The Enforcement Policy of a Self-Regulatory Organization

Peter M. DeMarzo[*], Michael J. Fishman[**], and Kathleen M. Hagerty**

First Version: October 1998
This Version: April 2, 1999

<u>Disclaimer: This is a preliminary and incomplete draft.</u>

Abstract

The federal government delegates various aspects of financial market regulation to self-regulatory organizations (SROs) such as the New York Stock Exchange and the National Association of Securities Dealers. We model one regulatory task of an SRO, the enforcement of rules designed to prevent the SRO's members from cheating customers. Specifically, we focus on the determination of an SRO's optimal policy for investigating agents who may have defrauded customers and the penalties associated with fraud.

We model contracting/enforcement as a two-tier problem. First, an SRO chooses its enforcement policy, consisting of a specification of the likelihood that an agent is investigated for fraud and a penalty schedule. We assume that the SRO's objective is to maximize the welfare of its members, the agents. Taking the SRO's enforcement policy as given, agents compete with one another to handle customer transactions. They compete by offering contracts promising (outcome-contingent) payoffs that maximize customers' expected utility. When choosing an enforcement policy, the SRO anticipates the competition among its members. Indeed, we show that the SRO's optimal enforcement policy is designed to mute this competition. In doing so, an SRO chooses a more lax enforcement policy than would be preferred by customers. Investigations for cheating are less frequent and penalties are lower than what a customer would choose. Moreover, a decrease in investigation cost might lead an SRO to actually investigate less. Enforcement will become more vigorous, however, as a customer's alternatives to dealing with an agent of the SRO improve. A general conclusion of the analysis is that control of the enforcement policy governing contracts confers substantial market power to a group of otherwise competitive agents. In fact, we show that if agents are risk neutral, control of the enforcement policy is equivalent to agents behaving as monopolists. We also investigate the effect of government oversight on the self-regulatory process. We show that in equilibrium the threat of governmental enforcement will lead to more rigorous enforcement by the SRO, to the benefit of customers. Moreover, this benefit is achieved even without actual governmental enforcement taking place.

[*] Haas School of Business, University of California, Berkeley
[**] Kellogg Graduate School of Management, Northwestern University

1. Introduction

In the U.S., the Securities and Exchange Commission (SEC) has primary responsibility for regulating securities markets. The SEC, however, delegates significant regulatory authority to self-regulatory organizations (SROs), securities industry organizations that are owned and operated by their members. Examples of SROs include the National Association of Security Dealers, the New York Stock Exchange, the Chicago Board Options Exchange, and regional stock and option exchanges. Among the tasks of SROs is to design rules governing their members' practices. In addition, SROs are responsible for enforcing their own rules as well as federal securities laws. This enforcement activity includes conducting disciplinary proceedings and imposing sanctions on members for violations of their own rules or federal laws.[1]

Given that SROs are run for the benefit of their members, how vigorously will an SRO enforce its rules? That is, how likely is an SRO to investigate one of its members for a rule violation? What penalties will an SRO set for members who violate the rules? Will the enforcement policy chosen by an SRO coincide with the policy that is preferred by their customers? These are among the questions addressed in this paper.

We use the costly-state-verification model of Townsend (1979) and Mookherjee and Png (1989). In this model, agents facilitate transactions for customers and an agent might lie about the payoff from the transaction; most importantly, he might report that the payoff is lower than it really is. Such fraud can be deterred in two ways: (i) by the threat of an investigation and penalty; and (ii) by an incentive contract that pays the agent more when he reports high payoffs.

In this setting, an enforcement policy, consisting of a specification of the likelihood of an investigation and a penalty schedule for fraud, is chosen by the SRO. We take the objective of the SRO to be one of maximizing the expected utility of its members, the agents. Taking the SRO's enforcement policy as given, agents compete with one another to handle customer transactions. They compete by offering contracts promising (outcome-contingent) payoffs that maximize customers' expected utility.[2]

When choosing an enforcement policy, the SRO anticipates the competition among its members. Indeed, we show that the SRO's optimal enforcement policy is designed to mute this competition. In doing so, an SRO chooses a more lax enforcement policy than would be preferred by customers. Investigations for fraud will be less frequent and penalties will be lower than what a customer would choose (though the revelation principle applies and so in equilibrium there is no fraud). The intuition is that an agent can be induced to be truthful by both the "carrot" of a bonus when he reports a high payoff and the "stick" of the threat of investigation and penalty. Agents prefer to be motivated by the carrot rather than by the stick. By choosing a relatively lax enforcement

[1] See National Association of Securities Dealers (1996) and Phillips (1997) for general discussions of SRO enforcement and Frankhauser, Gardner, McNally, and Leatherwood (1997) for details of SRO enforcement procedures.

[2] We do not explicitly model competition among SROs. We do examine the effects of changes in customers' reservation utility (their best alternative to contracting with a member of the SRO), which can be interpreted as a change in the competitive environment.

policy, an SRO induces customers to offer more of a carrot, i.e., higher bonuses for reporting high payoffs.

We also show that a decrease in the cost of an investigation might lead an SRO to actually investigate less. Enforcement becomes more vigorous, however, as a customer's alternatives to dealing with an agent of the SRO improve.

A general conclusion of the analysis is that control of the enforcement policy governing contracts confers substantial market power to a group of otherwise competitive agents. In fact, we show that if agents are risk neutral, control of the enforcement policy is equivalent to agents behaving as monopolists.

In typical analyses of contracting, one party to a contract chooses all relevant terms: outcome-contingent payoffs and an enforcement policy consisting of a specification of the likelihood that an agent is investigated for fraud and a penalty schedule. In practice, however, parties to a contract do not choose the enforcement policy that governs contracts. Instead, parties negotiate contracts taking the enforcement environment as given. The novel aspect of our analysis is the incorporation of this feature of contracting. We model contracting/enforcement as a two-tiered problem in which an institution determines the enforcement policy taking into account the effect of that policy on the contracts that are subsequently created. Our analysis accounts for the separate determination of (i) the enforcement policy governing contracts; and (ii) individual contracts between parties. We then highlight how the institution that chooses the enforcement policy can affect the division of rents between contracting parties.

In our analysis, the institution determining the enforcement policy governing contracts is an SRO that represents one party to the transaction. But of course in many situations, enforcement policies are chosen by the government. While the government's objectives differ from those of an SRO (perhaps the government is best modeled as seeking to maximize a weighted average of the expected utilities of the parties to a contract), the general problem is the same. In designing the environment governing contract enforcement, the incentives of the parties who will negotiate contracts must be anticipated.

The literature on self regulation typically deals with two issues. One issue is whether self-regulatory organizations face enough competition to induce them to regulate themselves in a socially efficient manner. See, for example, Pirrong (1995) and Mahoney (1997) who focus on exchanges. In our analysis, we assume that while the agents of an SRO compete with one another, the SRO itself has market power. The other issue in the literature is how an SRO that has market power exercises that power. For example, Leland (1979), Shaked and Sutton (1981) and Gehrig and Jost (1995) examine a self-regulating profession's incentives to limit membership through quality standards. Saloner (1984) examines an exchange's incentive to directly restrict the size of its membership. These papers find that SRO's set standards that are too high; that is, a social planner could increase efficiency by relaxing standards. In our analysis membership size plays no role. An SRO exercises market power through the way it enforces its rules. In contrast with the previous literature, we show that SRO's chooses weaker enforcement than that which would be chosen by a social planner.

The model is described in Section 2. Section 3 presents two benchmarks for comparison. One involves fully competitive agents and has the customer choosing both the contract and the enforcement policy. The other involves a monopolistic agent who chooses both the contract and the enforcement policy. Section 4 analyzes an SRO's optimal enforcement policy and the resulting contract between the agent and the customer. In Section 5 [**incomplete**] extends the model to account for heterogeneous customers. With heterogeneous customers, the SRO's enforcement policy also determines the extent of customer participation, i.e., the fraction of customers that will do business with the SRO's agents. Here we also consider the effect of competition between SRO's. Section 6 we present a two-state illustration of the analysis. Section 7 [**incomplete**] examines the possibility of government oversight of an SRO. We consider the case in which the government's objective is to maximize customers' expected utility. We show that if the cost of investigating an agent is higher for the government, then government oversight reduces but does not eliminate the market power of agents in an SRO. Section 8 [**incomplete**] explores the role of reputation in a dynamic setting. This expands the scope of penalties to include the possible loss of future rents. Section 9 [**not in this draft**] contains concluding remarks. Proofs of the Propositions appear in the Appendix.

2. The Model

There are a variety of ways to model a customer-agent relation in which the agent might cheat the customer. We use the costly-state-verification model of Townsend (1979) and Mookherjee and Png (1989) in which only the agent observes the outcome of the project, and may misreport this outcome to the customer. For a cost, it is possible to verify the agent's report, and detect the occurrence of fraud.

There are three players: the customer, the agent, and the self-regulatory organization (SRO). The customer hires the agent. The cash flow from using the agent's services is a random variable W that has support $\Omega \subset \Re_+$. Assume that Ω has a minimum element, denoted $\underline{w}$. The agent privately observes this cash flow and reports the cash flow to the customer. This is the source of the moral hazard problem. The agent might lie about the cash flow and keep some of it for himself. The agent can be investigated at a cost of $c \geq 0$. An investigation reveals the realization w of W.

We assume that the customer is risk neutral though the agent may be risk averse. This eliminates consideration of a contract that provides the customer with insurance. So in the absence of the moral hazard problem, the optimal contract simply involves the customer hiring the agent for a fixed fee. The customer's best alternative to transacting with the agent is represented by an expected payoff of α.

Let $u(y)$ denote the agent's utility as a function of final wealth y, where u is increasing, concave, and $u(0) = 0$. We assume that the agent has zero initial wealth and faces a limited liability constraint; his net income on the transaction with this customer cannot be less than zero. These assumptions have two implications. First, it bounds the size of the penalty that can be imposed on the agent. Second, though the agent behaves competitively, he cannot compete away all rents by paying a customer to do business with him.

A contract between the agent and the customer is represented by the function z, where $z(r)$ specifies the payment to be made by the agent to the customer if the agent reports that $W = r$. The SRO's enforcement policy specifies an investigation strategy and a penalty schedule. The investigation strategy is represented by the function p, where $p(r)$ specifies the probability that the agent is investigated given a report of r. If an investigation takes place, monetary penalties are assessed according to a penalty schedule x, where $x(w,r)$ specifies the penalty to be paid by the agent given the report r and an actual payoff of w.[3]

In the standard contracting literature, the customer and the agent directly negotiate the contract and enforcement policy (z,p,x). We instead consider an environment in which the enforcement policy (p,x) is first set by the SRO, and then the customer and the agent negotiate the contract z.

We assume the SRO acts to maximize the agent's expected utility. The SRO conducts investigations according to the strategy p, incurring cost c whenever an investigation takes place. The SRO collects penalties according to the schedule x from the agent. We restrict the SRO to set penalties $x \geq 0$; that is, we do not allow the SRO to pay the agent. Finally, the SRO charges a transaction fee t that is paid by the customer and used to finance expected enforcement costs net of penalties.

Given an enforcement policy, the customer and agent negotiate the contract, z. Assume the agent behaves competitively (effectively, the customer makes a take-it-or-leave-it offer).[4] The customer and the agent each act to maximize their own expected utility. Anticipating the behavior of the customer and the agent, the SRO chooses the enforcement policy (p,x) and transaction fee t to maximize the agent's expected utility.

To summarize, the timing of the problem is as follows:

1. The SRO chooses an enforcement policy (p,x) and transaction fee t.
2. Taking (p,x,t) as given, the customer offers a contract z to the agent, which the agent can accept or reject.
3. If the agent rejects the contract, the agent receives 0 and the customer receives α. If the agent accepts the contract, the customer pays t and the problem continues.
4. A cash flow w is realized.
5. The agent chooses a cash flow, r, to report and gives the customer the corresponding payoff, $z(r)$.
6. Given the reported cash flow r, the SRO investigates the agent with probability $p(r)$. If an investigation occurs, the SRO pays c and receives the penalty $x(w,r)$ from the agent (subject to the agent's resource constraint).

[3] Besides monetary sanctions SROs can also suspend or expel members and they can bar individuals from the industry. We consider only monetary sanctions.

[4] We will see later that this is without loss of generality since we can trace out the Pareto frontier by varying the customer's reservation payoff α.

Note that the assumption that the SRO collects the penalty, instead of the customer, does not matter. If instead the customer collected the penalty, the SRO would simply set the transaction fee to finance enforcement costs gross of penalties. Otherwise, the analysis would be unchanged.

This representation of the moral hazard problem can be interpreted as capturing a number of potentially fraudulent activities. For instance, consider the case in which the agent is a broker who executes trades for a customer. The customer does not directly observe the prices at which orders are filled and the broker may cheat by reporting that buy (sell) orders were filled at higher (lower) prices. Alternatively, consider the case in which the agent manages a customer's trading account. The agent may cheat by churning the account and collecting excessive brokerage fees. Of course, another possibility is that the agent simply steals the money.

We now analyze the subgame perfect equilibrium of this game.

2.1 The Agent's Problem

Given (z,p,x), for each outcome of W the agent must choose a report. We denote this choice by a reporting strategy r, where $r(w)$ is the agent's report if outcome w occurs. In this case, the agent must pay the customer $z(r(w))$. Since the agent has only w available, the reporting strategy is feasible for the agent if and only if it satisfies

$$\text{(AF)} \quad z(r(w)) \le w \ \text{ for all } w.$$

Given outcome w and report r, an investigation occurs with probability $p(r)$ and penalty $x(w,r)$. Thus, the agent's expected utility is given by

$$v(w,r|z,p,x) \equiv p(r)\, u(\max[w - z(r) - x(w,r), 0]) + (1-p(r))\, u(w - z(r)).$$

This reflects the fact that in the event of an investigation, the maximum penalty the agent can pay is the residual profit, $w - z(r)$.

Since the agent chooses a report to maximize this expected utility, in equilibrium the reporting strategy must satisfy the incentive constraint

$$\text{(AIC)} \quad v(w,r(w)|z,p,x) \ \ge\ v(w,s|z,p,x) \text{ for all } w \text{ and } s \text{ such that } \mathrm{z}(s) \le w.$$

2.2 The Customer's Problem

Having characterized the agent's problem, we now consider the problem faced by the customer when choosing a contract z to offer the agent. Since the agent has a reservation payoff of 0, and since any contract offers a non-negative payoff to the agent, the agent will accept any offer. Thus, the customer's problem is to choose the contract z that maximizes the customer's expected payoff taking into account the optimal reporting strategy of the agent. Taking the enforcement policy (p,x) as given, the customer's problem can be represented as the following mechanism design problem:

$$\text{CP}(p,x): \qquad \max_{z,r} \qquad E[\, z(r(W))\,]$$

$$\text{subject to}$$

$$\text{(AF)} \quad z(r(w)) \ \le\ w \ \text{ for all } w,$$

$$\text{(AIC)} \quad v(w,r(w)|z,p,x) \geq v(w,s|z,p,x) \text{ for all } w,s \text{ such that } z(s) \leq w.$$

The solution to CP(p,x) yields the best contract for the customer. Thus, any equilibrium must satisfy the customer's incentive constraint corresponding to this subgame:

$$\text{(CIC)} \quad (z,r) \text{ solves CP}(p,x).$$

In addition to collecting the payments $z(r(W))$, the customer must also pay the transaction fee t to the SRO. Since the customer has an outside opportunity of α, the customer chooses to contract with the agent only if the customer's individual rationality constraint is satisfied:

$$\text{(CIR)} \quad E[\, z(r(W))\,] - t \geq \alpha.$$

2.3 The SRO's Problem

The SRO chooses (p,x,t) to maximize the agent's expected utility. The transaction fee t must be chosen sufficient to cover the expected investigation cost net of penalties incurred by the SRO. This yields the SRO's budget constraint:

$$\text{(RB)} \quad t \geq E[\, p(r(W))\, (c - \min[x(W, r(W)), W - z(r(W))])\,],$$

where $\min[x(W, r(W)), W - z(r(W))]$ is the actual penalty recovered given the resource constraint of the agent.[5]

Given this constraint together with the characterization of the customer's and agent's optimal strategies, the SRO's problem can be written as the following mechanism design problem:

$$
\begin{aligned}
\text{SRP:} \quad & \max_{z,r,p,x,t} \quad E[\, v(W,r(W)|z,p,x)\,] \\
& \text{subject to} \\
& \text{(CIC)} \quad (z,r) \text{ solves CP}(p,x), \\
& \text{(CIR)} \quad E[\, z(r(W))\,] - t \geq \alpha, \\
& \text{(RB)} \quad t \geq E[\, p(r(W))\, (c - \min[x(W,r(W)), W - z(r(W))])\,].
\end{aligned}
$$

We have assumed that the SRO's objective is to maximize the agent's expected utility. Note, though, that the most of the analysis would be the same if the enforcement policy and transaction fee were chosen by a regulator (say the government) whose objective put weight on both the agent's and customer's expected utility. In this case, the regulator chooses some point on the Pareto frontier. Hence, to characterize the solutions for this more general case, it suffices to solve SRP for alternative values of α.

[5] Note that the regulator only balances the expected budget, rather than state-by-state. This corresponds to the presumption that the regulator overseas many such transactions and can rely on the Law of Large Numbers.

3. Contracting Environments

Before analyzing self-regulation, it is useful to put the problem in perspective by comparing our approach to models of contracting in the existing literature. Previous analyses assume that one party, either the customer or the agent, chooses both the contract and the enforcement policy. With perfect competition among agents, and the customer choosing everything, the problem can be written as:

$$\max_{z,r,p,x,t} \quad E[\, z(r(W))\,] - t \tag{1}$$

subject to (AF), (AIC), (CIR) and (RB).

With a monopolistic agent choosing everything the problem can be written as:

$$\max_{z,r,p,x,t} \quad E[\, v(W,r(W)|z,p,x)\,] \tag{2}$$

subject to (AF), (AIC), (CIR) and (RB).

In a sense, our approach is intermediate to these two, with the agent choosing (p,x,t) and the customer choosing (z,r). Hence we might expect both the agent's and customer's expected utility to be intermediate to that found in the typical approaches.

First compare SRP to the monopolistic agent problem. Clearly since (AF) and (AIC) are constraints to (CIC), the agent does no better under SRP. We will show, however, that if the agent is risk neutral, his expected utility is the same under SRP. If the agent is risk averse, his expected utility is lower under SRP. We also show that despite the presence of the customer incentive compatibility constraint, (CIC), in SRP, the customer's expected utility may actually be lower under SRP as compared to the monopolistic agent problem.

Now compare SRP to the perfectly competitive agent problem in which the customer chooses everything. It is clear that the customer fares worse under SRP. In fact, we will show that under SRP, the customer may receive no more than his reservation utility, leaving all of the rents for the agent.

After deriving the solution to SRP, we will provide a more detailed comparison with these two benchmark cases.

4. Optimal Contracts and Enforcement

In this section, we characterize the solution to SRP. We show that the ability of an SRO to choose the enforcement policy conveys substantial market power to the otherwise competitive agents. Even though the customer makes a take-it-or-leave-it contract offer to the agent, if the customer's reservation utility is not too low (specifically, $\alpha \geq \underline{w}$), the customer receives his reservation utility and the agent receives all of the rents.

To characterize the solution we proceed as follows. First we show that contract always entails the customer receiving at least $\underline{w}$. Next we show that the solution to SRP can be characterized by solving a simpler problem, the one that would be faced by a monopolistic agent who is constrained to give the customer at least $\underline{w}$.

Proposition 1 shows that we can restrict attention to contracts with $z(r) \geq \underline{w}$. Since the agent will accept any contract in which he receives at least 0, the customer will demand to be paid at least the minimum possible cash flow.

Proposition 1. Suppose (z,r) solves CP(p, x). Then $z(r) \geq \underline{w}$.

This and all other proofs are in the Appendix. The intuition is straightforward: given any contract z, the customer will earn a higher payoff offering $z' = \max(z,\underline{w})$.

This result yields an additional constraint that must be satisfied by any solution to SRP. This motivates consideration of the following modification of SRP, in which we drop the customer's incentive constraint:

SRP′: $\quad \max_{z,r,p,x,t} \quad E[\, v(W,r(W)|z,p,x)\,]$

subject to (AF), (AIC), (CIR), (RB) and

(ZW) $\quad z(r) \geq \underline{w}$.

Comparing SRP′ with SRP, we have replaced the customer's incentive constraint (CIC) with (AF), (AIC) and the new constraint (ZW) which places the lower bound $\underline{w}$ on z. Note that since (CIC) implies (AF) and (AIC) by definition, and (ZW) by **Proposition 1**, SRP′ is less constrained than SRP.

Note also that this problem is similar to the optimal contracting problem of a monopolist agent considered in Section 3. It differs by the addition of the constraint (ZW). Hence SRP′ is more constrained than the pure monopolist.

Thus, SRP′ lies between the monopolist's problem and our original problem SRP. We now characterize the solution to SRP′, with the ultimate goal of showing it to be, in fact, equivalent to SRP. That is, the full import of the letting the customer choose z (the (CIC) constraint) is captured by the much simpler (ZW) constraint.

With the removal of the customer's incentive constraint, the problem SRP′ looks like a standard optimal contracting problem. Thus, we can apply the standard technique of the mechanism design literature, the revelation principle, to restrict ourselves to contracts in which the agent reports the outcome truthfully:

Proposition 2. For the solution to SRP′, we can without loss of generality restrict attention to contracts in which $r(w) = w$.

Again, we leave the proof to the Appendix but the intuition is standard: given any contract, we can construct a new contract such that when the agent reports W, the outcomes are as if the agent reported $r(W)$ under the original contract.

We now consider the optimal penalty structure under SRP′. Under truth-telling, the agent only incurs the penalties $x(W,W)$. In order to satisfy the agent's incentive constraint and prevent the agent from misreporting, it is clearly optimal to impose the maximum penalty if the agent is ever caught lying. Thus, $x(W,r)$ should be large enough so that agent consumes 0 if $r \neq W$.

What about the penalty $x(W,W)$ imposed when the agent tells the truth? If this penalty is positive, then in equilibrium the agent faces the risk of an investigation and a penalty. Since the agent is risk averse, the agent is better off if we instead reduce his compensation (i.e., raise z) by the expected penalty, and impose no penalty when the truth is told. This leads to the following characterization of SRP′:

Proposition 3. The solution to SRP′ is equivalent to the solution to

$$
\begin{aligned}
\text{SRP}^*: \quad & \max_{z,p} \quad E[\, u(W - z(W))\,] \\
& \text{subject to} \\
& (\text{AF}^*) \quad z(w) \le w, \\
& (\text{AIC}^*) \quad u(w - z(w)) \ge (1-p(w'))\, u(w - z(w')) \\
& \qquad\qquad \text{for all } w' \text{ such that } z(w') \le w, \\
& (\text{ZW}^*) \quad z(w) \ge \underline{w}, \\
& (\text{CIR}^*) \quad E[\, z(W) - p(W)\, c\,] \ge \alpha\,,
\end{aligned}
$$

with $r(w) = w$, $t = E[\, p(W)\, c\,]$, $x(w,w) = 0$ and $x(w,w') = w - z(w')$ for $w' \neq w$. Similarly, the solution to the monopoly agent's problem is the solution to the above without the constraint (ZW^*).

Given this simplification of SRP′, we now address its relation with SRP. Recall that SRP differs from SRP′ by allowing the customer to propose z. If the customer would choose the z at the solution of SRP′, then the two problems coincide. Otherwise, the solution to SRP′ is not feasible for SRP, and SRP is inferior to SRP′.

When choosing z, the customer would like to raise the payments received from the agent as much as possible. In doing so, the customer is constrained by the fact that if $z(w)$ is too large, the agent's incentive constraint will be violated and the agent will misreport if w occurs. Thus, through the incentive constraint, the enforcement policy restricts the rent extraction the customer can attain.

This leads to the following intuition. In a standard optimal contracting problem, we can in general make off-the-equilibrium path punishments as large as possible, as in **Proposition 3**. This is because there is no cost to having the incentive constraints strictly satisfied. However, in this context using a large penalty may give the customer the ability to increase z. To limit the customer's flexibility, it is optimal to use the weakest penalties possible subject to satisfying (AIC^*). To this end, given (z,p) we define the *weakest possible penalty* x^* such that the agent's incentive constraints are satisfied. That is, define x^* such that $x^*(w,w') = 0$ if $z(w') \ge z(w)$ and

$$u(w - z(w)) = (1-p(w'))u(w - z(w')) + p(w')u(w - z(w') - x^*(w,w')), \tag{3}$$

if $z(w') < z(w)$.

Thus, if the customer proposes any increase in $z(w)$, the agent will choose to lie. The penalty schedule x^* therefore imposes the biggest limitation on the customer's choice of z.

We now have our main result characterizing SRP, where we show that by relaxing the penalties to x^*, the solution to SRP* is also a solution to SRP.

> **Proposition 4.** The solution to SRP is equivalent to the solution to SRP*, with $r(w) = w$, $t = E[\, p(W)\, c\,]$ and $x = x^*$.

Thus, SRP* characterizes the solutions to SRP. This is rather striking in that the problem SRP* seems to ignore the fact that the customer has complete bargaining power (equivalently, agents behave competitively) over the payment scheme z. The result shows that the only equilibrium implication of this bargaining power is the constraint (ZW*); beyond that, the SRO's control of the enforcement policy conveys complete monopoly power.

The following proposition highlights this by examining the equilibrium payoff of the customer:

> **Proposition 5.** If $\alpha \leq \underline{w}$, then the solution to SRP* is given by $z \equiv \underline{w}$ and $p \equiv 0$, and the customer's expected payoff is $\underline{w}$. If $\alpha > \underline{w}$, then the CIR* constraint binds and the customer's expected payoff is α.

Thus, the customer's ability to propose the payment scheme z guarantees a minimum payoff of at least $\underline{w}$ for the customer. Beyond that, however, the customer is held to his reservation payoff and is unable to extract additional rents from the agent.

The remaining results of this section characterize further properties of the solution to SRP*. First, we consider the optimal investigation policy. Intuitively, the agent has the greatest incentive to cheat by reporting an outcome that requires a low payment z. Hence it is these outcomes that require the most vigorous enforcement to prevent cheating. This leads to an investigation policy that is inversely related to the payment schedule z.

> **Proposition 6.** At a solution (z,p) to SRP*, define
>
> $$p^*(w) = \max_{z(w') \geq z(w)} 1 - u(w' - z(w'))/u(w' - z(w)).$$
>
> Then $p(W) = p^*(W)$. Thus, p is weakly decreasing in z.

Finally, we consider the relation between the optimal payment scheme z and the outcome w of the project. Obviously, if it were feasible, it would be optimal to use a constant payment scheme z since then the deadweight costs of enforcement could be completely avoided. However, if $\alpha > \underline{w}$, then such a scheme is infeasible since the agent will not be able to make the payment if the outcome is too low. This forces down the payment scheme for low outcomes. Thus, it is natural to expect that in the optimal scheme, z is increasing in w.

> **Proposition 7.** If the agent is risk averse and we allow for mixed strategies, then z is weakly increasing in w. If the agent is risk neutral, then z is weakly increasing and concave, and no mixed strategy is used.

Recall that our key result, **Proposition 4**, states that the SRO's problem is equivalent to the problem faced by a monopolistic agent except for the addition of the constraint (ZW^*). It is not clear, however, whether or not a monopolistic agent would always choose to violate the (ZW^*) constraint. If not, then the SRO's problem is equivalent to the agent having full monopoly power.

To gain some intuition, first consider the case $\alpha < \underline{w}$. The optimal solution for a monopolistic agent is to offer the customer an expected payoff of α, whereas under self-regulation and competition over z, the agent is forced to pay $\underline{w}$. Thus, monopoly is strictly better than self-regulation for the agent in this case.

Next suppose $\alpha > \underline{w}$. To keep enforcement costs as low as possible, the contract z should be as "flat" as possible. A constant z is infeasible, however, since the (AF) constraint implies $z \leq \underline{w}$ in the worst state. But clearly, choosing $z = \underline{w}$ in the worst state is optimal from the standpoint of reducing the agent's incentive to cheat and report the worst state.

Thus $z(\underline{w}) = \underline{w}$ is optimal in lowering enforcement costs. However, if the agent is risk averse, the agent also cares about smoothing consumption across states. The disadvantage of choosing $z(\underline{w}) = \underline{w}$ is that it yields zero consumption for the agent in the worst case. This will not be optimal if the agent's marginal utility of income is sufficiently high at zero.

> **Proposition 8.** If the agent is risk neutral and $\alpha \geq \underline{w}$, then the solution to SRP coincides with the solution when the agent is a monopolist, as in (2). If $\alpha < \underline{w}$ or the agent is risk averse and $u'(0) = \infty$, then the monopoly solution in (2) is strictly better than the solution to SRP.

We conclude our analysis with a comparative static result regarding the probability of an investigation under SRP. We show that the likelihood of an investigation is increasing in the reservation payoff of the customer. This is intuitive, since the more the agent must pay the customer, the greater the agent's incentive to lie and report a low state.

More surprisingly, we also show that the probability of an investigation increases with the cost of an investigation. That is, the SRO's "demand function" for investigations is *upward* sloping. The intuition for this result is as follows. The payments z made by the agent must be sufficient to cover both the customer's reservation payoff α and the expected cost of enforcement. Increasing the investigation cost raises enforcement costs and therefore increases the expected payment necessary from the agent. Thus, this leads to the same comparative static as an increase in α.

Proposition 9. If the agent is risk neutral, then at a solution to SRP the expected probability of an investigation, $E[\ p(W)\ ]$, is increasing in c and α.

We conclude this section with numerical examples.

Example 1. Suppose W is drawn uniformly from the set {100, 150, ..., 500}. Let $c = 100$, and $\alpha = 200$. Finally, let the agent be risk averse with utility $u(x) = x^{1/2}$. Then the optimal payment schedule $z(W)$ is illustrated below, together with the investigation policy $p(W)$ (represented in % on the same scale). Finally, the dashed lines represent the constraints on z imposed by (AIC*). Where z is below the dashed lines, the penalties x^* are below the maximum possible. Thus in this case, if the outcome is 150 or 200, and the agent reports a lower outcome, a weaker than maximal fine is imposed.

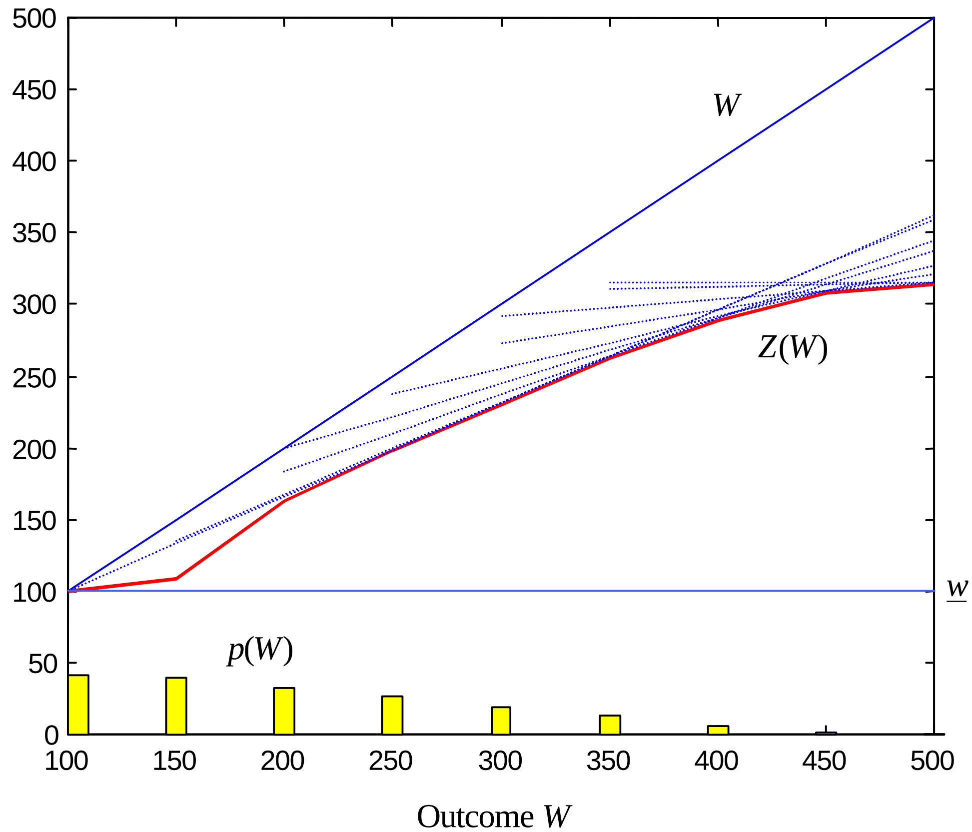

Figure 1. The Solution to SRP for Example 1.

Example 2. We now modify example 1 by making the agent risk neutral. In that case, by **Proposition 7**, the schedule z is concave. Equivalently, the agent's payment, $w - z(w)$, is convex in the payoff w. This is illustrated in Figure 2.

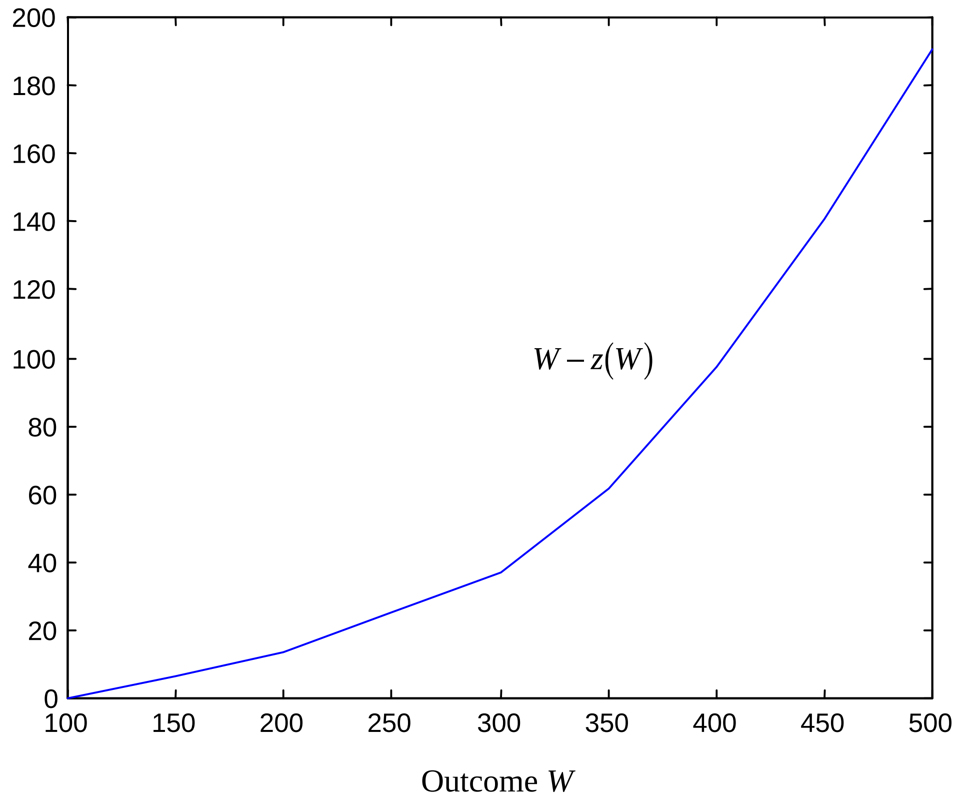

Figure 2. Payments to a Risk Neutral Agent.

5. Heterogeneous Customers

Thus far, we have taken the customer's opportunity cost α as fixed. In this section we allow for a heterogeneous customer population. Let $F(\alpha)$ represent the fraction of the customer population with opportunity cost below α. In this case, the enforcement policy chosen by the SRO ex-ante will determine the fraction of the population that will be willing to transact with an agent. The SRO thus faces a tradeoff between the expected rents of an agent per transaction and the volume of trade.

Once the SRO has set an enforcement policy (p,x,t), competition between agents implies that each customer i (independent of her opportunity cost α_i) will negotiate the payment schedule (z,r) that is the solution to CP(p,x). This schedule yields an expected payoff of $\alpha = E[z(r(W)) - t]$. Thus, only customers with $\alpha_i \leq \alpha$ will transact with the agent, resulting in a volume of $F(\alpha)$. Thus, among all enforcement policies that lead to the same expected payoff α for the customer, the volume of trade is identical, and so the SRO will choose the policy that maximizes the agent's expected utility. This, of course, is precisely the solution to SRP given α.

Therefore, the SRO's decision problem can be decomposed as follows. First, the SRO chooses an α determining the fraction of the customer population that will be served. Then the enforcement policy is chosen as the solution to SRP given α. If we denote by $V(\alpha)$ the agent's expected utility at the solution to SRP given α, then the SRO chooses α to solve

$$\max_{\alpha}\ F(\alpha)\, V(\alpha). \tag{4}$$

5.1 SRO Competition

Allowing for the endogenous determination of α in this fashion permits us to consider a simple model of SRO competition. We suppose that individual customers have a choice between two SRO-governed exchanges at which they can transact. Though each exchange offers similar trading opportunities, they are not perfect substitutes for each other. We assume that these differences manifest as differences in the ultimate payoffs for customers, but not as differences in the outcome of the transaction itself.

As a concrete example, a customer may wish to invest in the technology stock. This can be done by trading a stock listed on the NASDAQ exchange, or by trading a different stock listed on the NYSE. We might then interpret the customer-broker contract as covering the execution quality of the purchase. Given differing enforcement policies of the exchanges, customers might in general receive better execution on one exchange or the other. However, their beliefs about the returns on the alternative stocks might still lead them to trade on the exchange offering worse average execution.

Formally, we suppose SRO's 1 and 2 each choose an enforcement policy consistent with some α_1 and α_2 respectively. Agent i has a preference for trading on exchange 2 given by the amount ε_i. Thus, agent i will trade on exchange 1 if

$$\alpha_1 \geq \alpha_2 + \varepsilon_i.$$

If we assume that SRO's choose their enforcement policies simultaneously, we can then solve for a Nash equilibrium in enforcement. Let F represent the population distribution of ε_i. Then the equilibrium conditions are given by

$$\alpha_1^* \in \arg\max_{\alpha_1} F\left(\alpha_1 - \alpha_2^*\right) V\left(\alpha_1\right),$$

$$\alpha_2^* \in \arg\max_{\alpha_2} \left[1 - F\left(\alpha_1^* - \alpha_2\right)\right] V\left(\alpha_2\right).$$

If the distribution of ε_i is symmetric about zero, then $\left[1 - F\left(\alpha_1^* - \alpha_2\right)\right] = F\left(\alpha_2 - \alpha_1^*\right)$ and it is natural to consider symmetric equilibria with $\alpha_1^* = \alpha_2^* = \alpha^*$. This yields the first order condition:

$$\frac{V'\left(\alpha^*\right)}{V\left(\alpha^*\right)} = 2F'(0).$$

[Show that as F concentrates around 0, α^* goes to competitive solution.]

6. A Two-State Illustration of Self-Regulation

In this section, we simplify the model by supposing the cash flow, W, takes on one of two values, $\Omega = \{w_1, w_2\}$, where $w_2 > \alpha > w_1$. By **Proposition 1** and the agent's feasibility constraint, (AF), we have

$$z(w_1) = w_1.$$

Using **Proposition 6**, investigation probabilities are given by

$$p(w_1) = 1 - u(w_2 - z(w_2))/u(w_2 - w_1)$$

and

$$p(w_2) = 0.$$

The former equation, relating $p(w_1)$ and $z(w_2)$, comes from the agent's incentive-compatibility constraint, (AIC). Using (3), we have that the penalty for falsely reporting $W = w_1$ is

$$x(w_2, w_1) = w_2 - w_1,$$

the highest penalty possible. There is no penalty for truthfully reporting $W = w_1$,

$$x(w_1, w_1) = 0.$$

Since $p(w_2) = 0$, all other penalties are irrelevant.

Proposition 5 states that the customer's individual rationality constraint, (CIR), is binding if $\alpha > w_1$, so we have

$$z(w_2) = [\alpha - \Pr(W{=}w_1)(w_1 - p(w_1)\, c)]/\Pr(W{=}w_2).$$

The transaction fee equals the expected investigation cost and is given by

$$t = \Pr(W{=}w_1)\; p(w_1)\, c.$$

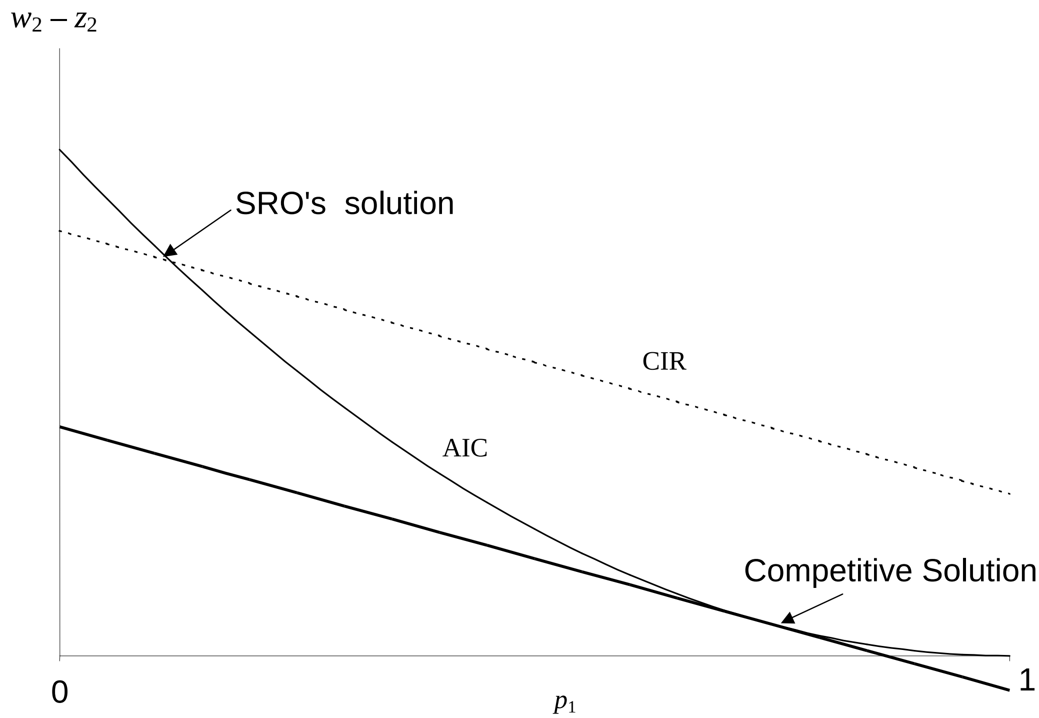

Figure 3. Competitive versus SRO contract.

Figure 3 illustrates the equilibrium choice of $z(w_2)$ and $p(w_1)$. Inducing the agent to truthfully report $W = w_2$ requires a high agent payoff $w_2 - z(w_2)$ and/or a high probability $p(w_1)$ of an investigation given a report of $W = w_1$. This is the region above (AIC) in Figure 3. The customer participates in the transaction if his payoff $z(w_2)$ is high enough and/or the investigation probability $p(w_1)$ (and expected investigation cost) is low enough. This is the region below (CIR) in Figure 3.

The customer takes the investigation probability $p(w_1)$ as given and chooses the lowest agent payoff $w_2 - z(w_2)$ that induces the agent to report truthfully. The agent prefers to be induced to tell the truth via the "carrot" of a high payoff rather than via the "stick" of a high probability of investigation. Hence the SRO's preference is to set a low $p(w_1)$ and thus induce the customer to offer a high $w_2 - z(w_2)$. Indeed, the SRO would like to set $p(w_1) = 0$ if it would lead the customer to satisfy (AIC) by setting $w_2 - z(w_2) = w_2 - w_1$. It would not though; instead the customer would not participate since the customer's individual-rationality constraint would not be satisfied. The customer's individual-rationality constraint limits how low the SRO can set $p(w_1)$. Hence the intersection of (AIC) and (CIR) determines the values of $z(w_2)$ and $p(w_1)$.

Figure 3 also shows the outcome that would prevail with fully competitive agents and the customer choosing everything, as in (1). As is the case with self-regulation, the customer would choose $z(w_1) = w_1$, $p(w_2) = 0$, $x(w_2,w_1) = w_2 - w_1$, $x(w_1,w_1) = 0$, and $t = \Pr(W{=}w_1)\, p(w_1)\, c$. The difference is in the choice of $z(w_2)$ and $p(w_1)$. The customer maximizes his expected payoff subject to the agent incentive-compatibility constraint. The solution lies at the tangency between the customer's indifference curve and (AIC).

Compare the self-regulation outcome to the fully competitive outcome. With self-regulation, enforcement is more lax; that is, the investigation probability is lower. And with self-regulation, the agent's payoff $w_2 - z(w_2)$ is higher. So the agent's expected utility is higher with self-regulation and the customer's expected utility is lower.

Now compare the self-regulation outcome to the outcome that would prevail with a monopolistic agent who chooses everything, as in (2). As in the case with self-regulation, a monopolistic agent would choose $p(w_2) = 0$, $x(w_2,w_1) = w_2 - w_1$, $x(w_1,w_1) = 0$, and $t = \Pr(W{=}w_1)\ p(w_1)\ c$. A possible difference between these problems, though, involves the monopolistic agent's choice of $z(w_1)$. If, like the self-regulation case, $z(w_1) = w_1$ is chosen, then the optimal $z(w_2)$ and $p(w_1)$ will also be the same as with self-regulation. If, however, $z(w_1) < w_1$ is chosen, then $z(w_2)$ and $p(w_1)$ will be different too. A monopolistic agent might choose $z(w_1) < w_1$ because of risk aversion. If his marginal utility at zero consumption is sufficiently high, then he will choose $z(w_1) < w_1$ and ensure himself positive consumption in both states. Figure 4 shows the effect of choosing $z(w_1) < w_1$ on the agent's incentive-compatibility constraint (AIC) and on the customer's individual-rationality constraint (CIR). Choosing $z(w_1) < w_1$ leads (AIC) to twist out. With $z(w_1)$ lower and $p(w_1)$ unchanged, reporting $W = w_1$ becomes more attractive so a higher agent payoff $w_2 - z(w_2)$ is required to induce the agent to truthfully report that $W = w_2$. Choosing $z(w_1) < w_1$ leads (CIR) to shift down. With $z(w_1)$ lower and $p(w_1)$ unchanged, a higher customer payoff when $W = w_2$ is required to induce the customer to participate in the transaction. These changes in (AIC) and (CIR) imply that if the monopolistic agent chooses $z(w_1) < w_1$, then he will also choose a higher investigation probability than will an SRO. In turn, the customer will choose a lower agent payoff for reporting $W = w_2$.

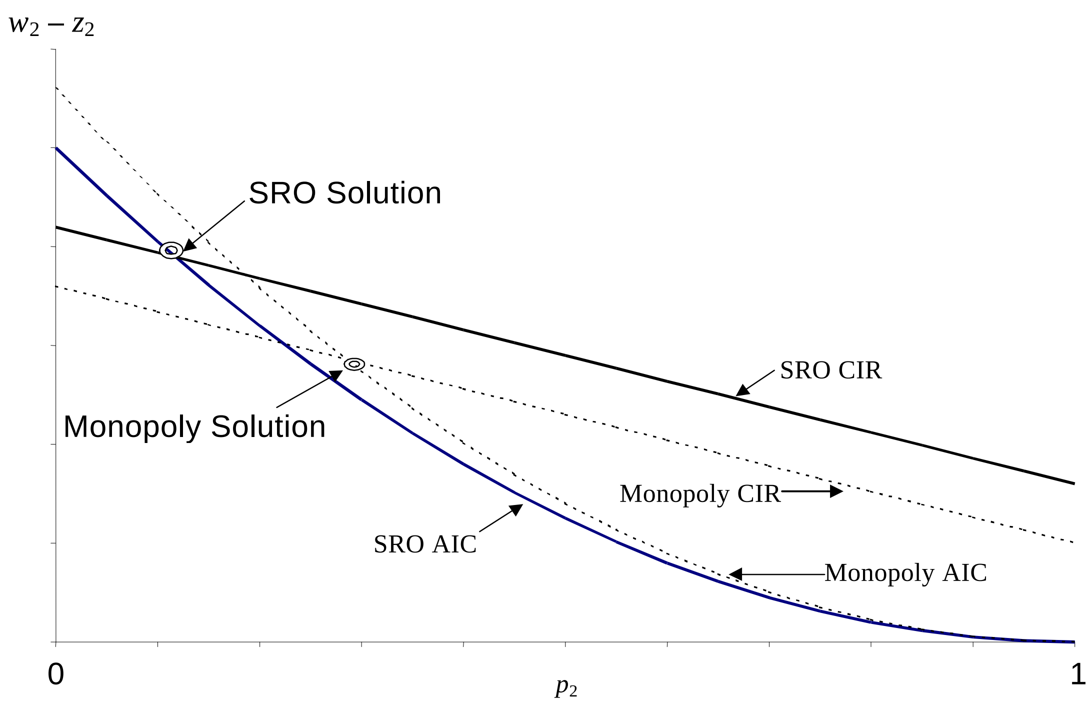

Figure 4. Monopoly versus SRO solution.

Now consider some comparative statics, beginning with the investigation cost, *c*. A change in *c* has opposite effects in the fully competitive and self-regulatory cases. In the fully competitive case, an increase in *c* leads the customer to choose a lower investigation probability $p(w_1)$ and a higher agent payoff $w_2 - z(w_2)$ (the customer uses more "carrot" and less "stick"). With self-regulation, an increase in the investigation cost leads the SRO to choose a *higher* investigation probability, inducing the customer to choose a lower agent payoff. See Figure 5. An increase in *c* leads to steeper customer indifference curves; an increase in $p(w_1)$ is more costly so indifference requires a larger decrease in $w_2 - z(w_2)$. With self-regulation, the customer individual-rationality constraint, (CIR), is binding. So an increase in the investigation cost (and thus transaction fee), holding the enforcement policy and contract fixed, results in a violation of (CIR). To satisfy (CIR) the SRO must choose an investigation probability that is more preferred by the customer, and the customer prefers a higher $p(w_1)$. And with a higher $p(w_1)$ the customer will choose a lower $w_2 - z(w_2)$.

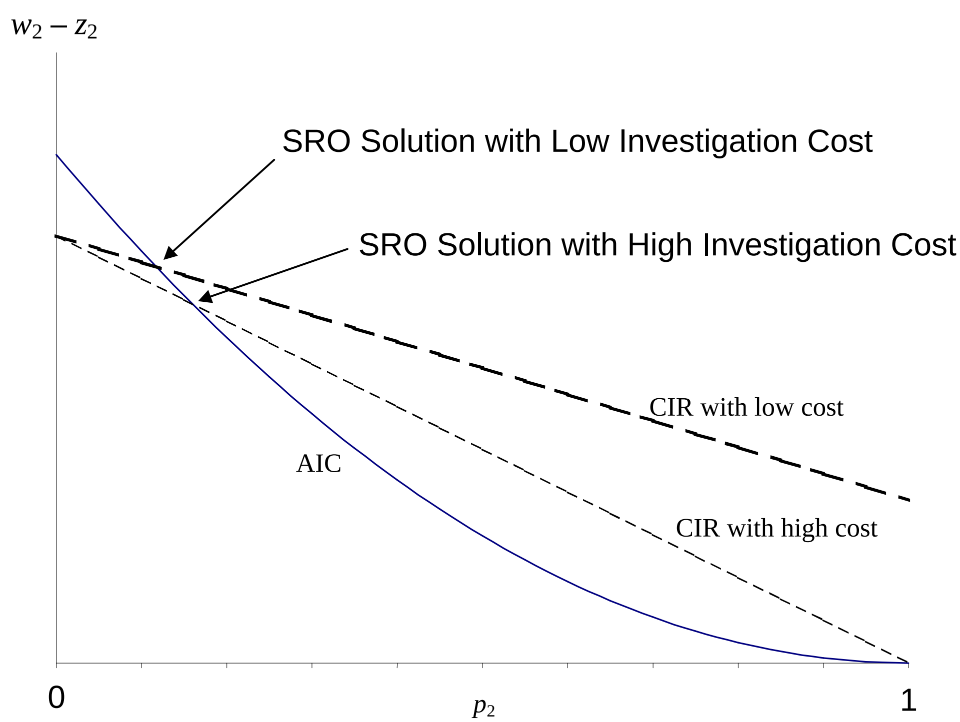

Figure 5. High versus Low Investigation Cost.

An increase in $\Pr(W=w_1)$ has the same effect as an increase in the investigation cost. With self-regulation, an increase in the likelihood of the low cash flow leads the SRO to choose a higher investigation probability $p(w_1)$ which induces the customer to choose a lower agent payoff $w_2 - z(w_2)$. By contrast, in the fully competitive case, an increase in $\Pr(W=w_1)$ leads to a lower $p(w_1)$ and a higher $w_2 - z(w_2)$. A change in the customer's reservation utility, α, leads an SRO to choose a higher investigation probability which leads the customer to choose a lower agent payoff.

7. Government Oversight

We now consider the possibility of government oversight of the SRO's enforcement policy. Oversight is modeled as an additional tier on the contracting problem where the government can also investigate the agent and impose penalties. First, the SRO chooses its enforcement policy. Second, taking the SRO's enforcement as given, the government chooses its enforcement policy. Here we suppose that the government can impose its own monetary penalties in addition to any imposed by the SRO, and can conducted its own investigation if the SRO does not.[6] Third, taking the SRO and government enforcement policies as given, the customer offers a contract to the agent. We assume that the government's objective is to maximize the expected utility of the customer.

Presumably members of an SRO have greater expertise and better information than government regulators and so can more easily determine whether another member has cheated. So we assume that the government's cost of investigating the agent is higher than the SRO's cost. Let c_g denote the government's investigation cost, where $c_g \geq c$. [7]

We will develop the model with government oversight for the two-state case discussed in Section 6. That is, $\Omega = \{w_1, w_2\}$. This allows us to immediately focus on equilibria in which the agent tells the truth. For if the agent reports the same outcome for both w_1 and w_2, then by feasibility the customer can receive at most w_1. Thus, in an equilibrium in which the customer receives more than w_1, the agent makes distinct reports, which we can label w_1 and w_2 without loss of generality.[8]

We can simplify the problem further as follows. Given an equilibrium in which the agent reports truthfully, it is optimal for the government to impose the maximum penalty if the agent is caught lying. This only strengthens the incentives for truth-telling, allowing the customer to extract the highest payment possible from the agent if w_2 occurs. Moreover, it is not optimal for the government to impose any additional penalty if the agent is found to have told the truth. This imposes risk on the agent, reducing the agent's incentive to report the truth and reducing the payment the customer can demand from the agent in equilibrium.

Since the government will always impose the maximum penalty for a cheating, the SRO is no worse off by simply specifying the maximum penalty initially. Also by a similar reasoning, the SRO will not impose a penalty for a truthful report. Thus we can restrict attention to only consider the penalty x_m that imposes the maximum penalties for fraud, i.e., $x_m(w,r) = w$ for $r \neq w$ and $x_m(w,r) = 0$ for $r = w$.

Given these observations, the problem with government oversight in the two-state case reduces to a selection of an investigation probabilty by the SRO and the government. After a report r by the agent, let $p_s(r)$ and $p_g(r)$ denote the probability of an investigation by the SRO and the government, respectively. We assume that the investigations are

[6] Note that we limit the analysis to civil penalties. An additional role of the government that we do not consider is the imposition of non-pecuniary penalties, such as jail time.
[7] This assumption justifies the existence of the SRO. If the government's investigation cost were lower than the SRO's cost, it would be Pareto optimal to eliminate the SRO and have all investigations performed by the government
[8] Note that in the general case, truth-telling in equilibrium cannot be automatically assumed. The revelation principle cannot be applied since there is no party that can commit to a mechanism.

sequential, with the SRO deciding first whether to investigate. Since investigations are costly, the government will never conduct a redundant investigation. Thus, the probability that the agent is investigated given that he reports r is given by $p(r) \equiv p_s(r)+p_g(r)$.[9]

The transaction fees used to fund the SRO's and government's enforcement activities are denoted t_s and t_g, respectively. The customer pays a total of $t \equiv t_s+t_g$ to participate in this market.

To summarize, the problem with government oversight is as follows:

1. The SRO chooses an investigation strategy p_s and a transaction fee t_s (and the SRO will choose the maximum penalty schedule, x_m).
2. Taking the SRO's enforcement policy and transaction fee, (p_s,x_m,t_s), as given, the government chooses its investigation schedule p_g and its transaction fee t_g (and the government will choose the maximum penalty schedule, x_m).
3. Taking the overall enforcement policy, (p,x_m,t), as given, the customer offers a contract z to the agent, which the agent can accept or reject.
4. If the agent rejects the contract, the agent receives 0 and the customer receives α. If the agent accepts the contract, the customer pays t and the problem continues.
5. A cash flow w is realized.
6. The agent chooses a cash flow, r, to report and gives the customer the corresponding payoff, $z(r)$.
7. Given the reported cash flow r, the agent is investigated with probability $p(r)$. If the SRO investigates the agent, the SRO pays c and assesses the penalty $x_m(w,r)$ on the agent. If the government investigates the agent, the government pays c_g and assesses the penalty $x_m(w,r)$ on the agent.

The customer's problem and the agent's problem are the same as before. Taking (p,x_m) as given, the customer's problem is given by CP(p,x_m). As noted above, with a two-state distribution for W, the solution to this problem entails truthful reporting by the agent. Let r_T denote the truthful reporting strategy, $r_T(W)=W$. Then we can rewrite the customer incentive constraint as

$$\text{(CIC}'\text{)} \quad (z,r_T) \text{ solves CP}(p,x_m).$$

The specification of the government's and SRO's problems will reflect truthful reporting and their budget constraints reflect the fact that no penalties will be imposed.

[9] This is equivalent to the following. First, the SRO investigates with probability $p_s(r)$. Conditional on the SRO not investigating, the government then investigates with probability $p' = p_g(r)/(1-p_s(r))$. This leads to a total investigation probability $p(r) = p_s(r) + (1-p_s(r))p' = p_s(r) + p_g(r)$, with the unconditional probability of a government investigation of $(1-p_s(r))p' = p_g(r)$.

7.1 The Government's Problem

The government takes the SRO's enforcement policy, (p_s,x_m,t_s), as given and chooses its own investigation strategy and transaction fee, (p_g,t_g), to maximize the customer's expected utility (the government will choose the maximum penalty schedule, x_m). The government's transaction fee, t_g, must be sufficient to cover the government's expected investigation cost. This yields a government budget constraint:

$$\text{(GB)} \quad t_g \geq E[\ p_g(W)\ c_g\ ].$$

The government's problem can be written as the following mechanism design problem:

$$\text{GP}(p_s,t_s): \qquad \max_{z,p_g,t_g} \qquad E[\ z(W)\ ] - t_g - t_s$$

$$\text{subject to (CIC}') \text{ and (GB).}$$

7.2 The SRO's Problem

Anticipating the government oversight, the SRO chooses p_s and t_s to maximize the agent's expected utility (the SRO will choose the maximum penalty schedule, x_m). The SRO's problem with oversight can be written as the following mechanism design problem:

$$\text{SRPO:} \qquad \max_{z,p_s,p_g,t_s,t_g} E[\ v(W,W|z,p,x_m)\ ]$$

subject to (CIR),

(GIC) (z,p_g,t_g) solves $\text{GP}(p_s,t_s)$,

(RB′) $t_s \geq \text{E}[\ p_s(W)c\ ]$.

[Add discussion here.]

> **Proposition 10.** Let p^{srp} be the investigation probability at the solution to SRP given investigation cost c. Let p^{cg} be the investigation probability given investigation cost c_g if agents were perfectly competitive as in (1). Then at the solution to SRPO, $p_g = 0$ and $p = p_s = \max[\ p^{srp}, p^{cg}\ ]$.

Thus, with government oversight, no actual enforcement by the government will be observed. Government oversight will be effective and will increase the customer's expected payoff, however, as long as an increased probability of investigation would be efficient given a cost of c_g.

[Add Figure showing solution.]

8. Reputation in a Dynamic Setting

In this section we extend our previous results to a dynamic setting in which an agent sees new customers over time. In this setting the SRO can also penalize the agent by restricting his ability to transact with customers in the future. That is, since agents earn positive expected rents per transaction, penalties can now be increased up to include the discounted value of the anticipated future rents.

We show that when investigation costs are sufficiently high, it is optimal for an SRO to suspend an agent who reports a low outcome without conducting an investigation. If investigation costs are low, however, the SRO will not suspend an agent unless an investigation reveals the agent cheated, in which case the agent is fired (permanently suspended). In this case, the SRO's problem is essentially SRP above, but with the ability to assess higher penalties.

To be completed.

9. Concluding Remarks

To be added.

10. Appendix

To be completed.

Proof of Proposition 1:

Define z' such that $z'(r) = \max(z(r), \underline{w})$. Then note that $v(w,r|z',p,x) \leq v(w,r|z,p,x)$ for all w, r.

Define $Q = \{w : z(r(w)) < \underline{w}\}$. Then let

$$r'(w) = r(w) \qquad \text{if } w \notin Q \text{, and}$$

$$r'(w) = \operatorname{argmax}_r \; v(w,r|z',p,x) \qquad \text{if } w \in Q.$$

Suppose $w \notin Q$ so that $z(r(w)) \geq \underline{w}$. Then, if $z(s) \leq w$,

$$v(w,r'(w)|z',p,x) = v(w,r(w)|z',p,x) = v(w,r(w)|z,p,x) \geq v(w,s|z,p,x) \geq v(w,s|z',p,x).$$

Hence, (z',r') satisfies (AIC) and is feasible for CP(p,x).

Finally, $z'(r'(w)) = z'(r(w)) = z(r(w))$ if $w \notin Q$. However, for $w \in Q$, $z'(r'(w)) \geq \underline{w} > z(r(w))$. Thus, $E[\, z'(r'(W)) \,] > E[\, z(r(W)) \,]$ unless $Q = \emptyset$.

Hence if z solves CP, we must have $Q = \emptyset$. ◆

Proof of Proposition 2:

Suppose (z,r,p,x,t) is feasible for SRP'. Then we show that there is a feasible (z',r',p',x',t) with the same payoff and $r'(w) = w$.

Define (z',p',x') via

$$z'(w) = z(r(w)),\ \ x'(w,s) = x(w,r(s)) \ \text{ and } \ p'(w) = p(r(w)).$$

Clearly, $z'(w) = z(r(w))$ implies that (AF), (CIR) and (ZW) are satisfied.

Next, note that

$$v(w,s|z',p',x') = p'(s)\, u([w - z'(s) - x'(w,s)]^+) + (1 - p'(s))\, u(w - z'(s))$$

$$= p(r(s))\, u([w - z(r(s)) - x(w,r(s))]^+) + (1 - p(r(s)))\, u(w - z(r(s)))$$

$$= v(w,r(s)|z,p,x).$$

Thus if $z'(s) = z(r(s)) \leq w$, we have

$$v(w,s|z',p',x') = v(w,r(s)|z,p,x) \leq v(w,r(w)|z,p,x) = v(w,w|z',p',x'),$$

so that (AIC) is satisfied.

Finally, note that

$$p'(w)(c - \min[x'(w,w),\, w - z'(w)]) = p(r(w))(c - \min[x(w,r(w)),\, w - z(r(w))]),$$

so that (RB) is unchanged.

Hence, (z',r',p',x',t) is feasible for SRP′ and since $v(w,w|z',p',x') = v(w,r(w)|z',p',x')$, the agent's payoff is unchanged. ◆

Proof of Proposition 3:

First we show that if (z',r',p',x',t') is feasible for SRP′, there is (z,p) feasible for SRP* with a higher payoff for the agent. Then we show that if (z,p) is feasible for SRP*, then (z,r,p,x,t) is feasible for SRP′ and has the same payoff for the agent.

<u>Step 1:</u> Suppose (z',r',p',x',t') is feasible for SRP′.

By Proposition 2, we assume $r'(w) = w$. Define $p(w) = p'(w)$ and

$$z(w) = z'(w) + p'(w)\min[x'(w,w),\ w - z'(w)].$$

Note that $x'(w,w) \geq 0$ implies that $z(w) \geq z'(w) \geq w$ and (ZW*) holds. Also, $z(w) \leq z'(w) + w - z'(w) = w$ and (AF*) holds.

Next note that

$$z(w) - p(w)\,c = z'(w) - p'(w)(c - \min[x'(w,w),\ w - z'(w)]).$$

Thus,

$$E[\ z(W) - p(W)\,c\ ] \geq E[\ z'(W)\ ] - t \geq \alpha,$$

and (CIR*) holds.

Finally, since u is concave,

$$\begin{aligned} u(w - z(w)) &= u(w - z'(w) - p'(w)\min[x'(w,w),\ w - z'(w)]) \\ &\geq p'(w)\,u([w - z'(w) - x'(w,w)]^+) + (1-p'(w))\,u(w - z'(w)) \\ &= v(w,w|z',p',x'). \end{aligned}$$

Thus, the agent's payoff is higher. It remains to check (AIC*).

If $z(s) \leq w$ then

$$\begin{aligned} u(w - z(w)) &\geq v(w,w|z',p',x') \geq v(w,s|z',p',x') \\ &= p'(s)\,u([w - z'(s) - x'(w,s)]^+) + (1-p'(s))\,u(w - z'(s)) \\ &\geq p(s)\,u(0) + (1-p(s))\,u(w - z(s)), \end{aligned}$$

and (AIC*) follows since $u(0) = 0$.

<u>Step 2:</u> Suppose that (z,p) is feasible for SRP*.

Define (r,x,t) as in the statement of the proposition. Clearly, (AF) and (ZW) are satisfied. Since $x(w,w) = 0$ and $r(w) = w$,

$$p(r(w))(c - \min[(x(w,r(w)),\ w - z(r(w))]) = p(w)\,c.$$

Thus $t = E[\, p(W)\, c\,]$ satisfies (RB) and (CIR).

Next note that

$$v(w,r(w)|z,p,x) = u(w - z(w)),$$

so that the agent's payoff is unchanged.

Thus, it remains to show that (AIC) is satisfied. Since

$$v(w,s|z,p,x) = p(s)\, u(0) + (1-p(s))\, u(w - z(s)) = (1-p(s))\, u(w - z(s)),$$

this follows from (AIC*). ◆

Proof of Proposition 4:

Recall that SRP′ is better than SRP (we dropped the customer's incentive constraint). Thus, to prove the result we need only show that (z,r,p,x^*,t) satisfies (CIC), or equivalently, (z,r) solves CP(p,x^*).

First note that x^* is well-defined and that for $z(w') < z(w)$, $0 < x^*(w,w') \leq w - z(w')$. Define z by $\underline{z}(w) \equiv \underline{w}$. By the definition of x^*, for $z(w') < z(w)$,

$$v(w,w|z,p,x^*) = v(w,w'|z,p,x^*) \leq v(w,w'|\underline{z},p,x^*).$$

Also, for $z(w') \geq z(w)$,

$$v(w,w|z,p,x^*) = u(w - z(w)) \leq u(w - \underline{w}) = v(w,w'|\underline{z},p,x^*).$$

Thus, we have shown that for all w',

$$v(w,w|z,p,x^*) \leq v(w,w'|\underline{z},p,x^*).$$

Now suppose the customer proposes an alternative (z',r') satisfying (AF) and (AIC) given (p,x^*). By (AF), $z'(r'(w)) \leq w$, so by **Proposition 1** we can assume $z'(r'(\underline{w})) = \underline{w}$.

Let $\underline{r} = r'(\underline{w})$. Then for $w \geq \underline{w}$,

$$v(w,r'(w)|z',p,x^*) \geq v(w,\underline{r}|z',p,x^*) = v(w,\underline{r}|\underline{z},p,x^*).$$

Thus, by the previous result we have

$$v(w,r'(w)|z',p,x^*) \geq v(w,w|z,p,x^*) \quad \text{for all } w \geq \underline{w},$$

so that the agent must be better off under (z',r') than under (z,r). We now show that this implies the customer must be worse off.

We can rewrite the above as,

$$p(r'(w))\, u(w - z'(r'(w)) - x^*(w,r'(w))) + (1-p(r'(w))\, u(w - z'(r'(w))) \geq u(w - z(w)).$$

Since u is concave, this implies

$$z'(r'(w)) + p(r'(w))\, x^*(w,r'(w)) \leq z(w).$$

Since $x^* \geq 0$, this implies $z'(r'(w)) \leq z(w)$, and the customer is worse off for all $w \geq \underline{w}$.

Thus, (z,r) solves CP(p,x^*). ◆

Proof of Proposition 5:

The case $\alpha \leq w$ is immediate. Next suppose $\alpha > w$ and (CIR*) does not bind.

Then define $z'(w) = \min[z(w), b]$ for b such that $E[\ z'(W) - p(W)c\ ] = \alpha$; i.e., such that (CIR*) binds.

Clearly (AF*) and (ZW*) hold for (z',p). For (AIC*), if $z'(w') < z'(w)$, then $z'(w') = z(w')$. Hence,

$$u(w - z'(w)) \geq u(w - z(w)) \geq (1-p(w'))\ u(w - z(w')) = (1-p(w'))\ u(w - z'(w')),$$

and (AIC*) holds.

Thus, (z',p) is feasible for SRP*, and since $z'(W) \leq z(W)$ and $E[\ z'(W)\ ] < E[\ z(W)\ ]$, it has a higher payoff for the agent. Therefore, at any solution to SRP*, (CIR*) must bind. ◆

Proof of Proposition 6:

Suppose (z,p) is feasible. Then from (AIC*), for $w \geq z(w')$,

$$u(w - z(w)) \geq (1-p(w'))\ u(w - z(w'))$$

which is equivalent to

$$p(w') \geq 1 - [\ u(w - z(w)) \ /\ u(w - z(w'))\ ].$$

Thus, (AIC*) is equivalent to (after switching w and w'),

$$p(w) \geq \max_{w' \geq z(w)} 1 - [\ u(w' - z(w')) \ /\ u(w' - z(w))\ ].$$

Note that for $w' = w$,

$1 - [\ u(w' - z(w')) \ /\ u(w' - z(w))\ ] = 0.$

If $w' \geq z(w)$ and $z(w') < z(w)$, then

$$1 - [\ u(w' - z(w')) \ /\ u(w' - z(w))\ ] < 0.$$

Therefore, (AIC*) is equivalent to

$$p(w) \geq p^*(w) \equiv \max_{z(w') \geq z(w)} 1 - [\ u(w' - z(w')) \ /\ u(w' - z(w))\ ].$$

Suppose $p(W) > p^*(W)$ with positive probability. Then (z,p^*) is also feasible for SRP*, and entails lower investigation probabilities. Since $E[\ p(W)\ ] > E[\ p^*(W)\ ]$, (CIR*) does not bind for (z,p^*). By **Proposition 5**, this implies that (z,p^*) and hence (z,p) cannot solve SRP*. Finally, $p^*(w)$ decreasing in $z(w)$ is immediate. ◆

Proof of Proposition 7:

Case 1: the agent is risk neutral.

Then (AIC*) becomes

$$w - z(w) \geq (1-p(w'))\,(w - z(w')) \text{ for } z(w') \leq w.$$

This is equivalent to

$$z(w) \leq p(w')\,w + (1-p(w'))\,z(w') \text{ for all } w',$$

or,

$$z(w) \leq z^*(w) = \min_{w'} p(w')\,w + (1-p(w'))\,z(w').$$

It is easy to see that z^* is increasing and concave in w. We will show that at an optimum, $z(W) = z^*(W)$. Suppose instead $z(W) < z^*(W)$ with positive probability, then we can define z' by

$$z'(w) = \min[z^*(w), b],$$

for b such that $E[\, z'(W)\,] = E[\, z(W)\,]$. This implies $z(W) > b$ with positive probability. Also define $p'(w) = p(w)$ if $z'(w) < b$ and $p'(w) = 0$ if $z'(w) = b$. This implies $E[\, p'(W)\,] < E[\, p(W)\,]$.

To see that (z',p') satisfies (AIC*),

$$z'(w) \leq z^*(w) = \min_{w'} p(w')\,w + (1-p(w'))\,z(w')$$

$$\leq \min_{w'} p'(w')\,w + (1-p'(w'))\,z^*(w').$$

Given this, it is straightforward to check that (z',p') is feasible for SRP*, and the payoffs of the agent are unchanged. Moreover, since $E[\, p'(W)\,] < E[\, p(W)\,]$, (CIR*) does not bind. Hence, by Proposition 5, (z',p') and hence (z,p) is not optimal. Thus, at an optimum we must have $z(W) = z^*(W)$.

Case 2: the agent is risk averse.

To be completed.

◆

11. References (Incomplete!)

Frankhauser, M. M., Gardner, L. M., McNally, B. F., and Leatherwood, L. A., "Enforcement by Self-Regulatory Organizations: A Growing Impact on Broker-Dealers and their Personnel," in The Securities Enforcement Manual: Tactics and Strategies, ed. by R. M. Phillips. Chicago: American Bar Association, 1997.

Gehrig, T. and Jost, P-J., "Quacks, Lemons, and Self Regulation: A Welfare Analysis," Journal of Regulatory Economics, 1995, 7, 309-325.

Leland, H. E., "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards," Journal of Political Economy, 1979, 87, 1328-1346.

Mahoney, P. G., "The Exchange as Regulator," Virginia Law Review, 1997, 83, 1453-1500.

Mookherjee, D. and Png, I. P. L., "Optimal Auditing, Insurance, and Redistribution," Quarterly Journal of Economics, 1989, 104, 399-415.

National Association of Securities Dealers, "Securities Regulation in the United States," NASD, 1996.

Phillips, R. M., "Overview: The Multilayered Securities Enforcement System," in The Securities Enforcement Manual: Tactics and Strategies, ed. by R. M. Phillips. Chicago: American Bar Association, 1997.

Pirrong, S. C., "The Self-Regulation of Commodity Exchanges: The Case of Market Manipulation," Journal of Law and Economics, 38, 141-206.

Saloner, G., "Self-Regulating Commodity Futures Exchanges," in The Industrial Organization of Futures Markets, ed. by R. W. Anderson. Lexington: D. C. Heath, 1984.

Shaked, A. and Sutton, J., "The Self-Regulating Profession," Review of Economic Studies, 1981, 48, 217-234.

Townsend, R., "Optimal Contracts and Competitive Markets with Costly State Verification," Journal of Economic Theory, 1979, 22, 265-293.