GAMES AND ECONOMIC BEHAVIOR 2, 247–272 (1990)

Competition on Many Fronts: A Stackelberg Signaling Equilibrium*

JERRY GREEN

Department of Economics, Harvard University, Cambridge, Massachusetts 02138

AND

JEAN-JACQUES LAFFONT

California Institute of Technology and GREMAQ, Toulouse, France

Received November 17, 1988

An economic agent, the incumbent, is operating in many environments at the same time. These may be locations, markets, or specific activities. He is informed of the particular conditions relevant to each situation. His action in each case is observable by another agent, the entrant, who does not have the private information. Because the incumbent is in operation in many environments simultaneously, the entrant has the ability to discern the exact statistical relationship between the incumbent's action and information, and we assume that he cannot commit not to draw this inference. At each location the entrant must choose one of two actions, which we call "attack" or "no attack." This paper concerns the interplay between these sets of decisions. We characterize the optimal actions of the incumbent under the assumption that he can commit to his decisions, and thus that he will behave as a Stackelberg leader by manipulating the inferences drawn by the entrant. The solution obtained is compared with and contrasted to the Bayesian perfect equilibria of a game where both players move simultaneously. That game is a more appropriate model of the "no commitment" case. Thus, the value of the possibility of commitment and manipulation of the entrant's beliefs is assessed. Some applications are discussed. © 1990 Academic Press, Inc.

* Support from the National Science Foundation, Grants IST-8310118 and IST-8507291, and Commissariat du Plan is gratefully acknowledged. We thank the referees for their helpful and detailed comments.

0899-8256/90 $3.00

1. Introduction

This paper presents a model of competition, or potential competition, between two agents that takes place simultaneously on many fronts. One economic agent, the incumbent, is operating on all these fronts. He faces the possibility that on each front he will be "attacked" by the other agent. We give an example of such situations below, in the context of further specifications of the model. Suffice it to say, for the present, that the "fronts" may be a multiplicity of products being produced by a firm, the locations of economic, or even military, activity, the specific services provided to a variety of clients, or many other similar situations. The "attack" can be, for example, a military attack, or it can represent entry into direct economic competition against the incumbent in a market, or it can represent a legal action taken against the clients of the incumbent based on observations of their actions.

The fronts are distinguished from each other by a characteristic, or set of characteristics, known to the incumbent but unknown to the potential competitor. This characteristic, to be denoted θ, plays three roles in the model.

First, there is an action, x, to be taken by the incumbent on each front. The payoff to the incumbent, if he is not attacked on that front, is given by $u(x, \theta)$. Thus, there would be a desire to tailor the action x to the characteristic θ, but for the fact that it would allow the attacker to make accurate inferences about θ by virtue of his observations of x.

Second, θ affects the value of making an attack to the potential competitor. If he attacks on a front whose characteristic is θ, he gains $v(\theta)$. This may represent the expected value of an attack whose actual result is uncertain, but where the probability of the success depends on θ. Alternatively, the result of the attack may be independent of θ, but the value of having attacked, for example, the post-entry duopoly profit, may depend on θ. One should interpret $v(\theta)$ as the value of attack, net of any direct costs of doing so.

Third, θ may affect the cost that the attack, if made, would impose on the incumbent. This is represented by $w(\theta)$. It comprises the direct costs of a defense, if one is attempted, and the expected costs of the result of the attack, for example, the loss of market share and the change in market conditions resulting from a duopolistic post-entry situation.

We assume that there are many fronts and that, therefore, the empirical distribution of θ across the fronts is the same as the a priori belief, $F(\theta)$, held by the potential attacker with respect to each given front. *We also assume that the incumbent chooses x at each front in advance of the attacker's choice.* Moreover we assume that the incumbent knows that the attacker will be inferring the value of θ from his observation of x using

Bayes' rule, based on the prior F and a knowledge of how the incumbent's choice of x depends on θ,[1] and that the attacker cannot commit to do otherwise (e.g., ignore this inference). These assumptions lead us to examine the Stackelberg equilibrium of the game in which the incumbent is the "leader."[2]

The problem we solve in this paper is the optimization problem of the incumbent described above, under some further assumptions about u, v, and w. The result we obtain is quite a strong one. For despite the complexity of this problem, and its nonstandard nature as an optimization problem, we can show that the incumbent will select $x(\theta)$ nonstochastically for each θ, and that the function $x(\theta)$ can be described quite simply. In addition we can characterize the set of θ's at which an attack will take place, and the complementary set on which an attack is avoided.

For purposes of comparison we also analyze the Bayesian perfect equilibria of the same model,[3] in which each front is controlled by an agent who optimizes given a knowledge of his own θ. This is the appropriate model when commitment to act according to a given behavioral rule, the choice of x as a function of θ, is impossible to enforce and each "front" optimizes independently.

Our results demonstrate a striking qualitative difference between the Stackelberg and Bayesian Perfect equilibria. The latter, as is well known, involve a combination of separating and pooling. The characteristics that are pooled form an interval in the middle of the characteristics space. Optimal strategies in the Stackelberg case also involve pooling—but of quite a different nature. For an interval in the characteristics space there are pairs of values, one vulnerable to attack and the other not, for which the principal will choose the same action. Only these two characteristics are pooled together at this action. Thus the optimal strategy uses a whole

[1] Attackers can represent either a single entrant who can enter on all fronts or a continuum of potential local entrants. We assume that attackers get to know the incumbent's strategy by sampling. In the second case above, sampling must be done jointly.

[2] In games of complete information, Fudenberg and Levine (1989) show how the Stackelberg equilibrium payoffs can emerge in repeated games with a long-run player playing against a sequence of short-run opponents. This can happen when the Stakelberg behavior belongs to the domain of the entrant's beliefs. Our large number assumptions in a game of incomplete information, combined with the inability of the entrants to commit not to learn the incumbent's strategy, is analogous to their assumptions that the sequence of entrant observe all prior plays and the discount factor is close to one.

[3] Our model has the temporal structure of a signaling model (see Spence, 1974), but the informed player controls his actions as a function of the relevant specific information instead of a separate economic agent in control of each action. Most of the literature on signaling models (see Milgrom and Roberts, 1982; Kreps and Wilson, 1982) has focused on the Bayesian perfect equilibria of this model. In view of the large number of equilibria obtained, efforts have concentrated on finding convincing selection criteria (see Kohlberg and Mertens, 1986; Cho and Kreps, 1987; Banks and Sobel, 1987).

range of actions to pool the continuum of pairs of characteristics in the pooled interval. Furthermore, we characterize the domain of characteristics θ for which the ability to commit modifies the incentive compatibility conditions that would be faced in Bayesian perfect equilibrium.

The problem is set up in Section 2. Section 3 offers a discussion in the context of an industrial organization application. Section 4 states the main results and provides further commentary on the relationship between the solution of the incumbent's problem and the Bayesian perfect equilibrium. Proofs, which are long, are deferred to appendices. Section 5 gives a brief numerical example.

2. The Problem

We assume that the domain of the parameter θ is a bounded interval $\theta \in \Theta = [\theta_{\min}, \theta_{\max}]$ of real numbers. The distribution function of θ is assumed to be atomless and is denoted F. Its density is denoted f; it is, for simplicity, assumed to be continuously differentiable on Θ and strictly positive. The set of possible actions is assumed to be the real line.

A *strategy* for the incumbent is a stochastic kernel $s(\cdot|\theta)$, which is a measure over the real line for each $\theta \in \Theta$. This allows for randomized choices of x, although as we show, they are not used at the optimum.

The reaction of the potential attackers depends on their belief about θ given the observations x. Let $H(\cdot|x)$ be the conditional distribution over θ that would be obtained by Bayes' rule. If an attack is made, the expected payoff to the attacker depends only upon whether or not θ exceeds a critical value $\bar{\theta} \in \Theta$. If $\theta > \bar{\theta}$, the attacker gains an amount $v_+ > 0$, if $\theta \leq \bar{\theta}$, the attacker loses an amount $-v_-(v_- < 0)$. One interpretation of this is that the attack succeeds or fails according to this condition. Once an attack has succeeded, however, the payoff to the attacker is independent of θ, and of the associated decision x. Thus the expected payoff to an attack will be

$$v_+ H(\Theta_+|x) + v_- H(\Theta_-|x), \tag{2.1}$$

where $\Theta_+ = \{\theta|\theta > \bar{\theta}\}$, $\Theta_- = \{\theta|\theta \leq \bar{\theta}\}$. An alternative notation that will sometimes be used is to define the function

$$\begin{aligned} v(\theta) &= v_+ \qquad \text{if } \theta > \bar{\theta} \\ &= v_- \qquad \text{if } \theta \leq \bar{\theta}, \end{aligned} \tag{2.2}$$

and then the expected value of an attack is just $\int v(\theta) dH(\theta|x)$, often denoted $\int v dH$.

The alternative to attacking is not to attack, and the value of not attacking is normalized to be zero. As is typical in the incentives literature, we assume that the agent attacks only if (2.1) is strictly positive. The incumbent presumes that the attacker will form his beliefs according to the Bayesian method described above. Thus the incumbent assumes that in choosing his strategy he is able to manipulate the attacker's beliefs. Instances in which this is a plausible model of the incumbent's behavior are described in the next section.

We assume that for each θ the incumbent's utility is derived from two sources. First, the action x is payoff relevant to him and he experiences a utility $u(x, \theta)$ if x is the decision associated with θ. We assume that u is twice differentiable, strictly concave in x, and $u_{x\theta} > 0$. Moreover, for each θ there is a unique value of x, denoted $x^*(\theta)$, that maximizes $u(x, \theta)$. It follows that $x^*(\theta)$ is a continuously increasing function. The importance of this condition is that the potential attacker can unambiguously learn all values of θ, unless nonoptimal actions are taken in an attempt to hide them.

Second, there is a disutility to being attacked. The level of the disutility depends on whether or not $\theta > \bar{\theta}$. If $\theta > \bar{\theta}$ is attacked it is w_+ and if $\theta \leq \bar{\theta}$ is attacked it is w_-. Following the interpretation mentioned above, one could say that defending against an unsuccessful attack costs w_-, but the loss incurred in a successful attack is w_+. Although it might be natural to assume $w_+ > w_-$, we do not need that hypothesis below.

Let

$$w(\theta) = w_+ \quad \text{if } \theta > \bar{\theta} \qquad (2.3)$$
$$= w_- \quad \text{if } \theta \leq \bar{\theta}.$$

The assumptions (2.2) and (2.3) that the value and cost of an attack depend only on its success or failure, and are otherwise independent of θ and of the action x, are admittedly very special. If they were weakened we would still preserve the structure of the optimization problem and the great qualitative differences of Stackelberg equilibria and Bayesian perfect equilibria that are stressed by our analysis. But the exact characterization of the Stackelberg equilibrium that we are able to obtain under these conditions would certainly not hold. In particular, the nature of the pooling that would occur might be much more complex that we obtain in Theorem 1.

Consider a strategy $s(\cdot|\theta)$. Let $A_s \subseteq R$ be the set of $x \in R$ such that (2.1) is strictly positive. We call A_s the set of *attacked* values, given the strategy s.

The incumbent's utility can then be written as

$$\int_\Theta \left\{ \int_R u(x, \theta)\, ds(x|\theta) - \int_{x \in A_s} w(\theta) ds(x|\theta) \right\} dF(\theta). \qquad (2.4)$$

We study the problem of maximizing (2.2) by the choice of the strategy s. The complexity (and nonlinearity) of this problem is due to the particular nature of the dependence of A_s on s.

A *Stackelberg signaling equilibrium* (SE) is a strategy $s^*(\cdot|\cdot)$ which maximizes (2.4).

3. Examples

There are several key ingredients in the model which determine its domain of applicability. The attacker draws his inference about the incumbent as a Bayesian statistician. He moves after the incumbent and it is common knowledge that he can learn the incumbent's behavioral rule and optimize against it. Thus he cannot, for example, threaten attacking in any way other than would be dictated by independent optimization at each front.

An example of this might be a retailer or a bank who is operating at many locations, initially in the absence of any competition. The incumbent may be the first firm to have expanded into a new area. It is reasonable to suppose that it will soon learn the profitability of each of its locations. At more profitable locations it might be optimal to expand the hours of business, increase staff or enlarge its physical facility.

But the bank knows that if it were to do so it would be giving future competitors the knowledge of the quality of each location. The competitor could learn the relationship between the observable attributes of the incumbent's locations and their underlying quality by entering at a sample of locations and drawing the appropriate inferences. Then it could target its entries at all the other locations accordingly. As the total number of locations is very large, the mistakes made in the initial sample are insignificant in the total payoff.

Because the incumbent's characteristics are fixed once and for all, there is no scope for the entrant to try to manipulate the incumbent by engaging in any nonmyopic behavior. Moreover, even after the initial entry (attack) at various cites, it turns out that the incumbent's optimal strategy is unchanged. Therefore, when the next entrant (if any) is present, he will not find it profitable to attack anywhere. All values of the observable for which there is a positive expected benefit of attack have already been attacked by the previous entrant.

Another example that is well modeled by the Stackelberg equilibrium concerns the preparation of income tax returns by a firm that handles the returns for many clients, each of whom may be audited by the IRS at some time in the future. Returns must be filed for the previous tax year but are not examined by the IRS until several years later. The IRS knows which returns have been prepared by this tax preparation firm and, as it begins to examine them it can note those characteristics of the return that indicate that it is worth auditing (audits are costly). Because of the large number of returns involved, the statistical relationship between observable characteristics and audit potential can be estimated very precisely using only a negligible proportion of all returns filed by this firm. For the remaining returns, the optimal audit policy can be followed. Thus, when the income tax preparation firm decides on its policy for how to prepare returns with underlying characteristics that are known, prior to audit, only by itself (and its client) it will behave as Stackelberg leader. The IRS cannot commit not to draw the inference about audit potential, for an announcement to this effect would not be credible—it is common knowledge that once the returns are filed, and several years have passed so that they cannot be changed, the IRS will have every incentive to audit them in a way that will produce the largest payoff at that time.

To be sure, some of the specific assumptions of Section 2—such as the invariance of the costs of an attack to the incumbent to the characteristics of the front in question—may not be satisfied in a particular application. Nevertheless, we believe that the strategic situation studied here, where the incumbent can make a commitment to his strategy and the attacker cannot, does characterize many competitive situations where "many fronts" are involved.

4. Statement of Results

In Theorem 1, we describe the qualitative features of an optimal strategy if an optimal strategy exists. Theorem 2 proves the existence of an optimal strategy under our assumptions. To contrast the Stackelberg signaling equilibrium with the Bayesian perfect equilibria (Proposition 1) we first characterize the Bayesian perfect equilibria using a mild condition on out of equilibrium expectations (Theorem 3).

Theorem 1. *If $s^*(\cdot|\cdot)$ is a Stackelberg signaling equilibrium, it is almost everywhere equal to a strategy $s(\cdot|\cdot)$ such that there exists an interval $\bar{T} = (a, b] \subset \Theta$, possibly degenerate and containing $\bar{\theta}$, and increasing functions $y(\cdot)$: $\bar{T} \cap \Theta_- = (a, \bar{\theta}] \to \mathbb{R}$ and $z(\cdot)$: $\bar{T} \cap \Theta_+ = (\bar{\theta}, b] \to \mathbb{R}$ such that*

(1) *for $\theta \notin \bar{T}$, $s(\cdot|\theta)$ is concentrated at $x = x^*(\theta)$;*

(2) *for $\theta \in \bar{T} \cap \Theta_-$, $s(\cdot|\theta)$ is concentrated at $x = y(\theta)$;*

(3) *for $\theta \in \bar{T} \cap \Theta_+$, $s(\cdot|\theta)$ is concentrated at $x = z(\theta)$;*

(4) $v_- f(\theta)\dot{z}(\tilde{\theta}) + v_+ f(\tilde{\theta})\dot{y}(\theta) = 0$ *a.e. with $y(\theta) = z(\tilde{\theta})$ for any $\theta \in \bar{T} \cap \Theta_-$;*

(5) *lim inf*$_{\theta\to a}$ $y(\theta)$ = *lim inf*$_{\theta\to\bar\theta}$ $z(\theta)$; $y(\bar\theta) = z(b)$;

(6) *the set of attacked values is* $A = \{\theta\colon \theta \in \Theta_+$ *and* $\theta \notin \bar{T}\}$.

Proof. See appendix.

Let us call S the subset of strategies that can be described by a 3-tuple $\bar{T} = (a, b] \subset \Theta$, $y(\cdot)$, $z(\cdot)$ with $y(\cdot)$ and $z(\cdot)$ increasing and satisfying (4), (5) of Theorem 1.

With Theorem 1 we can reduce the existence problem to the existence of a solution to (2.2) in S. Then, the optimization problem (2.2) can be rewritten.

$$\max_{(a,b,y(\cdot),z(\cdot))\in S} \int_{\theta_{\min}}^{a} u(x^*(\theta), \theta)f(\theta)d\theta + \int_{a}^{\bar\theta} u(y(\theta), \theta)f(\theta)d\theta \tag{4.1}$$
$$+ \int_{\bar\theta}^{b} u(z(\theta), \theta)f(\theta)d\theta + \int_{b}^{\theta_{\max}} (u(x^*(\theta), \theta) - w_+)f(\theta)d\theta.$$

We show below that there exists a solution to program (4.1) and therefore, from Theorem 1, that there exists a solution to program (2.2).

THEOREM 2. *There exists a solution to program* (4.1) (*which is not necessarily unique*).

Proof. Let $\psi(\theta) = z^{-1}(y(\theta))$ (well defined because z is increasing).

Conditions (4), (5) in Theorem 1 imply

$$\dot\psi(\theta) = -\frac{v_-}{v_+}\cdot\frac{f(\theta)}{f(\psi(\theta))} \tag{4.2}$$

$$\bar\theta = \psi(a). \tag{4.3}$$

This is a differential equation in $\psi(\cdot)$ with a boundary condition. Since $f(\cdot)$ is differentiable and bounded below by a strictly positive number, there exists from the fundamental existence theorem of the theory of differential equations (Pontriagin, 1962) a differentiable solution $\psi_a^*(\theta)$ defined on $[\theta_{\inf}, \theta_{\max}]$.

From (4.2), ψ_a^* is increasing in θ with a derivative bounded below by a strictly positive number.

Moreover, the differentiability of ψ_a^* in a follows from the differentiability of a solution with respect to the initial condition (Pontriagin, 1962) and from the differentiability of the solution in θ.

From (4.2), (4.3),

$$b = \psi_a^*(\bar\theta) = \psi_a^*(\psi_a^*(a)). \tag{4.4}$$

The maximization problem can now be rewritten:

$$\max_{(a,z(\cdot))} \left\{ \int_{\theta_{\min}}^{a} u(x^*(\theta), \theta) f(\theta) d\theta + \int_{a}^{\bar{\theta}} u(z(\psi_a^*(\theta)), \theta) f(\theta) d\theta \right.$$
$$\left. + \int_{\bar{\theta}}^{\psi_a^*(\psi_a^*(a))} u(z(\theta), \theta) f(\theta) d\theta + \int_{\psi_a^*(\psi_a^*(a))}^{\theta_{\max}} (u(x^*(\theta), \theta) - w_+) f(\theta) d\theta \right\}. \tag{4.5}$$

Let us now change variables in the second integral of (4.5): $\eta = \psi_a^*(\theta)$. This integral becomes

$$\int_{\psi_a^*(a)}^{\psi_a^*(\bar{\theta})} u(z(\eta), \psi_a^{*-1}(\eta)) f(\psi_a^{*-1}(\eta)) \frac{d\eta}{\psi_a^{*\prime}(\psi_a^{*-1}(\eta))}. \tag{4.6}$$

Substituting the running variable θ for η and using (4.3) and (4.4), (4.6) becomes

$$\int_{\bar{\theta}}^{b} u(z(\theta), \psi_a^{*-1}(\theta)) f(\psi_a^{*-1}(\theta)) \frac{d\theta}{\psi_a^{*\prime}(\psi_a^{*-1}(\theta))}.$$

But $f(\psi_a^{*-1}(\theta))/\psi_a^{*\prime}(\psi_a^{*-1}(\theta)) = -(v_+/v_-)f(\theta)$ from (4.2). Maximization with respect to $z(\cdot)$ reduces to

$$\max \int_{\bar{\theta}}^{b} [u(z(\theta), \theta) - \frac{v_+}{v_-} u(z(\theta), \psi_a^{*-1}(\theta))] f(\theta) d\theta. \tag{4.7}$$

For any a, there exists a solution to (4.7) because $u(\cdot, \theta)$ has by assumption a solution for any $\theta \in \Theta$ (we are maximizing a weighted average of two such functions for every θ).

Moreover, since u is strictly concave in z this solution is defined by

$$\frac{\partial u}{\partial z}(z, \theta) - \frac{v_+}{v_-} \frac{\partial u}{\partial z}(z, \psi_a^{*-1}(\theta)) = 0.$$

The solution is increasing since $u_{x\theta} > 0$ and $\psi_a^{*\prime} > 0$ and differentiable from the inverse function theorem.

There exists a solution in a as we are maximizing a continuous function in a compact set. However, as (4.5) is not concave in a, there may be multiple solutions. ■

Some simple results are seen directly in Fig. 4.1. The set of fronts that are attacked are those with the upper extreme values of θ. These are the values that are hardest to protect in the sense that pooling them with θ's

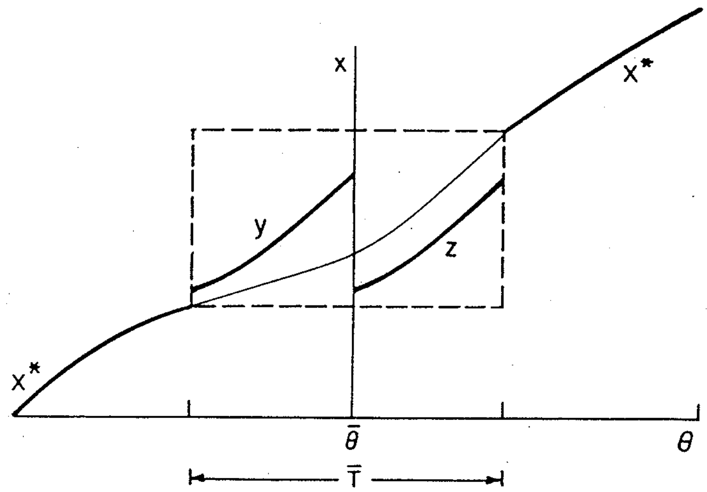

FIG. 4.1. The Stackelberg equilibrium.

below $\bar{\theta}$ would require large deviations from $x^*(\theta)$ than for the lower values, in $\bar{T}$, which are pooled and protected from attack. It is natural that the protection is afforded to the fronts that are less costly to protect.

The protection from attack in $\bar{T}$ requires an increase in x, above $x^*(\theta)$ for the fronts that cause the protection, and a decrease in x for those that are protected. Moreover, the two fronts that are pooled together choose an x that is between their respective values of $x^*(\theta)$. Again, this seems quite natural.

It is interesting to compare the solution above to the Bayesian perfect equilibria of the same game. The Bayesian perfect equilibrium concept would correspond to applications where each front is controlled by a separate agent who optimizes given his own θ, taking the pattern of inference used by potential attackers as given.

A *Bayesian perfect equilibrium*[4] is a pair of functions $\bar{x}(\theta)$: $\Theta \to \mathbb{R}$ and $\bar{\delta}(x)$: $\mathbb{R} \to \{0, 1\}$, such that

(i) $\theta \in \Theta$, $\bar{x}(\theta) \in \text{argmax}_x[u(x, \theta) - \bar{\delta}(x)w(\theta)]$

(ii) $F(\theta|x)$ is the revision of $F(\theta)$ using Bayes' rule whenever possible given x and $\bar{x}(\cdot)$

(iii) $\bar{\delta}(x) = 1(0) \Leftrightarrow v_- \int_{\theta \leq \bar{\theta}} dF(\theta|x) + v_+ \int_{\theta > \bar{\theta}} dF(\theta|x) > (\leq) 0.$

[4] We give the definition only in the case of pure strategies. It can be shown that equilibria always involve pure strategies.

We limit the number of equilibria by restricting out of equilibrium beliefs. We say that beliefs $\tilde{F}$ out of equilibrium are *plausible*[5] if $\int_{\theta>\bar{\theta}} d\tilde{F}(\theta|x) < \int_{\theta>\bar{\theta}} dF(\theta|x')$ for any $x < x'$, where the value x is not taken at the equilibrium while x' is taken. This plausibility requirement expresses the idea that the attacker knows that $x^*(\theta)$ is strictly monotonic and therefore believes that an unused value of x would be associated with a lower value of θ than that known to be associated with higher x, in equilibrium.

THEOREM 3. *If $\hat{\theta} > \bar{\theta}$ and $(\hat{\theta}, \hat{x})$ satisfy*

(a) $u(x^*(\hat{\theta}), \hat{\theta}) - w_+ = u(\hat{x}, \hat{\theta})$

(b) $\int_{x^{*-1}(\hat{x})}^{\bar{\theta}} v_- dF(\theta) + \int_{\bar{\theta}}^{\hat{\theta}} v_+ dF(\theta) \leq 0,$

then

$$\begin{aligned} \bar{x}(\theta) &= x^*(\theta) && \text{for } \theta < x^{*-1}(\hat{x}) \\ \bar{x}(\theta) &= \hat{x} && \text{for } \theta \in [x^{*-1}(\hat{x}), \hat{\theta}] \\ \bar{x}(\theta) &= x^*(\theta) && \text{for } \theta > \hat{\theta} \end{aligned}$$

is a (plausible) Bayesian perfect equilibrium, and conversely. Moreover, the set of attacked values is $A = \{\theta: \theta \in \Theta$ and $\theta > \hat{\theta}\}$.

Proof. Available from the authors.

In any plausible Bayesian perfect equilibrium, the conjectures about values of x that will be attacked are as follows. For $x > \hat{x}$ the incumbent agent believes that an attack will take place. Because of this, the values in the interval $(\hat{x}, x^*(\hat{\theta}))$ are not chosen by any θ. Instead, all θ in $(x^{*-1}(\hat{x}), \hat{\theta})$ pool at $\hat{x}$, the highest value of x, which escapes attack. For $\theta > \hat{\theta}$ the corresponding x is set at $x^*(\theta)$ where attacks actually do occur.

In the Bayesian perfect equilibrium which maximizes the expected payoff over all θ, (b) in Theorem 3 holds with equality (see Fig. 4.2).

PROPOSITION 1. *In the Stackelberg signaling equilibrium, there are some θ for which the payoff is lower than what they would receive in the Bayesian perfect equilibrium that maximizes the expected payoff over all θ.*

Proof. There are two cases. Either the SE pooling set, $\bar{T}$, is included in the best BPE pooling set or $\bar{T}$ contains the best BPE pooling set. This is because there must be equal weighted mass on each side of $\bar{\theta}$ for both pooling sets, as $\int v dH = 0$.

In the second case (Fig. 4.3), agents θ, $\theta \in (\theta_1, \theta_2)$ prefer the best BPE allocation because they get their first best.

[5] This monotonicity restriction is similar to the one found in Kreps and Wilson (1982).

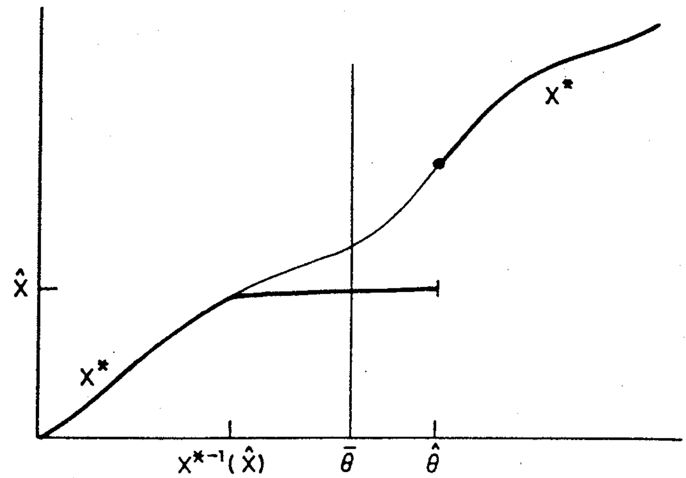

FIG. 4.2. A Bayesian perfect equilibrium.

In the first case (Fig. 4.4), agents θ, $\theta \in (\theta_1, \theta_2)$ prefer the best BPE allocation. The reason is as follows. In the BPE, agent θ_2 is indifferent between A and B. At B he is attacked; at A he is not. Take $\theta \in (\theta_1, \theta_2)$. In the SE he is attacked. The action ξ is now closer to his first best than it was at θ_2 for agent θ_2. Therefore he strictly prefers the action ξ. ■

We have argued above that in the SE the incumbent was able to commit himself to a given strategy because he was moving first. It is nevertheless interesting to know where the incentive constraints would be violated in a SE. We must choose plausible expectations about attack for the values of x which are not chosen. Suppose (Fig. 4.5) that for $x \in (x_1, x_2)$ no attack is expected and for $x \in (\bar{x}, x_3)$ attack is expected.

All those in $[\theta_{\min}, \theta_1]$ are satisfying incentive constraints since they obtain their first best and are not attacked. All those in (θ_1, θ_2) do not

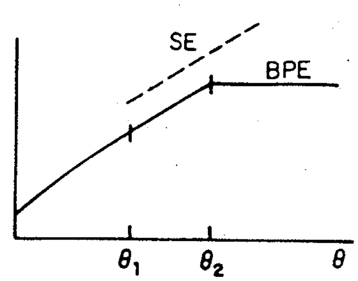

FIGURE 4.3

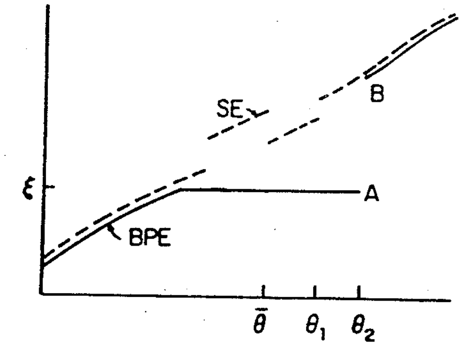

FIGURE 4.4

satisfy incentive constraints, because they could choose their first best and not be attacked (since for $x \leq \bar{x}$, there is no attack).

All those in $[\theta_2, \theta_3]$ do not satisfy incentive constraints; they can move closer to their first best, for example, to $\bar{x}$, and not be attacked.

If $\theta = \theta_3$ the incentive constraint is satisfied in general strictly (if not one could improve the SE). For θ in (θ_3, θ_4), in general a nondegenerate interval, the incentive constraint is violated; these values of θ would prefer to choose $\bar{x}$ and avoid attack rather than $x^*(\theta)$ where they are attacked. For $\theta > \theta_4$ the incentive constraints are satisfied at $x^*(\theta)$.

5. AN EXAMPLE

Let $\Theta = [0, 1]$, $\bar{\theta} = \frac{1}{2}$, $\bar{T} = (a, b]$, and let us denote the optimal $\bar{x}$ as $y(\cdot)$ if $a \leq \theta < \bar{\theta}$ and as $z(\cdot)$ for $\bar{\theta} < \theta < b$.

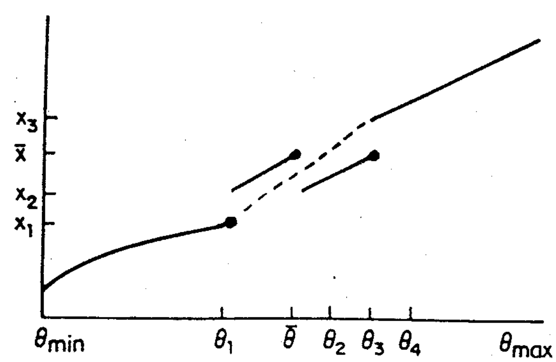

FIGURE 4.5

Consider now the following example:

$$u(x, \theta) = -(x - \theta)^2;$$
$$v_+ = -v_-, \; w_+ = w, \; w_- = 0,$$
$$F \text{ uniform.}$$

The condition implied by the theorem for pooled values is

$$f(\bar{\theta})\dot{y}(\theta) = f(\theta)\dot{z}(\bar{\theta}) \text{ for } y(\theta) = z(\bar{\theta}) \text{ for a.e. } \theta \in \bar{T} \tag{5.1}$$

$$\liminf_{\theta \to a} y(\theta) = \liminf_{\theta \to 1/2} z(\theta), \qquad y(\tfrac{1}{2}) = z(b). \tag{5.2}$$

From the proof of Theorem 2 we have

$$f(\psi(\theta))\dot{\psi}(\theta) = f(\theta).$$

For the uniform distribution

$$f(\theta) = f(\psi(\theta)) = 1; \; x^*(\theta) = \theta \text{ and } \psi(\theta) = \theta + k.$$

From $\liminf_{\theta \to a} y(\theta) = \liminf_{\theta \to 1/2} z(\theta)$ we get $k = \frac{1}{2} - a$; from $y(\frac{1}{2}) = z(b)$ we get $b = 1 - a$. The optimization problem of the principal can then be reduced to

$$\max \left\{ - \int_a^{1/2} [y(\theta) - \theta]^2 \, d\theta - \int_{1/2}^{1-a} [z(\theta) - \theta]^2 \, d\theta - aw \right\}$$
$$\text{subject to } y(\theta) = z\left(\theta + \frac{1}{2} - a\right) \text{ and (5.2).}$$

Changing variables in the second integral ($\bar{\theta} = \theta - \frac{1}{2} + a$) and using (5.2), we get

$$\max_{(a, y(\cdot))} \left\{ - \int_a^{1/2} \left[(y(\theta) - \theta)^2 + \left(y(\theta) - \theta + \frac{1}{2} - a \right)^2 \right] d\theta - aw \right\}.$$

The first-order conditions are

$$(y(\theta) - \theta) + \left(y(\theta) - \theta + \frac{1}{2} - a \right) = 0$$

$$(y(a) - a)^2 + \left(y(a) - \frac{1}{2}\right)^2 - 2\int_a^{1/2}\left(y(\theta) - \theta + \frac{1}{2} - a\right) d\theta = w,$$

yielding

$$a = \frac{1}{2} - \sqrt{\frac{2w}{3}}$$

$$y(\theta) = \theta + \sqrt{\frac{w}{6}}.$$

For any $w > 0$, some area is protected. As w reaches $\frac{3}{8}$ everything is protected (see Fig. 5.1).

It is intuitively clear that as soon as $w > 0$, the gain of some little pooling is of the first order and the loss is only of the second order. As w becomes very large everything is protected since the loss from full protection is finite. Figure 5.2 gives the comparison of the incumbent's levels of utility in the BPE (which is the best for the principal) and in the Stackelberg signaling equilibrium, for $w = \frac{3}{32}$. The BPE is calculated by using the uniformity of F, so that $\bar{\theta}$ is the midpoint of the pooled set, and determining the length of the pooled set from (a) of Theorem 3.

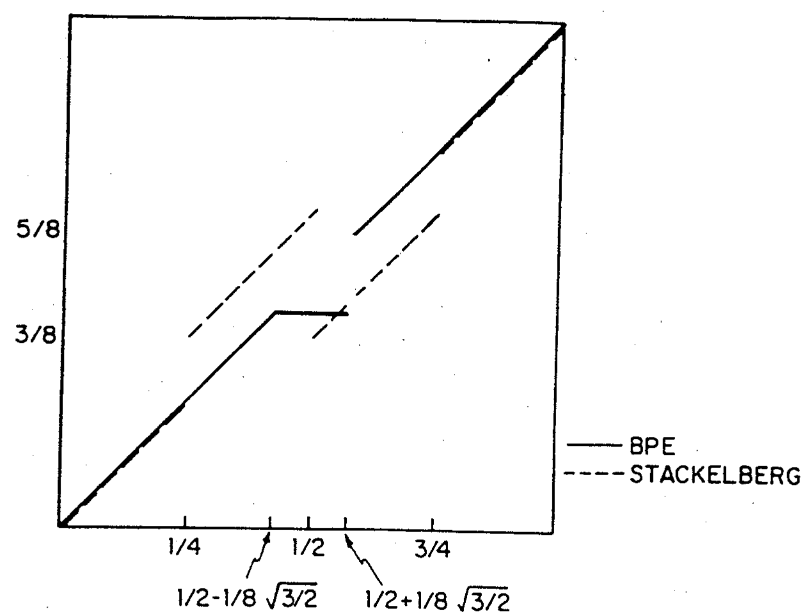

FIGURE 5.1

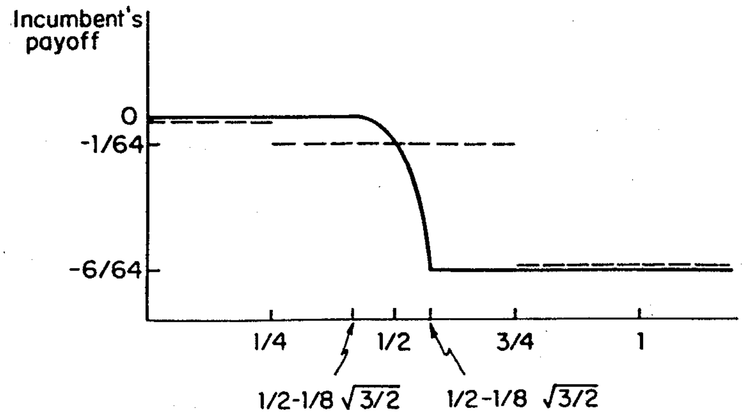

FIGURE 5.2

APPENDIX

Characterization of the Stackelberg Signaling Equilibrium: Proof of Theorem

Some further terminology is useful in the arguments below. Given a strategy s let M be the induced marginal distribution of x.

A subset $S \subseteq \mathbb{R}$ will be said to be *identified* if for M-almost every $x \in S$, $H(\theta|x)$ is a measure degenerate at a single point. A subset S will be said to be *pooled* if for M-almost every $x \in S$, $H(\theta|x)$ is not such a degenerate measure. The maximal subsets of identified and pooled values are denoted I and P, respectively.

All of the distributions and sets of observed values of x described above are determined by the strategy s. Where it is desirable to make this dependence explicit, we subscript the corresponding value by s, for example, H_s, A_s, etc. The first step is to show that no pooled value is attacked, i.e., $M(P \cap A) = 0$.

LEMMA 1. *There can be no atoms of M in $P \cap A$.*

Proof. Let $\bar{x}$ be such an atom. Then for a nonnull subset $T \subseteq \Theta$, $\bar{x}$ is an atom of $s(\cdot|\theta)$ for $\theta \in T$. Define

$$T_+^\varepsilon = \{\theta \in T | \theta > \bar{\theta} \text{ and } \bar{x} < x^*(\theta) - \varepsilon\}$$
$$T_-^\varepsilon = \{\theta \in T | \theta > \bar{\theta} \text{ and } \bar{x} > x^*(\theta) + \varepsilon\}.$$

For ε sufficiently small, at least one of T_+^ε and T_-^ε must be nonnull. Without loss of generality, assume it is T_+^ε. Take $x > \bar{x}$ such that $x - \bar{x} < \varepsilon$

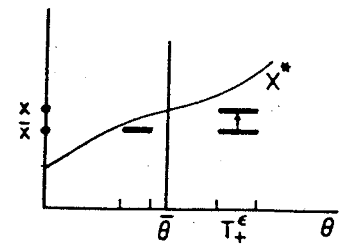

FIGURE A.1

and that x is not an atom of M. (This is possible because there are at most a countable number of atoms.) Then modify s to s' by replacing the atom at $\bar{x}$ with an atom of equal mass at x, for all $\theta \in T_+^\varepsilon$. Under the strategy s' the increased value of x with positive probability, in the direction of the optimum, will cause $\int u dG_{s'} > \int u dG_s$. The value of $\int_\Theta \int_{A_s} w(\theta) ds(x|\theta) dF(\theta)$ will not increase because $M(A_{s'}) \leq M(A_s)$. Thus s' is a superior strategy to s. The case in which T_-^ε is nonnull is symmetrically treated (see Fig. A.1).

LEMMA 2. $M(P \cap A) = 0$.

Proof. By the above, we know that M is nonatomic on $P \cap A$. Assume that $M(P \cap A)$ is positive. Consider the joint distribution $G(\theta, x)$ restricted to $x \in P \cap A$, denoted by $G^1(\theta, x)$. Let $H^1(\theta|x)$ be the conditional distribution of θ, defined from G^1, on $P \cap A$. Let $C^\varepsilon = \{(\theta, x) \in \Theta \times (P \cap A) \mid |x - x^*(\theta)| < \varepsilon\}$. If, for all $\varepsilon > 0$, $G^1(C^\varepsilon) = M(P \cap A)$, then $H^1(\theta|x)$ would be degenerate at $\theta = x^{*-1}(x)$ for each $x \in P \cap A$. This would contradict the fact that they are pooled values (see Fig. A. 2).

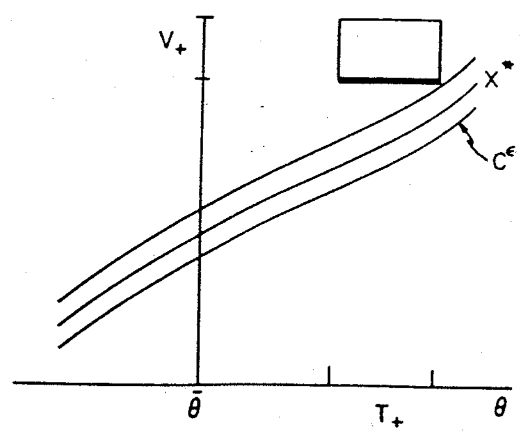

FIGURE A.2

Thus there exists $\varepsilon > 0$ and $\eta > 0$ such that $M(P \cap A) - G^1(C^\varepsilon) > \eta$. It follows that either there exists subsets $T_+ \subseteq \Theta$ and $V_+ \subseteq P \cap A$ such that for each $\theta \in T_+$, $v \in V_+$, $v > x^*(\theta) + \varepsilon$, and $G^1(T_+ \times V_+) > 0$, or else there are T_- and V_- with $v < x^*(\theta) - \varepsilon$ and $G^1(T_- \times V_-) > 0$. Without loss of generality we can consider the former case. Following a method similar to that used in the preliminary lemma above, replace s by the strategy s' that assigns a point mass of $s(V_+|\theta)$ at a given $\bar{x} \in V_+$ arbitrarily close to $\inf_x x \in V_+$ for every $\theta \in T_+$, and such that $s'(V_+ \cap \{x|x > \bar{x}\}|\theta) = 0$. As above, this improves the efficiency of the strategy s with respect to $\int udG$ while not increasing A_s because $\bar{x}$ is already an attacked value and therefore not increasing $\int_\Theta \int_{A_s} wdG$. ■

We then show that for identified values, the incumbent is choosing the action that maximizes $u(x, \theta)$.

LEMMA 3. *For M-almost every $x \in I$, $H(\theta|x)$ is a point mass concentrated at $\theta = x^{*-1}(x)$.*

Proof. Let $\phi(x)$ be the value of $\theta \in \Theta$ corresponding to the observation of $x \in I$. Let $\Phi \subseteq \Theta \times I$ be its graph, that is,

$$\Phi = \{(\theta, x)|\theta = \phi(x), x \in I\}.$$

Let $X^* = \{(\theta, x)|x = x^*(\theta), \theta \in \Theta\}$. We want to show that $G(\Phi \backslash X^*) = 0$.

We follow the same procedure as in Lemma 2 above. If $G(\Phi \backslash X^*) > 0$, then there must be an $\varepsilon > 0$ such that $G(\Phi \backslash N_\varepsilon(X^*)) > 0$. We then can find T contained in either Θ_- or Θ_+ and $V \subseteq I$ such that $G((T \times V) \cap (\Phi \backslash N_\varepsilon(X^*)) > 0$ and either $(\theta, x) \in (T \times V)$ implies $x > x^*(\theta) + \varepsilon$ or $(\theta, x) \in T \times V$ implies $x < x^*(\theta) - \varepsilon$. Thus there are four possible cases, as $T \times V$ is above or below X^* and to the right or left of $\bar{\theta}$. In any case, a superior strategy, s', can be found by assigning all the mass in $T \times V$ to a single point $x \in I$ which is selected arbitrarily close to the extreme of V closer to X^* (for example, see Fig. A.3).

This change creates a pooled value, x, such that $H_{s'}(T|x) = 1$ and $M_{s'}(\{x\}) > 0$.

If $T \subseteq \Theta_-$, $x \notin A_s$, and if $T \subseteq \theta_+$, $x \in A_{s'}$. In either case, however, $A_s = A_{s'}$ and $\int_\Theta \int_{A_{s'}} wdG_s = \int_\Theta \int_{A_{s'}} wdG_{s'}$. Thus the change is beneficial because $\int\int udG$ is increased. ■

We next show in the optimal strategy all unattacked pooled values are on the margin of being attacked.

LEMMA 4. *For M-almost every $x \in P$, $\int vdH = 0$.*

Proof. (1) Assume that there is a nonnull subset $V \subseteq P$ with $x > x^*(\bar{\theta})$ for all $x \in V$ and such that $\int vdH(\theta|x) < 0$ for almost every $x \in V$. We can

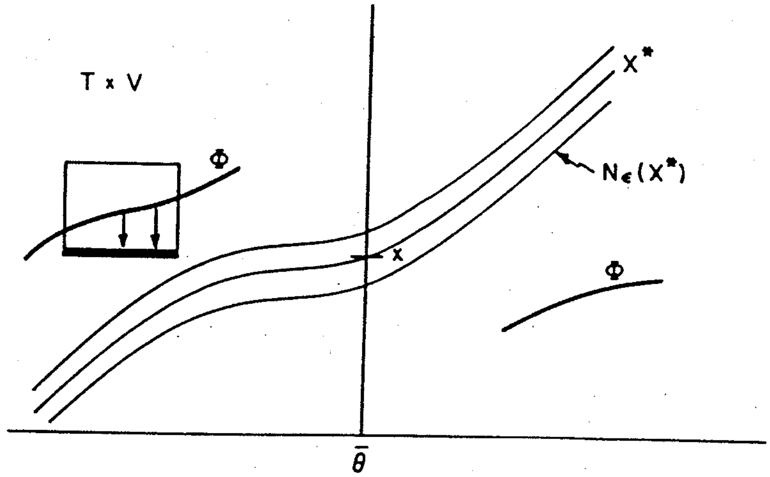

FIGURE A.3

find a further nonnull subset $V' \subseteq V$ such that $\int v dH(\theta|x) < \delta < 0$ for almost every $x \in V'$. Rewrite the last expression as

$$v_- H(\Theta_-|x) + v_+ H(\Theta_+|x) < \delta < 0.$$

Since $x > x^*(\bar{\theta})$ we know that $x > x^*(\theta)$ for all $\theta \in \Theta_-$. We improve the strategy s by introducing a small randomization which assigns a point mass at $x^*(\theta)$ to $\theta \in \Theta_-$, decreases $H(\Theta_-|x)$ slightly for $x \in V'$, and otherwise does not change s. This change improves the efficiency of the choice of x with respect to $\int u dG$ and, as its does not cause $\int v dH \leqq 0$ to be violated for any x where it was satisfied under s, it does not increase A or $\int_\Theta \int_A w dG$. Hence s was not optimal.

(2) Now consider the complementary case where $\int v dH < 0$ for $x \in V$, and $x < x^*(\bar{\theta})$ for all $x \in V$. As above, let V' be a nonnull subset of V with $\int v dH < \delta < 0$ for all $x \in V'$ and let $T \subseteq \Theta_+$ be such that $G(T \times V') > 0$.

For arbitrary $\varepsilon > 0$, partition V' into V_+^ε and V_-^ε, such that $x_+ > x_-$ for $x_+ \in V_+^\varepsilon$, $x_- \in V_-^\varepsilon$ and such that $M(V_-^\varepsilon) < \varepsilon$. Then modify the strategy s by setting, for $\theta \in T$:

$$s'(V_-^\varepsilon|\theta) = 0$$

$$s'(x|\theta) = s(x|\theta) \times \left(1 + \frac{s(V_-^\varepsilon|\theta)}{s(V_+^\varepsilon|\theta)}\right) \qquad \text{for } x \in V_+^\varepsilon.$$

This improves $\int u dG$ because $x^*(\theta) > x$ for $(\theta, x) \in T \times V'$. For ε sufficiently small and $x \in V_+^\varepsilon$, $\int v dH_s$ is still nonpositive, and thus still avoids attack. For $v \in V^{\varepsilon}_{-}$, $\int v dH_{s'} < \int v dH_s$. Thus s' improves $\int u dG$ and does not increase the probability of attack for any θ. ∎

$$\text{Let } T_I^- = \{\theta \in \Theta_- | s(I|\theta) > 0\}.$$

$$\text{Let } T_P^- = \{\theta \in \Theta_- | s(P|\theta) > 0\}.$$

LEMMA 5. *There exist* $\hat{\theta}_- \in \Theta_-$ *such that if* T_I^- *and* T_P^- *are nonnull, then* $\theta \in T_I^-$ *implies* $\theta \leqq \hat{\theta}_-$ *and* $\theta \in T_P^-$ *implies* $\theta \geqq \hat{\theta}_-$.

Proof. If the lemma were false, there would exist nonnull sets $T_P \subseteq T_P^-$ and $T_I \subseteq T_I^-$, and $\alpha > 0$ such that $\theta \in T_P$ and $\theta' \in T_I$ imply $\theta < \theta' - \alpha$ almost everywhere.

From the previous lemmas we know that, almost everywhere, for $\theta \in T_I$, $s(\cdot|\theta)$ has an atom at $x^*(\theta)$. Moreover, there is no other $\theta \in \Theta$ that has an atom at this value.

Consider the distribution of x given $\theta \in T_P$,

$$\mu(x|T_P) = \frac{\int_{T_P} s(x|\theta) dF(\theta)}{F(T_P)}.$$

As $\theta \in T_P$ is pooled with positive probability, there is a μ nonnull set, P', of pooled values. By Lemma 4 all pooled values have $\int v dH = 0$, hence, for all $x \in P'$, $H(\Theta_+|x) > 0$.

There are now two cases according to whether or not $\mu(P_+) > 0$, where

$$P_+ = P' \cap \{x | x \geqq x^*(\theta), \text{ for all } \theta \in T_I\}.$$

(see Fig. A.4).

Let us consider first the case where $\mu(P_+) > 0$. In this instance, G will assign positive mass to the rectangle $R_+ = T_P \times P_+$.

An improvement in s can be made by assigning some of the mass in R_+ to X^*, removing the same amount of mass from $X^* \cap \{(\theta, x) | \theta \in T_I^-\}$, and distributing it over $R'_+ = T_I \times P_+$ in such a way that M is unchanged. Because u is concave in x and $u_{x\theta} > 0$ we know that $\int\int u dG$ is improved, and at the same time A and $\int_\Theta \int_A w dG$ are invariant.

In the case $\mu(P_+) = 0$, G will assign zero mass to R_+. There will exist $\hat{\theta} \in T_I$ such that G will assign positive mass to $R_- = T_P \times (P' \cap \{x | x < x^*(\hat{\theta})\})$ (see Fig. A.5).

As R_- consists almost surely of pooled values, and $\int v dH = 0$ for all pooled values, G must assign positive mass to the rectangle $K = \{(\theta, x) | \theta \in \Theta_+, x < x^*(\hat{\theta})\}$. An improvement can be made, by virtue of convexity, by pooling some of the mass in K with values of θ in T_I such that $\theta > \hat{\theta}$. This would allow an increase in x toward X^* with positive probability. ∎

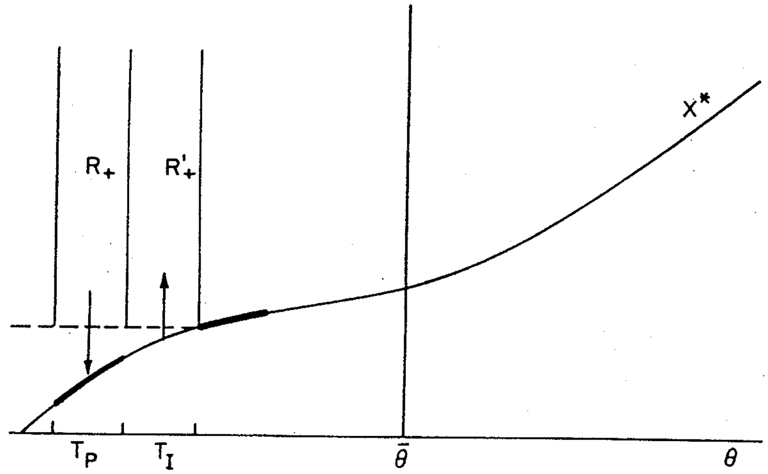

FIGURE A.4

On the right of $\bar{\theta}$ the situation is much the same, and we omit the proof. Define

$$T_I^+ = \{\theta \in \Theta_+ | s(I|\theta) > \theta\}$$
$$T_P^+ = \{\theta \in \Theta_+ | s(P|\theta) > \theta\}.$$

LEMMA 6. *There exists $\hat{\theta}_+$ such that if T_I^+ and T_P^+ are nonnull, then $\theta \in T_P^+$ implies $\theta \leqq \hat{\theta}_+$, and $\theta \in T_I^+$ implies $\theta \geqq \hat{\theta}_+$.*

Thus we have Fig. A.6.

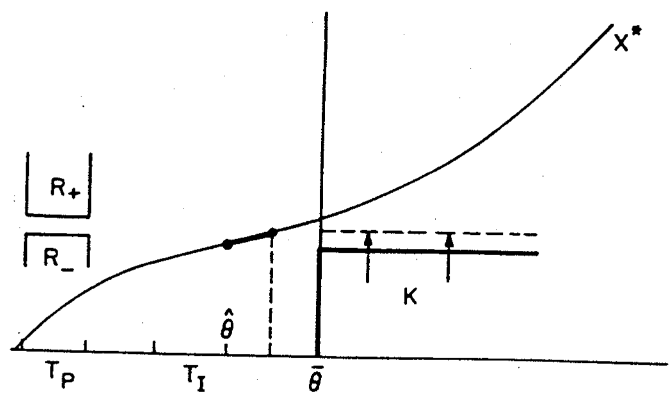

FIGURE A.5

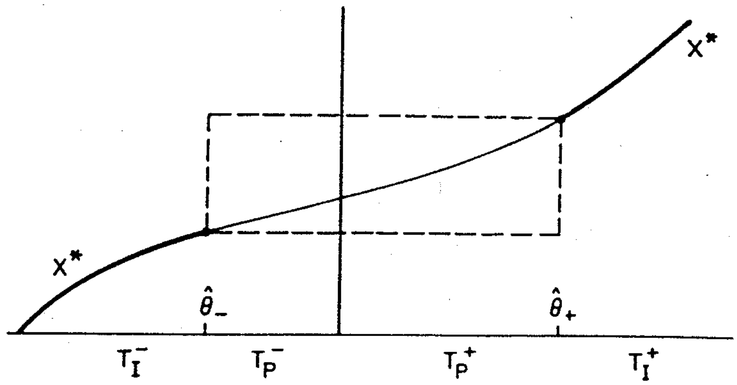

FIGURE A.6

Let

$$\overline{T} = (\hat\theta_-, \hat\theta_+]$$
$$\overline{P} = [x^*(\hat\theta_-), x^*(\hat\theta_+)].$$

LEMMA 7. $G(\overline{T} \times \overline{P}) = F(\overline{T})$.

If $G(\overline{T} \times \overline{P}) < F(\overline{T})$, then either $G(\overline{T} \times (x^*(\hat\theta_+), \infty)) > 0$ or $G(\overline{T} \times (-\infty, x^*(\hat\theta_-))) > 0$. In either case, the strategy can be improved by moving the corresponding mass toward $\overline{T} \times \{x^*(\hat\theta_+)\}$ or $\overline{T} \times \{x^*(\hat\theta_-)\}$.

We now show that on $\overline{T}$, s is nonstochastic and monotonic increasing over $\overline{T}_P$ and T_P^+.

LEMMA 8. *Let* $A = \{(\theta, x) \in \overline{T} \times \overline{P} \mid \theta > \bar\theta, x > x^*(\theta)\}$ *and* $B = \{(\theta, x) \in \overline{T} \times \overline{P} \mid \theta < \bar\theta, x < x^*(\theta)\}$. *Then* $G(A) = G(B) = 0$.

Proof. We show $G(A) = 0$, as the proof for B is completely analogous. If $G(A) > 0$, then $G(A') > 0$, where $A' = \{(\theta, x) \in \overline{T} \times \overline{P} \mid \theta < \bar\theta, x > x^*(\bar\theta)\}$. This is because $\int v dH = 0$ for all $x \in \overline{P}$. But then the strategy could be improved by moving a positive mass in A' and A downward toward X^*, in such a way as to maintain $\int v dH = 0$ (see Fig. A.7). ■

LEMMA 9. On $\overline{T}$, s is almost surely nonstochastic and monotonic nondecreasing over T_P^- and T_P^+.

Proof. Let us consider T_P^-. If s is stochastic or if s is nonstochastic but decreasing over the nondegenerate part of the domain, then there exist $\tilde\theta \in T_P^-$ and $\tilde x \in \overline{P}$ and rectangles C_1 and C_2 such that

$$(\theta, x) \in C_1 \textit{ implies } \theta < \tilde\theta \text{ and } x > \tilde x$$
$$(\theta, x) \in C_2 \textit{ implies } \theta > \tilde\theta \text{ and } x < \tilde x$$

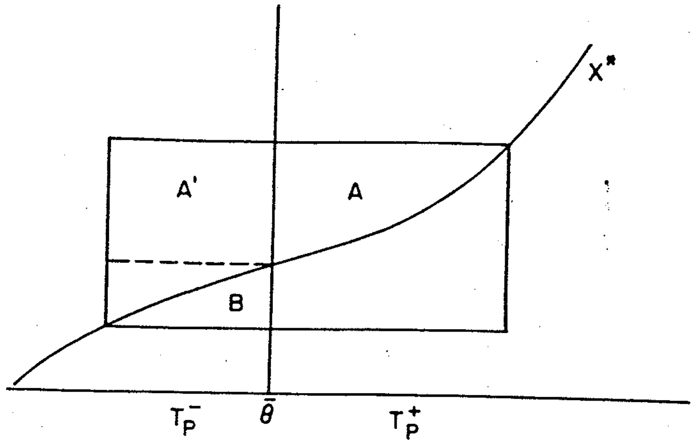

FIGURE A.7

and with $G(C_1)$ and $G(C_2) > 0$. For any $\varepsilon > 0$, we can find subrectangles of C_1 and C_2, denoted D_1 and D_2, with positive mass and such that ε exceeds their diameters (see Fig. A.8).

Let these rectangles be given by the products

$$D_1 = [\theta_{11}, \theta_{12}] \times [x_{11}, x_{12}]$$

$$D_2 = [\theta_{21}, \theta_{22}] \times [x_{21}, x_{22}].$$

Let $\alpha = (x_{12} + x_{11})/2 - (x_{22} + x_{21})/2$ be the distance between the centers of D_1 and D_2 in their x-coordinate, denoted respectively x_1 and x_2.

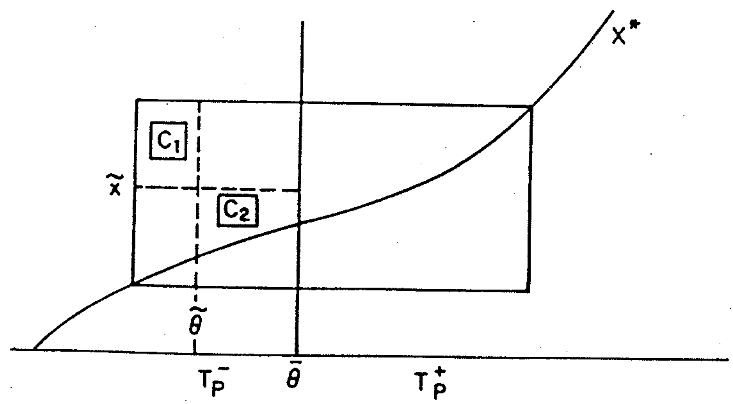

FIGURE A.8

Consider a pair of distributions on D_1 and D_2 with equal mass and dominated by G. Denote them by ψ_1 and ψ_2. Let $\hat{G} = G - (\psi_1 + \psi_2)$.

We consider a modification of the strategy that will be shown to be beneficial. It involves moving the distribution ψ_1 downward and ψ_2 upward, and leaving the residual $\hat{G}$ unchanged. This modification is now described in three steps.

Step 1. Concentrate the distributions ψ_1 and ψ_2 on the segments $\{(\theta, x) \in D_1 | x = x_1\}$ and $\{(\theta, x) \in D_2 | x = x_2\}$, respectively. This will result in a loss of at most $2\psi_1(D_1) \cdot \varepsilon/2 \cdot \bar{u}_x$, where $\bar{u}_x = \sup_{(\theta,x)\in D_1} |u_x|$. Let the resulting distributions be denoted $\psi_{1\theta}$ and $\psi_{2\theta}$; they are just the marginal distributions of ψ_1 and ψ_2 over Θ.

Step 2. Translate the resulting distributions downward and upward, respectively, by the distance α. This changes the utility by

$$\int_0^{\alpha} \left[-\int_{\theta_{11}}^{\theta_{12}} u_x(\theta, x_1 - \xi) d\psi_{1\theta} + \int_{\theta_{21}}^{\theta_{22}} u_x(\theta, x_2 + \xi) d\psi_{2\theta} \right] d\xi.$$

Note, however, that $x_1 - x_2 = \alpha$, so that for each ξ, $x_1 - \xi = x_2 + \alpha - \xi$. Therefore the change in utility can be written

$$\int_{x_1}^{x_2} \left[\int_{\theta_{11}}^{\theta_{12}} u_x(\theta, x) d\psi_{1\theta} - \int_{\theta_{21}}^{\theta_{22}} u_x(\theta, x) d\psi_{2\theta} \right] dx.$$

For each $x \in [x_2, x_1]$ the bracketed expression can be bounded above by

$\psi_1(D_1)(\theta_{21} - \theta_{12}) u_{\theta x}$,

where $u_{\theta x} = \{\inf u_{\theta x}(\theta, x) \colon \theta \in [\theta_{11}, \theta_{22}] \text{ and } x \in [x_{21}, x_{22})\}$.

Thus the change in utility from the translations defined in this step is bounded below by

$$\alpha \psi_1(D_1)(\theta_{21} - \theta_{12}) \mathbf{u}_{\theta x}.$$

Step 3. Redistribute the mass which has not been shifted to the interval $[(\theta_{11}, x_2), (\theta_{12}, x_2)]$ in such a way that its marginal distribution over x duplicates the marginal distribution of the original D_2. Likewise for the other segment and D_1. As in Step 1, since these involve movements of at most $\varepsilon/2$, the loss is bounded by $\psi_1(D_1) \cdot \varepsilon/2 \cdot \bar{u}_x$.

Clearly, as ε can be taken arbitrarily small, the gain obtained in Step 2 can be made to outweigh the potential losses in Steps 1 and 3. ■

LEMMA 10. *On T_P^- and T_P^+, s is strictly increasing.*

Proof. Suppose, to the contrary, that s is constant, x, over a nondegenerate subinterval $T_x^= \subseteq T_P^-$. Then s must also be concentrated at x over a subinterval $T_x^+ \subseteq T_P^+$. This strategy can be improved upon following a method analogous to that used in the last lemma, in Steps 1 and 2:

First observe that if s is optimal then the level of x cannot be advantageously varied.

Thus

$$\int_{T_P^+ \cup T_P^-} u_x(x, \theta) dF(\theta) = 0.$$

Because $u_{x\theta} > 0$, we can find a pair of subintervals of equal mass, $\hat{T}_x^- \subseteq T_x^-$ and $\hat{T}_x^+ \subseteq T_x^+$ such that

$$|u_x(x, \theta)| > |u_x(x, \theta')| \qquad \text{for } \theta \in \hat{T}_x^-, \theta' \in \hat{T}_x^+.$$

Then, by changing s to have slightly lower common value of $\hat{T}_x^- \cup \hat{T}_x^+$ the payoff can be improved. ∎

Proof of Theorem 1. Condition (1) follows from Lemmas 3, 5, 6. Conditions (2), (3) follow from Lemma 10.

To prove (4), we use Lemma 4 to write

$$v_- \int_{\theta_-} dF(\theta|x) + v_+ \int_{\theta_+} dF(\theta|x) = 0 \qquad \text{for } x \in P.$$

From Lemma 9 we know that s is almost surely nonstochastic, and that $F(\theta|x)$ is concentrated on two values, $\theta = y^{-1}(x)$ and $\tilde{\theta} = z^{-1}(x)$. From Lemma 10, y and z are increasing. Therefore the two integrals in the last equation are almost everywhere the densities $f(\theta)/\dot{y}(\theta)$ and $f(\tilde{\theta})/\dot{z}(\tilde{\theta})$.

This completes the proof of (4).

Part (5) follows from the monotonicity guaranteed in Lemma 9 and the domains of definition of y and z proven in parts (2) and (3). ∎

REFERENCES

BANKS, J., AND SOBEL, J. (1987). "Equilibrium Selection in Signaling Games," *Econometrica* **55**, 647–661.

CHO, I., AND KREPS, D. M. (1987). "Signaling Games and Stable Equilibria," *Quart. J. Econ.* **102**, 179–221.

FUDENBERG, D., AND LEVINE, D. (1989). "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica* **57**, 759–778.

KOHLBERG, E., AND MERTENS, J. F. (1986). "On the Strategic Stability of Equilibria," *Econometrica* **54,** 1003–1037.

KREPS, D., AND WILSON, R. (1982). "Reputation and Imperfect Information," *J. Econ. Theory* **27,** 253–279.

MILGROM, P., AND ROBERTS, J. (1982). "Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis," *Econometrica* **50,** 443–460.

PONTRIAGIN, L. S. (1962). *Ordinary Differential Equations* (Translated by Leonas Kacinskas and Walter B. Counts). Reading MA: Addison–Wesley.

SPENCE, M. (1974). *Market Signaling*. Cambridge: Harvard Univ. Press.