THE USE OF AN AGENT IN A SIGNALLING MODEL

Bernard Caillaud[*]

and

Benjamin Hermalin[†]

N° 9136

December 1991[‡]

THE USE OF AN AGENT IN A SIGNALLING MODEL

Abstract

In many models in finance and industrial organization, a firm's manager signals information. Typically, any agency problems are assumed away. In contrast, we assume there is an agency problem and explore the consequences. We derive conditions under which models with agency problems yield *qualitatively* similar predictions to models without agency problems. These conditions are stronger than the standard assumptions in signalling models, which raises doubts about the robustness of the earlier literature to agency problems. Fortunately, in some well-studied models, our conditions are met. We also show that, in some circumstances, the principal does better by hiring an agent to signal for her than by signalling herself. We even show that it is possible that the principal's expected profit is *increasing* in the cost of agency !

Key words: Signalling, principal-agent contracting theory, incomplete information
JEL Classification: 026

DELEGATION DANS LES MODELES DE SIGNALISATION

Résumé

Dans de nombreux modèles en finance ou en économie industrielle, un manager essaie de signaler l'information de son entreprise. On néglige alors généralement les problèmes de délégation à l'intérieur de l'entreprise. Cet article, au contraire, prend en compte ces problèmes de délégation et en analyse les conséquences. On trouve des conditions sous lesquelles les modèles de signalisation avec délégation conduise *qualitativement* à des prédictions similaires à celles des modèles sans délégation. Ces conditions, plus restrictives que les conditions standard des modèles de signalisation sans délégation permettent de relativiser la robustesse des résultats de la littérature. Heureusement, dans certains modèles classiques, ces conditions sont vérifiées. On montre aussi que, dans certaines circonstances, le principal a strictement intérêt à déléguer à un agent la tâche de signalisation plutôt que de l'accomplir lui-même. Il est même possible que le profit espéré du principal soit croissant avec le coût de délégation !

Mots Clés: signalisation, principal-agent, délégation, information incomplète.
Nomenclature JEL: 026.

1/ INTRODUCTION

An often considered situation in the finance and industrial organization literatures is that of a firm whose manager must signal information about the firm.[1] Although agency problems are widely believed to cause managers not to take profit-maximizing actions, these literatures typically ignore such problems.[2] This paper, in contrast, assumes that agency problems do exist; and it asks what economic significance they have in the context of signalling models: What are the circumstances in which they are significant? How are they significant? Do they always harm the principal; or can she sometimes do better by employing an agent than by managing the firm herself?

To answer these questions, we develop a conventional signalling model (*the principal game*) in Section 2. The manager of a firm (there, the principal) privately learns the characteristics of the firm (its "type"). She, then, takes an observable action that directly affects profits. This action conveys information to a third party about the firm's characteristics. The third party, in turn, responds to this information, which also affects the firm's profits. For example, the R & D decisions of a start-up's manager affect the profitability of the firm and affect a potential investor's beliefs about the firm's future profits. These beliefs, in turn, determine the cost of financing the start-up. As a second example, the manager of an

---

[1] See Harris and Raviv [1991] for a survey of signalling models in finance. Examples in industrial organization include Milgrom and Roberts [1982a, 1982b, and 1986] and Aghion and Bolton [1987]. Gertner *et al.* [1988] considered simultaneous signalling in both the capital and product market.

[2] Some financial signalling models (e.g., Ross [1977] or Glazer and Israel [1990]) allow for an agent-manager. Three important features, however, differentiate these models from ours: first, there are, essentially, no agency costs in those models, whereas agency costs play a central role here; second, the principal can signal only by employing an agent, whereas we allow for the possibility that she could signal herself; third, managerial compensation contracts are observable by outside parties, whereas we assume outside parties cannot observe these contracts.

incumbent monopoly takes actions (e.g., investments or first-period pricing) that affect profits and influence the beliefs of a potential entrant about the profitability of entry. These beliefs, in turn, influence his entry decision.

Under the "standard" assumptions -- in particular, monotonicity and single crossing of indifference curves for different types -- we obtain well-known results for signalling games (see, e.g., Cho and Sobel [1990]). First, better types cannot take smaller actions than worse types; and, second, there exists a unique stable equilibrium outcome in the sense of Kohlberg and Mertens [1986], namely the *least-cost separating equilibrium*. As is typical, this equilibrium can exhibit *upward distortions* in the principal's actions: To distinguish themselves from worse types, better types take greater actions than they would under full information.

In Section 3, we consider the agent-managed firm (*the agent game*). Now, it is an agent who privately learns the characteristic of the firm. We assume that he 1) bears some of the consequences of his actions (the cost of signalling); 2) is free to quit after learning the firm's type; and 3) is employed under a contract that is *not* observable to the third party. We assume the agent game satisfies a *screening condition*: Better-type agents find a given level of the signal less costly than worse-type agents; moreover, this cost is increasing faster in the level of the signal for worse-type agents than for better-type agents.

Given the screening condition, the agent game and the principal game are similar in that the size of the action (signal) is non-decreasing with type in all equilibria (Proposition 2). The games, however, differ in that the principal must bear certain agency costs in the agent game: To prevent better-type agents from mimicking worse-type agents, the principal must pay the better-type agents an *information rent*. This information rent, in turn, creates a motive for the principal to *distort downward* the actions different types of agent choose. This downward distortion offsets, to some degree, the upward distortion from signalling to the third party;

2

which benefits the principal. This benefit, in turn, can exceed the agency costs, so the principal can actually expect a greater profit when she employs an agent than when she manages the firm herself (Proposition 3). The general intuition is that the costs inherent in employing an agent can serve *indirectly* to commit the principal to actions to which she could not otherwise commit herself if she played for herself.[3]

This last result is closely related to Katz [1991], which also develops this general intuition. Katz, however, does so through a different class of games than those considered here. One difference is that Katz is *primarily* concerned with models in which the principal wishes to use the agency contract to force her agent to play a "tough" response to the third party's action. Here, instead, the agent moves before the third party. Another difference is that, unlike Katz, we focus on games exhibiting asymmetric information between the firm's manager and the third party.[4]

We explore this intuition further through a simple limit-pricing example in the spirit of Milgrom and Roberts [1982]. We show that there are potentially a broad range of parameter values such that an incumbent monopolist (the principal) would do better to hire a retailer (the agent) than to retail her output herself. We also show, perhaps surprisingly, that the principal can do better the greater are the costs of agency. The intuition is that the greater are the costs of agency, the greater is the number of actions to which the principal can, now, indirectly commit; and this gain in commitment can outweigh the greater cost of agency.

Under the screening condition, the agent game is similar to a game in which the

---

[3] Were contracts observable by the third party, the principal could use the contract to commit *directly*. That we are able to show the principal does better with an agent in some settings is, therefore, more striking because of our assumption of *private* contracts.

[4] In his general analysis of games in which agency does *not* matter, Katz does allow for *ex ante* asymmetric information. When he considers games in which agency does matter, however, there is no *ex ante* asymmetric information. For a further discussion of the relation between our paper and Katz's, see his Section 3E.

principal manages the firm herself, but in which her payoffs have been modified. We refer to this game as the *modified-principal game*. Provided this modification preserves single crossing -- a property we call *agent single-crossing* -- principal games and agent games are qualitatively similar in the following senses: 1) any outcome of a perfect Bayesian equilibrium (PBE) of the agent game is an outcome of a PBE of the modified-principal game and *vice versa* (Proposition 4); and 2) the stable equilibrium of the two games is the same, namely the least-cost separating equilibrium (Proposition 5). So, from this *qualitative* perspective, some signalling games are "robust" with respect to possible agency problems.

We emphasize that these "robustness" results are dependent on the screening condition and the agent single-crossing property. In general, these conditions are not implied by the usual assumptions for signalling games. That is, implicit in any assumption that agency problems can be ignored in signalling models are potentially strong assumptions about the information structure and the underlying technology. On the other hand, however, these conditions do follow from the usual assumptions in some well-studied cases (e.g., the assumptions of the original Spence [1973] model).

To underscore the potential *lack* of robustness, we consider a simple example of entry deterrence through investment in cost reduction. In this example, the screening condition fails, but the standard assumptions hold. The consequence is fairly dramatic: The principal and agent games yield exactly opposite predictions![5]

2/ THE PRINCIPAL-MANAGED FIRM

A principal owns a firm. Here, she is the manager; later, in Section 3, we consider the game in which she hires an agent to be the manager.

The game begins with the principal learning information relevant to the firm.

---

[5] A result that should, in some sense, have been expected from Guesnerie and Laffont [1984].

4

This is her private information. More generally, it is the private information of whomever manages the firm (principal or agent). The principal then chooses an action $x \in \mathbb{R}_+$ -- for example, output, price, quality of product, or an attribute of financial structure. We let $x = 0$ stand for the decision not to engage in any activity. Next, a third party observes this action; forms posterior beliefs, denoted by $\mu(x)$, about the principal's private information conditional on his observation of x; and, based on these beliefs, responds by choosing an action $a \in \mathbb{R}$.

The *gross* benefit to the principal of x and a is $B(x,a,n)$, where n -- the *type* -- is an index over the class of possible information. It is common knowledge that n belongs to $\{1, ..., N\}$ and prior beliefs are $\mathrm{Prob}\{\hat{n} = n\} = h_n$, with cumulative distribution $H_n = \mathrm{Prob}\{\hat{n} \le n\} = \sum_{j=1}^{n} h_j$. Let $h \equiv (h_1, h_2, ..., h_N)$ denote an element of the N-dimensional simplex $\mathscr{S}^N$. Choosing x costs the principal $\phi(x,n)$, where $\phi(0,n) = 0$ for all n (since $x = 0$ corresponds to no activity at all). Thus, the net benefit to the principal of x and a is $V(x,a,n) \equiv B(x,a,n) - \phi(x,n)$, which we assume to be continuous in (x,a).

The third party's utility is such that, conditional on his beliefs μ, his best response correspondence is a *continuous function* of (x,μ), which we denote by $A(x,\mu)$. We assume that $A(x,\mu) > A(x,\mu')$ if μ first-order stochastically dominates μ'. Loosely, the better the firm is believed to be, the greater is the response from the third party. As a notational convenience, let $A_n(x)$ be the third party's optimal response when his posterior belief is that the firm's type is *certainly* n. It follows that $A_n(x)$ is increasing in n for all x.

To be consistent with our interpretation that high values of n generate high responses from the third party, we concentrate on *monotonic signalling games* (see Cho and Sobel [1990]): The principal benefits from higher responses by the third party and, therefore, she has an incentive to signal high values of n. Formally, $B(x,a,n)$ is increasing in a for all (x,n).

We also impose a regularity assumption: For all (μ,n), $V\left(\cdot,A(\cdot,\mu),n\right)$ is strictly quasi-concave on $\mathbb{R}_+$ and such that for all n, $\lim\limits_{x\to+\infty} V(x,A_N(x),n) = -\infty$.   Define $x_n^*(\mu)$ to be the Stackelberg level of action for each exogenously given belief μ; i.e.,

$$x_n^*(\mu) = \underset{x \in \mathbb{R}_+}{\operatorname{Argmax}} \; V(x,A(x,\mu),n) \; .$$

Let $x_n^{FI} = x_n^*(\delta_n)$, where δ_n denotes the Dirac mass concentrated on type n, and the superscript "FI" stands for *full-information* Stackelberg optimal action.

We also assume, crucially, that the usual *single-crossing property* holds: A given indifference curve in action-response $(x-a)$ space for one type can cross a given indifference curve for another type at most once.   Formally:

SCP (Single-Crossing Property):   For all a, $\hat{a}$, x, $\hat{x}$, n and j, if $n > j$ and $\hat{x} < x$,

$$V(\hat{x},\hat{a},n) \geq V(x,a,n) \quad\Longrightarrow\quad V(\hat{x},\hat{a},j) > V(x,a,j) \; .$$

SCP implies that $x_n^*(\mu) > x_j^*(\mu)$ for all μ if $n > j$.

Our solution concept is the strong version of Perfect Bayesian Equilibrium (PBE) put forth by Fudenberg and Tirole [1991].[6]   As typical, there are many equilibria here, some separating equilibria that involve perfect signalling of the principal's private information, some pooling and (more generally) hybrid equilibria that involve the possibility of different types of principals choosing the same action with positive probability.   A well-known property of any PBE is that the principal's action is non-decreasing with her type n:

---

[6]   This rules out beliefs for the third party in which he hypothesizes that a deviation by the principal will be followed by a deviation by the agent.   Hence, an uninformed principal cannot "signal" what she does not know.   Were the space of mechanisms finite, this solution concept would be equivalent to sequential equilibrium (Kreps and Wilson [1982]); see Fudenberg and Tirole for the details.

Lemma 1: *Under* **SCP,** *in any PBE of the game, any action taken with positive probability by an n-type principal is not less than any action taken with positive probability by a j-type principal if* n > j.

The proof, which is by a revealed-preference argument, is standard and, so, omitted.

We will not provide a full description of the set of equilibria; instead we focus on only those that are *stable* in the sense of Kohlberg and Mertens [1986]. Given our assumptions, the only stable PBE is the so-called *least-cost separating equilibrium*, which is defined by (a) each type's choosing x_n^S, where x_n^S solves

$$\max_x \ V(x, A_n(x), n) \tag{1}$$

$$\text{such that } V(x, A_n(x), n-1) \leq V(x_{n-1}^S, A_{n-1}(x_{n-1}^S), n-1)$$

and (b) given action x, beliefs are that the principal is type n $(1 < n < N)$ if $x_n^S \leq x < x_{n+1}^S$, type 1 if $x < x_2^S$, and type N if $x \geq x_N^S$. Using the SCP, it is straightforward to prove that x_n^S is strictly increasing in n and can be sustained as a PBE under these beliefs.

In this equilibrium, each type is either taking her full-information Stackelberg action or an action *above* it, but as close to it as possible given the no-mimicking constraint. This possible *upward distortion* is typical of signalling models: To prove she is better than the types worse than she, an n-type principal may have to "work harder" than she would under full information.

Proposition 1: *Given our assumptions, the least-cost separating PBE exists and is the only stable PBE.*

The proof follows from Cho and Sobel [1990] (Propositions 4.4, 4.5, and connected discussion). To gain some intuition for Proposition 1, consider the model with only two types and a separating PBE *other than* the least-cost separating PBE. Since the equilibrium is separating, the low type must choose x_1^{FI}. Imagine a

deviation by the high type to $x_2^s + \varepsilon$, where ε is any arbitrarily small positive number. To sustain the candidate PBE, the third party must hold posterior beliefs upon seeing $x_2^s + \varepsilon$ that put sufficient weight on the principal's being the low type. But from (1), even if the third party thought the principal were the high type and responded accordingly, a low-type principal would still do better not to deviate. Thus, it seems unreasonable for the third party to put any weight on the low type upon witnessing this deviation; particularly since the high type *would* have incentive to deviate in this manner if she thought she would be perceived as high. If, consequently, we require the third party to hold reasonable beliefs upon witnessing $x_2^s + \varepsilon$ -- i.e., that the high type chose this action -- then the candidate PBE collapses. Similar reasoning shows that pooling or hybrid PBE are also supported by unreasonable out-of-equilibrium beliefs.

3/ THE AGENT-MANAGED FIRM

Now we assume that the principal employs an agent to manage the firm. The new timing is: The principal and agent engage in contract negotiations that fix a menu of contracts (a *mechanism*). Once employed, the agent learns the firm's type, n. Recall that n is the private information of the player who manages the firm -- in this case, the agent. Next, the agent chooses a compensation contract from the menu of contracts or he quits. Assuming he does not quit, he then chooses an action, x, observable to the third party. This action is the only verifiable variable on which a contract can be based (in particular, we assume (secretly) selling the firm to the agent is not possible).

The cost, or disutility, of choosing x is still $\phi(x,n)$ if the firm's type is n, but, now, it is borne by the agent. The agent's utility is his compensation minus

this cost. We normalize the agent's reservation utility to be zero.[7]

A crucial assumption is that *contract negotiations are private*. By this we mean that the third party can observe neither the mechanism offered to the agent, nor the communication stage between the principal and the agent. Caillaud and Hermalin [1989] and Katz [1991] have shown that the PBE of principal-agent contracting games in the presence of third parties are sensitive to how much of the contract the third parties can observe. It is our view that, in most settings, people cannot credibly commit to make their negotiations and the contract they sign public. Public negotiations could be quite costly to organize because outside observers cannot be sure that the parties to the negotiations have not conversed privately or entered into other secret agreements. Moreover, even though people may sometimes negotiate in public and release copies of their contracts to the public (e.g., when there are legal reporting requirements), they generally cannot commit not to renegotiate secretly.[8] It seems, therefore, important to consider situations where outside parties observe neither the contract negotiations, nor the ultimate agreement.[9]

By *the revelation principle*, we can restrict attention to direct truthful revelation mechanisms of the form $(\mathbf{x}, \mathbf{u}) = \left[(x_1, u_1), \ldots, (x_N, u_N) \right]$, where x_n is the

---

[7] Reservation utilities are type-*independent* since types refer here to characteristics of the firm. Replacing this assumption with type-*dependent* reservation utilities could alter our results, particularly if reservation utilities were increasing with type, as this would reduce the downward distortion in actions under the second-best mechanism. For a further discussion of these issues, see, e.g., Caillaud *et al.* [1990].

[8] However, a situation with secret renegotiation of a publicly disclosed contract is not *always* equivalent to a situation of an initially private contract (see Caillaud *et al.* [1990]).

[9] In this model, observing the ultimate agreement would be equivalent to observing the compensation actually paid to the agent (in addition to x). Caillaud and Hermalin [1989] considers a model in which the third party also observes the agent's (monetary) compensation: Similar results obtain, but the analysis is more complicated because it involves signalling over two dimensions.

action required and u_n is the agent's utility *provided he chooses* x_n.[10] In other words, the agent's *monetary* compensation if he chooses x_n is $u_n + \phi(x_n, n)$. Truthful revelation means that given n, the agent must do best by truthfully reporting his type as n: For all n and j,

$$u_n \geq u_j + \phi(x_j, j) - \phi(x_j, n).$$

As the agent can quit after learning n, equilibrium mechanisms must be *interim individually rational*: For all n, $u_n \geq 0$.[11] We call a mechanism that is interim individually rational and incentive compatible *feasible*.

An *action profile* is an N-tuple of action targets assigned to each type by a mechanism: $\mathbf{x} = (x_1, ..., x_N)$. $\mathbf{x}$ is *feasible* if there exist utilities $\mathbf{u} = (u_1, ..., u_N)$, such that the mechanism $(\mathbf{x}, \mathbf{u})$ is feasible.

In equilibrium, a mechanism offered with positive probability can be completely defined by its action profile -- the associated profile of utilities follows from the minimization of the principal's expected wage bill subject to the feasibility constraints. Any PBE can, thus, be stated in terms of a strategy (mixed or pure) over feasible action profiles; and a response function $a(\cdot) \equiv A(\cdot, \mu(\cdot))$: $\mathbb{R}_+ \to \mathbb{R}$ that characterizes the equilibrium reaction by the third party given the best-response function $A(\cdot)$ and Bayesian-consistent equilibrium beliefs $\mu(\cdot)$.

For $n \in \{1, 2, ..., N-1\}$, we define $R_n(x) \equiv \phi(x, n) - \phi(x, n+1)$. This is the

---

[10] Since the third party's response, a, is a function of x, there is no reason for the principal to make the agent's compensation also contingent on a. Note that the revelation principle applies because the third party only observes the action x; it might not apply if the third party could also observe what mechanism was offered because, in that case, the principal might wish to commit not to learn n (see Caillaud and Hermalin [1989]).

[11] Since the mechanism $(\mathbf{x}, \mathbf{u}) = (0, 0)$ is always possible and yields the agent his reservation utility, the principal is always (at least weakly) better off inducing even the worst-type agent not to quit. Thus, there is no loss of generality in restricting attention to mechanisms that are interim individually rational for all types.

information rent type *n+1* receives in return for not mimicking its neighboring type *n*. We assume:

SC (Screening Condition): $\forall n \in \{1, 2, \ldots, N-1\}$, $R_n(\cdot)$ is positive, strictly increasing, and convex on $\mathbb{R}_+$.

The screening condition states that it costs an *n*-type agent more to choose a given action than the better *(n+1)*-type agent; and this cost difference is increasing with the action. The screening condition is, for example, satisfied if cost is increasing and convex in the difference between output (the action) and some productivity parameter, where better types have greater productivity parameters. As we will, however, show it is possible to construct models in which the screening condition is not met. We are now in a position to prove (all proofs are in the appendix):

Proposition 2: *Under* **SC**, *an action profile,* $(x_1, \ldots, x_N)$, *is feasible if and only if* x_n *is non-decreasing in type n. Moreover, if* $(\hat{x}_1, \ldots, \hat{x}_N)$ *is an action profile played with positive probability in an equilibrium where the third party's response is* a($\cdot$) *and the prior is* h, *then*

$$(\hat{x}_1, \ldots, \hat{x}_N) \in \underset{x}{\text{Argmax}} \sum_{n=1}^{N} h_n \left(V\left(x_n, a(x_n), n\right) - \frac{1-H_n}{h_n} \cdot R_n(x_n) \right) \tag{2}$$

$$\textit{subject to } x_1 \leq \ldots \leq x_N.$$

Note, from (2), that the profit produced by an *n*-type agent is equal to the profit an *n*-type principal would earn if she managed the firm herself, $V\left(x_n, a(x_n), n\right)$, minus an amount, $\frac{1-H_n}{h_n} \cdot R_n(x_n)$, that corresponds to the increment in the information rent (agency cost) that must be paid to the better types to induce them to reveal their types truthfully. Because of this rent effect -- which is present for all but the best type -- the principal will desire *ceteris paribus* that each type of agent (except the best) take a *smaller* action than the principal would choose if she

11

managed the firm herself. That is, as is typical, the use of an agent creates a tendency toward a *downward distortion* in action.

4/ SIGNALLING CAN PROVIDE A MOTIVE TO EMPLOY AN AGENT

Signalling can create a tendency toward an upward distortion of the action. Employing an agent can counteract this tendency. One might, therefore, wonder if the benefit of counteracting the upward distortion due to signalling can exceed the cost of agency. As we will show, under certain circumstances, the principal can expect a greater profit by hiring an agent than by managing the firm herself.

In the principal game, the stable equilibrium is independent of the prior, provided it lies in the relative interior of $\mathcal{S}^N$. When the principal hires an agent, however, Proposition 2 clearly shows that the agency cost varies with the prior, and, thus, so does the set of stable equilibria.[12] Comparing the principal's expected profits in the two games is, therefore, difficult in general. We will first prove that, comparing stable equilibria, the principal is better off in the agent game than in the principal game if her prior is sufficiently concentrated on one type n. Admittedly, the use of stability for almost degenerate priors may be regarded by some as problematic, but our result simply shows that there is a set of possible priors for which hiring an agent is *ex ante* optimal for the principal. We, then, show that this set of priors can be relatively large through an example, which applies our analysis to a simple model of limit pricing.

The principal gains most from employing an agent when there is little uncertainty as to the agent's type, n. In this case, the principal can nearly

---

[12] It has been our experience from seminars and discussions, that a number of economists view the invariance of (many) signalling games' stable equilibria with respect to the priors to be an unappealing property of these games. We, therefore, note that this suggests an additional "bonus" from considering signalling models with agents: The stable equilibria of these games *will* vary with the priors!

eliminate the agent's information rent by approximately "shutting down" the types that are worse than n; i.e., by requiring very small actions from types $j < n$. Moreover, since the probability that the types better than n are paid an information rent is small, the downward distortion in n's action will be small. Hence, type n should produce profit close to the *full-information profit for type* n; that is, in the agent game, the principal's *ex ante* expected profits will be close to $V(x_n^{FI}, A_n(x_n^{FI}), n)$ for priors that are concentrated around type n. In contrast, in the principal game, *ex ante* expected profits are close to $V(x_n^{S}, A_n(x_n^{S}), n)$ for priors that are concentrated around n. *Ex ante*, the principal, thus, does better by hiring an agent if $x_n^{S} > x_n^{FB}$; i.e., if signalling implies *ex post* inefficiencies for type n in the principal game. This is the argument behind:

Proposition 3: *Assume* SCP *and* SC *are met. Consider any type* n *other than the worst type (i.e.,* n $\geq$ 2) *such that the least-cost separating action of type* n, x_n^{S}, *is greater than the full-information action of type* n, x_n^{FI}. *Then, there exists a subset of* $\mathcal{P}^N$ *containing* δ_n, *with non-empty relative interior,*[13] *such that, for any prior* h *in that relative interior, the principal's expected profit is greater in any stable PBE of the agent game than in the stable PBE of the principal game.*

Intuitively, when the principal manages the firm, a separating PBE can entail distortion at the top because the worse-type principals do not internalize the externality that their potential mimicry of better-type principals imposes on the better-type principals. Moreover, even though the principal recognizes this externality *ex ante*, there is nothing she can do about it because she cannot commit not to engage in opportunistic mimicry if she proves to be a low type. On the other

---

[13] If $\mathcal{D}$ is a subset of $\mathcal{P}^N$ and $\mathcal{O} \subset \mathcal{P}^N$ is an open set, then $\mathcal{D} \cap \mathcal{O}$ is a *relatively* open set in $\mathcal{D}$. The union of all relatively open sets in $\mathcal{D}$ is the *relative interior* of $\mathcal{D}$.

hand, when the agent manages the firm, the principal can *indirectly* commit not to behave so opportunistically. Having a low type mimic a high type is more expensive when the agent manages the firm because of the information rent. Consequently, the third party is willing to accept a smaller action as "proof" that a type is high. Moreover, as the prior probability that an agent is a given type increases, this "level of proof" decreases; which further reduces the signalling distortion. Indeed, for a sufficiently high probability, there is no distortion and the type produces the full-information Stackelberg optimal action.

Moreover, as a given type becomes more likely, the smaller become the action targets of types worse than the given type. So the information rent earned by the given type falls. In the limit, it disappears, so the principal's profit in the equilibrium of the agent game approaches the profit she would receive *if* she managed the firm. Combining the reduced-signalling-distortion effect and the reduced-rent effect, we see that the principal's expected profit is, therefore, approaching the full-information optimum profit as the prior becomes more concentrated.

To illustrate the general reasoning behind Proposition 3, consider a simple example of limit pricing and entry deterrence, which shows that the intuition given above generalizes beyond concentrated priors. There are two periods. In period 1, the market is served by an incumbent monopolist. Prior to period 2, an entrant chooses to enter ($a = 0$) or stay out ($a = 1$). Suppose that the incumbent can employ a retailer (the agent) for the first period or she can retail her output herself. Assume the market inverse demand curve is $p = d - x$ in the first period, where x is the quantity supplied to the market. Assume x is publicly observable and, moreover, that x is the only variable on which a retail contract can be based. Assume that the incumbent's *retail* cost, c^I, is the private information of the retailer (principal or agent as the case may be); it is a constant per unit cost; and it is drawn from $\{c_1, c_2\}$, where $d > c_1 > c_2$ and $\mathbb{P}\{c^I = c_1\} \equiv h$. For a c_n-type incumbent, let M_n

14

denote second-period monopoly profit and let D_n denote second-period duopoly profit. Define $g \equiv 2\sqrt{M_1 - D_1}$. The variable g is a convenient normalization of the gain enjoyed by a high-cost incumbent from successfully deterring entry. Assume $c_2 > c_1 - g$ (this ensures costly signalling in the principal game) and assume $d - c_2 \geq g$ (this simplifies matters in the agent game).

Assume that the entrant's fixed costs of entry are uncertain from the incumbent's perspective and are distributed continuously on $\mathbb{R}_+$; so the probability of entry given the entrant's beliefs, $1-a(\mu)$, is continuous and increasing in μ ($\mu = \text{Prob}\{c^I = c_1\}$). Assume there is no entry if the incumbent is believed to be low cost and there is certain entry if the incumbent is believed to be high cost; i.e., $a(0) = 1$ and $a(1) = 0$. In terms of our general model:

$$V(x,a,n) = x(d - x - c_n) + aM_n + (1-a)D_n.$$

Note it satisfies our assumptions, including SCP, if $(M_1-D_1)(d-c_2) > (M_2-D_2)(d-c_1)$.[14]

Consider the principal game. Simple monopoly pricing yields $x_1^S = \frac{1}{2}(d - c_1)$. From (1), $x_2^S = \frac{1}{2}(d - c_1 + g) > \frac{1}{2}(d - c_2) = x_2^{FI}$; so signalling is distortionary. Let $Z^P(h)$ be the incumbent's expected profit when she does her own retailing:

$$Z^P(h) = h\left[\frac{(d - c_1)^2}{4} + D_1\right] + (1-h)\left[\frac{(d - c_2)^2}{4} - \frac{(c_2 - c_1 + g)^2}{4} + M_2\right].$$

Now, consider the game in which the incumbent employs a retailer for the first period. Using (2), let $k \equiv c_1 + \frac{1-h}{h}\left(c_1 - c_2\right)$ be the *virtual* cost of employing the more costly retailer (i.e., the c_1-type retailer). Anticipating the results of Section 5, the agent game is equivalent to a principal game in which the high-cost principal's retail cost is k; in particular, the stable equilibrium is the same. So,

---

[14] In the appendix, we present a model from which these assumptions follow from first principles.

by simple monopoly pricing, $x_1^A = 0$ if $k \geq d$; and $x_1^A = \frac{1}{2}(d - k)$ if $k < d$. Note that $k \geq d$ if and only if

$$h \leq \frac{c_1 - c_2}{d - c_2} \equiv \underline{h} \ .$$

Define $m(h)$ to be the virtual profit produced by the high-cost retailer: $m(h) = 0$ if $h \leq \underline{h}$; and $m(h) = \frac{1}{4}(d - k)^2$ if $h > \underline{h}$. From (1), $x_2^A = \max(x_2^{FI}, \tilde{x})$, where $\tilde{x}$ is the larger root of

$$m(h) + D_1 = \left(d - k - \tilde{x}\right)\tilde{x} + M_1 \ .$$

Calculations reveal that there exists an $\bar{h} \equiv (c_1 - c_2)/g \geq \underline{h}$, such that $x_2^A = x_2^{FI}$ if $h \leq \bar{h}$; and such that $x_2^A = \frac{1}{2}(d - k + g) > x_2^{FI}$ if $h > \bar{h}$. Let $Z^A(h)$ be the incumbent's expected profit when she hires an agent to do her retailing:

$$Z^A(h) = \begin{cases} h[m(h) + D_1] + (1-h)\left[\dfrac{(d - c_2)^2}{4} + M_2\right] & , \text{ if } h \leq \bar{h} \\[3em] h\left[\dfrac{(d - k)^2}{4} + D_1\right] + (1-h)\left[\dfrac{(d - c_2)^2}{4} - \dfrac{(c_2 - k + g)^2}{4} + M_2\right] & , \text{ if } h > \bar{h} \ . \end{cases}$$

Before comparing the incumbent's expected profit in the two games, note that the cost of the signalling distortion (when present) is captured by the terms of the form $\frac{1}{4}(c_2 - \gamma + g)^2$ (where γ is c_1 or k depending on the game). These costs are increasing in g: The greater is the gain from concealing from the entrant the fact that the incumbent has a high retail cost, the greater is the temptation to have the high-cost type mimic the low-cost type; hence, the greater is the distortion necessary for the low-cost type to prove its type. Note that the cost of signalling is always greater in the principal game, which accords with our intuition that using an agent reduces these costs. Moreover, this cost increases faster in g in the

principal game than in the agent game. So, $g = d - c_2$ minimizes $Z^P(h) - Z^A(h)$. That is, the principal benefits most from employing an agent when the signalling distortion is greatest in the principal game.

To see how far we can get from concentrated priors, let $g = d - c_2$. Then $\bar{h} = \underline{h}$ and,

$$Z^P(h) - Z^A(h) = (2h - 1)\left[\frac{(d - c_1)^2}{4} - \frac{(d - k)^2}{4} \cdot 1_{(h > \underline{h})}\right] ,$$

where $1_{(h > \underline{h})}$ is an indicator variable. From this expression, it follows that, for all $h < \frac{1}{2}$, the principal's expected profit is greater when she employs an agent.

This example also allows a further illustration of the gains of hiring an agent.[15] Suppose that, when hiring an agent, the principal bears a fraction $(1-\beta)$ of the retail cost, and the agent a fraction β. This change only affects the virtual cost: $\hat{k} \equiv c_1 + \frac{1-h}{h}\beta(c_1 - c_2) < k$. That is, the cost of agency (the informational rent) is increasing in β. Continue to suppose $g = d - c_2$, and, in addition, suppose $c_1 < \frac{1}{2}d$. Then, following the same steps as before,

$$\underline{h}(\beta) = \bar{h}(\beta) = \frac{\beta(c_1 - c_2)}{g - (1-\beta)(c_1 - c_2)} < \frac{1}{2}$$

and, for any $h > \underline{h}(\beta)$, $\hat{k} < d$ and $x_2^A = \frac{1}{2}(d - \hat{k} + g) > x_2^{FI}$. As before, for $h < \frac{1}{2}$, the principal does better by employing an agent. Moreover, we have

$$\frac{\partial Z^A}{\partial \beta} = \frac{(1 - 2h)(1 - h)}{2h}(d - \hat{k})(c_1 - c_2) > 0$$

for all $h \in \left(\underline{h}(\beta), \frac{1}{2}\right]$. But this says that *reducing the informational rent to be paid to the agent can actually hurt the principal when it is optimal for her to hire an*

---

[15] We thank Hal Varian for suggesting the discussion that follows.

agent. The intuition is that hiring an agent is beneficial because the costs inherent to the agency relation allow the principal to commit to actions to which she could not otherwise commit (here, it is a high-cost incumbent committing not to produce the more efficient levels of output for the low-cost incumbent). When the costs of agency are reduced, the commitment value of agency falls; and, for certain parameter values, this loss in commitment value exceeds the savings from lower agency costs.

Continuing along these lines, return to the assumption that retail costs are borne fully by the agent. Now, however, let the agent be more efficient than the principal in retail distribution: If the principal's cost is c, then the agent's cost is βc, where $\beta \in (0,1)$. Continue to suppose $g = d - c_2$ and $c_1 < \frac{1}{2}d$, but suppose, now, that $c_2 = 0$. Eliminating improvement for the good type (i.e., setting $c_2 = 0$) keeps the analysis short because we can trivially avoid a second-order effect that does not, in any case, reverse our general conclusion. It is readily shown that there exists a non-empty interval $\left[\underline{h}(\beta), \frac{1}{2}\right]$, $\underline{h}(\beta) = \beta c_1/d < \frac{1}{2}$, such that the principal does better by employing an agent for all h in that interval, and

$$\frac{\partial Z^A}{\partial \beta} = \frac{(1 - 2h)c_1}{2h}\left(d - \frac{\beta c_1}{h}\right) > 0$$

for all h in that interval. The last inequality shows that *when the principal does better by employing an agent, the principal can prefer to hire a less-able agent over a more-able agent.* The intuition, again, is that by lowering the costs of agency, the commitment value of employing an agent is reduced; and the loss of commitment value can exceed the savings from lower agency costs.

5/ AN EQUIVALENCE RESULT BETWEEN THE AGENT GAME AND THE PRINCIPAL GAME

In this section, we further analyze the agent game, with an emphasis on comparing the agent game to the principal game. In light of Proposition 2, we may

define the following:

$$V^A(x,a,n) \equiv V(x,a,n) - \frac{1 - H_n}{h_n} \cdot R_n(x).$$

We also assume that the *Agent Single-crossing Property* is met:

ASCP: For all a, $\hat{a}$, x, $\hat{x}$, n and j, if $n > j$ and $\hat{x} < x$,

$$V^A(\hat{x},\hat{a},n) \geq V^A(x,a,n) \;\Rightarrow\; V^A(\hat{x},\hat{a},j) > V^A(x,a,j)$$

Note that, except if there are only two types, **ASCP** is *not* implied by **SCP**. Compared to **SCP**, **ASCP** requires that the agency cost is not great enough to reverse the principal game's ranking of marginal rates of substitution between actions and the third party's responses across the different types.

The notation $V^A(\cdot,\cdot,\cdot)$ reflects the closeness of the principal game and the agent game under **SC** and **ASCP**. Indeed, we will refer to a game in which the principal plays for herself, but where her profit function is $V^A(\cdot,\cdot,\cdot)$, as the *modified-principal game*. What we will show is that, given these two assumptions, the agent game is equivalent in two senses to the modified-principal game. First, an outcome is part of a PBE for the agent game if and only if it is an outcome of a PBE of the modified-principal game (Proposition 4). Second, stable outcomes are the same in the agent game and the modified-principal game (Proposition 5).

The first step in deriving these results is to show that an equilibrium action profile in the agent game is incentive compatible from the perspective of the modified-principal game.

Lemma 2: *Under* **SC** *and* **ASCP,** *an action profile,* $(\hat{x}_1, \ldots, \hat{x}_N)$, *played with positive probability in a PBE of the agent game in which the third party's equilibrium response is $a(\cdot)$ must be incentive compatible for the modified-principal game in the sense that, for all types and $\forall y \in \mathbb{R}_+$,*

$$V^A(\hat{x}_n,a(\hat{x}_n),n) \geq V^A(y,a(y),n).$$

In the modified-principal game, the set of possible deviations is greater than in the agent game because, when the principal plays for herself, she is not restricted to deviations that are feasible action profiles. What Lemma 2 shows is that any action profile played with positive probability in the agent game remains incentive compatible for the modified-principal game despite the greater set of possible deviations. We can, now, prove:

Proposition 4: *Assume* SC *and* ASCP *are met. Then, an outcome is part of a PBE in the agent game if and only if it is part of a PBE in the modified-principal game.*[16]

Next, we turn to stable PBE of the agent game. Define the least-cost separating equilibrium of the agent game as the analog of the least-cost separating equilibrium of the modified-principal game: (a) the principal implements with probability one the action profile, $\mathbf{x}^{AS} \equiv (x_1^{AS}, ..., x_N^{AS})$, where x_n^{AS} solves:

$$\max_{x} V^A(x, A_n(x), n)$$

such that $V^A(x, A_n(x), n-1) \leq V^A(x_{n-1}^{AS}, A_{n-1}(x_{n-1}^{AS}), n-1)$,

and (b) given action x, beliefs are such that $a(x) = A_n(x)$ if $x_n^{AS} \leq x < x_{n+1}^{AS}$ $(1 < n < N)$, $a(x) = A_1(x)$ if $x < x_2^{AS}$, and $a(x) = A_N(x)$ if $x \geq x_N^{AS}$. From Section 2, we know that $\mathbf{x}^{AS}$ and $a(\cdot)$ constitutes a (stable) PBE of the modified-principal game. Therefore, from Proposition 4, they constitute a PBE of the agent game.

More importantly, the least-cost separating equilibrium for the agent game is the only stable PBE of the agent game:

Proposition 5: *Assume* SC *and* ASCP *are met, then the only stable PBE of the agent game is the least-cost separating equilibrium for the agent game.*

Stability is extremely powerful in the agent game since it yields a unique

---

[16] To be precise, the only mixed-strategy PBE we consider are those in which the strategies have *finite supports*.

outcome. The proof relies on Proposition 6B of Kohlberg and Mertens [1986]; i.e., on the so-called *Never Weak Best Response* criterion. Other well-known refinements prove too weak to eliminate unwanted equilibria here.

Qualitatively, the stable equilibria of the principal game and the agent game are the same: both are least-cost separating. In a sense, Propositions 4 and 5 are robustness results because typically articles in industrial organization or finance are concerned only with such *qualitative* aspects of equilibria. Under **SC** and **ASCP**, characteristics that are true about the equilibria of a family of signalling games continue to be true about the equilibria of those games with an agent. More realistic models of the firm -- which recognize the existence of a principal-agent relation -- will, thus, preserve the qualitative nature of the aforementioned articles' results if they satisfy **SC** and **ASCP**.

In light of the last remark, one might wonder how sensitive *non*-separating equilibria might be to the introduction of an imperfect agent. Here, we will only compare the set of action levels that can be sustained in a (full) pooling equilibrium in the principal game and in the agent game. In the principal game, this set is $M_1 \cap ... \cap M_N$, where

$$M_n = \left\{ x \mid V(x,A(x,h),n) \geq \underset{y}{Max}\ V(y,A_1(y),n) \right\}.$$

In in the agent game, the set is $M_1^A \cap ... \cap M_N^A$, where

$$M_n^A = \left\{ x \mid V^A(x,A(x,h),n) \geq \underset{y}{Max}\ V^A(y,A_1(y),n) \right\}.$$

In both cases, the strategies should be clear. Under **SCP**, **SC**, and **ASCP**, these intersections are the intervals [min M_N, max M_1] and [min M_N^A, max M_1^A], respectively. From the definition of V^A, it follows that $M_N = M_N^A$ and max $M_1^A \leq$ max M_1. Hence, under **SCP**, **SC** and **ASCP**, the set of actions that can be supported in a pooling equilibrium of the agent game is a subset of the actions that can be supported in a pooling equilibrium of the principal game; indeed, the former set can be empty even though the latter set is not. The conclusion is, therefore, that a specific analysis of the

pooling equilibria in a signalling game might not be robust to the introduction of an imperfect agent. In a loose sense, using an agent in a signalling game induces less pooling and, as seen above, can allow for more efficient separation.

6/ SPECIAL CASES

Most of the above results rely on **SC** and **ASCP**. For the general problem considered here, these two properties are not implied by the "standard" signalling-game assumptions. In particular, **ASCP** can prove to be a strong assumption. For instance, it is stronger than just assuming the *Monotonic Hazard Rate Property* in addition to **SCP**, **SC**, and the convexity of $R_n(\cdot)$. In two important cases, however, the principal game implies important conditions about the agent game.

If there are only two types, then **ASCP** follows from **SCP** and **SC**, since, for all a, $\hat{a}$, and $x > \hat{x}$, $R_1(x) > R_1(\hat{x})$, and therefore:

$$V^A(\hat{x},\hat{a},2) \geq V^A(x,a,2) \Rightarrow V(\hat{x},\hat{a},2) \geq V(x,a,2) \Rightarrow V(\hat{x},\hat{a},1) > V(x,a,1)$$

$$\Rightarrow V(\hat{x},\hat{a},1) - R_1(\hat{x})(1-H_1)/h_1 > V(x,a,1) - R_1(x)(1-H_1)/h_1$$

Thus, references to **ASCP** can be dropped from the statements of Propositions 4 and 5 for the two-types case. This reinforces the robustness result obtained in the previous section, inasmuch as the applied signalling literature tends to focus on two-type models. On the other hand, it underlines the sensitivity of these models to the assumption of just two types.

The second special case arises when the gross benefit function, $B(x,a,n)$, is *independent* of n. This is an important case because in many models only the cost of signalling is dependent on the signaller's type (e.g., as in the original Spence [1973] model). In this case, **SCP** implies that for all n, a, $\hat{a}$, and $x > \hat{x}$,

$$B(\hat{x},\hat{a}) - \phi(\hat{x},n+1) \geq B(x,a) - \phi(x,n+1) \Rightarrow B(\hat{x},\hat{a}) - \phi(\hat{x},n) > B(x,a) - \phi(x,n),$$

which, in turn, implies

$$\phi(\hat{x},n+1) - \phi(x,n+1) > \phi(\hat{x},n) - \phi(x,n),$$

22

and, therefore, **SC**. This proves:

Proposition 6: *The results of Lemma 2 and Propositions 3-5 hold if the assumption that the gross benefit function,* B(x,a,n), *is independent of* n *replaces* **SC**.

Proposition 6 is of particular interest for our limit-pricing model example. If the retail costs are entirely borne by the agent, then the gross benefits for the principal only involve the value of sales and production costs; and, hence, are independent of type. In this example, then, **SCP** implies **SC**: The results of Proposition 3 (e.g., the value of hiring a retailer) depend only on **SCP** being met.

7/ WHAT IF THE SCREENING CONDITION IS NOT MET?

Consider the following model of entry deterrence through cost reduction. Prior to the market opening, an incumbent monopolist can make publicly observable R&D investment. Based on its observation of the incumbent's R&D investment, a potential rival decides whether to compete against the incumbent ($a = 0$) or stay out of the market ($a = 1$). Assume that any competition is Cournot competition.

Formally, suppose an incumbent monopolist has an initial probability of having low constant marginal cost, $\underline{c}$, equal to θ_n. If her constant marginal cost is not $\underline{c}$, it is $\bar{c} > \underline{c}$. She can increase the probability of having marginal cost $\underline{c}$ to $x\theta_n$ by expending $\frac{1}{2}(x+\theta_n)^2$ for $x \in (1,1/\theta_N)$ (the cost of $x = 1$ -- no investment -- is zero). Note that the greater is her initial probability of having low marginal cost, the greater is her marginal expenditure for increasing this probability. Thus, **SC** does not hold.

The realization of the incumbent's marginal cost occurs after the rival's decision to compete. Once realized, the incumbent's marginal cost is common knowledge. The potential rival's constant marginal cost, c^E, is also common knowledge. Assume the inverse market demand is linear, with a large enough intercept

23

and that the difference in cost, $\bar{c} - \underline{c}$, is not too small (formally, if $p = d - sq$, then we require that $d > \bar{c} + \underline{c} - c^E + \frac{9}{4}s(\bar{c} - \underline{c})^{-1}$). It is, then, straightforward to show that, in x-a space, the principal's marginal rate of substitution at any point is decreasing in θ. So SCP is satisfied.

In this model, if the principal played for herself, then the stable PBE would be the least-cost separating equilibrium and $x_1^S < ... < x_N^S$ (from Proposition 1). If, however, she employed an agent, then any feasible action profile, $(x_1, ..., x_N)$, must satisfy $x_1 \geq x_2 \geq ... \geq x_N$ because SC does not hold. When SC fails, the principal game and the agent game can lead to radically different predictions (e.g., as here, where the principal game predicts that the investment in cost reduction is strictly increasing with type, while the agent game predicts that it is non-increasing with type). As we noted in Section 5, the parameterization of private information is crucial: The signalling and the agency problems need to be compatible, in the sense that SCP and SC both hold, if signalling games are to be robust with respect to possible agency problems.

8/ CONCLUSION

This paper analyzed what happens when an agency problem is introduced into a standard (monotonic) signalling model. In particular, we asked and answered three questions: first, how does the game with an agent -- the agent game -- differ from the game in which the principal manages the firm herself -- the principal game? Second, how significant are any differences? Third, can the principal achieve a better outcome by using an agent compared to when she manages the firm herself?

The answer to the first question is that the two games will always differ, if only because the presence of agency costs in the agent game results in a downward distortion in the level of the signal (the action). An even greater difference can exist if the screening condition is not met in the agent game: Whereas the level of

the signal is strictly increasing with type in the principal game, the level of the signal will be non-increasing with type in the agent game.

If, however, the screening condition is met and another condition, the agent single-crossing property, is met, then the two games are qualitatively the same: equilibrium properties of the principal game hold true (at least if they are generic) in the agent game (at least generically); moreover, the stable equilibria are the same -- the least-cost separating equilibria are the only stable equilibria for the two games. For games meeting these conditions, there may be little economic difference between signalling games with agents and without agents.

The answer to the third question is yes: In cases in which there is little uncertainty about the firm's type or in which the distortions from signalling are large, the principal can do better by employing the agent to manage her firm than by managing it herself. This is because the downward distortion in the level of the signal due to agency costs tends to cancel beneficially the upward distortion in the level of the signal due to a signalling motive; and, in these cases, this benefit exceeds the cost of employing an agent. Indeed, because this benefit can be increasing in the cost of agency, the principal's expected profit can even be *increasing* in the cost of agency.

There are, admittedly, questions left unanswered. For instance, what are the effects of agency in *non*-monotonic signalling games? We speculate that for agency not to have a qualitative impact on these games, the signalling and agency problems must be "aligned" in terms of the information structure and underlying technology; but it is difficult to characterize in general what aligned means in these games. We also expect that the principal can continue to do better with an agent in some situations because of the strength of the general argument that the costs inherent in agency provide commitment benefits that exceed these costs.

REFERENCES

Aghion, P. and Bolton, P. [1987]: "Contracts as a Barrier to Entry," *American Economic Review*, 77, pp. 388-401.

Caillaud, B. and Hermalin, B. [1989]: "The Role of Outside Considerations in the Design of Compensation Schemes," Working Paper No. 89-121, University of California.

Caillaud, B., Jullien, B., and Picard, P. [1990]: "On Precommitment Effects between Competing Agencies," *Document de travail* N° 9033, CEPREMAP, Paris.

Cho, I.-K., and Sobel, J. [1990]: "Strategic Stability and Uniqueness in Signaling Games," *Journal of Economic Theory*, 50, pp. 381-413.

Fudenberg, D. and Tirole, J. [1991]: "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, pp. 236-260.

Glazer, J. and Israel, R. [1990]: "Managerial Incentives and Financial Signaling in Product Market Competition," *International Journal of Industrial Organization*, 8, pp. 271-280.

Gertner, R., Gibbons, R., and Scharfstein, D. [1988]: "Simultaneous Signalling to the Capital and Product Markets," *The RAND Journal of Economics*, 19, pp. 173-190.

Guesnerie, R. and Laffont, J.-J. [1984]: "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *Journal of Public Economics*, 25, pp. 329-369.

Harris, M. and Raviv, A. [1991]: "Financial Contracting Theory," unpublished manuscript, University of Chicago.

Katz, M. [1991]: "Game-Playing Agents: Unobservable Contracts as Precommitments," *The RAND Journal of Economics*, forthcoming.

Kohlberg, E. and Mertens, J.-F. [1986]: "On the Strategic Stability of Equilibria," *Econometrica*, 54, pp. 1003-1037.

Kreps, D. and Wilson, R. [1982]: "Sequential Equilibria," *Econometrica*, 50, pp. 863-894.

Milgrom, P. and Roberts, J. [1982a]: "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, pp. 280-312.

Milgrom, P. and Roberts, J. [1982b]: "Limit Pricing and Entry Under Incomplete Information: An Equilibrium Analysis," *Econometrica*, 50, pp. 443-460.

Milgrom, P. and Roberts, J. [1986]: "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 94, pp. 796-821.

Ross, S. [1977]: "The Determination of Financial Structure: The Incentive Signalling Approach," *The Bell Journal of Economics*, 8, pp. 23-40.

Spence, A.M. [1973]: "Job Market Signalling," *The Quarterly Journal of Economics*, 87, pp. 355-374.

APPENDIX

Proof of Proposition 2

We first show that an action profile $(x_1, \ldots, x_N)$ is feasible only if it is non-decreasing. Consider $n > j$. From incentive compatibility:

$$\phi(x_n, j) - \phi(x_n, n) \geq u_n - u_j \geq \phi(x_j, j) - \phi(x_j, n) \tag{A.1}$$

or

$$\sum_{k=j}^{n-1} R_k(x_n) \geq \sum_{k=j}^{n-1} R_k(x_j). \tag{A.2}$$

Since $R_k(\cdot)$ is an increasing function for all k, (A.2) can be satisfied if and only if $x_n \geq x_j$. This and (A.1) also imply that u_n is non-decreasing for any feasible mechanism.

Define the *adjacent incentive* compatibility constraints to be

$$\forall n, \quad u_n \geq u_{n+1} - R_n(x_{n+1})$$

$$u_{n+1} \geq u_n + R_n(x_n).$$

Next, we show that if the *adjacent* incentive compatibility constraints hold and the action profile is non-decreasing, then all the incentive compatibility constraints hold. Consider any $j < n - 1$;

$$u_n - u_j = \sum_{k=j}^{n-1}\left(u_{k+1} - u_k\right) \geq \sum_{k=j}^{n-1} R_k(x_k) \geq \sum_{k=j}^{n-1} R_k(x_j) = \phi(x_j, j) - \phi(x_j, n)$$

and

$$u_n - u_j = \sum_{k=j}^{n-1}\left(u_{k+1} - u_k\right) \leq \sum_{k=j}^{n-1} R_k(x_{k+1}) \leq \sum_{k=j}^{n-1} R_k(x_n) = \phi(x_n, j) - \phi(x_n, n)$$

where the rightmost inequalities in the above expressions follow because x_k is non-decreasing and $R_k(\cdot)$ is an increasing function. The argument is similar for $j > n + 1$. Expression (A.1) follows.

Now, we show that any non-decreasing action profile is feasible. Let $u_1 = \phi(x_1, 1)$ and, for $n > 1$, $u_n = u_{n-1} + \phi(x_{n-1}, n-1) - \phi(x_{n-1}, n)$. Since the profile is non-decreasing, $R_{n-1}(x_{n-1}) \leq R_{n-1}(x_n)$, so the adjacent incentive compatibility constraints are met for all n. But from above, all the incentive compatibility

constraints are met. Hence the profile is feasible.

Finally, we prove the last part of the proposition. Given $a(\cdot)$, if $x \equiv (x_1, ..., x_N)$ is a feasible action profile played with positive probability in equilibrium, then there exists $u \equiv (u_1, ..., u_N)$ such that

$$(x,u) \in \underset{(\hat{x},\hat{u})}{\text{Argmax}} \sum_{n=1}^{N} h_n \left[V(\hat{x}_n, a(\hat{x}_n), n) - \hat{u}_n \right] \tag{P}$$

$$\text{subject to } \hat{u}_n \geq 0, \; \forall n$$

$$\hat{u}_n \geq \hat{u}_j + \phi(\hat{x}_j, j) - \phi(\hat{x}_j, n), \; \forall n, j.$$

We are done, if we can show that the u that solves (P) given a feasible action profile x is such that

$$\sum_{n=1}^{N} h_n u_n = \sum_{n=1}^{N} (1 - H_n) R_n(x_n). \tag{A.3}$$

Equation (A.3) will, in turn, hold if we can show that u is such that:

$$u_1 = 0 \text{ and, for } n > 1, \; u_n = \sum_{j=1}^{n-1} R_j(x_j), \tag{A.4}$$

since, then,

$$\sum_{n=1}^{N} h_n u_n = \sum_{n=2}^{N} h_n u_n = \sum_{n=2}^{N} h_n \sum_{j=1}^{n-1} R_j(x_j) = \sum_{j=1}^{N-1} R_j(x_j) \sum_{n=j+1}^{N} h_n = \sum_{n=1}^{N-1} (1 - H_n) R_n(x_n)$$

To prove (A.4), let λ_n be the Lagrange multiplier on the nth individual rationality constraint; i.e., on $\hat{u}_n \geq 0$. From above, the set of all incentive compatibility constraints can be replaced by the adjacent incentive compatibility constraints. Let α_n be the Lagrange multiplier on the upward adjacent incentive compatibility constraint; i.e., on

$$\hat{u}_n \geq \hat{u}_{n+1} - R_n(\hat{x}_{n+1}).$$

Let β_n be the Lagrange multiplier on the downward adjacent incentive compatibility constraint; i.e., on

$$\hat{u}_{n+1} \geq \hat{u}_n + R_n(\hat{x}_n).$$

The first-order conditions with respect to $\hat{u}_n$ are

$$-h_n + \lambda_n + \alpha_n - \alpha_{n-1} - \beta_n + \beta_{n-1} = 0, \quad 2 \le n \le N - 1;$$

$$-h_1 + \lambda_1 + \alpha_1 - \beta_1 = 0;$$

$$-h_N + \lambda_N - \alpha_{N-1} + \beta_{N+1} = 0.$$

Summing the first order conditions yields

$$\sum_{n=1}^{N} \lambda_n = 1.$$

Hence $\lambda_n > 0$ for at least one n. From above, u_n is non-decreasing in n, so there exists n_0 such that $\forall n > n_0$, $u_n > 0$ and $\lambda_n = 0$, and $\forall n \le n_0$, $u_n = 0$ (and $\lambda_n \ge 0$). Moreover $n_0 \ge 1$ since all λ_n's sum to one, so that $u_1 = 0$. Consider $n \le n_0$; we have

$$0 = u_n - u_{n-1} \ge R_{n-1}(x_{n-1}) \ge 0.$$

Hence, $R_{n-1}(x_{n-1}) = 0$. Thus, for all $n \le n_0$,

$$\sum_{j=1}^{n-1} R_j(x_j) = 0.$$

So (A.4) holds for all $n \le n_0$. Suppose $n_0 < N$. Define $\gamma_n \equiv \beta_n - \alpha_n$. Clearly, $\gamma_{N-1} = h_N$ and $\forall n > n_0$,

$$\gamma_{n-1} = \gamma_n + h_n = 1 - H_{n-1} > 0.$$

Hence, $\beta_{n-1} > 0$, $\forall n > n_0$, and, thus,

$$u_n = u_{n-1} + R_{n-1}(x_{n-1}) = u_{n_0} + \sum_{j=n_0}^{n-1} R_j(x_j) = \sum_{j=1}^{n-1} R_j(x_j).$$

So (A.4) holds for all $n > n_0$.

■

Proof of Proposition 3

Consider any type n other than the worst type (i.e., $n > 1$); for any $\Omega \subseteq \mathscr{S}^N$, let $\overset{\circ}{\Omega}$ denote the relative interior of Ω. Define recursively:

$$\bar{x}_1 = \text{Max} \left\{ y \in \mathbb{R}_+ \mid V\left(y, A_N(y), 1\right) \ge V\left(x_1^{FI}, A_1(x_1^{FI}), 1\right) \right\}$$

and

$$\bar{x}_j = \text{Max} \left\{ y \in \mathbb{R}_+ \mid V\left(y, A_N(y), j\right) \ge V\left(\bar{x}_{j-1}, A_1(\bar{x}_{j-1}), j\right) \right\} \text{ for } j > 1 .$$

Given our assumptions, all $\bar{x}_j$ are finite. Moreover, $\bar{x}_j$ is increasing in j. First, we prove that the principal can restrict her choice to a compact set of action profiles, $\mathcal{X} \equiv \prod_{j=1}^{N} [0, \bar{x}_j]$:

Claim 1: $\forall h \in \mathcal{S}^N$, any feasible profile x, such that $\exists j$ $x_j > \bar{x}_j$, is strongly dominated for all the third party's sequentially rational responses.

Proof: By induction. Suppose $x_1 > \bar{x}_1$. The principal would do better proposing the same mechanism except with x_1^{FI} instead of x_1. This follows from the definition of $\bar{x}_1$ and the fact that $a(x_1) \leq A_N(x_1)$ and $a(x_1^{FI}) \geq A_1(x_1^{FI})$, where $a(\cdot)$ is any sequentially rational response by the third party.

Suppose $x_j \leq \bar{x}_j$ $\forall j \leq p-1$, but $x_p > \bar{x}_p$. Define the profile x':

$$x_j' = x_j \text{ if } j \neq p \text{ and } x_p' = \bar{x}_{p-1} .$$

Since x' only involves lowering x_p and $x_{p-1} \leq \bar{x}_{p-1}$, x' is feasible and $R_p(x_p') < R_p(x_p)$. Moreover, whatever the response $a(\cdot)$,

$$V\left(x_p', a(x_p'), p\right) \geq V\left(x_p', A_1(x_p'), p\right) = V\left(\bar{x}_{p-1}, A_1(\bar{x}_{p-1}), p\right) > V\left(x_p, A_N(x_p), p\right) \geq V\left(x_p, a(x_p), p\right).$$

Thus, x' strictly dominates the profile x.

□

Next, we prove that when h is close to δ_n any profile x that induces types j smaller than n to take actions significantly different from zero, is strongly dominated for any sequentially rational response function for the third party. Consequently, for h close to δ_n, all PBE of the agent game will have types $j < i$ nearly shutdown.

Claim 2: $\forall \varepsilon > 0$, $\exists \eta > 0$ such that if $h \in \mathcal{S}^N$ and $\|h - \delta_n\| < \eta$, then any feasible profile x such that $\exists j$, $j < n$, $x_j \geq \varepsilon$, is strongly dominated for any sequentially rational response functions for the third party.[1]

---

[1] $\| \cdot \|$ denotes the Euclidean norm.

Proof: Suppose the claim were false. Then necessarily, there exists an $L > 0$ such that $\forall \eta > 0$, $\exists h^\eta \in \mathcal{P}^N$ and $\exists x^\eta \in \mathcal{X}$ such that

$$\|h^\eta - \delta_n\| < \eta \quad \text{and} \quad \sum_{j=1}^{n-1} R_j(x_j^\eta) > L > 0, \tag{A.5}$$

and x^η is not strongly dominated. Facing any sequentially rational response, $a(\cdot)$, and given prior h, the principal's expected profit from profile x, is:

$$Z(x,h,a) \equiv \sum_{j=1}^{N} \left(h_j V(x_j, a(x_j), j) - (1-H_j) R_j(x_j) \right).$$

Define for the same prior and the same response:

$$\hat{Z}(x,h,a) \equiv \sum_{j=n}^{N} \left(h_j V(x_j, a(x_j), j) - (1-H_j) R_j(x_j) \right) - \sum_{j=1}^{n-1} R_j(x_j).$$

By (uniform) continuity on $\mathcal{X}$, $\exists \eta_0 > 0$ such that $\forall h \in \mathcal{P}^N$, if $\|h - \delta_n\| < \eta_0$, then for $\forall x \in \mathcal{X}$,

$$\left| \hat{Z}(x,h,a) - Z(x,h,a) \right| \leq L/3 \tag{A.6}$$

In particular, consider h^0 and x^0 satisfying (A.5) for $\eta = \eta_0$, and define $\hat{x}^0$ as follows:

$$\hat{x}_j^0 = x_j^0 \text{ if } j \geq n; \quad \text{and} \quad \hat{x}_j^0 = 0 \text{ if } j < n.$$

(A.6) implies that, for any $a(\cdot)$,

$$Z(\hat{x}^0, h^0, a) \quad > \hat{Z}(\hat{x}^0, h^0, a) - L/3$$

$$= \hat{Z}(x^0, h^0, a) + 2L/3$$

$$\geq Z(x^0, h^0, a) + L/3,$$

and, therefore, for any $a(\cdot)$, $Z(\hat{x}^0, h^0, a) > Z(x^0, h^0, a)$: x^0 is strongly dominated by $\hat{x}^0$, a contradiction.

□

Next, we will concentrate on *stable* PBE and prove that, in any stable PBE, the principal's expected profit is close to, or above, the full information profit corresponding to type n.

Claim 3: $\forall \varepsilon > 0$, $\exists \xi > 0$ such that if $\|h - \delta_n\| < \xi$ and $h \in \mathcal{P}^N$, then for any stable PBE characterized by $a^h(\cdot)$ and any profile x played on the equilibrium path,

$$Z(x,h,a^h(.)) \geq V(x_n^{FI}, A_n(x_n^{FI}), n) - \varepsilon$$

Proof: Recall that $x_n^{FI} > 0$; from Claim 2, then, $\exists \eta_0 > 0$, such that if $h \in \mathcal{P}^N$ and $\|h - \delta_n\| < \eta_0$, then any profile x such that $x_j \geq x_n^{FI}$, for some $j < n$, is an *inferior* response for any component of the PBE; hence, $x_j < x_n^{FI}$ for all $j < n$.

Suppose the claim were false. There, then, exists an $L > 0$ such that $\forall \eta > 0$, there exists an $h^\eta \in \mathcal{P}^N$ satisfying

$$\|h^\eta - \delta_n\| < \eta ; \tag{A.7}$$

and an associated stable PBE characterized by $a^\eta(\cdot)$, in which a profile $x^\eta \in \mathcal{X}$ is played with positive probability satisfying

$$Z(x^\eta, h^\eta, a^\eta) < V(x_n^{FI}, A_n(x_n^{FI}), n) - L. \tag{A.8}$$

So, by (uniform) continuity of Z on $\mathcal{X}$, $\exists \xi \in (0, \eta_0)$ such that, $\forall h \in \mathcal{P}^N$ and for any PBE characterized by $a^h(\cdot)$ and a profile $x \in \mathcal{X}$, $\|h - \delta_n\| < \xi$ implies

$$\left| Z(x,h,a^h) - V(x_n, a^h(x_n), n) \right| \leq L/2$$

In particular, for h^ξ, $a^\xi(.)$, and any x^ξ satisfying (A.7) and (A.8),

$$Z(x^\xi, h^\xi, a^\xi) \geq V\left(x_n^\xi, a^\xi(x_n^\xi), n \right) - L/2.$$

Now, suppose $a^\xi(x_n^{FI}) \geq A_n(x_n^{FI})$ and consider the feasible deviation $\hat{x}$ such that $\hat{x}_j = x_n^{FI}$ for all $j > n$:

$$Z(\hat{x}, h^\xi, a^\xi) \geq V(x_n^{FI}, a^\xi(x_n^{FI}), n) - L/2 \geq V(x_n^{FI}, A_n(x_n^{FI}), n) - L/2$$

$$> Z(x^\xi, h^\xi, a^\xi) + L/2$$

and, therefore, $Z(\hat{x}, h^\xi, a^\xi) > Z((x^\xi, h^\xi, a^\xi))$: $\hat{x}$ would be a profitable deviation, which is a contradiction. Therefore, $a^\xi(x_n^{FI}) < A_n(x_n^{FI})$.

Consider the component of equilibria that yields the proposed outcome such that $a^\xi(x_n^{FI}) < A_n(x_n^{FI})$. From the first paragraph of this proof, only mechanisms that have

$j \geq n$ choose x_n^{FI} can be weak best responses (for the principal); which is inconsistent with $a^\xi(x_n^{FI}) < A_n(x_n^{FI})$. So, after deleting never-weak-best-response action profiles, the component of equilibria does not contain any equilibria: It is, thus, not stable.

□

The rest of the proof is straightforward: The principal's expected profit merely approaches $V(x_n^S, A_n(x_n^S), n)$ in the principal game, but, by assumption, this is strictly smaller than $V(x_n^{FI}, A_n(x_n^{FI}), n)$.

■

A Limit-Pricing Example

Consider the following primitive assumptions for the limit-pricing example: The entrant's *marginal* cost equals c_1 and $d/2 > c_1 > 0 = c_2$. If there is entry, then second-period competition is Cournot. Finally, second-period inverse demand is

$$p = c_1 + 3d/\sqrt{5} - x .$$

It is straightforward to show that all the assumptions made in the course of the limiting-pricing example in the text are satisfied by this (admittedly simple) model.

Starting from the same primitive assumptions, but assuming the inverse demand is $p = d - x$ in *both* periods, one can also show that the principal does better by employing an agent even if the priors are non-degenerate. For example, if $d = 40$, $c_1 = 4$, and $h = .3$, then the principal would do better to hire an agent than to manage the firm herself. More generally, if $h^*(d, c_1)$ is the greatest h such that the principal does better with an agent, then, with time-*invariant* demand,

$$\lim_{d \to \infty} h^*(d, c_1) = \sqrt{5}\left(3 + \sqrt{5}\right)^{-1} \approx .427 .$$

This result is obtained by considering the difference $Z^P(h) - Z^A(h)$ and noting that $g = \sqrt{\frac{5}{9}} \times (d - c_1)$ for time-invariant demand. So, for Cournot competition and time-invariant demand, the results would be only slightly less dramatic than those in

the text.

Proof of Lemma 2

Take any n and y such that $\hat{x}_{n-1} \le y \le \hat{x}_{n+1}$. [To avoid endpoint problems, define $\hat{x}_t = 0$, $t < 1$, and $\hat{x}_t = \infty$, $t > N$.] Since the action profile $(\hat{x}_1, ..., \hat{x}_{n-1}, y, \hat{x}_{n+1}, ..., \hat{x}_N)$ is feasible, it must be that $V^A(y,a(y),n) \le V^A(\hat{x}_n,a(\hat{x}_n),n)$.

The rest of the proof follows by induction: We seek to show that if, $\forall n$, $\forall y \in [\hat{x}_{n-t},\hat{x}_{n+t}]$, $V^A(y,a(y),n) \le V^A(\hat{x}_n,a(\hat{x}_n),n)$, then, $\forall n$, $\forall y \in [\hat{x}_{n-t-1},\hat{x}_{n+t+1}]$, $V^A(y,a(y),n) \le V^A(\hat{x}_n,a(\hat{x}_n),n)$. Suppose this were false. In particular, suppose for some n and $y \in (\hat{x}_{n+t},\hat{x}_{n+t+1}]$ that

$$V^A(y,a(y),n) > V^A(\hat{x}_n,a(\hat{x}_n),n) \qquad (A.9)$$

then

$$V^A(y,a(y),n) > V^A(\hat{x}_{n+t},a(\hat{x}_{n+t}),n).$$

Since $y > \hat{x}_{n+t}$, it then follows, from **ASCP**, that

$$V^A(y,a(y),n+t) > V^A(\hat{x}_{n+t},a(\hat{x}_{n+t}),n+t).$$

But this violates the first step of the proof. Now, suppose (A.9) is true for some $y \in [\hat{x}_{n-t-1},\hat{x}_{n-t})$; then

$$V^A(y,a(y),n) > V^A(\hat{x}_{n-t},a(\hat{x}_{n-t}),n).$$

Since $y < \hat{x}_{n-t}$, it then follows, from **ASCP**, that

$$V^A(y,a(y),n-t) > V^A(\hat{x}_{n-t},a(\hat{x}_{n-t}),n-t).$$

But this, again, violates the first step of the proof. Induction follows.

∎

Proof of Proposition 4

Only If Part. Consider the T action profiles, $[x^1, ..., x^T]$, played with positive probability in a PBE of the agent game, where x^t is the action profile $(x_1^t, ..., x_N^t)$. Let $\sigma \equiv (\sigma_1, ..., \sigma_T)$ be the principal's equilibrium mixed strategy weights. Let $a(\cdot)$ be the third party's equilibrium response function, where $a(\cdot)$ is

A-8

sequentially rational.

Now, consider the modified-principal game. Let $\Pi \equiv [\pi^1, \ldots, \pi^N]$ denote the matrix of strategies, where π^n is the n-type principal's strategy,

$$\pi^n \equiv \left(x_n^1, \ldots, x_n^T; \sigma_1, \ldots, \sigma_T \right) ;$$

i.e., she chooses x_n^t with probability σ_t.[2] Consider an x played with positive probability: In both the agent game and the modified-principal game the Bayesian posterior, $\text{Prob}\{n = j | x\}$, is the same. Thus, $a(\cdot)$, which was sequentially rational in the agent game, must also be sequentially rational in the modified-principal game. Finally, given $a(\cdot)$, any $x_n^t \in \pi^n$ is also an element of $\text{Argmax}_x V^A(x,a(x),n)$ from Lemma 2. Thus, no type of principal has an incentive to deviate from her strategy π. Therefore $\left[\Pi, a(\cdot) \right]$ is a PBE of the modified-principal game.

If Part. Let $[\pi^1, \ldots, \pi^N]$ and $a(\cdot)$ be a PBE of the modified-principal game, where $\pi^n = \left(x_n^1, \ldots, x_n^{T(n)}; \rho_n^1, \ldots, \rho_n^{T(n)} \right)$, where ρ_n^m is the probability the principal chooses x_n^m. For any two types, n and j, $n > j$, we have, from Lemma 1, that $x_n^t \geq x_j^m$, for any two actions taken with positive probability by the two types. Thus, any action profile formed by taking an element from the support of each type's strategy is feasible. There are, in fact, $T = \prod_{n=1}^{N} T(n)$ possible action profiles that can be formed in this way. Assign the probability σ_t to the tth such profile, where, if the tth profile is $\left(x_1^{m(1)}, \ldots, x_N^{m(N)} \right)$, then $\sigma_t = \rho_1^{m(1)} \times \ldots \times \rho_N^{m(N)}$. Since $x_n^m \in \text{Argmax}_x V^A(x,a(x),n)$, it follows that, given $a(\cdot)$, the principal (a) likes all T action profiles equally; and (b) has no incentive to deviate by implementing any other profile. Moreover, for any x played with positive probability, $\text{Prob}\{n = j | x\}$ is the same in both the modified-principal game and the agent game in which the principal plays the strategy $(\sigma_1, \ldots, \sigma_T)$. Thus, $a(\cdot)$, which was sequentially

---

[2] For notational convenience -- and without loss of generality -- we make no note of the possibility that x_n^t may equal x_n^m for some n, m, and t.

rational in the modified-principal game, must also be sequentially rational in the agent game.

■

Proof of Proposition 5

Consider any PBE of the agent game. Define $V_n^* = V^A(x_n, a(x_n), n)$, where the nth type agent is induced to choose x_n with positive probability. From Lemma 2, the definition of V_n^* is unambiguous. We also have from Lemma 2 that if $n > j$ and x_n and x_j are actions that the two types are, respectively, induced to choose with positive probability, then $x_n \geq x_j$. Define the *pseudo-indifference* curve ρ_n^*:

$$\rho_n^* \equiv \{(y,a) \mid V^A(y,a,n) = V_n^*\}.$$

Let us consider a PBE in which *all* action profiles x implemented with positive probability are "least cost" for all $j > n_0$, n_0 fixed and $N \geq n_0$, in the sense that $a(x_j) = A_j(x_j)$ and

$$V_j^* = \max_x V^A(x, A_j(x), j) \tag{A.10}$$
$$\text{such that } V^A(x, A_j(x), j-1) \leq V_{j-1}^*,$$

but where $\hat{x}$ is *not* least cost for n_0; i.e.,

$$V_{n_0}^* < \max_x V^A(x, A_{n_0}(x), n_0) \tag{A.11}$$
$$\text{such that } V^A(x, A_{n_0}(x), n_0-1) \leq V_{n_0-1}^*.$$

[Note that n_0 can equal N, in which case no restrictions are imposed on the action profiles.] We first prove that the third party's beliefs satisfy:

Claim 1: $\forall n \leq n_0$, $a(x_n) \leq A_{n_0}(x_n)$.

Proof: Suppose not. Fix $n \leq n_0$ such that $a(x_n) > A_{n_0}(x_n)$. Sequential rationality implies that there exist types greater than n_0 who also choose x_n with positive probability in equilibrium. Let j denote the greatest of these types. Since equilibrium actions are non-decreasing with type, we, then, have that $\forall k$, $n < k < j$,

A-10

Figure A.1

Proof of Proposition 5
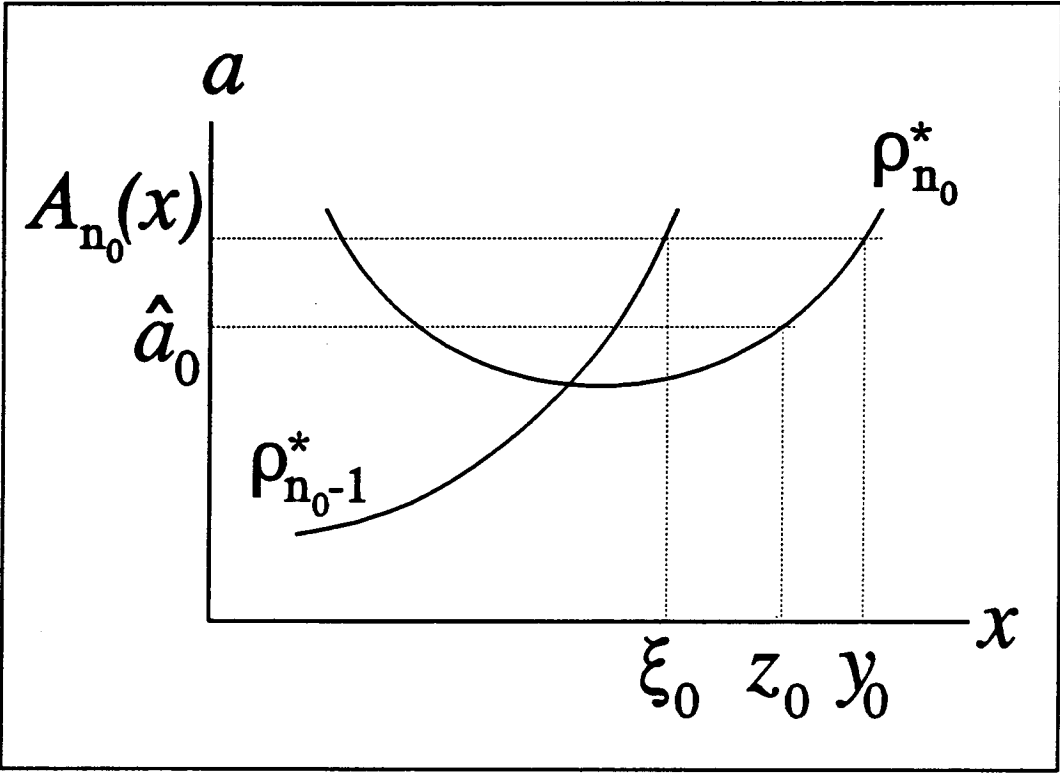
that $x_k = x_n$ for all x_k played with positive probability. Hence, $j = n_0 + 1$. But, for all actions x_j played by j with positive probability, $a(x_j) = A_j(x_j)$ by assumption. This, however, is impossible for x_n: It is played by no types better than j and it is played with positive probability by a type worse than j.

$\square$

Claim 1 allows us to define

$$y_0 = \max \{y \in \mathbb{R}_+ \mid V^A(y, A_{n_0}(y), n_0) = V^*_{n_0}\},$$

$$\xi_0 = \max \{y \in \mathbb{R}_+ \mid V^A(y, A_{n_0}(y), n_0-1) = V^*_{n_0-1}\}$$

(See Figure A.1). Moreover, $a(x_{n_0}) \leq A_{n_0}(x_{n_0})$; so if $y_0 < x_{n_0}$, then

$$V^*_{n_0} = V^A(y_0, A_{n_0}(y_0), n_0) > V^A(x_{n_0}, a(x_{n_0}), n_0) ,$$

a contradiction. So $x_{n_0} \leq y_0$. Let us compare y_0 and ξ_0:

Claim 2: $\xi_0 < y_0$ and $V^*_{n_0-1} > V^A(y_0, A_{n_0}(y_0), n_0-1)$.

Proof: Suppose $\xi_0 > y_0$. Then $x_{n_0} \leq y_0 < \xi_0$ and

$$V^A(x_{n_0}, a(x_{n_0}), n_0) \geq V^A(\xi_0, A_{n_0}(\xi_0), n_0).$$

ASCP then implies

$$V^A(x_{n_0}, a(x_{n_0}), n_0-1) \geq V^A(\xi_0, A_{n_0}(\xi_0), n_0-1) = V^*_{n_0-1},$$

which violates Lemma 2 for n_0-1. Suppose $\xi_0 = y_0$; then Lemma 2 implies:

$$V^*_{n_0-1} = V^A(y_0, A_{n_0}(y_0), n_0-1) \geq V^A(x_{n_0}, a(x_{n_0}), n_0-1).$$

As $x_{n_0} < y_0$, ASCP implies

$$V^A(y_0, A_{n_0}(y_0), n_0) \geq V^A(x_{n_0}, a(x_{n_0}), n_0) = V^*_{n_0}.$$

This, in turn, implies $x_{n_0} = y_0$. But, then, x_{n_0} would be least-cost separating -- a contradiction. Hence $\xi_0 < y_0$ and, thus, $V^*_{n_0-1} > V^A(y_0, A_{n_0}(y_0), n_0-1)$.

$\square$

Next, we prove that, all pseudo-indifference curves ρ_j^* corresponding to the proposed PBE for types $j \neq n_0$, lie *above* $(y_0, A_{n_0}(y_0))$ -- see Figure A.1.

Claim 3: $\forall j < n_0$, $V_j^* > V^A(y_0, A_{n_0}(y_0), j)$.

Proof: ASCP implies that ρ_j^* intersects the curve $\{a = A_{n_0}(x)\}$ at an output level no larger than ξ_0, since, otherwise, ρ_j^* would not intersect $\rho_{n_0-1}^*$ below ξ_0 and one would have $V_j^* < V^A(\hat{x}_{n_0-1}, a(\hat{x}_{n_0-1}), j)$. Claim 2 then implies Claim 3.

$\square$

Claim 4: $\forall j > n_0$, $V_j^* > V^A(y_0, A_{n_0}(y_0), j)$.

Proof: From (A.10) and (A.11), $\hat{x}_j \geq \xi_j \geq \hat{x}_{j-1}$, where

$$\xi_j \equiv \max \{y \mid V^A(y, A_j(y), j-1) = V_{j-1}^*\}.$$

Hence, ρ_j^* lies strictly above ρ_{j-1}^* for $y < \xi_j$. Claim 4 follows.

$\square$

Consider now $z_0 = y_0 - \varepsilon$ such that $V^A(z_0, A_{n_0}(z_0), n_0) > V_{n_0}^*$ and $\forall j \neq n_0$, $V^A(z_0, A_{n_0}(z_0), j) < V_j^*$. Claims 3 and 4 and the fact that x is not least-cost separating for n_0 ensure this construction is possible (see Figure A.1). Then, it follows that $a(z_0)$ must be such that $A_1(z_0) \leq a(z_0) \leq \hat{a}_0$, where $(z_0, \hat{a}_0) \in \rho_{n_0}^*$. By construction, $\hat{a}_0 < A_{n_0}(z_0)$. So, $\forall j \neq n_0$,

$$V^A(z_0, a(z_0), j) \leq V^A(z_0, \hat{a}_0, j) < V^A(z_0, A_{n_0}(z_0), j) < V_j^*. \tag{A.12}$$

For n_0, however, if $a(z_0) = \hat{a}_0$, $V^A(z_0, a(z_0), n_0) = V_{n_0}^*$. Hence the only weak best responses to equilibria in the component $a(z_0) \in [A_1(z_0), \hat{a}_0]$ are contracts that require only type n_0 to produce z_0. Deleting never-weak-best responses, we have $a(z_0) = A_{n_0}(z_0) > \hat{a}_0$. This, however, upsets all the equilibria in the component.

The proposed outcome is not stable.

Thus no such n_0 can exist, which means we are left with x^{AS} as the only stable equilibrium outcome. It is straightforward to verify that no similar construction is possible for x^{AS}.

■