91-F-3

ON BELIEF BASED REFINEMENTS IN SIGNALING GAMES

by

George J. Mailath
University of Pennsylvania

Masahiro Okuno-Fujiwara
University of Tokyo and University of Pennsylvania

and

Andrew Postlewaite
University of Pennsylvania

January 1991

ON BELIEF BASED REFINEMENTS IN SIGNALING GAMES*

by

George J. Mailath
University of Pennsylvania

Masahiro Okuno-Fujiwara
University of Tokyo and
University of Pennsylvania

and

Andrew Postlewaite
University of Pennsylvania

January 28, 1991

*This work was begun while Postlewaite was visiting University of Tokyo under a grant by the Japan Society for the Promotion of Science. Their support and support from the National Science Foundation is gratefully acknowledged. This is an extension and revision of an earlier paper by Okuno-Fujiwara and Postlewaite [1987] that was circulated under the title "Forward Induction and Equilibrium Refinement," CARESS Working Paper #87-01. We are grateful for comments made by Jeff Banks, Ken Binmore, Georg Noldeke, Motty Perry, and Joel Sobel.

1. INTRODUCTION

There are many economic problems which, when modelled as games of incomplete information, give rise to many (often infinitely many) sequential equilibria. This multiplicity of equilibria severely limits the usefulness of the model in understanding the underlying economic problem and in predicting the outcome that would arise should the underlying problem change slightly. For some problems however, many of the equilibria seem implausible because of the beliefs associated with some disequilibrium information sets (i.e., disequilibrium beliefs). A number of refinements of the set of sequential equilibria have been proposed that attempt to formalize plausible restrictions on the disequilibrium beliefs (see, e.g., McClennan [1985], Kreps [1984], Cho and Kreps [1987], Banks and Sobel [1986], Cho [1986], Farrell [1985] and Grossman and Perry [1986]). While some of these papers deal with general games, many restrict attention to a set of games that motivated much of the interest among economists in refinements, signalling games. In this paper, we also restrict attention to this class of games.

Many of the refinements mentioned above are based on the idea that some disequilibrium beliefs are implausible in that they put positive probability on some types of player I which, it is argued, are unlikely to send this disequilibrium message. Many of these refinements seem intuitively appealing and, for some examples at least, select equilibria that seem plausible. The aim of this paper is to suggest that both the logical foundations of these refinements and their performance for some games of greatest interest to economists are problematic. The logical difficulties we see stem from the difficulty in trying to specify the reasonableness of beliefs (or restrictions on them) separately from specifying the beliefs along the equilibrium path. The difficulties with the performance of the various refinements are due both to their lack of continuity with respect to perturbations of the prior beliefs and the efficiency properties of the equilibria they select.

The recent history of refinements in signaling games has been almost exclusively concerned with concepts that select the Riley separating equilibrium[1] (Grossman and Perry [1986], Farrell [1985], and Engers and Schwartz [1987] being the exceptions). While we hesitate to add to the already formidable literature on refinements of Nash equilibria, we believe the issues raised are significant enough to warrant another paper. We do not intend the concept introduced here to be a definitive description of play. We view a refinement as being a candidate that satisfies a restriction which we believe should be satisfied in order to be a plausible description of play. A refinement can be reasonable in some settings and

---

[1]The separating sequential equilibrium outcome that is Pareto undominated (according to the informed agent's payoffs) by any other separating equilibrium outcome is often called the Riley outcome (after Riley [1979]).

unreasonable in others. We do *not* claim that only undefeated equilibria (the solution concept introduced here) need be considered. Rather, we think that undefeated equilibria deserve study. In particular, where a (semi-)pooling equilibrium is undefeated, the current concentration on the Riley separating equilibrium is too narrow, neglecting behavior which may be more economically intuitive.

The search for *the* solution concept--one that applies to *all* games--is a search that is doomed to failure. While *conceptually* there is such a solution concept (e.g., a list of special cases and rules that apply to each of these cases), this conceptual solution is unsatisfactory. A general solution concept is of value precisely because it applies a unifying principle to the analysis of diverse phenomena. There will necessarily be different forces at work in different settings that determine what descriptions of play or outcomes are plausible. This paper is meant to provide additional understanding of these forces.

The paper is organized as follows. In the next section, we will discuss in some detail what we think are the difficulties with the logical foundations and the selections made by refinements that treat disequilibrium messages as signals.[2] In the following section, we present the formal model and the definition of an alternative refinement, *undefeated equilibria*, that addresses some of these difficulties. Section 4 contains proofs of the existence theorem for undefeated equilibrium. The set of undefeated equilibria is shown to have a continuous selection in Section 5. Section 6 discusses the relationship of this paper to the refinements literature and Section 7 provides concluding remarks.

2. FORWARD INDUCTION

In this paper we restrict attention to signalling games, a class of games that has been extensively analyzed and utilized in a wide range of economic problems. A signalling game is a game in which there are only two players, a sender and a receiver, who will be denoted player I and player II respectively. Player I possesses private information, modelled by identifying a type for player I with each different piece of information he might have. On the basis of his type, he sends a message to II. Player II, not knowing the true type of player I, upon observing the message player I sends takes an action (or reply). The payoffs to each player are determined by player I's true type (his private information), his message, and player II's action.

---

[2]In the context of a specific example, Engers and Schwartz [1987] have independently made some of the arguments made here.

A signalling game that has been extensively analyzed and has motivated much of the work on refinements is due to Spence [1973, 1974]; we will present a simple version of that model to illustrate some of the ideas and concepts that are important to this paper.[3]

Consider a world in which there are workers who (exogenously) have either high or low ability, and further that there are equal numbers of each ability. We will use t to denote the level of ability, with $t = 1$ for low ability and $t = 2$ for high. A worker of type t is worth exactly t to any firm. We assume that firms compete to hire workers and further, that there these firms operate in a competitive environment and make zero profits in equilibrium. Firms do not know an individual's ability, but know that the population is equally divided between workers of each ability level. Each individual knows his own ability, but cannot verify this to anyone else.

In the absence of any method of distinguishing workers, all workers will be paid an identical wage equal to the average ability of the workers in the population. Assume that workers differ in one respect, however: suppose that there is the possibility of acquiring education and that the disutility of acquiring education is lower for high ability workers than for low. More specifically, assume a worker's utility function over wage and education is given by $u_t(w,e) = w - e/t$.

The problem is then easily represented as a signalling game with a worker being the type I player with private information (his ability) who sends a message (chooses a level of education) to the firm. The firm -- player II -- takes an action after observing the education level: choose a wage to offer the worker as a function of the message sent. The firm is assumed to offer a wage equal to the worker's expected productivity; otherwise other firms can profitably lure away the worker by offering a higher wage.[4] As is well known, there are many Nash equilibria in this game. There is a continuum of separating equilibria, that is, equilibria in which the two types of workers choose different messages. This set can be characterized by a pair of education levels chosen by the two types (e_1,e_2), $e_1 \neq e_2$, $e_1 \leq 1$, $e_2 \leq 2$, $1 \leq e_2-e_1 \leq 2$ and a wage function of $w(e_1) = 1$, $w(e_2) = 2$ and $w(e) = 0$ for $e \neq e_1,e_2$.

Some of these equilibria seem implausible, namely those in which $e_1 > 0$. Here, a low ability worker is choosing a positive level of education because any reduction leads to a 0 wage, given the firm's strategy. But regardless of a worker's ability, a worker with $e = 0$ is still worth at least 1 to

---

[3]In our presentation that follows, we have drawn shamelessly from Kreps' [1990] presentation.

[4]Strictly speaking, we should model this as a game with two uninformed players (firms). The two firms would then engage in a first price auction for the worker. A single uninformed agent with payoffs $-(t-w)^2$ yields similar behavior on the part of the uninformed firm.

the firm. As a result, these equilibria are not sequential (Kreps and Wilson [1982]); only those separating equilibria that have $e_1 = 0$ can be sequential.

In addition to these separating equilibria, there is also a continuum of pooling equilibria, that is, equilibria in which both high and low workers are choosing the same level of education. Some, however, are not sequential equilibria by the same sort of argument as above. There is still a continuum of pooling sequential equilibria with $e_1 = e_2 = \bar{e}$, $0 \leq \bar{e} \leq ½$ and $w(\bar{e}) = 1½$, with $w(e) = 1$ for any $e \neq \bar{e}$; that is, the disequilibrium beliefs are that any education level $e \neq \bar{e}$ is chosen by a low ability worker.

Thus, even though the sequential equilibrium concept "refines away" some of the equilibria that seem unreasonable, there is still a large set of equilibria that remains. Further, the qualitative characteristics of the equilibria in which we might be interested differ substantially. For example, the $\bar{e} = 0$ pooling equilibrium Pareto dominates the other pooling equilibria.

Despite "passing" the test imposed by sequentiality, some of the equilibria described above seem less plausible than others. For example, consider the separating equilibrium with $e_1 = 0$ and $e_2 = 1.9$. Suppose the firm now sees an education level of 1.8. As mentioned above, the given equilibrium is a sequential equilibrium supported by beliefs this disequilibrium education level is chosen by a low ability worker. But is this reasonable? In equilibrium, a low ability worker received utility 1; choosing an education level of 1.8 would result in a utility no greater than 0.2 regardless of the firm's choice of action (since the highest wage the firm would pay for any possible beliefs it had about the worker's ability is 2). Since the firm will pay at most 2 for any worker drawn from a population with abilities of 1 or 2, only a high ability worker could possibly be better off following a choice of education 1.8. Hence, perhaps it is only reasonable that the firm should assume that the worker is high ability. Such beliefs, of course, result in a wage offer of 2, upsetting the proposed equilibrium.[5]

The idea that a player should take into account the fact that a disequilibrium move might have a particular effect on other players' beliefs about some aspect of the game is sometimes referred to as *forward induction* (see Kohlberg and Mertens [1986]). In the words of Kohlberg and Mertens, a disequilibrium move and the resulting subgame should be considered as a "very specific form of pre-play communication." A disequilibrium move should be interpreted as the player [effectively] sending the following message to other players:

---

[5]If (as in footnote 4) the firm is modelled as player II with payoff $-(t-w)^2$, then any wage offer not in [1,2] is strictly dominated. The preceding argument shows that the separating sequential equilibrium with $e_2 = 1.9$ does not survive iterative elimination of dominated strategies.

"Look, I had the opportunity to play the equilibrium strategy, and nevertheless I decided to play this move, and my move is already made. We both know that you can no longer talk to me, because we are in the game, and my move is made. So think now well, and make your decision."

While the concept of forward induction is appealing, one of the aims of this paper is to argue that there are serious difficulties in implementing forward induction in a formal equilibrium refinement. To illustrate some of the difficulties, we will consider a refinement of sequential equilibrium proposed by Grossman and Perry [1986], perfect sequential equilibria, that places restrictions on the beliefs that an agent could hold at disequilibrium information sets. Roughly speaking, Grossman and Perry restrict an agent finding himself at an information set which should not have been reached during the play of the game to "try to interpret the move as a signal by the player I." They test a given sequential equilibrium in the following manner. For each information set which is not reached in the given equilibrium, player II hypothesizes that the move was made by some set of types of player I and revises his prior according to Bayes' rule conditional upon player I being in the specified set of types. If his best response given these beliefs is preferred by precisely the prespecified set of types, the given sequential equilibrium is said to fail the Grossman-Perry test.

The idea that if a disequilibrium message can be interpreted as a signal from some set of types of the player with private information that would be better off by sending this disequilibrium message than they would have been by following the proposed equilibrium strategy is attractive. The difficulty is that some types may have more than one such disequilibrium message they could send. If players are fully rational such a type would presumably send that disequilibrium message that yields the highest utility among the disequilibrium messages that increase his utility. But this means that a receiver cannot assess the plausibility of interpreting a particular disequilibrium message in isolation; the types that might want to send such a message depends upon how other disequilibrium messages will be interpreted. The following example illustrates the difficulty.

Consider the game in Figure I. Player I has three types, each type having prior probability ⅓. Player I has four moves while player II has three. One equilibrium of this game is that player I plays strategy 4 regardless of his type, giving both player I and II payoffs of 2. A set of disequilibrium beliefs for player II which support this as an equilibrium have player II believing that strategy i is played by type i of player I and playing strategy i as his best response. These beliefs do not satisfy the Grossman-Perry test. If player II observes strategy 1, he could conjecture that it was played by types 1 and 3, giving rise to a probability distribution over the types of (½,0,½), that is, it is equally likely that the types 1 or 3 played this strategy but never type 2. With these beliefs, II's best response is to play strategy 2. The set of types who prefer this outcome to the proposed equilibrium

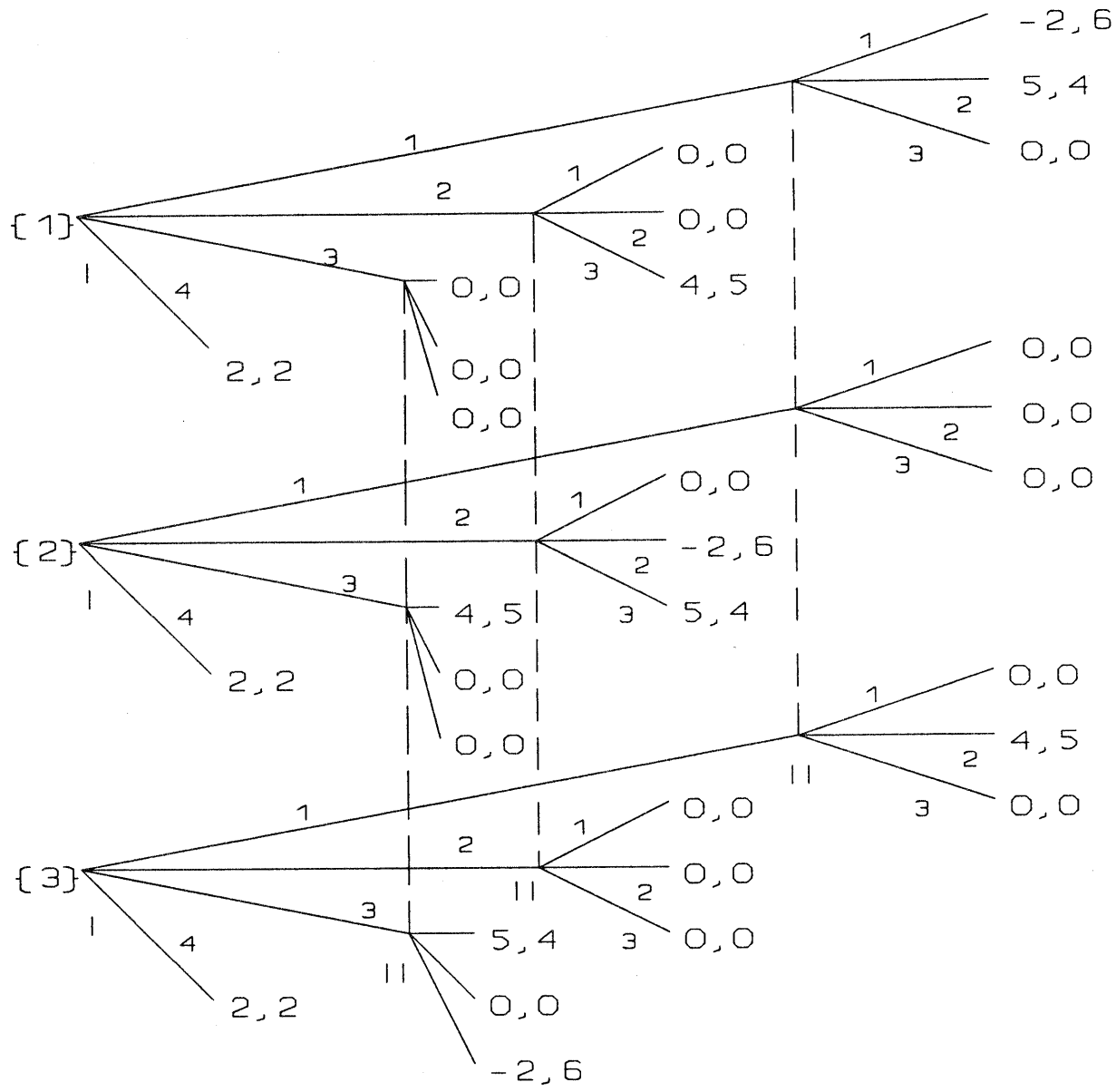

FIGURE 1

consists of types 1 and 3, the two types with positive probability in the proposed beliefs. Type 2 would strictly prefer the outcome in the existing equilibrium to this outcome. These beliefs, in the case that strategy 1 is played, rule out the beliefs that supported the original equilibrium described above. Since the posterior of (½,0,½) is the only credible posterior for player II, the proposed equilibrium is not perfectly sequential.

In a similar manner, player II could conjecture upon seeing strategy 2 that it was chosen by types 1 and 2 which would lead him to a probability distribution (½,½,0) in the case that strategy 2 was

chosen by player I. With these beliefs his optimal choice is strategy 3, which is preferred to the existing equilibrium by types 1 and 2 but less preferred by type 3. Thus the beliefs in the proposed equilibrium which follow player I's choosing the non-equilibrium strategy 2 also fail the Grossman-Perry test. Similarly, the beliefs of player II given strategy 3 by player I fail the test via a conjecture that this strategy is played by types 2 and 3.

But now suppose that player II were to follow the suggested logic. If the logic is compelling, player I ought to be able to calculate in the same way as player II does. What would player I do? Since disequilibrium strategies 1, 2, and 3 lead player II to choose strategies 2, 3, and 1 respectively, type 1 should play strategy 1 because this leads to the highest possible utility. It is true that playing strategy 2 also leaves type 1 better off than in the existing equilibrium, but we should assume that he would optimize. Similarly, type 2 would play strategy 2 and type 3 would play strategy 3. In summary, player II's original disequilibrium beliefs were ruled out by the Grossman-Perry test because of the existence of other "self-fulfilling" beliefs on II's part. *But*, and this is the crucial step, if II were to adopt these beliefs and behave optimally given these beliefs, the optimal behavior on I's part preceding II's move would not support these revised beliefs, but rather, would support II's original beliefs. Player II, who can calculate these changes in player I's choice of optimal strategies, might not be persuaded to change his original beliefs.

The argument above suggests that the mere existence of other "self-fulfilling" beliefs in the sense of Grossman-Perry is not enough to convince player II to change his beliefs, because changing his beliefs is bound to create further adjustments in player I's choice of strategies (if the analysis of the game is common knowledge), which requires further revisions in player II's beliefs, creating yet further change in I's optimal strategies and so on. If both players I and II are rational and if it is common knowledge that they are rational, player II would change his disequilibrium beliefs only when all types of the prescribed set of player I types prefer to choose the strategy after taking account of all the subsequent adjustments that will take place once such revisions in disequilibrium beliefs are made. In other words, the beliefs that are suggested by the refinement should be a steady state of the procedure which implied those beliefs.

We are not the first to point out potential logical difficulties in interpreting disequilibrium messages as signals. Cho and Kreps [1987] proposed a refinement of sequential equilibrium, the intuitive criterion, that was motivated by the idea that disequilibrium messages might be interpreted as signal. In their paper they mention a criticism by Stiglitz to that refinement that is in the spirit of the above example.

The intuitive criterion, however, differs from the perfect sequential refinement in several important respects. First, unlike perfectly sequential equilibria, the set of equilibria that pass the intuitive test is always nonempty for the set of signalling games. Second, although the intuitive criterion is motivated by, and defined in terms of, restrictions on the beliefs that the receiver might have following a disequilibrium message, it is a subset of the set of *stable* equilibria (Kohlberg and Mertens [1986]) which is defined for all finite games rather than only signalling games, and is defined without reference to restrictions on beliefs.

These features (among others) have lead to the application of the intuitive criterion to a number of economic problems. Before proceeding with a discussion of the merits of the intuitive criterion as a refinement of sequential equilibria for signalling games, we describe it for the case of two types.

Fix a sequential equilibrium and let m be an associated disequilibrium message (a message which no type of player I will send in the equilibrium.) Suppose there is a type (say t') who would never want to send this message because, regardless of the beliefs player II would form upon observing this message and regardless of the optimal response II would subsequently choose, this type would obtain a smaller payoff than had the original equilibrium been played. Suppose further that if II believes that m was chosen by the other type, then any best response by II yields a higher payoff than the original equilibrium payoff to that type. Then the original equilibrium is said to *fail the intuitive criterion*. Clearly, the original equilibrium must have been sustained because disequilibrium beliefs associated with m are not consistent with the above observation. Specifically, the original disequilibrium belief must necessarily have put positive probability on the type t'. Cho and Kreps [1987] show that the intuitive criterion selects the Riley separating equilibrium--the separating sequential equilibrium in which $e_2 = 1$.

This refinement is sometimes motivated as follows. Suppose we have a sequential equilibrium that fails the intuitive criterion, and suppose m is a disequilibrium message associated with the failure. Then the given equilibrium will "collapse" if some type of player I sends this message and if the message is interpreted as the sender making the following speech to player II. "I am sending disequilibrium message m and you should believe that I am of type t. If I were the other type t', I would never have sent this message because no matter how you interpret this speech I would have obtained higher payoff had I not sent this message. Moreover, any interpretation of this message which excludes this other type will lead to a higher payoff for the type t, which is the type I claim to be."[6]

---

[6]Paraphrased from Cho and Kreps [1987].

If the disequilibrium message were interpreted in this way, then the behavior suggested to the sender by the original sequential equilibrium is not likely to be realized.

This motivation is subject to a criticism similar, albeit more subtle, to that made in reference to perfect sequentiality above. Suppose that we have a particular sequential equilibrium that fails the intuitive test. If the receiver finds the intuitive criterion compelling, he would restrict his beliefs in a way that alter his action at some disequilibrium information sets. This, in turn would alter the message sent by some type(s) of player I, if player I is aware that player II finds the intuitive criterion compelling. In essence, the intuitive criterion stops here, concluding that this eliminates the candidate equilibrium from serious consideration. But, if player II knows that (some types of) player I will alter their messages, he should update his beliefs about player I's type if player I *does* send the equilibrium message. But if player II does this, in the determination of whether a particular type might benefit from sending some disequilibrium message, the relevant comparison is *not* with the utility that he would receive in the proposed equilibrium, but rather the utility that he would receive given that player II is thinking in this way.

To illustrate this criticism, consider the pooling equilibrium with education level $e = 0$ in the Spence game when $Pr(t = 1) = p < ½$; both types of worker receive an equilibrium payoff of 2-p. This equilibrium fails the intuitive criterion: Any education level $e^* > ½$ yields a payoff to the low productivity worker less than 1½. Furthermore, if the firms believe it is the high productivity worker sending e, the high productivity worker receives a payoff greater than 1½ for any $e \in (½,1)$. However, if all players understand the logic of the intuitive criterion, then the low productivity worker should expect that the firms, upon seeing the equilibrium and *not* a deviation, will conclude the worker has low productivity. Thus the low productivity worker is not guaranteed its equilibrium payoff. Rather, the low productivity worker would receive a payoff of 1, which is *less* than the payoff from deviating to e, when it is interpreted as a signal that the worker has high productivity. But this means that $e \in (½,1)$ can no longer be viewed as an unambiguous signal that the worker is of high productivity.[7,8]

Notice that this criticism also has the flavor that the out-of-equilibrium beliefs that are supposed to be motivated by forward induction logic lack a "global" consistency. Which education levels, then,

---

[7]Cho and Kreps [1987] respond to this by arguing that it is enough to cast doubt on the pooling equilibrium as a common knowledge description of self-enforcing play, and this has been done. Since the criticism undermines the original contention that a particular deviation is an unambiguous signal that the worker is of high productivity, we do not find this response convincing.

[8]It is interesting to note that this criticism is similar to the ideas motivating Riley's Reactive Equilibrium notion (Riley [1979]). See also Wilson [1977].

should be viewed as being unambiguous signals that the worker is of high productivity? One strong requirement is that the type who is believed to have not deviated should receive a smaller payoff from that message, no matter how it is interpreted, than from an alternative message which satisfies incentive compatibility. A very simple example of this is provided by the two type job market signaling game. The low productivity worker (weakly) prefers $\bar{e} = 0$ to $e \in (½,1]$, when $\bar{e}$ yields a wage of 1 and e yields a wage of 2 only when $e = 1$. It is important to observe that $e = 1$ is a separating equilibrium education choice for the high productivity worker, so that the pooling equilibrium is only eliminated if the high productivity worker earns a higher payoff in the separating equilibrium.

Besides the question of the logic of the motivation for the intuitive criterion, there is some question of the plausibility of the sequential equilibrium it selects in this simple game. As we pointed out, the pooling equilibria fail the intuitive test, leaving the separating equilibrium with $e_2 = 1$ as the single equilibrium passing the test. But this equilibrium yields a utility of 1 to the low productivity worker and 1½ to the high productivity worker, while the pooling equilibria with $\bar{e} = 0$ yields each worker utility 2-p (while leaving player II, the firm, equally well off). A crude but appealing refinement would be to select the equilibrium that Pareto dominates the other equilibria, if such an equilibrium exists, but this is clearly inconsistent with the intuitive criterion refinement.

Even if one rejects the Pareto refinement as being of little interest, there is still a question of the plausibility of the equilibria that the intuitive criterion selects for this simple game. Suppose we let the proportion of low ability workers in the population decrease. The set of separating equilibria will remain unchanged, with the equilibrium with $e_2 = 1$ being the sole separating sequential equilibrium passing the intuitive test. The pooling equilibria will change as the proportions of the two types changes, but they will still fail the intuitive criterion by the same reasoning as before.

While the equilibrium selected by the intuitive criterion remains unchanged for any positive proportion of low ability workers, the situation is quite different when there are *no* low ability workers. Then education serves no separating purpose and the only equilibrium of interest is the one in which the workers (who are all of high ability) all receive a wage of 2 while choosing education level 0.

But is it reasonable to believe that the outcome in an economy with 1,000,000 workers of high ability and 1 worker of low ability will differ significantly from that in an economy with only the 1,000,000 workers of high ability? If not, this discontinuity with respect to the probability distribution over the two types of player I is disturbing. For many problems, economists are not absolutely certain about the distribution of types in the problem they are studying; the models and the equilibria of those models can be useful only if the predicted outcome is not overly sensitive to the description of the environment.

We should note that we are discussing a type of continuity that differs from that in Kohlberg and Mertens' stability concept (Kohlberg and Mertens [1986]). The continuity considered there excludes perturbations in the support of the distribution over types. We will discuss this further below.

To sum up, while it may be appealing to consider disequilibrium messages as signals, the particular implementation given by the intuitive criterion is problematic from both a logical basis and on the basis of the resulting equilibrium selection. This does not mean that it is impossible to implement the underlying idea of treating disequilibrium messages as signals. The criticism was that, starting from a given equilibrium, adjusting the beliefs at some out-of-equilibrium information set cannot be done without simultaneously adjusting beliefs at other information sets, including some information sets on the equilibrium path. But we can make all these adjustments. Once all the subsequent adjustments are made, we must be at an equilibrium; if not some further adjustments should be contemplated. Hence, if players are to engage in the exercise of determining how to interpret disequilibrium messages as signals, and all players carry the forward induction to completion, we are led to a test as follows. Consider a proposed sequential equilibrium and an action for player I that is not played in the equilibrium. Suppose there is an alternative equilibrium in which some non-empty set of types of player I choose the given action and that that set is precisely the set of types who prefer the alternative equilibrium to the proposed equilibrium. The test requires player II's beliefs at that action in the original equilibrium to be consistent with this set. If the beliefs are not consistent, we say the second equilibrium *defeats* the proposed equilibrium. We study the properties of *undefeated equilibria* next.

3. UNDEFEATED EQUILIBRIA IN SIGNALING GAMES

3.1 MODEL AND DEFINITIONS

In this section we show that, in an economically important class of (continuous strategy set) signalling games, the set of undefeated equilibria is non-empty. We show this by proving that a particular sequential equilibrium, which we call the lexicographically maximum sequential equilibrium is always undefeated. The class of signalling games to which we restrict attention includes most of the asymmetric information situations analyzed in the literature, such as the signaling game of Spence, the insurance model of Rothschild and Stiglitz [1977], the limit pricing model (e.g., Milgrom and Roberts [1982]) and some litigation models (e.g., Banks and Sobel [1986]).

We restrict attention to pure strategy equilibria. This is for several reasons: While mixed strategies provide a nice mathematical structure for finite games through convex strategy spaces and represent the uncertainty that a player may have about the action choice of other players (this being

perhaps the best interpretation of mixed strategies), they are not an important element of the analysis of the class of games studied here. They are also not needed for existence and the uncertainty interpretation is not fundamental to the economics. Lastly, no new qualitative features are introduced by mixed strategy equilibria, and they are, for large degrees of uncertainty, "almost" pure (see Mailath [1990] for a precise discussion of this issue).

There are two players, I and II. We denote the set of types by $T = \{1,\ldots,n\}$, with the common knowledge prior probability of $t \in T$ given by $p(t)$. Player I, knowing his type chooses a message m from the set M. Player II seeing m, responds with a choice r from the set R. The payoff for type t of player I for the pair of moves $(m,r) \in M \times R$ is given by $u(m,r,t)$. The payoff to player II when I is of type t and (m,r) is chosen is given by $v(m,r,t)$. Player I's pure strategy is denoted $\mu: T \to M$ and player II's pure strategy is denoted $\rho: M \to R$. We will assume that v is strictly concave in r, so that player II will never randomize. The set of all probability distributions on a set X is denoted Δ_X. We write $\beta: M \to \Delta_T$ for II's belief function, assigning a probability distribution over T upon observing m, so that $\beta(t|m)$ is II's conditional belief that t sent m when he observes $m \in M$. Finally, we let BR: $M \times \Delta_T \to R$ denote the set of best responses to m given $\beta(m)$, i.e. $BR(m,\beta(m)) = \operatorname{argmax}_{r \in R} \sum_{t \in T} v(m,r,t)\beta(t|m)$.

Definition 1: $\sigma^* = (\mu^*,\rho^*,\beta^*)$ is a *pure strategy sequential equilibrium* if:

(D1.1) $\forall\, m \in M$ and $t \in T$: $\mu^*(m) \in \operatorname{argmax}_{m' \in M} u(m',\rho^*(m'),t)$;

(D1.2) $\forall\, r \in R$ and $m \in M$: $\rho^*(r) \in BR(m,\beta^*(m))$;

(D1.3) $\forall\, t \in T$ and $m \in M$: $\beta^*(t|m) = \dfrac{p(t)\mu^*(m|t)}{\sum_{t' \in T} p(t')\mu^*(m|t')}$ if the denominator is positive.

We denote the set of pure strategy sequential equilibria for the game G by PSE(G). With an abuse of notation we write $u(\sigma^*,t)$ for type t's expected payoff associated with σ^*.

Definition 2: $\sigma \equiv (\mu,\rho,\beta) \in \text{PSE(G)}$ *defeats* $\sigma' \equiv (\mu',\rho',\beta') \in \text{PSE(G)}$ if $\exists\, m \in M$ such that:

(D2.1) $\forall t \in T$: $\mu'(t) \neq m$, and $K \equiv \{t \in T \mid \mu(t) = m\} \neq \varnothing$;

(D2.2) $\forall t \in K$: $u(\sigma,t) \geq u(\sigma',t)$, and
$\exists t \in K$: $u(\sigma,t) > u(\sigma',t)$; and

(D2.3) $\beta'(t \mid m) \neq \dfrac{p(t)\pi(t)}{\sum_{t' \in T} p(t')\pi(t')}$ for any $\pi: T \to [0,1]$ satisfying
$t \in K$ and $u(\sigma,t) > u(\sigma',t) \Rightarrow \pi(t) = 1$, and
$t \notin K \Rightarrow \pi(t) = 0$.

Note that the expression in (D2.3) is the probability that t has sent the message, conditional on the sender's type being in K and allowing for the possibility that any type which is indifferent may have randomized. We say that $\sigma \in \text{PSE(G)}$ is *undefeated* if there does not exist $\sigma' \in \text{PSE(G)}$ that defeats σ. The strategy profile $\sigma \in \text{PSE(G)}$ *lexicographically dominates* (*ℓ-dominates*) $\sigma' \in \text{PSE(G)}$ if there exists $j \in T$ such that $u(\sigma,j) > u(\sigma',j)$ and for $t > j$, $u(\sigma,t) \geq u(\sigma',t)$.[9] The profile $\sigma \in \text{PSE(G)}$ is the *lexicographically maximum sequential equilibrium* (LMSE) if there does not exist $\sigma' \in \text{PSE(G)}$ that ℓ-dominates σ. Clearly, LMSE exist as long as PSE(G) is non-empty (as is the case here[10]). Finally, $\sigma \in \text{PSE(G)}$ is a *completely separating equilibrium* if $\forall t,t' \in T$, $\mu(t) \neq \mu(t')$ whenever $t \neq t'$.

Before describing the economic assumptions we impose, observe that the example described in the previous section illustrates that some assumptions are needed in order to get existence. In that example, there are three other sequential equilibria besides the one in which all three types of player I play strategy 4. Equilibrium 2 has types 1 and 3 of player I playing strategy 1 and type 2 playing strategy 4. The beliefs associated with the strategies not played in equilibrium are that strategy i is played by type i, i = 2,3. There are two other similar equilibria, equilibrium 3 in which types 1 and 2 play strategy 2 while type 3 plays strategy 4 and equilibrium 4 in which types 2 and 3 play strategy 3 while type 1 plays strategy 4. The disequilibrium beliefs are similar to those in the equilibrium described just above. None of these equilibria passes the test we proposed above. Equilibrium 2 defeats

---

[9]This particular ordering is justified by Assumption 2 below.

[10]See, for example, Mailath [1990].

equilibrium 3, equilibrium 3 defeats equilibrium 4, equilibrium 4 defeats equilibrium 2 and each of these defeats the equilibrium in which all types of player I play strategy 4.

We now confine our attention to the class of signalling games that satisfy the following four assumptions.

Assumption 1: Continuity and concavity.

(i) M and R are closed, convex subsets of $\Re$.

(ii) u and v are continuous in m and r.

(iii) v is strictly concave in r.

Remark: By A.1 (ii)-(iii), for any $q \in \Delta_T$, BR(m,q) is a single-valued continuous function on M. Hence u(m,BR(m,q),t) is well-defined.

Assumption 2: Stochastic dominance.

$\forall t \in T$, $\forall m \in M$, $\forall q,q' \in \Delta_T$, whenever q stochastically dominates q' (i.e., $\sum_{t' \leq t} q'(t') \geq \sum_{t' \leq t} q(t')$ $\forall t \in T$ and strict inequality holds for some $t \in T$), $u(m,BR(m,q),t) > u(m,BR(m,q'),t)$.

Assumption 3: Monotonicity.

$\forall m,m' \in M$, $\forall r,r' \in R$, $\forall t,t' \in T$, if

(i) $u(m,r,t) \geq u(m',r',t)$,

(ii) $m > m'$, and

(iii) $t' > t$, then

(iv) $v(m,r,t') > u(m',r',t')$.

For the next assumption, we need an additional definition. For any nonempty subset K of T, $q_K \in \Delta_T$ is called *the* K-*conditional belief* if:

$$q_K(t) = \begin{cases} \dfrac{p(t)}{\sum_{t' \in T} p(t')}, & \text{if } t \in K, \\ 0, & \text{otherwise.} \end{cases}$$

With an abuse of notation, we sometimes write the best response against m, when the belief is the K-conditional belief, by BR(m,K), i.e., $BR(m,K) = BR(m,q_K)$.

Assumption 4: Satiation.

$\forall t \in T$, $\forall m \in M$, and $\forall q \in \Delta_T$, if q is not the {n}-conditional belief and if $t \neq n$, then there exists $\bar{m}(m,BR(m,q),t) \in M$ such that for all $m' \geq \bar{m}(m,BR(m,q),t)$, $u(m',BR(m',\{n\}),t) < u(m,BR(m,q),t)$.

Assumption 1 is primarily a technical assumption. The second assumption states that all types of Player I prefer the (best) response of player II when player II believes I more likely to be of higher type. The third assumption is the "single-crossing" property. It says that if some type t prefers a message-response pair (m,r) to a second pair (m',r'), when m is greater than m', then any type higher than t will also prefer (m,r) to (m',r'). This captures the idea that higher messages are "easier" for higher types to send than for lower types. The last assumption, 4, says that sending very high messages is prohibitively costly in the sense that there is a message level such that no type, except possibly the highest type, would want to exceed even if the result was the most favorable possible beliefs on player II's part.

Theorem 1: Under A1-A4, the LMSE is undefeated.

Theorem 2: Under A1-A4, if the LMSE is completely separating, it is the only undefeated pure strategy sequential equilibrium.

Remark: If the LMSE is pooling, there may be multiple undefeated pure strategy sequential equilibria. A simple example is the Spence job market example with $u(e,w,t) = w - e/t$ and $T = \{1,2,3\}$. Let p_t be the probability of t. In the calculations we assume $p_1 = 0.35$, $p_2 = 0.20$, and $p_3 = 0.45$. The important inequalities are that $\frac{1}{3} < p_1 < p_3$ and $p_1 < \frac{1}{2}$. There are four PSEs which are candidates for lex max and undefeated: the separating equilibrium with education levels $\mu^1(.) = (0,1,3)$ and equilibrium utilities $u^1(.) = (1,1.5,2)$; the completely pooling equilibrium, with $\mu^2 = (0,0,0)$ and $u^2 = (2.10,2.10,2.10)$; the 1,2 pooling equilibrium, with $\mu^3 = (0,0,3.27)$ and $u^3 = (1.36,1.36,1.91)$; and the 2,3 pooling equilibrium with $\mu^4 = (0,1.10,1.10)$ and $u^4 = (1,1.85,2.13)$. While the 2,3 pooling is the

lex max outcome, the completely pooling outcome is undefeated. It is undefeated because while 3 prefers the 2,3 pooling outcome, 2 prefers the completely pooling outcome.

3.2 PROOFS

Lemma 1: (monotonicity implies reverse monotonicity) $\forall m,m' \in M$, $\forall r,r' \in R$, $\forall t,t' \in T$, if

(i) $u(m,r,t) \geq u(m',r',t)$,

(ii) $m < m'$, and

(iii) $t' < t$, then

(iv) $u(m,r,t') > u(m',r',t')$.

Proof: Suppose, contrary to the assertion, $u(m,r,t') \leq u(m',r',t')$. Then, by monotonicity, (ii) and (iii) imply $u(m',r',t) > u(m,r,t)$, contrary to (i). **Q.E.D.**

Lemma 2: $\forall m,m' \in M$, $\forall r,r' \in R$, $\forall t,t' \in T$, if

(i) $u(m,r,t) \geq u(m',r',t)$

(ii) $u(m',r',t') \geq u(m,r,t')$, and

(iii) $t' > t$, then

(iv) $m' \geq m$.

Proof: Suppose, contrary to the assertion, $m \geq m'$. Then by monotonicity, (i) and (iii) imply

$$u(m,r,t') \geq u(m',r',t')$$

contrary to (ii). **Q.E.D.**

Corollary: If $\sigma \equiv (\mu,\rho,\beta) \in PSE(G)$, then $\mu(t) \leq \mu(t')$ whenever $t \leq t'$.

Proof: Immediate from Lemma 2. **Q.E.D.**

Next we define games truncated from G through restricting player I's types to be a subset of the original set, T. Formally for any $j \in T$, let

$$T^j = \{1,\ldots,j\} \text{ and } p^j(t) = q_{T^j}.$$

A truncated game G^j is defined by substituting the subset T^j for T and the T^j-conditional belief p^j for p in the original game G. We shall denote the set of pure strategy sequential equilibria of G^j by $PSE(G^j)$.

The following property is important. Given any pure strategy equilibrium of the original game, $\sigma \in PSE(G)$, we define a j-truncated equilibrium σ^j for $j \in T$ by simply deleting those types higher

than j. It follows from the corollary to Lemma 2 that, as long as no type higher than j is sending the same message that j sends at σ, a j-truncated equilibrium σ^j is a pure strategy equilibrium of G^j.

The next lemma is the key lemma for our proof. In effect, we shall prove the following. Suppose we are given two pure strategy equilibria of the original game, σ and $\hat{\sigma}$. If for some $j \in T$ j's equilibrium payoff is no smaller at $\hat{\sigma}$ than at σ, then we can construct yet another pure strategy equilibrium in the $j+1$ truncated game in which every type less than or equal to j obtains a payoff at least equal to the payoff obtained at $\hat{\sigma}$ while $j+1$ obtains a payoff at least equal to the payoff obtained at σ.

Lemma 3: Suppose $\sigma \in PSE(G)$ and $\hat{\sigma} \in PSE(G^j)$ for some $j < n$. Suppose further that $u(\hat{\sigma},j) \geq u(\sigma,j)$. Let $H = \{t | t \geq j+1 \text{ and } \mu(t) = \mu(j+1)\}$ and $h = \max\{t | t \in H\}$. Then there exists $\sigma^* \in PSE(G^h)$ such that:

(L3.1) $u(\sigma^*,t) \geq u(\hat{\sigma},t)$ for all $t \leq j$, and

(L3.2) $u(\sigma^*,t) \geq u(\sigma,t)$ for all $t \in H$.

Proof: The proof is by construction. We shall write for all t: $\mu(t) = m_t$, $\rho(\mu(t)) = r_t$, $\hat{\mu}(t) = \hat{m}_t$, and $\hat{\rho}(\hat{\mu}(t)) = \hat{r}_t$.

Case 1: $u(\hat{m}_j,\hat{r}_j,j+1) \geq u(m_{j+1},r_{j+1},j+1) \equiv u(\sigma,j+1)$.

We shall only prove that there exists $\sigma^* \in PSE(G^{j+1})$ which satisfies (L3.1) for all $t \leq j$ and (L3.2) for $t = j+1$. For if this were the case, we can replace j in the original proposition by $j+1$ and $\hat{\sigma}$ by this new σ^*, and induction will prove the lemma.

Let $K = \{t | t \leq j \text{ and } \hat{m}_t = \hat{m}_j\}$ and let $k = \min\{t \in K\}$. By definition, $\hat{m}_j = \hat{m}_k$ and $\hat{r}_j = \hat{r}_k = BR(\hat{m}_j,K)$. Define

$$m_K^* \equiv \max\{m \in M | u(m,BR(m,K \cup \{j+1\}),k) \geq u(\hat{\sigma},k)\}, \text{ and}$$

$$r_K^* \equiv BR(m_K^*,K \cup \{j+1\}).$$

We show in Appendix 1 that m_K^* exists. Since the $K \cup \{j+1\}$-conditional belief stochastically dominates the K-conditional belief, $\hat{m}_j < m_K^*$.

Moreover, in view of the definition of m_K^*,

$$u(m_K^*,BR(m_K^*,K \cup \{j+1\},k) = u(m_K^*,r_K^*,k) = u(\hat{\sigma},k). \tag{1}$$

We now define $\sigma^* \equiv (\mu^*,\rho^*,\beta^*)$ in G^{j+1} as follows:

(a)	for all $t < k$:	$\mu^*(t) = \hat{\mu}(t)$,
(b)	for all $t \in K \cup \{j+1\}$:	$\mu^*(t) = m_K^*$
(c)	for all $m \leq \hat{m}_j$:	$\beta^*(m) = \hat{\beta}(m)$ and $\rho^*(m) = \hat{\rho}(m)$,
(d)	for $m = m_K^*$:	$\beta^*(m) = q_{K\cup\{j+1\}}$ and $\rho^*(m) = r_K^*$.
(e)	for all $m > \hat{m}_j$ but $m \neq m_K^*$:	$\beta^*(m) = q_{\{1\}}$ and $\rho^*(m) = BR(m,\{1\})$,

Namely, we preserve all equilibrium messages (a) and replies (c) associated with $\hat{\sigma}$ for the types smaller than k and messages no larger than $\hat{m}_j$. For those types in $K\cup\{j+1\}$, we assign the message m_K^* (b) and the associated reply r_K^* (d). For other messages, we assign the worst possible belief and the associated replies (e).

Clearly, assertion (L3.1) holds for all $t < k$. For $t \in K$, (L3.1) follows because of monotonicity and (1). Finally, for $t = j+1$,

$$u(\sigma^*,j+1) = u(m_K^*,r_K^*,j+1) \geq u(\hat{m}_j,\hat{r}_j,j+1) \geq u(\sigma,j+1),$$

where the first inequality follows from monotonicity and (1), and the second from the condition defining this case. Thus (L3.2) holds.

It remains to show that σ^* is indeed a sequential equilibrium of G^{j+1}. Condition (D1.3) is straightforward from the definition of conditional beliefs, and (D1.2) is also straightforward. The proof that (D1.1) is satisfied is in Appendix 1.

Case 2: $u(\hat{m}_j,\hat{r}_j,j+1) < u(\sigma,j+1) \equiv u(m_{j+1},r_{j+1},j+1)$.

By hypothesis, $u(\hat{\sigma},j) \geq u(\sigma,j)$, i.e., $u(\hat{m}_j,\hat{r}_j,j) \geq u(m_j,r_j,j)$; combining this inequality with the equilibrium condition for σ, we obtain

$$u(\hat{m}_j,\hat{r}_j,j) \geq u(m_{j+1},r_{j+1},j).$$

In view of this inequality and the inequality characterizing the current case, Lemma 2 implies $m_{j+1} > \hat{m}_j$.

Let $r_H^* = BR(m_H^*,H)$. Define

$$m_H^* = \max\{m\in M \mid u(m,BR(m,H),j+1) \geq u(\sigma,j+1)\}.$$

By the definition of m_H^*, it follows that:

$$u(m_H^*,r_H^*,j+1) = u(m_{j+1},r_{j+1},j+1) = u(\sigma,j+1). \tag{2}$$

As in the previous case, m_H^* is well defined.

Define $\sigma^* \equiv (u^*,\rho^*,\beta^*)$ in the following way:

(a)	for all $t \le j$:	$\mu^*(t) = \hat{m}_t$;
(b)	for all $t \in H$:	$\mu^*(t) = m_H^*$;
(c)	for all $m \le \hat{m}_j$:	$\beta^*(m) = \hat{\beta}(m)$ and $\rho^*(m) = \hat{\rho}(m)$;
(d)	for $m = m_H^*$:	$\beta^*(m) = q_H$ and $\rho^*(m) = r_H^*$;
(e)	for all $m > \hat{m}_j$ but $m \neq m_H^*$:	$\beta^*(m) = q_{\{1\}}$ and $\rho^*(m) = BR(m,\{1\})$.

Property (L3.1) follows trivially and (L3.2) follows from (2) for $j+1$ and from (2) and monotonicity for other $t \in H$. Using essentially the same arguments as in Case 1, σ^* is readily established as a sequential equilibrium of G^h. **Q.E.D.**

Corollary: Suppose $\sigma \in PSE(G)$ and $\hat{\sigma} \in PSE(G^j)$ for some $j < n$. Suppose further that $u(\hat{\sigma},j) \ge u(\sigma,j)$. Then there exists $\sigma^* \in PSE(G)$ such that:

(i) $u(\sigma^*,t) \ge u(\hat{\sigma},t)$ for all $t \le j$, and

(ii) $u(\sigma^*,t) \ge u(\sigma,t)$ for all $t > j$.

Proof: Repeated application of Lemma 3 yields the result. **Q.E.D.**

Proof of Theorem 1:

Let σ be the LMSE and suppose, contrary to the supposition, there exists $\hat{\sigma} \in PSE(G)$ that defeats σ via (m,K) and let $j = \max\{t \in K\}$. The restriction of $\hat{\sigma}$ on $T^j = \{1,2,\ldots,j\}$ is obviously a PSE of G^j and $u(\hat{\sigma},t) \ge u(\sigma,t)$ for all $t \in K$ with strict inequality for at least one $t \in K$ by the definition of defeat. Thus, by the corollary to the previous lemma, there exists $\sigma^* \in PSE(G)$ such that:

$$u(\sigma^*,t) \ge u(\hat{\sigma},t) \text{ for all } t \le j; \text{ and}$$
$$u(\sigma^*,t) \ge u(\sigma,t) \text{ for all } t > j.$$

But then σ^* ℓ-dominates σ, contradicting the hypothesis that σ is LMSE. **Q.E.D.**

Proof of Theorem 2:

In view of Theorem 1, we only have to show that, if the LMSE is a separating equilibrium, there is no undefeated pure strategy equilibrium which is not the LMSE. So suppose, contrary to our assertion, there exists an undefeated $\sigma \in PSE(G)$ which is ℓ-dominated by the LMSE, σ^*. Then there exists $j \in T$ such that:

(1) $u(\sigma^*,j) > u(\sigma,j)$,

(2) $u(\sigma^*,t) \geq u(\sigma,t)$ for all $t > j$.

We shall assume that $\mu^*(j)$ is an out-of-equilibrium message under σ, for otherwise we can construct $\sigma^{**} \in PSE(G)$ slightly perturbed from σ^* by assigning a new message $\mu^{**}(j)$ to j from the open interval $(\mu^*(j), \mu^*(j+1))$ and a new reply $\rho^{**}(\mu^{**}(j)) \in BR(\mu^{**}(j),\{j\})$, still preserving

(1') $u(\sigma^{**},j) > u(\sigma,j)$, and

(2') $u(\sigma^{**},t) \geq u(\sigma,t)$ for all $t > j$.

Note that we can construct the above perturbed equilibrium because the LMSE, σ^*, is a separating equilibrium. Otherwise, a change in message $\mu^*(j)$ may affect the equilibrium payoff of other types, in particular, of type $j+1$.

Obviously, $\beta(\mu^*(j))$ is not $q_{\{j\}}$, for otherwise an equilibrium condition is violated. Since $\mu^*(j)$ is sent only by type j, σ^* defeats σ via $(\mu^*(j),\{j\})$, contrary to our supposition. **Q.E.D.**

4. CONTINUITY OF THE LEX MAX EQUILIBRIUM

As we discussed in Section 2, the separating equilibria (and thus the equilibria selected by stability) have the undesirable feature that generally there is a discontinuity in the prior probabilities as one of these probabilities goes to zero. To illustrate this problem, consider the simple Spence model with three types, p_t being the prior probability that the worker has productivity t. The Riley outcome has 1 selecting zero units of education (and so receiving a wage of 1), 2 selecting 1 unit of education (and so receiving a wage of 2) and 3 selecting 3 units of education (and so receiving a wage of 3). In this outcome, the two high productivity level workers' choices are distorted above their full information optimal education choices of zero. It is worth emphasizing that this is the (Kohlberg-Mertens) stable outcome for *all* completely mixed probability distributions over $\{1,2,3\}$. Consider now the case of p_2

= 0. Then the stable outcome has the worker of type 0 still choosing zero units of education, but the worker of productivity 3 choosing 2 (not 3) units of education. There is a discontinuity: the predicted outcome differs in a nontrivial way upon whether a type is believed to occur with an arbitrarily small probability or not at all.

We do not believe that the world is discontinuous in this way; that is, we believe that the economic analysis of the two cases should be nearly the same. In modeling an incomplete information environment, presumably the aspect that the analyst *and* the other players should feel *least* confident about is the specification of players' beliefs. Now, recall the motivation for using types. In a game of incomplete information, a player (II say) has beliefs over some payoff relevant aspect of another player (I say). Moreover, I has beliefs over II's beliefs, II has beliefs over I's beliefs over II's beliefs, etc. Types are a mathematically consistent way of representing these hierarchies. A model in which different types represent not only different values of the payoff relevant variable (such as productivities) but also infinite hierarchies of beliefs is of course completely intractable and economists have reasonably made the simplifying assumption that different types represent just different values of the payoff relevant variable, with the beliefs over these different values being common knowledge. This type of assumption is only reasonable if you believe that slight changes in beliefs do not have a dramatic impact on the outcome. This not only suggests the requirement that the equilibrium correspondence be upper hemicontinuous in beliefs (using the Prohorov metric on beliefs), but also directs attention to the role (or, hopefully, lack thereof) of the support. Since weak convergence of beliefs implies nothing about the behavior of the supports of those beliefs, and separating equilibria depend only on the supports, separating (and so stable) equilibria do not satisfy this requirement.

Interestingly, the sequential equilibrium correspondence also fails this requirement. If the set of types is {2,3}, then in any separating sequential equilibrium the worker of type 2 chooses zero units of education, rather than some level between 1 and 2 units. The failure in the case of sequential equilibrium is less troubling for two reasons. First, as we will show, there is a continuous selection from the sequential equilibrium correspondence (recall that this is not true for the separating sequential equilibrium or stable equilibrium correspondence). Second, the failure of upper hemicontinuity occurs only when the probability of the smallest type goes to zero. The failure is related to the distinction between an event occurring with probability one and its being common knowledge. That is, in the Spence example with workers having types {2,3}, if we allow the firms to believe that the worker has a productivity of 1 after an out-of-equilibrium message, then it is an equilibrium for the worker of type 2 to choose m = 1 and the worker of type 3 to choose m = 4 even when $p_1 = 0$.

It is worth observing that the convergence issue of interest here involves changes in the extensive form. As the support of the prior distribution changes, the number of initial nodes changes. The study of reputation games is also (at least partly) motivated by changes in an extensive form, with some nodes occurring with small probability.[11] While, with a long horizon, small probabilities of a "crazy" type can dramatically change the nature of equilibrium play (a la chain store paradox type of behavior), there is no discontinuity: The smaller the probability of a crazy player, the longer the horizon needed. In contrast, for the types of changes in extensive form here, there is a real discontinuity.

The next theorem asserts that, generically in p, the lex max equilibrium is continuous when u(m,BR(m,q),t) is a strictly quasiconcave function of m. Genericity rules out cases in which two different equilibria, one at least involving some pooling, can give rise to identical payoffs to some types without those types choosing the same actions in the two equilibria. The genericity is with respect to the *interior* of the simplex--we are not ruling out convergence to the boundary. Slightly weaker versions of the quasiconcavity assumption are used in Mailath [1987] and Cho and Sobel [1990]. We are aware of no instances where this assumption fails in applications. Since there exist nonlex max equilibria which are undefeated, lex max is a continuous selection from the undefeated correspondence. In the statement of the theorem, a superscript ℓ denotes lex max equilibrium choices.

Theorem 3: Suppose that u(m,BR(m,q),t) is strictly quasiconcave in m for all t and for all $q \in \Delta_T$. For all $T' \subset T$, there exists a closed set of measure zero $\Delta^*_{T'} \subset \mathrm{Int}(\Delta_{T'})$ such that if $\{p_k\} \subset \Delta_T \backslash \Delta^*_T$ and $p_k \to p_\infty \notin \Delta^*_{\mathrm{supp}(p_\infty)}$, then $\mu^\ell_k(t) \to \mu^\ell_\infty(t)$ as $k \to \infty$, $\forall t \in \mathrm{supp}(p_\infty)$.

Proof: See Appendix 2.

5. REFINEMENTS

The literature on refinements can be divided into two categories. The first category focusses on the possibility of mistakes to restrict behavior off the equilibrium path. The second assumes that players are rational and know (or operate under the strong presumption) that the other players are also rational. The refinement proposed in this paper falls within the second category.

Perhaps the most familiar class of refinements of sequential equilibrium is the class subsumed by the notion of strategic stability due to Kohlberg and Mertens [1986]. This refinement implies the intuitive criterion (sometimes called equilibrium domination, introduced by Kreps [1984]), D1 (Cho and Kreps

---

[11] See Kreps, Milgrom, Roberts, and Wilson [1982], Kreps and Wilson [1982], and Milgrom and Roberts [1982].

[1987]), divinity and universal divinity (Banks and Sobel [1987]). Cho and Kreps [1987] show that the intuitive criterion selects the Riley outcome in the Spence job market signaling game with two types and that D1 may be necessary with more than two types. Cho and Sobel [1990] show that stability is equivalent to D1 for the class of signaling games studied here.[12] Moreover, for this class D1 selects the Riley outcome.

Stability is defined by requiring a "good" outcome to have nearby good outcomes in games close by (where only payoffs, and not the extensive form, are being perturbed). Thus it is in the tradition of refining Nash equilibria by including trembles. Essentially, stability modifies the definition of trembling hand perfect by requiring equilibrium outcomes to be robust to *all* trembles. While the concept seems to be motivated by "near" irrationalities, it is not clear from Kohlberg and Mertens [1986] whether they intend the trembles to be taken as a positive description or as a computational convenience. If it is the latter, the best interpretation is perhaps that stability is a solution motivated by continuity considerations. (Here we are distinguishing between the solution and their list of desiderata.) Of particular interest for our purposes is their formal statement of forward induction or elimination of inferior strategies. While they do not list robustness to elimination of inferior strategies as a main requirement (p.1020), Kohlberg [1989] discusses it in such a way. Moreover, robustness to elimination of inferior strategies is the only property of stability that is necessary to obtain the refinements listed above. Like forward induction, this property is introspective.

Hillas [1990] investigates the implications of a requirement which has a superficial similarity to undefeated: Fix a sequential equilibrium and an out-of-equilibrium message. If this message is sent with positive probability in some other equilibria, then play and beliefs after that message should be consistent with these equilibria. Note that this is not motivated by a forward induction logic. In contrast, undefeated focuses only on those sequential equilibria that also give the types sending that message a higher payoff than in the original equilibrium.

6. CONCLUDING REMARKS

The definition of defeat might be altered in several interesting ways. In the test of whether an equilibrium is defeated, each disequilibrium strategy is conjectured to be a "signal" by some set of types that an alternative equilibrium is being played. If there are several alternative equilibria in which a given disequilibrium message is sent, player II will not be able to make an unambiguous comparison between

---

[12]In fact the class for this result is strictly larger. See their paper for details.

the proposed equilibrium and possible alternatives. The test of whether a given equilibrium is defeated or not could be strengthened by requiring that there be a *unique* alternative equilibrium in which the given disequilibrium message is played and which has the additional properties in the definition of defeat. This would make it less likely that an equilibrium is defeated and would make the set of undefeated equilibria larger. It would not solve the problem that for general games the set of undefeated equilibria may be empty, but it might make the concept of defeat applicable to a larger set of problems.

In the definition of defeat, an equilibrium is tested according to the beliefs held by player II at disequilibrium information sets. In this sense the refinement is similar to that introduced by Grossman and Perry. The test to which an equilibrium is put to determine whether it is defeated or not differs in that it may be used to generate a partial ordering on the set of sequential equilibria. In the case that the set of undefeated equilibrium is empty, we can form a set of sequential equilibria such that every sequential equilibrium is defeated by a sequential equilibrium in this set. A minimal set, that is, a set such that no proper subset has the property must trivially exist. There may be multiple minimal sets, but such a set would constitute a refinement which trivially must be non-empty. For some examples, this may reduce the set of equilibria in reasonable ways.

The fact that an equilibrium may be defeated by an equilibrium that is itself defeated is not a criticism of the undefeated equilibrium concept: Once an out-of-equilibrium message has been sent, considerations of whether the alternative equilibrium is itself susceptible to a deviation are irrelevant because there is no possibility of a different message being sent.

A third way in which the concept might be altered is to extend the definition to more general games. It should be straightforward to define the concept of defeat for games with more stages and with more than two players. For each history which is an equilibrium history except for a single move by some agent, we can ask whether there is an alternative equilibrium which is consistent with the history. If there is, we can again compare the payoffs to the types who prefer the alternative equilibrium to the proposed equilibrium to determine whether the disequilibrium message can be interpreted as a signal that the alternative equilibrium is being played by that set of types.

We should also note that undefeated is open to the following criticism: Suppose σ is defeated by σ' (with message m' and deviating types K') and by σ'' (with message m'' and deviating types K''). Moreover, suppose $K' \cap K'' \neq \varnothing$ and some types in K' strictly prefer σ' and some types in $K'' \cap K'$ strictly prefer σ''. In this situation, the interpretation of the message as a signal of a new equilibrium is weakened. One could alter the concept by only requiring beliefs in this situation to have support in $K' \cup K''$. This may or may not eliminate the original equilibrium.

These last comments underlie the statement of our purpose made in the introduction. The concept of undefeated proposed in this paper is not meant to be a grand solution to the problems of multiple equilibria in signalling games. Given the alterations suggested above, it is clear that undefeated is not a compelling solution to all such problems. Our goal was to discuss some of the difficulties presented by other generally used refinements and present a refinement that addressed some of those difficulties.

APPENDIX 1

Proof that m_K^* exists: Denote the set defining m_K^* by M_K. We shall show that M_K is non-empty and bounded from above.

Non-emptiness follows because $\hat{m}_j$ is in M. To see this, observe that assumption 2 implies:

$$u(\hat{m}_j,BR(\hat{m}_j,K\cup\{j+1\}),k) > u(\hat{\sigma},k),$$

since the $K\cup\{j+1\}$-conditional belief stochastically dominates the K-conditional belief.

The set M_K is bounded from above because, for any $m \geq \bar{m}(\hat{m}_j,\hat{r}_j,k)$,

$$u(m,BR(m,K\cup\{j+1\}),k) < u(m,BR(m,\{n\}),k) < u(\hat{\sigma},k).$$

The first inequality follows from the fact that the $\{n\}$-conditional belief stochastically dominates the $K\cup\{j+1\}$-conditional belief, while the second inequality holds from assumption 4. **Q.E.D.**

Proof that (D1.1) is satisfied: Consider the following three exhaustive possibilities:

(1) $t < k$:

If $m \leq \hat{m}_j$, then the fact that $\hat{\sigma}$ is itself a sequential equilibrium implies $u(\sigma^*,t) \geq u(m,\rho^*(m),t)$.

If $m > \hat{m}_j$, but $m \neq m_K^*$,

$$u(\sigma^*,t) = u(\hat{\sigma},t) \geq u(m,BR(m,\hat{\beta}(m)),t) \geq u(m,BR(m,\{1\}),t) \equiv u(m,\rho^*(m),t),$$

where the first inequality holds due to $\hat{\sigma}$ being a sequential equilibrium and the second due to $\hat{\beta}(m)$ (weakly) stochastically dominating the $\{1\}$-conditional belief.

Finally, if $m = m_K^*$,

$$u(\sigma^*,t) = u(\hat{\sigma},t) \geq u(\hat{m}_k,\hat{r}_k,t) > u(m_K^*,r_K^*,t),$$

where the first inequality follows from $\hat{\sigma}$ being itself a sequential equilibrium and the second inequality from $\hat{m}_j < m_K^*$, (1) and Lemma 1.

(2) $t \in K$:

If $m \leq \hat{m}_j$,

$$u(\sigma^*,t) = u(m_K^*,r_K^*,t) \geq u(\hat{\sigma}_t,t) \geq u(m,\hat{\rho}(m),t) = u(m,\rho^*(m),t),$$

where the two equalities follow from the definition of σ^* and the first inequality follows from (L3.1). The second inequality follows from the equilibrium condition of $\hat{\sigma}$.

If $m > \hat{m}_j$, but $m \neq m_K^*$,

$$u(\sigma^*,t) \geq u(\hat{\sigma},t) \geq u(m,BR(m,\hat{\beta}(m)),t) \geq u(m,BR(m,\{1\}),t) = u(m,\rho^*(m),t),$$

where the first inequality follows from (L3.1), the second from the equilibrium condition $\hat{\sigma}$, and the third from the fact that $\hat{\beta}(m)$ (weakly) stochastically dominates $q_{\{1\}}$.

(3) $t = j+1$:

For all $m \leq m_K^*$, monotonicity and the results for $t \in K$ imply (D1.1).

For all $m > m_K^*$, (D1.1) follows from (L3.2), σ an equilibrium of G and the observation that the associated belief is the worst possible belief. **Q.E.D.**

APPENDIX 2

Proof of Theorem 3: We first use the genericity of p to argue that if $\sigma,\sigma' \in PSE(G(T'))$, σ lex max, $u(\sigma,t) = u(\sigma',t)\ \forall t \geq t''$, $t,t'' \in T'$, then $\mu(t) = \mu'(t)\ \forall t \geq t''$. This is immediate if σ and σ' induce the same partition over the type space T', because of strict quasiconcavity. If the equilibria involve different partitions pooling, then perturbing the probabilities will then eliminate the tie, since the posteriors will change.

Suppose σ is lex max, σ' is another equilibrium with $u(\sigma,t) = u(\sigma',t)\ \forall t \geq t''+1$, $u(\sigma,t'') > u(\sigma',t'')$ and $p(t''+1) > 0$. Since, by the previous paragraph, the sets of types pooling with $t''+1$ in σ and in σ' are identical, t'' cannot be pooling with $t''+1$.[13]

We consider the case of $p_k(t) \to 0$ for at most one type (the case of many types then follows immediately) and denote that type by t'. If $p_k(t) \to 0$ for some type t', let $t_l \equiv \max\{t \in T : t < t'\}$, $t_u \equiv \min\{t \in T : t > t'\}$, and $T' \equiv T \backslash \{t'\}$. If $p_k(t) \not\to 0$ for all $t \in T$, then set $T' \equiv T$. Denote by $G_k(T')$ the game with prior distribution p_k, conditioning on T', and by $G_\infty(T')$ the game with prior distribution p_∞, conditioning on T'. Finally, for $k = 1,2,\ldots,\infty$, if q_K^k is the K-conditional belief based on the prior p_k, $BR_k(m,K) \equiv BR(m,q_K^k)$ and $k = \infty$ is understood if the subscript is absent.

Claim 1: For all $\sigma \in PSE(G_\infty(T'))$ such that $u(\sigma,t) > \max u(m,BR(m,\{1\}),t)$ for all $t \in T'$, $t \neq 1$, for all $\epsilon > 0$ there exists k^* so that for $k > k^*$, there exists $\sigma_k \in PSE(G_k(T))$ with $|\mu(t) - \mu_k(t)| < \epsilon$ and $|u(\sigma,t) - u(\sigma_k,t)| < \epsilon\ \forall t \in T'$.

Proof: The equilibrium σ partitions T' into $T'_1 \cup \ldots \cup T'_m$, where $t,t'' \in T'_i$ iff $\mu(t) = \mu(t'')$ and $t \in T_i$, $t'' \in T_j$, $i < j$ implies $t < t''$. Let $t_i = \max T_i$ so that $t_{i-1}+1 = \min T_i$. In what follows, it is understood that if $T' = T$, then the comments with respect to t' should be ignored. Suppose $t_u \in T'_i$. If $\mu(t_{i-1}) = \mu(t_i)$, (so that $t_l \in T'_{i-1}$) then $[\mu(t_{i-1}),\mu(t_i)) = \varnothing$ and either Case I or Case II applies. If $t' > n$, then trivially $t_l \in T'_{i-1}$ and Case III is the relevant case.

Case I: Either $T' = T$, or $u(\mu(t_u),BR(\mu(t_u),T'_i),t') > u(m,BR(m,\{1\}),t')\ \forall m \in [\mu(t_{i-1}),\mu(t_i))$, $t' < n$ and if $t_l \in T'_{i-1}$ then $u(\mu(t_u),BR(\mu(t_u),T'_i),t') \geq u(\mu(t_l),BR(\mu(t_l),T'_{i-1}),t')$.

Fix $\delta > 0$. Consider the following specification of μ_k in which t' pools with T'_i (so that $T_i = T'_i \cup \{t'\}$ and $T_j = T'_j$ for $j \neq i$):

(a) $\forall t \in T_1$: $\mu_k(t) = \text{argmax}_m\{u(m,BR_k(m,T_1),1) : |m - \mu(t)| < \delta\}$; and

---

[13] This implies that $\mu(t''+1)$ maximizes $u(m,BR(m,q_K),t''+1)$: The value of $\mu(t''+1)$ is not fixed by the t'' incentive constraint, since $u(\sigma,t'') > u(\mu(t''+1),\rho(\mu(t''+1)),t'')$ (if equality holds, then $u(\sigma',t'') < u(\mu(t''+1),\rho(\mu(t''+1)),t'') = u(\mu'(t''+1),\rho'(\mu(t''+1)),t'')$, a contradiction).

(b) $\forall t \in T_j$, $j \geq 2$: $\mu_k(t) = \min \{|m-\mu(t_j)| : u(m,BR_k(m,T_j),t_{j-1}) \leq u(\mu_k(t_{j-1}),BR_k(\mu_k(t_{j-1}),T_{j-1}),t_{j-1})$ and $u(m,BR_k(m,T_j),t_{j-1}+1) \geq u(\mu_k(t_{j-1}),BR_k(\mu_k(t_{j-1}),T_{j-1}),t_{j-1}+1)\}$.

We now argue that the profile σ_k given by using Bayes rule for $\mu_k(t)$, $t \in T$, and setting $\rho_k(m) = BR_k(m,\{1\})$ for $m \notin \mu_k(T)$ is a sequential equilibrium for δ small and k large:[14]

Note first that monotonicity implies that the profile is Nash. Monotonicity also implies that the two incentive constraints on t_j and t_j+1 in σ $(u(\mu(t_{j+1}),BR(\mu(t_{j+1}),T_{j+1}),t_j) \leq u(\sigma,t_j)$ and $u(\sigma,t_j+1) \geq u(\mu(t_j),BR(\mu(t_j),T_j),t_j+1))$ cannot be simultaneously binding. Thus for k large, the analogous two constraints can be satisfied for m close to $\mu(t)$. By continuity, $u(\sigma_k,t) > \max u(m,BR_k(m,\{1\}),t)$ for all $t \in T'\backslash\{1\}$ and $u(\mu(t_u),BR_k(\mu(t_u),T_i),t') > u(m,BR_k(m,\{1\}),t')$ for all $m \in [\mu(t_{i-1}),\mu(t_i))$ for k large. By construction, $u(\sigma_k,1) \geq \max u(m,BR_k(m,\{1\}),1)$, so that σ_k is sequential.

Case II: $T' \neq T$, $u(\mu(t_u),BR(\mu(t_u),T'_i),t') > u(m,BR(m,\{1\}),t')$ $\forall m \in [\mu(t_{i-1}),\mu(t_i))$, $t' < n$, $t_l \in T'_{i-1}$ and

$$u(\mu(t_u),BR(\mu(t_u),T'_i),t') < u(\mu(t_l),BR(\mu(t_l),T'_{i-1}),t').$$

This is just like Case I, except that t' is pooled with T'_{i-1}.

Case III: $T' \neq T$ and either $\exists m \in [\mu(t_{i-1}),\mu(t_i))$ such that $u(\mu(t_u),BR(\mu(t_u),T'_i),t') \leq u(m,BR(m,\{1\}),t')$ or $t' = n$.

Define $\mu_k(t)$ for $t \leq t_l$ as in Case I. The type t' separates from t_l and t_u at $\mu_k(t') = \text{argmax}\ \{u(m,BR_k(m,\{t'\}),t') : m \geq \mu_k(t_l),\ u(m,BR_k(m,\{t'\}),t_l) \leq u(\mu_k(t_l),BR_k(\mu_k(t_l),T_{i-1}),t_l)\}$. Define $\mu_k(t)$ for $t \geq t_u$ as in Case I with $T_i = T'_i$. We now argue that the profile σ_k given by using Bayes rule for $\mu_k(t)$, $t \in T$, and setting $\rho_k(m) = BR_k(m,\{1\})$ for $m \notin \mu_k(T)$ is a sequential equilibrium for all k.

The incentive constraint for t_l is satisfied by construction. It is also easy to see that $u(\sigma_k,t') \geq u(\mu_k(t_l),BR_k(\mu_k(t_l),T_{i-1}),t')$.

If the incentive constraint for t_l is not binding, then t' prefers $\mu_k(t')$ to any out-of-equilibrium message m (this is clear for $m > \mu_k(t_l)$ and is implied by monotonicity for $m < \mu_k(t_l)$). Suppose then that the constraint is binding (i.e., holds with equality) and $u(\sigma_k,t') < u(m,BR_k(m,\{1\}),t')$ for some message m. If $m < \mu_k(t')$, monotonicity implies that $u(\mu_k(t'),BR_k(\mu_k(t'),\{t'\}),t_l) < u(m,BR_k(m,\{1\}),t_l)$, yielding a contradiction. So suppose $m > \mu_k(t')$. Since $u(\sigma_k,t') < u(m,BR_k(m,\{1\}),t') < u(m,BR_k(m,\{t'\}),t')$, we have $u(m,BR_k(m,\{t'\}),t_l) > u(\sigma_k,t_l)$ (by the definition of $\mu_k(t')$). But $u(\mu_k(t_l),BR_k(\mu_k(t_l),\{t'\}),t_l) > u(\sigma_k,t_l)$, which implies (by strict quasiconcavity) $u(\mu_k(t'),BR_k(\mu_k(t'),\{t'\}),t_l) > u(\sigma_k,t_k)$, contradicting the definition of $\mu_k(t')$.

---

[14]We need to allow for $\mu_k(1) \neq \mu(1)$, since it is possible that $u(\mu(1),\rho(\mu(1)),1) = \max u(m,BR(m,\{1\}),1)$ and p_k (conditional on T_1) may be first order stochastically dominated by p (conditional on T_1).

Finally, we need to show that $\mu_k(t')$ does not adversely constrain the choice of t_u. If $u(\mu_k(t'), BR_k(\mu_k(t'),\{t'\}),t_u) \leq u(\sigma_k,t_u)$, then this is clear. So suppose $u(\mu_k(t'),BR_k(\mu_k(t'),\{t'\}),t_u) > u(\sigma_k,t_u)$. But then $u(\mu_k(t'),BR_k(\mu_k(t'),\{t'\}),t') > u(\mu(t_u),BR_k(\mu(t_u),T_i),t')$ and since by construction $u(\mu_k(t'), BR_k(\mu_k(t'),\{t'\}),t') > u(m,BR(m,\{1\}),t')$ for $m \geq \mu_k(t_l)$, we can apply the same arguments as in Case II to show that no type $t \geq t_u$ wishes to deviate from σ_k. **Q.E.D.**

Claim 2: If σ is lex max of G(T'), then $u(\sigma,t) > \max u(m,BR(m,\{1\}),t)$ for all $t\in T'$, $t \neq 1$.

Proof: Suppose $u(\sigma,t^*) = u(m^*,BR(m^*,\{1\}),t^*)$ for some m^* and $t^* \neq 1$. Using the same notation as in the previous claim, suppose $t^*\in T_i$. Observe that by the corollary to Lemma 3, an equilibrium σ' of $G(\{t\leq t_i\})$ with $u(\sigma',t_i) > u(\sigma,t_i)$ or $u(\sigma',t^*) > u(\sigma,t^*)$ and $u(\sigma',t) = u(\sigma,t)$ for $t^*<t\leq t_i$ contradicts σ being lex max. We now construct such an equilibrium to obtain the contradiction.

Suppose first that $m^* < \mu(t_i)$, in which case monotonicity implies $t^* = t_{i-1}+1$ and so $m^* > \mu(t_{i-1})$. Define a strategy profile σ' on $G(\{t\leq t_i\})$ by $\mu'(t) = \mu(t)$ for $t \leq t_{i-1}$, $\mu'(t^*) = \text{argmax}\{u(m, BR(m,\{t^*\}),t^*) : m \geq \mu(t_{i-1}), u(m,BR(m,\{t^*\}),t_{i-1}) \leq u(\sigma',t_{i-1})\}$, and $\mu'(t) = \min \{m\geq\mu(t_i) : u(m,BR(m,K),t^*) \leq u(\sigma',t^*)\}$ for all $t \in K \equiv T_i\backslash\{t^*\}$ (where σ' is given by using Bayes rule for $\mu'(t)$, $t\in T'$, and setting $\rho'(m) = BR(m,\{1\})$ for $m\notin\mu'(T')$).

We now argue that $u(\sigma',t^*) > u(\sigma,t^*)$: For suppose not and let $m \geq \mu(t_{i-1})$ solve $u(m,BR(m, \{t^*\}),t_{i-1}) = u(\sigma',t_{i-1})$. Quasiconcavity implies that if $u(\sigma',t^*) < u(\sigma,t^*)$ then $m^* < m$. But if $m^* < m$, $u(m,BR(m,\{t^*\}),t^*) \leq u(\sigma',t^*) \leq u(\sigma,t^*) = u(m^*,BR(m^*,\{1\}),t^*)$ implies $u(m^*,BR(m^*,\{1\}),t_{i-1}) > u(m,BR(m,\{t^*\}),t_{i-1}) = u(\sigma',t_{i-1}) = u(\sigma,t_{i-1})$, which contradicts $\sigma\in PSE(T')$.

Clearly, t^* does not want to send $\mu'(t_i)$. If $t^* = t_i$ then σ' is the desired equilibrium of $G(\{t\leq t_i\})$. Suppose $t^* \neq t_i$ and $\mu'(t_i) > \mu(t_i)$. Then $u(\mu'(t_i),BR(\mu'(t_i),K),t^*) = u(\sigma',t^*)\}$, so that we again have σ' an equilibrium of $G(\{t\leq t_i\})$ and moreover, $u(\sigma',t_i) > u(\sigma,t_i)$ (otherwise t_i prefers $\mu(t_i)$, which implies $u(\mu(t_i),BR(\mu(t_i),T_i),t^*) > u(\sigma',t^*)\} > u(m^*,BR(m^*,\{1\}),t^*)$, a contradiction).

If $t^* < t_i$ and $\mu'(t_i) = \mu(t_i)$, it is possible that $u(\mu'(t^*),BR(\mu'(t^*),\{t^*\}),t) > u(\sigma',t)$ for some $t\in K$. In this case, define a profile σ'' in which t^* pools with t^*+1 at $\mu''(t^*) = \text{argmax}\{u(m,BR(m,\{t^*,t^*+1\}),t^*) : m \geq \mu(t_{i-1}), u(m,BR(m,\{t^*,t^*+1\}),t_{i-1}) \leq u(\sigma',t_{i-1})\}$. If σ'' is sequential then this is the desired equilibrium. If it is not, pool t^*+2 with t^* and t^*+1. Proceeding in this way yields the desired equilibrium.

Suppose now that $m > \mu(t_i)$, in which monotonicity implies $t^* = t_i$. Suppose that $u(\mu(t_i),\rho(\mu(t_i)), t_{i-1}+1) > u(\mu(t_{i-1}),\rho(\mu(t_{i-1})),t_{i-1}+1)$. Then increasing $\mu(t_i)$ by ϵ still gives an equilibrium of $G(\{t\leq t_i\})$ and quasiconcavity implies that t_i has a higher payoff, contradiction. So, suppose $u(\mu(t_i),\rho(\mu(t_i)),t_{i-1}+1) \doteq u(\mu(t_{i-1}),\rho(\mu(t_{i-1})),t_{i-1}+1) \equiv u'$. Define μ' by $\mu'(t) = \mu(t)$ for $t \leq t_{i-1}$, $\mu'(t_{i-1}+1)$ solves min

$\{m \geq \mu(t_{i-1}) : u(m,BR(m,\{t_{i-1}+1\}),t_{i-1}+1) \geq u'\}$, $\mu'(t) = \min\{m \geq \mu'(t_{i-1}+1) : u' \geq u(m,BR(m,K'), t_{i-1}+1)\}$ for $t \in K' \equiv T_i \backslash \{t_{i-1}+1\}$. Since $t_{i-1}+1$ is indifferent over the three messages $\mu'(t_{i-1})$, $\mu'(t_{i-1}+1)$, and $\mu'(t_i)$, the implied σ' is a Nash equilibrium. Now $t > t_{i-1}+1$ strictly prefer $(\mu'(t_i),\rho'(\mu'(t_i)))$ to $(\mu(t_i),\rho(\mu(t_i)))$ because $t_{i-1}+1$ is indifferent between the two. Thus the profile is sequential in $G(\{t \leq t_i\})$ and t_i strictly prefers σ'. **Q.E.D.**

Suppose σ_k is the lex max equilibrium of $G_k(T)$ and that $\{\sigma_k\}$ does not converge to the lex max equilibrium of $G_\infty(T')$. Then there is an $\epsilon > 0$ and a subsequence, $\{\sigma_{k_m}\}$, which is at least ϵ distant from the lex max equilibrium. Since the simplex is compact, there is a convergent subsubsequence with limit σ_∞. Relabel so that this subsequence is denoted $\{\sigma_k\}$. Denoting the lex max equilibrium of $G_\infty(T')$ by σ', we have that there exists a type t^* such that $u(\sigma_\infty,t) = u(\sigma',t)$ for $t > t^*$ and $u(\sigma_\infty,t^*) < u(\sigma',t^*)$. Now, t^* does not pool with any type $t > t^*$ in σ'. Thus $\sigma'|_{\leq t^*}$ is an equilibrium of $G_\infty(\{t \leq t^*\})$. But by the claims, there exists an equilibrium arbitrarily close to $\sigma'|_{\leq t^*}$ in $G_k(\{t \leq t^*\})$. Applying the corollary to Lemma 3 shows that σ_k cannot be lex max. **Q.E.D.**

REFERENCES

Banks, J. and J. Sobel [1987], "Equilibrium Selection in Signalling Games," *Econometrica*, 55, 647-661.

Cho, I.-K. [1987], "A Refinement of the Sequential Equilibrium Concept," *Econometrica*, 55, 1367-1389.

Cho, I.-K. and D. Kreps [1987], "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179-221.

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

Engers, M. and M. Schwartz [1987], "On the Uniqueness of Signalling Equilibria," mimeo, University of Virginia and Georgetown University.

Farrell, J. [1985], "Credible Neologisms in Games of Communication," mimeo.

Grossman, S. and M. Perry [1986], "Perfect Sequential Equilibrium," *Journal of Economic Theory* 39, 97-119.

Hillas, J. [1990], "Sequential Equilibria and Stable Sets of Beliefs," mimeo, SUNY.

Kohlberg, E. [1989], "Refinement of Nash Equilibrium: The Main Ideas," Working Paper 89-073, Harvard Business School.

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

Kreps, D. [1984], "Signalling Games and Stable Equilibria," mimeo.

Kreps, D. [1990], *A Course in Microeconomic Theory*, Princeton University Press, Princeton.

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

Mailath, G. [1987], "Incentive Compatibility in Signaling Games with a Continuum of Types," *Econometrica*, 55, 1349-1365.

Mailath, G. [1990], "Signaling Games," *Recent Developments in Game Theory*, forthcoming.

McClennan, A. [1985], "Justifiable Beliefs in Sequential Equilibrium," *Econometrica* 53, 889-904.

Milgrom, P. and J. Roberts [1982], "Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis," *Econometrica* 50, 443-459.

Okuno-Fujiwara, M. and A. Postlewaite [1987], "Forward Induction and Equilibrium Refinement," CARESS Working Paper #87-01, University of Pennsylvania.

Riley, J. [1979], "Informational Equilibrium," *Econometrica*, 47, 331-359.

Rothschild, M. and J.E. Stiglitz [1977], "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 80, 629-49.

Spence, M. [1973], "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355-374.

Spence, M. [1974], *Market Signalling*, Harvard University Press, Cambridge.

Wilson, C. [1977], "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 16, 167-207.