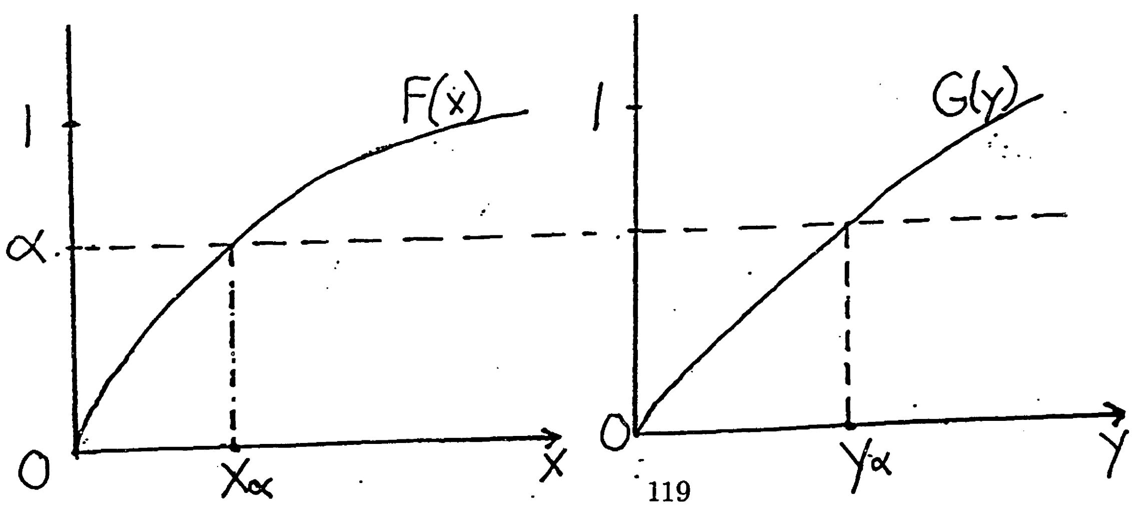

Figure 3: Illustration of Monotone Decision Problem (MDP) Conditions.

States: $\Omega = \{\omega_1, \omega_2, \omega_3\}$. $\mathcal{X} = \{x^L, x^H\}$. The posteriors generated by $\tilde{x}$ are illustrated in the diagram.

Case 1: U_2 is set of supermodular functions.

$F(\omega|x)$ satisfies (MDP) only if $F(\omega|x^H) \succ_{FOSD} F(\omega|x^L)$ for all $x^H > x^L$.

Given $F(\omega|x^L)$ as illustrated in diagram, only posteriors within *solid* lines satisfy (MDP) restriction.

Thus, $\tilde{x}$ *is admissable.*

Case 2: U_2 is set of functions with single crossing incremental returns.

$F(\omega|x)$ satisfies (MDP) only if $F(\omega|x^H) \succ_{MLR} F(\omega|x^L)$ for all $x^H > x^L$.

Given $F(\omega|x^L)$ as illustrated in diagram, only posteriors within *dotted* lines satisfy (MDP) restriction.

Thus, $\tilde{x}$ *is not admissable.*

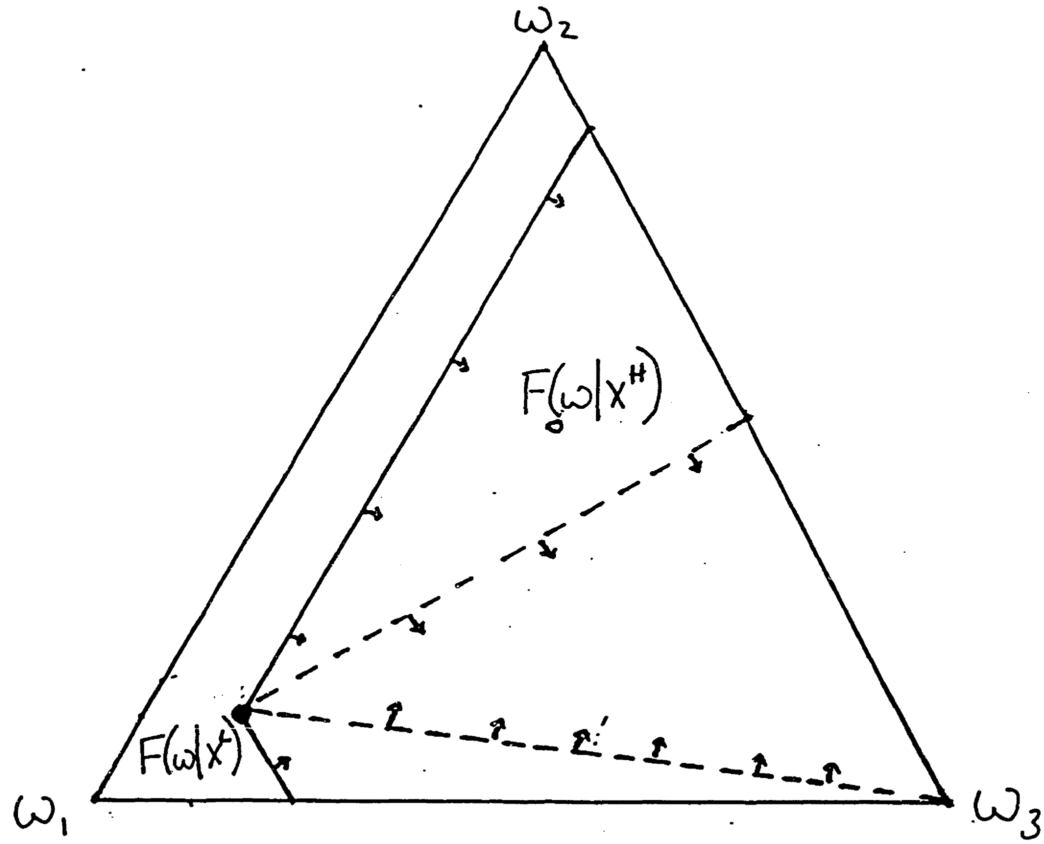

Figure 4: Supermodular Monotone Information Order with 3 states and 2-point signal distributions.

States: $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Assume $\omega_1 < \omega_2 < \omega_3$. Prior: $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Three signals, $\tilde{x}$, $\tilde{y}$, and $\tilde{z}$, where $\tilde{z} \succ_{MIO-SPM} \tilde{y} \succ_{MIO-SPM} \tilde{x}$.

No two signals are ordered by sufficiency.

Each signal has two (equally likely) possible realizations which satisfy *MDP-SPM*.

Robustness: If $F(\omega | x^L)$ is anywhere within the dotted region, $\tilde{y} \succ_{MIO-SPM} \tilde{x}$.

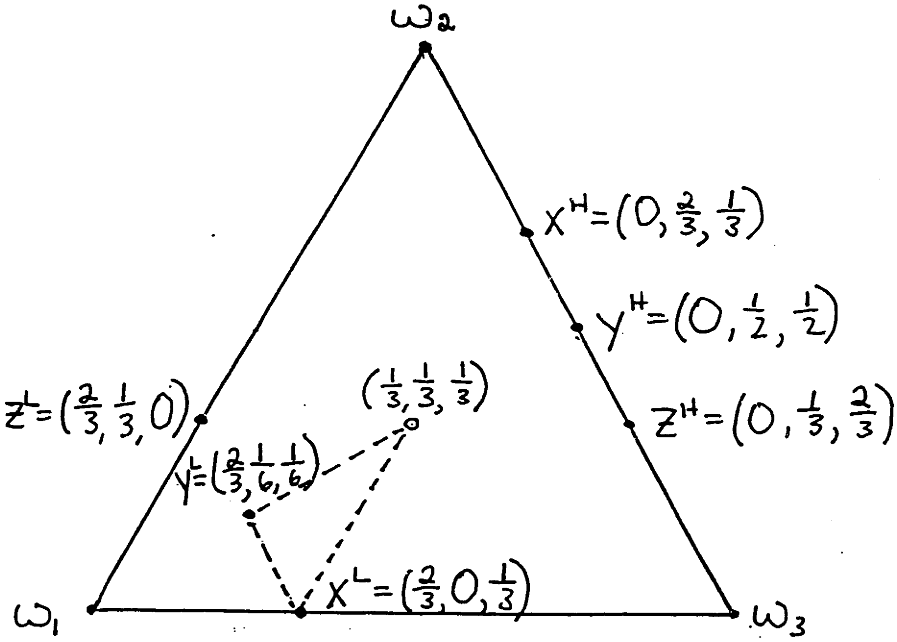

Figure 5: Illustration of the Supermodular Monotone Information Order.

States: $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Assume $\omega_1 < \omega_2 < \omega_3$.

Two signals, $\tilde{y}$ and $\tilde{z}$, where $\tilde{z} \succ_{MIO-SPM} \tilde{y}$.

Moving from $\tilde{y}$ to $\tilde{z}$ moves probability weight *onto*

(*Low signal*,ω_2) and (*High signal*,ω_3), and away from

(*High signal*,ω_2) and (*Low signal*,ω_3).

Change in Probability of event from $\tilde{y}$ to $\tilde{z}$

	ω_1	ω_2	ω_3
High Signal Realization	No Change	–	+
Low Signal Realization	No Change	+	–

Chapter 4

The Relationship between Information and Trade: A Look at the Lemons Market

4.1 Introduction

This chapter poses a basic question: do greater information asymmetries reduce the potential for realizing gains from trade? It looks for an answer in the context of the classic asymmetric information model, Akerlof's (1970) market for lemons. Sellers have some amount of private information, while buyers are uninformed. There is no potential for screening or signalling, nor any mechanism for bargaining — a price is posted and buyers and sellers decide whether or not to enter the market. I consider how the potential for trade depends on the quality of the seller's information, and hence the degree of information asymmetry between the two sides of the market.

This question might seem to have a rather obvious answer — namely, the worse the adverse selection, the less trade. And at an intuitive level, it seems easy to understand the idea that some markets are particularly plagued by adverse selection. But relative to what? Would the market become more efficient if there was slightly less information asymmetry? Could it become *less* efficient? While one can obtain answers under special assumptions (normality, some fraction of sellers perfectly informed), little work directly tackles the question of how changes in information affect the possibilities for trade.[1]

A first, and perhaps counterintuitive, observation is that trade need not decrease as the degree of information asymmetry increases. This is because there are two competing effects:

[1] Persico's (1997) paper on auction markets probably comes closest. Glosten and Milgrom (1985) have a comparative static result on the effect of information in bid-ask markets that gets at one aspect of the problem.

for a fixed supply curve, the demand curve will shift down as information asymmetry increases — the "buyer's curse" intensifies; however, the supply curve shifts as well. Information sorts sellers: roughly, more information drives some sellers away from the market, but it also drives some sellers towards the market. Following a change in the information structure, it may be possible to find a new price at which the latter effect dominates and the possibilities for trade improve vis a vis the earlier situation. For simple partition information structures, a very sharp characterization can be obtained: when new information arrives on the market margin, it may drive some sellers out of the market and reduce trade; if the new information is deep in the market, the price remains unchanged and trade is unaffected; and if sellers who are out of the market acquire new information, trade may increase as sellers who get bad news are pushed towards the market. This same logic extends to general information structures using the appropriate notion of "more information." I also obtain the weaker, but quite general, result that first best trade, once lost, can never be restored by increasing the information asymmetry between buyers and sellers. The last section of the chapter looks briefly at the incentive of buyers and sellers to acquire information. Both buyers and sellers prefer having more information so long as they can obtain it covertly. Overt information acquisition is much more complicated and remaining ignorant may be the best policy.

4.2 The Model

An indivisible good, with quality $\omega \in \Omega$, can potentially be traded between a buyer and seller who have nondecreasing values $b(\omega)$ and $s(\omega)$. The prior distribution of quality $F(\omega)$ is commonly known, but the seller observes a private signal x, where $x|\omega \sim F(x|\omega)$. I make two assumptions: first, that gains from trade always exist; and second, that the seller's signal is positively related to the true quality.

Assumption 1 $b(\omega) \geq s(\omega)$ for all ω.

Assumption 2 $E[b(\omega)|x]$, $E[s(\omega)|x]$ are nondecreasing in x, for all b, s nondecreasing, i.e. $F(\omega|x^H)$ first order stochastic dominates $F(\omega|x^L)$ for any $x^H > x^L$.[2]

The first assumption is basically for convenience, but the second is important. Essentially, it means that the signals are stochastically ordered — this makes it possible to proceed without specific functional forms for b, s and also to use a definition of information that "fits" the problem.

[2]Recall that $G \succ_{FOSD} F$ if and only if $G(\omega) \leq F(\omega)$ for all ω. The point of this condition is that signals can be stochastically ordered.

The market structure is minimal: I consider the possibility of trade at a posted market price p. The seller will sell at price p if and only if

$$E[s(\omega)|x] \leq p \tag{IC$_S$}$$

The buyer, taking into account the seller's decision, will buy at p if and only if

$$E\left[b(\omega) \mid E\left[s(\omega)|x\right] \leq p\right] \geq p \tag{IC$_B$}$$

Trade is *feasible* at p if the buyer enters the market. Let P be the set of feasible prices. In terms of a demand schedule, demand is 1 if (IC_B) is satisfied, and 0 if it fails, and P is the set of prices at which there is positive demand. It is well-known (Wilson, 1980) that demand need not be downward sloping.

Denote by $\alpha(p)$ the proportion of sellers in the market at a price p (alternatively, the probability that a seller will enter the market at price p):

$$\alpha(p) = \sup\left\{\alpha : E\left[s(\omega) \mid F(x) = \alpha\right] \leq p\right\}$$

The function $\alpha(p)$ can be thought of as a supply curve. By Assumption 2, it is increasing in p. The *(maximum) extent of trade* given information structure F is $\alpha^* = \sup_{p \in P} \alpha(p)$, which will be the measure of market performance.

Let me make two observations. First, the measure of trade is *not* the price at which trade occurs, but rather the maximum proportion of goods that can be traded. If $b - s$ is constant, this is proportional to the realized welfare gains. Second, upward sloping demand in this model means that market equilibrium need not be not unique (see Wilson (1980) or Mas-Colell, Whinston and Green (1995)). The highest feasible price (at which the maximum extent of trade is realized) is the constrained pareto efficient equilibrium.

4.3 More Private Information May Increase Trade

Consider an example in the spirit of Akerlof. A seller and a buyer can potentially trade a good of uncertain quality; the good is equally likely to be a lemon ($\omega = L$), a melon ($\omega = M$), or a peach ($\omega = H$). The buyer and seller have values:

$$b = \begin{cases} 14 & \text{if } \omega = L \\ 28 & \text{if } \omega = M \\ 42 & \text{if } \omega = H \end{cases} \qquad s = \begin{cases} 0 & \text{if } \omega = L \\ 20 & \text{if } \omega = M \\ 40 & \text{if } \omega = H \end{cases}$$

where the buyer always has higher value than the seller, and everyone likes higher quality.

If buyer and seller are equally clueless, then

$$E[b] = 28 > 20 = E[s]$$

means that full trade can take place at any price between 20 and 28.

Suppose now that the seller is partially informed. She knows a lemon when she sees one, but can't distinguish a melon from a peach. It is now impossible to find a price at which lemons, melons and peaches can all be sold, since

$$E[b] = 28 < 30 = E[s \mid \omega \in \{M, H\}].$$

Only the market for lemons is active, at a price between 0 and 14. As in Akerlof's model, adverse selection reduces the amount of trade.

What if the seller becomes still more perceptive and can identify quality exactly? Peaches cannot be traded at any price, but at a price between 20 and 21, both lemons *and* melons can be exchanged.

$$E[b \mid \omega \in \{L, M\}] = 21 > 20 = E[s \mid \omega = M].$$

The market has expanded in the face of greater information asymmetry! Evidently, the relationship between information asymmetry and market efficiency is non-monotonic. But as the next section will demonstrate, it is non-monotonic is a very intuitive way.

4.4 Partition Information

The story about lemons, melons and peaches is an example of a (noiseless) partition information structure. The seller learns for certain that the type of good she owns is from some subset of possible types, but she learns nothing new about the relative likelihoods within this subset. Better information means a finer partition. Partitioning a group of sellers spreads out their reservation prices in a very specific way: it lowers the reservation price of the low value sellers in the group, and raises the reservation price of the high value sellers in the group. In the example, the first partition involves sellers *on the market margin:* this reduces trade by separating off the sellers with higher quality goods. On the other hand, the second partition separates sellers who are *out of the market*: this increases trade by lowering the reservation price of melon owners and bringing them back into the market.

In fact, it is possible to prove a general result to this end. Let the initial set of signals be $X = \{x_1, ..., x_n\}$, where $x = x_i$ corresponds to $\omega \in \Omega_i$; I assume that the subsets Ω_i are ordered, $\Omega_{i+1} > \Omega_i$ (meaning that any element in Ω_{i+1} is greater than every element in Ω_i), and that they form a partition: $\Omega = \cup_{i=1}^{n} \Omega_i$, and $\Omega_i \cap \Omega_j = \emptyset$. Let $\alpha^*(X)$ be the maximum amount of trade given this information; where $\alpha^* = \Pr[\omega \in \cup_{i=1}^{I^*} \Omega_i]$. Sellers with signals $x_1, ..., x_{I^*}$ are active in the market, those with signals $x_{I^*+1}, ..., x_n$ are not. Suppose $X' = \{x_1, ..., x_k', x_k'', ..., x_n\}$ is identical to X, except that Ω_k is partitioned for some k, i.e. $\Omega_{k'} \cup \Omega_{k''} = \Omega_k$, $\Omega_{k'} \cap \Omega_{k''} = \emptyset$. If $k < I^*$, this is additional information *in the market;* if $k = I^*$, this is information *on the margin*; and if $k > I^*$, the information is *out of the market.*

Proposition 4.1 ***(Partition Information and Trade).*** *More information* in the market *can not affect the maximal amount of trade. More information* on the margin *can only reduce trade. More information* out of the market *can only increase trade.*

Proof. For α^* to be that maximal amount of trade given X, it must be that

$$E[b \mid x \leq x_{I^*}] \geq E[s \mid x = x_{I^*}], \tag{4.1}$$

$$E[b \mid x \leq x_j] < E[s \mid x = x_j], \qquad \forall j > I^*. \tag{4.2}$$

If $k < I^*$, then (4.1) and (4.2) hold unchanged under X', so the maximal amount of trade is unchanged. If $k = I^*$, then (4.2) holds, so trade may not increase. But to sustain the same level of trade requires $E[b|x \leq x_k''] \geq E[s|x = x_k''] \geq E[s|x = x_k]$ where the first inequality is clearly a stronger condition than (4.1). If it fails, then trade must decrease. Finally, let $k > I^*$. Now, (4.1) holds identically; trade will increase if and only if $E[b|x \leq x_k'] \geq E[s|x = x_k']$ which is not ruled out by (4.2). *Q.E.D.*

4.5 General Analysis

Tackling the general problem demands an appropriate concept of "more information." It turns out that the relevant concept of information in this setting is the supermodular monotone information order introduced in the previous chapter. To review briefly: Consider a decision maker who takes action $a \in A$ based on a signal $x|\omega$ in order to maximize $E[\pi(\omega, a)|x]$, where π is a supermodular payoff function. Recall that a function $\pi(\omega, a)$ is supermodular if $\pi(\omega', a') + \pi(\omega, a) \geq \pi(\omega', a) + \pi(\omega, a')$ for all $\omega' \geq \omega, a' \geq a$, i.e. if $\pi_{\omega a} \geq 0$. Let $G(\omega, y)$ and $F(\omega, x)$ be signal distributions, with common prior $f(\omega)$ on ω.

Theorem 4.1 *Suppose (A2) holds for F, G. All supermodular decision-makers prefer signal distribution G to F if for all $q \in [0, 1]$,*

$$F(\omega|F(x) \leq q) \succ_{FOSD} G(\omega|G(y) \leq q) \tag{MIO-spm}$$

If F^θ is a family of signal distributions continuously indexed by θ, satisfying (A2) for all θ, this condition is necessary for all spm DMs to prefer $F^{\theta+d\theta}$ to F^θ.

In the above theorem, $\succ_{FOSD}$ signifies first order stochastic dominance. The condition (MIO-spm) compares the average posterior beliefs under signals in the bottom q^{th} percentile. If posterior beliefs are on average more pessimistic after obtaining a low signal under G than after obtaining a low signal under F (which conversely means they are more optimistic after obtaining a high signal), then G is more information than F. Under (A2), this condition is weaker, i.e. it orders more signal distributions, than Blackwell's (1953) sufficiency criterion, which is a standard statistical notion of information. It is also weaker that Lehmann's (1988) efficiency ranking of information.[3]

The Athey-Levin information property is closely related to the degree of adverse selection in a market. One can interpret (MIO-spm) as saying the expected value of the object to the buyer conditional on a proportion q of sellers tendering (i.e. those with the lowest signals) is decreasing in the amount of private information. To see this, recall that the expectation of *any* increasing function of a random variable decreases when the distribution shifts down by FOSD. In other words, if one fixes the number of sellers who must come to market and increases the information asymmetry in the sense of Athey–Levin, the quality of goods available shifts down in the sense of first order stochastic dominance. The buyer's curse intensifies!

Of course, this is only half the story. It says that for a fixed supply curve, an increase in information shifts the demand curve down. To compare two market equilbria, however, it is necessary to account for changes in the supply curve. Suppose some amount of trade $\alpha \in (0, 1]$ is exactly feasible at a price $p_{a,L}$ when private information is poor, i.e.

$$E\left[b(\omega) \mid F(x) \leq \alpha\right] = p_{a,L} = E\left[s(\omega) \mid F(x) = \alpha\right].$$

If the level of private information increases, then trade α will not feasible at any price $p > p_{a,L}$, because

$$E[b(\omega) \mid G(y) \leq \alpha] \leq E[b(\omega) \mid G(y) \leq \alpha] = p_{a,L}.$$

Nevertheless, it may be possible to *lower* the price and still induce the same number of sellers to enter the market, i.e. it may well be the case that

$$E[s(\omega) \mid F(x) = \alpha] > E[s(\omega) \mid G(y) = \alpha].$$

In this case, there is a horse race between two competing effects—the intensified winner's curse which lowers demand and the selection effect which might raise supply. The outcome turns on the condition identified above: whether the new information has greater impact *on the margin*

[3] G is Blackwell sufficient for F if and only if *all* decision makers prefer G to F, or equivalently if F a noisy version of G. G is higher than F in Lehmann's efficiency order if $G^{-1}(F(x|\omega)|\omega)$ is increasing in ω, or if all decision makers with single crossing payoff functions prefer G to F. Lehmann's order can be obtained by requiring (MIO-spm) to hold for all prior distributions $F(\omega)$.

or *out of the market*. Before fleshing this out, I first note that it is possible to make a weaker statement about private information and the possibility for realizing *all* the gains from trade. I add an additional assumption for this result.

Assumption 3 $E[s(\omega)|H(x) = \alpha]$ is continuous in α for $H \in \{F, G\}$.

Proposition 4.2 ***(Full Trade is Nonrestorable)*** *Suppose $G \succ F$ in MIO-spm. If full trade is not feasible with F, it is not feasible with F.*

Proof. If full trade ($\alpha = 1$) is not feasible at F, then there exists some $\overline{\gamma} > 0$ such that for all $\gamma < \overline{\gamma}$, $E[b(\omega)|F(x) \leq 1-\gamma] < E[s(\omega)|F(x) = 1-\gamma]$. Since $b(\omega)$ and $s(\omega)$ are both increasing in ω, Theorem 1 says that for any $\epsilon \in [0,1]$, $E[b(\omega)|G(y) \leq 1-\epsilon] \leq E[b(\omega)|F(x) \leq 1-\epsilon]$, and also that $E[s(\omega)|G(y) > 1-\epsilon] \geq E[s(\omega)|F(x) > 1-\epsilon]$. (The second inequality uses the fact that for any $\epsilon \in [0,1]$ and any $H(z) \in \{F(x), G(y)\}$, $(1-\epsilon)E[s(\omega)|H(z) \leq 1-\epsilon]+\epsilon E[s(\omega)|H(z) > 1-\epsilon] = E[s(\omega)]$, which is constant). By continuity (A3) and monotonicity (A2) of the conditional expectation, there is some $\overline{\epsilon} > 0$ such that for all $\epsilon < \overline{\epsilon}$, $E[s(\omega)|G(y) = 1-\epsilon] \geq E[s(\omega)|F(x) = 1-\epsilon]$. Now, choose $0 < \delta < \min(\overline{\gamma}, \overline{\epsilon})$. It follows that $E[b(\omega)|G(y) \leq 1-\delta] \leq E[b(\omega)|F(x) \leq 1-\delta] < E[s(\omega)|F(x) = 1-\delta] \leq E[s(\omega)|G(y) = 1-\delta]$; full trade is infeasible at G. *Q.E.D.*

Proposition 4.2 describes a very weak notion of decreasing trade. The more interesting question is what happens to the maximal amount of trade. Consider looking for conditions under which trade will decrease in response to greater information asymmetry. If a level of trade α is infeasible with poor seller information F, i.e.

$$E[b \mid F(x) \leq \alpha] \leq E[s \mid F(x) = \alpha], \tag{4.3}$$

it should be the case that α is infeasible with better seller information G ($G \succ F$),

$$E[b \mid G(y) \leq \alpha] \leq E[s \mid G(y) = \alpha]. \tag{4.4}$$

If the set of feasible prices with F is convex (equilibrium is unique), then (4.3) implies (4.4) is a necessary as well as sufficient condition for information to decrease trade.

Proposition 4.3 ***(Information Decreases Trade).*** *Suppose $G \succ F$ in MIO-spm. Trade is decreasing in information if*

(i) $b(\omega) - s(\omega)$ *nondecreasing in* ω.

(ii) $G(\omega|G(y) = \alpha) - F(\omega|F(x) = \alpha) \succ_{FOSD} G(\omega|G(y) \leq \alpha) - F(\omega|F(x) \leq \alpha)$, $\forall \alpha \in [0,1]$.

Proof. I show that if a level of trade α is infeasible at F, it is infeasible at G. If α is infeasible at F^L, then $E[b|F(x) \leq \alpha] < E[s|F(x) = \alpha]$, or equivalently

$$E[s \mid F(x) = \alpha] - E[s \mid F(x) \leq \alpha] > E[b - s \mid F(x) \leq \alpha]. \tag{4.5}$$

Since $(b - s)(\omega)$ is nondecreasing in ω, then $E[b - s|F(x) \leq \alpha] \geq E[b - s|G(y) \leq \alpha]$, i.e. the right hand side of (4.5) is decreasing in the amount of private information. Similarly, since $s(\omega)$ is increasing, the left hand side of (4.5) is increasing in the amount of private information. This implies

$$E[s \mid G(y) = \alpha] - E[s \mid G(y) \leq \alpha] > E[b - s \mid G(y) \leq \alpha], \tag{4.6}$$

or re-arranging, that $E[b|G(y) \leq \alpha] < E[s|G(y) = \alpha]$; trade α is infeasible at G. *Q.E.D.*

Condition (i) requires that the gains from trade be positively related to quality. If this holds, then for a fixed amount of trade α, the gains realized by the market decrease in information because sellers self-select more accurately. The second condition is roughly analagous to the partition condition. It says that the impact of information *on the margin* (i.e. for $F(x) = \alpha$) is greater (stochastically) than the impact *in the market* ($F(x) \leq \alpha$). And note that Bayes' rule implies that the impact *in the market* is directly proportional to the impact *out of the market*:

$$F(\omega) = \alpha F(\omega \mid F(x) \leq \alpha) + (1 - \alpha) F(\omega \mid F(x) > \alpha).$$

where the left hand side is independent of the signal informativeness.

The conditions under which trade is increasing in information are similar. A sufficient condition is that whenever trade in feasible at F it also be feasible at G — in other words, the reverse of (4.3) implies the reverse of (4.4).

Proposition 4.4 ***(Information Increases Trade).*** *Trade is increasing in information if*

(i') $b(\omega) - s(\omega)$ *nonincreasing in* ω.

(ii') $G(\omega|G(y) = \alpha) - F(\omega|F(x) = \alpha) \prec_{FOSD} G(\omega|G(y) \leq \alpha) - F(\omega|F(x) \leq \alpha)$, $\forall \alpha \in [0, 1]$

The proof and the interpretation are directly analagous. Note that if the gains from trade are commonly known ($b - s$ is constant), then the effect of information turns exactly on where the information impacts the market.

When the price region is convex, these conditions in Propositions 4.3 or 4.4 are "tight" in that if either (i) or (ii) is relaxed (or (i') or (ii')), it is possible to construct a counterexample with some b, s, G, F that satisfy the other assumptions. So it is interesting to know when the price region is convex (apart from the direct question of equilibrium uniqueness). For a given information structure F, this means that trade being infeasible at α implies that trade is infeasible for any $\alpha' > \alpha$ (i.e. if (4.3) holds for some α, it holds for all $\alpha' > \alpha$).

Proposition 4.5 ***(Conditions for Convex Price Region).*** *Let F be given information structure. The set of feasible prices P is convex if*

(I) $b(\omega) - s(\omega)$ *nonincreasing in* ω.

(II) $F(\omega|F(x) = \alpha) - F(\omega|F(x) \leq \alpha)$ *is decreasing in* α.

Proof. Very similar to Proposition 4.3. *Q.E.D.*

Conditions (I) and (i) are both satisfied only for the special case where $b-s$ is constant, which implies that Proposition 3 should be regarded only as a sufficient condition for information to decrease trade.

Equilibrium uniqueness in adverse selection markets is often remarked upon, so let me digress briefly on the conditions in Proposition 4.5. A more restrictive condition that implies (II) is that $E[s|F(x) = \alpha]$ is convex. Given our assumption that $s(\omega)$ is nondecreasing, this means requiring the distribution function $F(\omega|F(x) = \alpha)$ to be convex, a strong assumption. Equilibrium uniqueness is more plausible if one is willing to assume that $s(\omega)$ is convex, in which case a sufficient condition for (II) is that the joint density of (ω, x) be *totallly positive of order 3* (Karlin, 1968) — it turns out that many common distributions are TP_3.

To digress somewhat further, just as a stronger assumption on values allows a weaker distributional assumption for equilibrium uniqueness, it is possible to make alternate assumptions on values when comparing information. If valuations are linear — $s(\omega) = \omega$, $b(\omega) = a + b\omega$, where $b \geq 0$ — the FOSD conditions above give way to conditions on first moments. For example, condition (ii) in Proposition 4.3 can be weakened to

$$E[\omega|G(y) = \alpha] - E[\omega|F(x) = \alpha] \geq E[\omega|G(y) \leq \alpha] - E[\omega|F(x) \leq \alpha], \qquad \forall \alpha \in [0, 1].$$

Along the lines of the previous chapter, a different concept of informativeness would also be appropriate.

4.6 Further Notes

4.6.1 Information Acquisition

If either the buyer or the seller can acquire information *covertly* without drawing the attention of the other party or affecting the market price, they always prefer to obtain a higher signal as ranked by (MIO-spm). To see this, suppose the market price is p and that the buyer can covertly obtain a signal z that is informative about ω, but is independent of the seller's signal x conditional on ω.[4] The buyer's belief prior before observing z is that the good's quality has

[4]That is, if the joint signal/state distribution is $F(x, y, \omega)$, then $F(x|y, \omega) = F(x|\omega)$.

distribution $F(\omega|E[s(\omega)|x) \leq p)$. Write this as $H(\omega)$ and let $H(\omega|z)$ denote the buyer's beliefs about the good's quality after observing z. Assume that observing a high signal z is good news about quality, $z' > z$ implies that $H(\omega|z') \succ_{FOSD} H(\omega|z)$.

After observing z, the buyer will face a single-person decision problem: whether to trade ($t = 1$) or not ($t = 0$). He solves:

$$\max_t \int_\Omega t\,(b(\omega) - p)\,dH(\omega|z) = \max_t \int_\Omega B(\omega, t; p) dH(\omega|z).$$

Because $b(\omega)$ is increasing in ω, the buyer's objective function $B(\omega, t; p)$ is supermodular in (ω, t). Since the signals are ordered by FOSD, Theorem 4.1 applies. The buyer always prefers signals that are more informative, as ranked by (MIO-spm).[5] An analagous argument applies for the seller.

When information acquisition is *overt*, so that the other party can observe information acquisition, the situation is much more complicated. For instance, it is easy to construct examples where the seller prefers to remain completely ignorant about quality — since complete ignorance leads to full trade, just imagine a situation where acquiring private information causes the whole market to unravel. One general principle is that the seller will *never* acquire additional information if she knows for sure that it will not change her decision about whether to enter the market. But of course, the buyer will always adjust his beliefs to compensate for altered seller decisions. So in general, information will have non-zero value precisely when acquiring it will change the market equilibrium. This means, for instance, that under a partition infomation structure, no seller deeply *in the market* cares at all about acquiring information.

4.6.2 Other Mechanisms for Trade

The concepts discussed in this paper are relevant for other trading institutions, although general results appear difficult to obtain. For example, suppose the seller has complete information, i.e. she knows ω, and the buyer can obtain a signal z which is affiliated with ω, and then make a take-it-or-leave-it offer. If $v(\omega), s(\omega)$ are nondecreasing, then the buyer prefers signals that are more informative in the sense of Lehmann (1988), which is a stronger version of (MIO-spm) (though still a weakening of Blackwell).[6] Signals are more informative in Lehmann's order if (A2) and (MIO-spm) hold for all possible prior distributions $F(\omega)$. Lehmann's concept of information has also been applied to auction markets by Persico (1997).

[5] Of course, the buyer also prefers signals that are just less noisy in the sense of Blackwell. But the (MIO-spm) ranks more signal distributions and is tight in the sense that if H and H' are close, all buyers prefer H' to H if and only if $H' \succ_{MIO-spm} H$.

[6] To see this, observe that the buyer's problem is $\max_p \int_\Omega \{v(\omega) 1_{\{s(\omega)\leq p\}} - p\} dF(\omega|z)$. One can verify that $u(\omega, p) = v(\omega) 1_{\{s(\omega)\leq p\}} - p$ is bivariate single crossing (i.e. $u_p(\omega)$ crosses zero at most once from below). The claim then follows from Lehmann (1988).

4.7 References

Akerlof, G. (1970): "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84, 488–500.

Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Glosten, Lawrence and Paul Milgrom (1985): "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71-100.

Karlin, Samuel (1968): *Total Positivity.* Stanford: Stanford University Press.

Lehmann, E. (1988): "Comparing Location Experiments," *Annals of Statistics*, 16, 521–533.

Mas-Colell, Andreu, Michael Whinston and Jerry Green (1995): *Microeconomic Theory.*

Persico, N. (1997): "Information Acquisition in Auctions," Mimeo, UCLA.

Wilson, C. (1980): "The Nature of Equilibrium in Markets with Adverse Selection," *Bell Journal of Economics*, 11, 108–130.

THESIS PROCESSING SLIP

FIXED FIELD: ill. ______ name ______

index ______ biblio ______

► COPIES: (Archives) Aero (Dewey) Eng Hum

Lindgren Music Rotch Science

TITLE VARIES: ►☐ ______

NAME VARIES: ►☐ ______

IMPRINT: (COPYRIGHT) ______

► COLLATION: 133p

► ADD: DEGREE: ______ ► DEPT.: ______

SUPERVISORS: ______

NOTES:

cat'r: date:

page:

► J192

► DEPT: Econ

► YEAR: 1999 ► DEGREE: Ph.D.

► NAME: LEVIN, Jonathan David