UC Berkeley

Other Recent Work

Title

Deviations, Dynamics and Equilibrium Refinements

Permalink

<https://escholarship.org/uc/item/40s882v6>

Authors

Rabin, Matthew
Sobel, Joel

Publication Date

1993-05-01

Peer reviewed

UNIVERSITY OF CALIFORNIA AT BERKELEY

Department of Economics

Berkeley, California 94720

Working Paper No. 93-211

Deviations, Dynamics and Equilibrium Refinements

Matthew Rabin

University of California at Berkeley

and

Joel Sobel

University of California at San Diego

May 1993

Key words: Deviations, Dynamics, Equilibrium, Intuitive Criterion, Never a Weak Best Response, Signaling Refinements, Stiglitz Critique

JEL Classification: B49, C72, C73

We thank Eddie Dekel-Tabak and Joe Farrell for early feedback, and participants of theory workshops at Berkeley, University of Iowa, and especially the University of Chicago for helpful comments. Each author also thanks the National Science Foundation for financial support.

Abstract

Many standard solution concepts rule out those Nash equilibria that are susceptible to deviations. We propose a framework for considering not only which equilibria are not susceptible to deviations, but also which equilibria are likely to persist in the long run because they are repeatedly deviated *to*. We call such equilibria *recurrent*. We explore which equilibria are recurrent based on the deviations underlying each of several prominent signaling refinements. We show that the set of recurrent equilibria based on Cho and Krep's (1987) intuitive criterion and Kohlberg and Mertens's (1986) NWBR criterion are precisely what those papers already predict. In contrast, we show that applying our framework to cheap-talk refinements proposed by Farrell (1993) and Matthews, Okuno-Fujiwara, and Postlewaite (1991) can 1) make those solution concepts more realistic, 2) guarantee existence, and 3) guarantee meaningful communication in at least one class of games where it is not guaranteed by either Farrell or MOP.

Contact: Joe Sobel, Department of Economics, University of California at San Diego, La Jolla, CA 92093-0508.

1. Introduction

It is a basic tenet of non-cooperative game theory that if players settle upon predictable play, this play will be a Nash equilibrium. Yet many standard solution concepts propose that some Nash equilibria are themselves susceptible to deviations. Such deviation-based refinements have been developed in the cheap-talk literature (see Farrell (1993) and Matthews, Okuno-Fujiwara, and Postlewaite (1991)), and most prominently in the signaling literature (see Banks and Sobel (1987) and Cho and Kreps (1987)). While these refinements identify equilibria that are susceptible to deviations, they leave an important question unanswered: If players are likely to deviate from some equilibrium, what will they deviate to? We feel this question is important because some equilibria that are subject to deviations may still persist in the long run because they may be repeatedly deviated to. It seems inappropriate to omit such recurrent equilibria from the solution concepts with which we make predictions in economic models.

In this paper, we develop a framework to address this question, and use this framework to critique existing solution concepts. We then develop new solution concepts by adopting intuitions for when deviations occur from existing refinements, but using our framework to develop the dynamic implications of such intuitions.

Signaling refinements have been criticized because proposed deviations from equilibria often rely on an apparent inconsistency: some players defect from an equilibrium while others continue to believe in that equilibrium. Our approach to deviations takes this criticism, often called the *Stiglitz critique*, into account. We begin in Section 2 by defining for each equilibrium a *deviation correspondence*, which describes a set of the strategies that might

follow a deviation from that equilibrium. To assure that they are consistent with common knowledge, we construct deviation correspondences by starting with the proposed deviation, and iteratively adding best responses by each player until no more strategies need be added. We apply deviation correspondences to the arguments presented in Cho and Kreps (1987), and show that the Stiglitz critique does indeed have force--rational players might respond to hypothesized deviations from an equilibrium by continuing to play that equilibrium.¹

We complete our framework in Section 3 by characterizing certain broad properties of dynamics that (given a deviation correspondence) allow us to characterize a set of equilibria that may persist in the long run. We assume that players repeatedly play a game, and have a tendency to settle down on equilibrium behavior.² We assume that deviations may upset an equilibrium, but that following a deviation, play equilibrates to an equilibrium in the deviation correspondence. We define as *recurrent* those equilibria that are likely to be played repeatedly in the long run.³

To illustrate our approach, suppose that there are four equilibria, A, B, C, and D. Suppose that: A is not subject to deviations, B and C tend to deviate to each other but not to A or D, and that D either deviates to B or

¹ But we also show that the Stiglitz critique does not always recover unstable equilibria. For earlier arguments and examples of when the Stiglitz critique does and does not have force, see Mailath (1988).

² As is standard when modeling the equilibration process, we assume no strategic foresight by the players, so that every round of play is treated as a one-shot game.

³ Our exclusive focus on equilibrium rather than non-equilibrium outcomes makes sense in environments in which deviations are relatively rare and equilibration is relatively quick. We feel that some such story underlies the equilibrium-refinement literature. Foster and Young (1990) and Kandori, Mailath, and Rob (1992) develop explicitly dynamic models which assume that play will generally be in equilibrium, but that equilibrium behavior will periodically be interrupted by mutations.

does not deviate at all. The standard refinement literature would have us limit attention to equilibrium A, because it is the only fully stable equilibrium.⁴ But this seems inappropriate--if either equilibrium B or C occur, then equilibrium A will never occur again. Our notion of plausible long-run equilibria would therefore include B and C. Although a deviation from D is not a sure thing, however, we do exclude D on the grounds that because a deviation is plausible it is likely to occur eventually, after which play will never again return to D.

Despite our conceptual criticism of current refinements, we show in Sections 3 and 4 that Cho and Kreps's (1987) intuitive criterion and Kohlberg and Mertens's (1986) never-a-weak-best-response criterion have the property that equilibria are recurrent if and only if they are not susceptible to deviations. This is because their deviation correspondences always contain an equilibrium that is stable; we therefore assume that play will eventually deviate to a stable equilibrium, at which point no unstable equilibrium will again be played. If one accepts our dynamic, long-run interpretation of these refinements, then they are robust to the Stiglitz critique.

In Section 5, however, we show that our framework does not generally justify the existing approach to refinements. Based on the arguments of Matthews, Okuno-Fujiwara, and Postlewaite (1991), we develop a solution concept for cheap-talk games called *recurrent mop*, and provide some examples where recurrent mop improves upon existing cheap-talk solution concepts. We also show that recurrent mop guarantees communication in games of partially common interest; other common cheap-talk refinements do not have this

⁴ Throughout the paper, we refer to equilibria as "stable" if they are not subject to deviations, and as "unstable" if they are subject to deviations.

property.⁵

While our framework indicates that existing signaling refinements may be too liberal in eliminating equilibria, Section 5 illustrates that our approach has attractions for those who like their solution concepts strong. We feel that many existing solution concepts with non-existence problems are "too strong" not because they have the wrong intuition for which equilibria are unstable, but rather because unstable equilibria are too readily eliminated from the solution concept. By focusing on recurrent equilibria rather than stable equilibria, we guarantee existence while applying intuitive notions of deviations that lead to severe existence problems in the standard framework.

We conclude the paper in Section 6 with a discussion of a potential application of our approach, as well as some ideas on how to extend our approach to allow for non-equilibrium behavior, and some conjectures about what we might learn from such an extension.

2. Deviation Correspondences and Expansions

While our framework could be applied more generally, we shall develop it for finite-action, two-player games, and often restrict our attention further to simple signaling games. Consider a 2-player game with players 1 and 2. Player i has finite pure-strategy set S_i , with the set of strategy profiles $S \equiv S_1 \times S_2$, and payoff functions (u_1, u_2) . For any finite set X , we denote by $\Delta(X)$ the set of probability distributions on X , so that a mixed strategy for player i is an element in $\Delta(S_i)$. Given a set of strategies $X_j \subseteq \Delta(S_j)$, we

⁵ Blume, Kim, and Sobel's (1992) solution concept does, however, yield the same result.

denote by $BR_i(X_j)$ the set of best responses in S_i to X_j . That is, $BR_i(X_j) = \{s_i^* \in S_i : u_i(s_i^*, x_j) \geq u_i(s_i, x_j) \text{ for all } s_i \in S_i\}_{x_j \in X_j}$. We extend u to mixed strategies in the obvious way. We denote by N the set of Nash equilibria.

We associate with every equilibrium γ a *deviation correspondence* $D(\gamma)$, which summarizes the players' beliefs about whether and how deviations will occur. We define the set of Nash-equilibrium *outcomes* contained in a deviation correspondence $D(\cdot)$ as $ND(\cdot)$. We assume that each set $ND(\cdot)$ is finite. While typically there can be an infinite number of equilibria supporting any one equilibrium outcome, we shall sometimes abuse terminology and use "equilibrium" to mean an equilibrium outcome.

While one could develop deviation correspondences in many ways, we construct them in this paper using only one particular algorithm.⁶ Begin with an equilibrium $\gamma \equiv (\gamma_1, \gamma_2)$. Suppose that, for each player i , there is a (possibly empty) set of allowable deviations $Q_i \subseteq \Delta(S_i)$. Then we define:

$$\Sigma_i(0) \equiv \begin{cases} \gamma_i & \text{if } Q_i = \emptyset \\ A_i & \text{if } Q_i \neq \emptyset \end{cases}$$

That is, if there are no allowable deviations, $\Sigma_i(0)$ is simply player i 's equilibrium strategy; if there are one or more allowable deviations, then $\Sigma_i(0)$ consists of these possible deviations. To define a deviation correspondence based on $(\Sigma_1(0), \Sigma_2(0))$, we define $\Sigma_i(n)$ iteratively by setting $\Sigma_i(n) \equiv BR_i(\Delta(\Sigma_{-i}(n-1)))$ if $\Sigma_{-i}(n-1) \neq \gamma_{-i}$ and by $\Sigma_i(n) = \gamma_i$ if $\Sigma_{-i}(n-1) = \gamma_{-i}$. That is, if we have added strategies to player $-i$, we then add all best

⁶ Because deviation correspondences create sets of strategies consistent with common knowledge of rationality without imposing the equilibrium assumption, our approach here is similar to that taken in Rabin (1989, forthcoming(b)), who develops an approach to combining behavioral assumptions with rationalizability. Game-theoretic applications using similar ideas include Cho (1992), Rabin (1990, forthcoming(a)), Watson (1992), and Zapater (1992).

responses by player i to all beliefs over the new strategies; if we have not added strategies to player $-i$ --and thus continue to assume that player $-i$ will play his equilibrium strategy--then we continue to assume that player i will play his equilibrium strategy (or any deviations we have previously posited for him). Because each S_i is finite, we know there exists an n^* such that for all i and all k , $\Sigma_i(n^*) = \Sigma_i(n^*+k)$. Then let $\Sigma_i^* \equiv \Sigma_i(n^*)$. We refer to this process of constructing deviation correspondences as *expansions*:

Definition 1:

For a given equilibrium $\gamma \equiv (\gamma_1, \gamma_2)$, and deviation sets $Q \equiv (Q_1, Q_2)$, let the *expansion* of (γ, Q) , denoted $Exp(\gamma, Q)$, be the sets of strategies (Σ_1^*, Σ_2^*) constructed as outlined above.

We shall throughout this paper simply set $D(\gamma)$ equal to $Exp(\gamma, Q)$. Every expansion will include a Nash equilibrium--given that we do not stop adding in strategies until all best responses are included, our assumptions allow us to invoke fixed-point theorems in establishing existence.

To illustrate our approach, we develop a deviation correspondence in simple signaling games motivated by Cho and Kreps's (1987) intuitive criterion. There is an informed Sender, who has private information drawn from a finite set T according to a common-knowledge distribution $\pi(\cdot)$. Her set of pure strategies consists of rules that assign to each $t \in T$ a message m , which is a member of a finite set M . The uninformed Receiver observes which message m is chosen, and then chooses an action a , which is a member of a finite set A . Players have utility functions $u_S(t, m, a)$ and $u_R(t, m, a)$.

Figure 1 is an example of a signaling game. The top number in each pair is the Sender's payoff; the bottom number is the Receiver's payoff. The Sender's private information is either t_1 or t_2 , and she can choose message m_1 or m_2 , after which the Receiver chooses his action. (For ease of reference, we

label the actions by the Receiver differently depending on which message they follow.)

There is a sequential equilibrium of the game in which both types of the Sender choose m_2 and the Receiver chooses a_2 given m_1 and r_1 given m_2 . We denote this equilibrium by $((m_2, m_2), (a_2, r_1))$. This outcome fails the intuitive criterion of Cho and Kreps (1987). They argue that because t_2 receives her highest feasible payoff in this equilibrium, she would not deviate from the equilibrium, no matter how she thought the Receiver would respond to m_1 . If the Receiver understands this, then he will interpret m_1 as being sent by t_1 --who does have a plausible reason for deviating--and will respond to m_1 with a_1 . This argument undermines the equilibrium: If the Receiver chooses a_1 given m_1 , then t_1 prefers to send m_1 rather than send m_2 as specified by the equilibrium. Thus it is concluded that the equilibrium $((m_2, m_2), (a_2, r_1))$ is not stable.

The Stiglitz critique says that this analysis stops too soon, because the Receiver should reason that if t_1 is sending message m_1 , then she will not be sending m_2 . Hence he should respond to m_2 with r_2 , which is the optimal response given that only t_2 is sending m_2 . Now t_2 *does* gain by sending m_1 . In this way, the Receiver's original conjecture that m_1 must be sent by t_1 creates a rationale for t_2 sending m_1 as well. This provides a reason for the Receiver to choose a_2 given m_1 , which in turn supports the original equilibrium.

To show that this argument is captured by our framework, we use the deviations implied by intuitive criterion to construct the "intuitive deviation correspondence." Let $BR_R(K, m)$ denote the set of best responses by Receiver to the message m if his beliefs are concentrated on the set of types $K \subseteq T$. For the equilibrium γ , let $u^*(t)$ be the payoff for type t of Sender,

and let M^* be the set of signals sent with probability zero. We say that the strategy is in Q_S if it involves all players besides some type t^* sending their equilibrium signals, and involves t^* sending the message $m^* \in M^*$, where $u(t^*, m^*, a) > u(t^*)$ for all $a \in BR_R(J, m^*)$, where $J \equiv \{t \mid \text{there exists } a \in BR_R(T, m^*) \text{ such that } u(t, m^*, a) \geq u(t^*)\}$.

We also define deviations by the Receiver. A strategy is in Q_R if it has the Receiver playing a best response to the message m^* for some $s \in Q_S$, and playing the original equilibrium strategy in response to all strategies $m \neq m^*$. This implies that if Q_S is not empty, Q_R does not include the Receiver's equilibrium strategy, and does include all strategies involving off-the-equilibrium-path best responses which motivate the Sender to deviate. If $Q_S = \emptyset$, then $Q_R = \emptyset$.

If we apply our framework to the pooling equilibrium of Figure 1, we see that $\Sigma_S(0) = Q_S = \{(m_1, m_2)\}$, and $\Sigma_R(0) = Q_R = \{(a_2, r_1)\}$. In the expansion process, $\Sigma_S(1) = \Sigma_S(0)$, but $\Sigma_R(1)$ adds in the strategy (a_2, r_2) , because perceiving a deviation will mean that Receiver will change his response on the equilibrium path. Depending on Sender's beliefs about the likelihood of the Receiver's two strategies in $\Sigma_R(1)$, therefore, the Sender might now prefer to play the strategy (m_1, m_1) . Thus, $\Sigma_S(2)$ contains (m_1, m_1) . In turn, this means that $\Sigma_R(3)$ contains (a_2, r_1) , which is optimal if the Sender chooses (m_1, m_1) and that $\Sigma_S(4)$ contains (m_2, m_2) , so the original equilibrium $((m_2, m_2), (a_2, r_1))$ is contained in (Σ_S^*, Σ_R^*) , the intuitive deviation correspondence.

Because the "unintuitive" equilibrium is contained in its own deviation correspondence, the Stiglitz critique applies in Example 1. Figure 2 demonstrates, however, that it does not always redeem an unintuitive equilibrium. The pooling equilibrium $((m_2, m_2), (a_2, r_2))$ fails the intuitive criterion in this example because t_1 (and only t_1) could gain by deviating.

Because this is the only deviation designated by the intuitive criterion, $\Sigma_S(0) = Q_S = \{(m_1, m_2)\}$ and $\Sigma_R(0) = Q_R = \{(a_1, r_2)\}$. In the iteration process, (a_1, r_3) will be included in $\Sigma_R(1)$, but nothing more is added into either player's best-response set. The deviation correspondence will therefore be $(\Sigma_S^*, \Sigma_R^*) = (\{(m_1, m_2)\}, \{(a_1, r_2), (a_1, r_3)\})$, which contains only the separating equilibrium to this game, $((m_1, m_2), (a_1, r_3))$, and not the original equilibrium. The unintuitive equilibrium does not therefore become plausible even if we assume common knowledge of the posited deviation, so the outcome here is not stable in any sense: Even when we take the Stiglitz critique fully into account, the pooling equilibrium does not survive a deviation.

These examples illustrate how our approach formalizes the Stiglitz critique, and suggest that we could use our framework to define a "Stiglitz-proof version" of the intuitive criterion or any other signaling refinement. Translating into our framework, the current approach in the literature is to eliminate any equilibrium γ such that Q is not empty, so that $D(\gamma) \neq \{\gamma\}$. The modified approach could be to eliminate an equilibrium γ only if $\gamma \notin D(\gamma)$.

Of course, whenever $D(\gamma) \neq \{\gamma\}$, one might still say that an equilibrium γ is unstable in the sense that players *might* deviate. In a sense, some of the debate over the merits of signaling refinements and the Stiglitz critique may be about which of these two notions of "stability" is appropriate--should we build our solution concepts around equilibria with the property $D(\gamma) = \{\gamma\}$, or based on equilibria with the property that $\gamma \in D(\gamma)$? In the next section, we argue that *neither* view takes a sufficiently dynamic view of deviations.

3. Dynamics (Sort of)

We now consider the dynamic implications of deviation correspondences. Because we are considering the implications of deviations, and since deviations might not lead immediately to alternative equilibria, we would expect to frequently observe *non-equilibrium* play. We shall nonetheless focus exclusively on equilibrium outcomes. Partly we justify this by supposing that the rate of re-equilibration following a deviation is fast relative to the frequency of deviations. But we are also simply carrying over the equilibrium assumption from the solution concepts we are critiquing and modifying.

To begin developing the dynamic implications of our framework, we suppose that after play deviates from some Nash-equilibrium outcome, γ , it will eventually equilibrate on some outcome $\gamma' \in ND(\gamma)$. If a deviation correspondence is *self-contained*, in the sense that for all equilibria γ , and for all $\gamma' \in ND(\gamma)$, $ND(\gamma') \subseteq ND(\gamma)$, this will imply that following a deviation from γ , play will never again equilibrate outside the set of $ND(\gamma)$.⁷ Lemma 1 establishes that all the deviation correspondences of this paper are self-contained:

Lemma 1:

Suppose that a deviation correspondence is constructed using expansions with the properties that, for all equilibria σ and deviations $(Q_1(\sigma), Q_2(\sigma))$ ⁸,

- 1) for either $i = 1$ or 2 , $Q_i(\sigma) \subseteq BR_i(\sigma_{-i})$,
- 2) for both $i = 1, 2$, $Q_i(\sigma) \subseteq BR_i(\sigma_{-i}) \cup BR_i(Q_{-i}(\sigma))$.

Then for all γ , and for all $\gamma' \in ND(\gamma)$, $ND(\gamma') \subseteq ND(\gamma)$.

⁷ Of course, to fully justify this supposition, we need to more construct deviation correspondences with respect to *non-equilibrium* outcomes, to be sure that, for all outcomes $\sigma \in D(\gamma)$, $D(\sigma) \subseteq D(\gamma)$.

⁸ Note that this lemma is meant to include the case where one or both sets $Q_i(\sigma)$ are empty, where we consider $\emptyset$ to be a subset of everything.

Proof:

Let $\Sigma_i(\cdot)$ denote the expansion starting from the equilibrium γ . Assume that $\gamma' \in ND(\gamma)$ and that $\gamma' \in (\Sigma_1(n), \Sigma_2(n))$ for all $n \geq N$. If $\Sigma_i(n) = \gamma_i$ for some i , then it follows from the definition of the expansion process that $\Sigma_{-i}(n+1) = \gamma_{-i}$ and therefore that $\gamma = \gamma'$, so that the result follows. Otherwise it follows that $BR_{-i}(\gamma_i) \subset \Sigma_{-i}(n+1)$ for $i = 1$ and 2 . Suppose without loss of generality that $Q_1(\gamma') \subseteq BR_1(\gamma'_2)$. It follows that $Q_1(\gamma') \subseteq BR_1(\gamma'_2) \subseteq \Sigma_1(n+1)$, and $Q_2(\gamma') \subseteq BR_2(\gamma'_1) \cup BR_2(Q_1(\gamma')) \subseteq \Sigma_2(n+1) \cup \Sigma_2(n+2) \subseteq \Sigma_2(n+2)$. It follows from this that $(Q_1(\gamma'), Q_2(\gamma')) \subseteq (\Sigma_1(n+2), \Sigma_2(n+2))$, which implies that $D(\gamma') \subseteq D(\gamma)$. The lemma follows directly. Q.E.D.

In all the theories of deviations we consider in this paper, the sets (Q_1, Q_2) meet the premise of Lemma 1. In simple signaling games, for instance, the set Q_R of the Receiver's deviations will always be best responses to the Sender's strategy in the equilibrium from which they are deviating, and the Sender's deviations will be best responses to the deviations allowed the Receiver.

A trivial class of equilibria that might persist are those from which no deviations will occur. We call these equilibria *focal*:

Definition 2:

An equilibrium γ is *focal* if $D(\gamma) = \{\gamma\}$.

When constructed by expansions, an equilibrium is focal if and only if $Q = \emptyset$; i.e., if the equilibrium is not subject to deviations. It is of course focal equilibria that have been emphasized in the refinement literature. The simplest type of equilibrium ignored by the refinement literature that can persist in the long run is a *quasi-focal* equilibrium:

Definition 3:

An equilibrium is *quasi-focal* if $ND(\gamma) = \{\gamma\}$.

If a deviation from a quasi-focal equilibrium occurs, play must return to the original equilibrium when it next equilibrates. Unless we are willing to abandon exclusive focus on equilibrium outcomes, ruling out quasi-focal equilibria is inappropriate.

More generally, what can we say about the set of equilibria that might occur in the long run? Let $E(0)$ be the set of all Nash equilibria in a game. Now let $E(1) \equiv \{\gamma \mid \text{there exists } \gamma' \in E(0) \text{ such that } \gamma \in ND(\gamma')\}$. $E(1)$ is the set of equilibria that are in the deviation correspondence of some other equilibrium. If an equilibrium is not in $E(1)$, it clearly cannot persist in the long run, because it is itself not robust to deviations, and would never be deviated to by another equilibrium. Likewise, we can say that any equilibrium not contained in $E(2) \equiv \{\gamma \mid \text{there exists } \gamma' \in E(1) \text{ such that } \gamma \in ND(\gamma')\}$ will not be around in the long run. This suggests the following notion of which equilibria are likely to persist. Let $E(t+1) \equiv \{\gamma \mid \text{there exists } \gamma' \in E(t) \text{ such that } \gamma \in ND(\gamma')\}$.

Definition 4:

An equilibrium γ is *pseudo-recurrent* if, for all t , $\gamma \in E(t)$.

If an equilibrium is *not* pseudo-recurrent, then there exists a finite number of equilibrated periods after which that equilibrium will never be observed.⁹ If it is pseudo-recurrent, then, starting with some equilibrium, that equilibrium might occur infinitely often.

As the reader may infer from our terminology, we feel that not all pseudo-recurrent equilibria will necessarily persist in the long run. Consider

⁹ In principle, the motivation we develop for quasi-focal equilibria suggests Definition 4 should be: An equilibrium γ is pseudo-recurrent if, for all T , there exists $t \geq T$ such that $\gamma \in E(t)$. For self-contained deviation correspondences, however, this is equivalent to the definition in the text.

the case where there are two equilibria, γ and γ' , and a deviation correspondence $D(\cdot)$ such that $ND(\gamma) = \{\gamma, \gamma'\}$ and $ND(\gamma') = \{\gamma'\}$. What might we expect in the long run? We know that once γ' occurs, γ will never occur again. Also, $ND(\gamma) = \{\gamma, \gamma'\}$ suggests that γ' is a possible outcome after γ has occurred. If we suppose that there is some probability p such that, after γ occurs, γ' occurs in the next equilibrated period with at least probability p , then we can be certain that in the long run play will converge to γ' . Example 1 discussed above illustrates precisely this case. The unintuitive outcome is pseudo-recurrent, but seems unlikely to persist in the long run, because its deviation correspondence contains the separating sequential equilibrium, which is focal.

To formalize the notion that some equilibria are unlikely, we imagine that the equilibration process is *unbiased*, in the sense that there is a lower bound on the probability that any given element of $ND(\gamma)$ will occur following an equilibrium γ . Unbiased processes permit us a simple characterization of long-run behavior. If re-equilibration following a deviation is always unbiased, then in the long run players will almost always play an equilibrium in a *minimal* set of equilibria:

Definition 5:

A set of equilibria $\mathcal{S}$ is *minimal* with respect to a deviation correspondence $D(\cdot)$ if for all $\gamma \in \mathcal{S}$, $ND(\gamma) = \mathcal{S}$.

Once play reaches some equilibrium in a minimal set, play will stay there. Moreover, every equilibrium in some minimal set is a potential long-run outcome, because it could be deviated to infinitely often from another equilibrium that will itself occur infinitely often. We thus have a simple characterization of which equilibria will with positive probability persist in the long run:

Definition 6:

An equilibrium γ is *recurrent* if it is contained in some minimal set.

In any game and any deviation correspondence we can solve for the set of recurrent equilibria by figuring out what are all the minimal sets. Much of our subsequent analysis, however, will take advantage of a simple fact that follows straight from definitions:

Lemma 2:

If there exists a quasi-focal equilibrium $\gamma' \in ND(\gamma)$ such that $\gamma' \neq \gamma$, then γ is not recurrent.

Because any quasi-focal equilibrium is trivially recurrent, Lemma 2 means that if the deviation correspondence from every equilibrium contains a quasi-focal equilibrium, then an equilibrium is recurrent if and only if it is quasi-focal. Indeed, precisely this argument holds with respect to the intuitive deviation correspondence:

Proposition 1:

Any equilibrium that fails the intuitive criterion is not recurrent with respect to the intuitive deviation correspondence.

Proof:

Applying Lemma 2, we need only show that, for all equilibria γ , there exists γ' surviving the intuitive criterion such that $\gamma' \in ND(\gamma)$. This result holds trivially if γ survives the intuitive criterion. Suppose that it does not, and that the expansion of γ is (Σ_S^*, Σ_R^*) . Now consider the hypothetical game in which the payoffs are the same as the game being examined, but in which only the messages and actions in (M, Q^*) are available to the players, where Q^* is the set of actions played with positive probability (on or off the equilibrium path) by some strategy in Σ_R^* . We know by the Cho and Kreps existence theorem that, with respect to this hypothetical game, there exists $\sigma \in (\Sigma_S^*, \Sigma_R^*)$ that passes the intuitive criterion.

To complete our proof, we claim that if an equilibrium survives the

intuitive criterion in the hypothetical game, then it also survives the intuitive criterion in the actual game. Suppose not. Then there exists an unsent message, m^* , of the type invoked in the intuitive criterion. But we know by definition that whenever an equilibrium σ fails the intuitive criterion, it must be that strategies $s_r \in Q_r(\sigma)$ must be best responses to the equilibrium strategy, σ_s , of the Sender. But then such strategies would be added in during the expansion process that was used to construct (Σ_s^*, Σ_r^*) , so that the best responses to m^* must be contained in Q^* , which contradicts the assumption that σ does survive the intuitive criterion in the hypothetical game.

Q.E.D.

We conclude from Proposition 1 that, interpreting the refinement within a dynamic framework, there is a sense in which the intuitive criterion is robust against the Stiglitz critique.

4. The Never-a-Weak-Best-Response Deviation Correspondence

In this section we apply our framework to the never-a-weak-best-response (hereafter, NWBR) criterion developed by Kohlberg and Mertens (1987) and named and applied in Cho and Kreps (1987). While it is defined more generally, the NWBR criterion has proved especially important in simple signaling games. In fact, Banks and Sobel (1987) and Cho and Kreps (1987) demonstrate that any sequential equilibrium that fails the NWBR criterion cannot satisfy signaling refinements such as the intuitive criterion, divinity, or universal divinity.

Our main conclusion is that, as with the intuitive criterion, the NWBR criterion "has it right" with respect to dynamic implications: for games whose extensive forms have generic payoffs, the set of recurrent equilibria is precisely the set of focal equilibria.

We limit attention to games in which each connected set of Nash

equilibria gives rise to a single payoff pair.¹⁰ We now present a version of the NWBR criterion that is sufficient for our purposes. Fix an equilibrium, γ , with equilibrium payoffs $u^*(\gamma)$. Let $C(\gamma)$ be the connected set of Nash equilibria that contains γ , and let $R_i(\gamma)$ be the set of pure strategies for player i that receive positive weight in some element of $C(\gamma)$. Since we assume that all elements of $C(\gamma)$ lead to the same payoff, it follows that if $\sigma_i \in R_i(\gamma)$, then there exists $\sigma_{-i} \in R_{-i}(\gamma)$ such that $u_i(\sigma_i, \sigma_{-i}) = u_i^*$. Now form sets (R_1^*, R_2^*) as follows. First delete some subset of the strategies in the original strategy set S that are weakly dominated with respect to the sets $(R_1(\gamma), R_2(\gamma))$. Next, with the strategies that remain from S iteratively remove weakly dominated strategies in any order. Different orders of iterative deletion can yield different sets, so there may be more than one such (R_1^*, R_2^*) .

Define the NWBR deviation sets by saying that a strategy $s_i \neq \gamma_i$ is an allowable deviation if there exists (R_1^*, R_2^*) such that $u_i(s_i, r_{-i}) > u_i^*$ for all $r_{-i} \in R_{-i}^*$. That is, we allow a deviation by a player if that player expects to gain relative to the equilibrium against any strategy in R_{-i}^* . R_{-i}^* consists of strategies that are weak best responses to the strategies that support the outcome. Let $Q'_i(\gamma)$ be the set of all allowable deviations of player i . If $Q'_j(\gamma) \neq \emptyset$ for at least one player j , then $Q_i(\gamma) = Q'_i(\gamma) \cup R_i(\gamma)$ for each i . If, for all i , there is no such s_i that leads to a gain relative to the equilibrium, then $Q_i(\gamma) = \emptyset$. The NWBR deviation correspondence is then defined using an expansion of these deviation sets, as defined in Section 2.

As with the intuitive deviation correspondence, Lemma 2 can be used to

¹⁰ We know from the work of Kreps and Wilson (1982) and Kohlberg and Mertens (1986) that in all finite games, the set of Nash equilibria consists of a finite number of connected sets. For extensive-form games that have generic payoffs, each connected set of Nash equilibria is associated with one outcome. Consequently, for almost all games of interest, we can associate the same payoffs with all equilibria within a component.

show that focal and recurrent equilibria are equivalent when applying the NWBR criterion. The logic of the argument is precisely that of the result in the previous section on the intuitive criterion: We show that if an equilibrium does not satisfy the NWBR criterion, then its deviation must contain a focal equilibrium relative to the NWBR criterion. By Lemma 2, therefore, only focal equilibria are recurrent.

Proposition 2:

Any equilibrium that fails the NWBR criterion is not recurrent with respect to the NWBR deviation correspondence.

Proof:

Suppose instead that the equilibrium γ fails the NWBR criterion. It follows from the existence theorem of Kohlberg and Mertens (1986) that the deviation game contains a stable set. Let ρ be an element of this set. If we can show that ρ is a focal equilibrium relative to the NWBR deviation correspondence (for the original game), then the proposition follows by Lemma 2. Proposition 6 of Kohlberg and Mertens (1986) guarantees that relative to the strategies in the deviation game, there are no allowable deviations from ρ . Further, it follows from the construction of the deviation game that ρ is an equilibrium of the original game and that the component of Nash equilibria containing ρ in the deviation game is precisely the same as the component in the original game, since any strategy in the original game but not in the deviation game fails to be a best response to any probability distribution over the other player's deviation-game strategies. Hence, ρ is a focal equilibrium of the original game. Q.E.D.

5. Cheap-Talk Games and Recurrent Mops

A major point of our framework is that there may be recurrent equilibria that are not focal, but our discussion of the intuitive and NWBR deviation

correspondences did not reflect this point; with those solution concepts, all recurrent equilibria were focal equilibria. In this section, we present in the context of cheap-talk games a deviation correspondence where not all recurrent equilibria are focal. Cheap-talk games are simple signaling games with the property that players' payoffs do not depend directly on the message sent by the Sender. In other words, the payoffs are representable by the utility functions $u_S(t,a)$ and $u_R(t,a)$.

We shall construct deviation correspondences based on deviations that are weakly credible in the sense of Matthews, Okuno-Fujiwara, and Postlewaite (1991). They partition the Sender's types into disjoint groups, where one group sticks with the equilibrium and each other group sends a different message. They require that each type t gets a higher payoff from the Receiver's optimal response to t 's group than what t could get if it imitated another group's message. Formally:

Definition 7:

For all $t \in T$, let $u_S^*(t)$ be the payoffs type t gets in equilibrium γ . Call a subset J of T a *self-signaling family* relative to γ if there exists a partition of J into J_i , $i = 1, \dots, j$, and actions a_i^* such that:

- i) $a_i^* \in \operatorname{argmax}_{a \in AR} \sum_{t \in J_i} u_R(t,a)\pi(t)$ for $i = 1, \dots, j$;
- ii) $u_S(t, a_k^*) > \max \{u_S(t), u_S(t, a_i^*)\}$ for $t \in J_k$ and $i \neq k$, and
- iii) $u_S(t, a_i^*) < u_S^*(t)$ for all i if $t \notin J$.

While inspired by MOP (1991), Definition 7 incorporates a more liberal notion of allowable deviations than do MOP.¹¹ While we find MOP's restrictions

¹¹ When there are multiple statements that might be believed, or if the Receiver has more than one optimal response to given beliefs, MOP do not assume that he will choose the action that the relevant types of the Sender prefer. We allow a deviation if *some* beliefs and optimal response by the Receiver benefits the relevant types. See MOP (1991) and Rabin (1990) for a discussion of this issue. We make a further restriction for simplicity that does not substantively change our results: While MOP present a definition

compelling if interpreted as a theory of which deviations *necessarily* occur in a given play of the game, our perspective leads us to consider the broader class of deviations that are likely to eventually occur.

Because the allowable deviations are not stated in terms of strategies, further structure is needed to construct deviation correspondences. We follow Farrell (1993) in assuming that in every equilibrium, and for every subset of types $X \subseteq T$, there exists a "neologism"--an unused message $m(X)$ that means "I am some type $t \in X$." We assume that there are a finite number of messages used in equilibrium, and a finite set of potential neologisms, $\{m(L)\}_{L \subseteq T}$. For a given equilibrium γ , Definition 7 can be used to construct sets (Q_S, Q_R) . For every self-signaling family of types J , partition $\{J_1, \dots, J_N\}$, and best responses to a_i^* to J_i meeting the criteria of Definition 7, let σ_R be the strategy that responds to the neologism $m(J_i)$ by action a_i^* and to every other message according to the equilibrium strategy. Then let Q_R be the set of all such σ_R , and let Q_S be the set of optimal responses to beliefs over the set Q_R .

Because of the nature of cheap-talk games, we must modify slightly the expansion process used to create deviation correspondences. We add strategies in the iterative process only if they are "strong" best responses--i.e., we add σ_i to $\Sigma_i(t)$ only if there exist *full-support* beliefs over $\Sigma_{-i}(t-1)$ for which the strategy is a best-response. This reasonable restriction was not needed in our earlier examples, but is needed for non-trivial results in

valid for mixed-strategy deviations, we only allow for pure-strategy deviations.

cheap-talk games because payoffs are not generic with respect to the message space.

We call the solution concept created by applying our framework to this deviation correspondence *recurrent mop*. The game in Figure 3 illustrates some implications of recurrent mop, and how it differs from existing solution concepts. Assuming that the four types of Sender are equally likely, there is a pooling equilibrium where all types of the Sender send the same messages and the Receiver always takes action G. The pooling outcome is not plausible according to the arguments of MOP: types t_1 and t_2 could improve their payoffs by jointly deviating with the message "I am either t_1 or t_2 ", and types t_3 and t_4 could jointly deviate by sending the message "I am either t_3 or t_4 ." For that reason, the pooling equilibrium is not focal with respect to the mop deviation correspondence. Moreover, the deviation correspondence will contain the only other sequential equilibrium of the game.

This other equilibrium is a partially pooling outcome in which types t_1 and t_2 pool and induce the action C, while types t_3 and t_4 pool and induce action F. This equilibrium is also not focal with respect to the mop deviation correspondence. This equilibrium is susceptible, for instance, to type t_1 deviating from the equilibrium by self-signaling herself, inducing the action A rather than C. Likewise, type t_3 is tempted to self-signal herself to induce the action D.

MOP's original solution concept is empty in this game. But the partially pooling equilibrium is a recurrent mop, because it is quasi-focal with respect to the mop deviation correspondence. That is, a deviation from the partially pooling equilibrium will never lead players back to the pooling equilibrium. Deviations from the partially pooling equilibrium involve the different types trying to separate themselves further, but nothing in the logic of responding

to these deviations suggests that types $\{t_1, t_2\}$ would wish to pool with types $\{t_3, t_4\}$.¹²

We feel that recurrent mop makes the right prediction in this game--the subsets of types $\{t_1, t_2\}$ and $\{t_3, t_4\}$ will separate from each other. By using recurrence as the standard for prediction-making, we formulate a solution concept that captures that fact. By contrast, the non-existence of MOP's (1991) *announcement-proof equilibrium* leaves the analyst agnostic about the possible outcomes in this game. We also note that the solution concept *neologism-proof equilibrium* developed by Farrell (1993) predicts precisely the opposite conclusion--that only the pooling equilibrium will occur. In our opinion, this is because he is overly conservative in determining what

¹² Formally, consider the partially pooling equilibrium in which types t_1 and t_2 send the message m_{12} and types t_3 and t_4 send the message m_{34} . The mop deviation set Q_R will include responding to the message $m(t_1)$ with action A, and message $m(t_3)$ with action D, and responding to m_{12} and m_{34} with actions C and F. The set Q_S will therefore contain only the strategy where t_1 sends $m(t_1)$, t_2 sends m_{12} , t_3 sends $m(t_3)$, and t_4 sends m_{34} . (Note here why it is important that we not include "weak" best responses during iteration; doing so would allow, for instance, the Receiver to respond to message $m(t_1)$ with action D because he thought the message would be sent with probability zero.) In the iterative process, the Receiver would now add in strategies where he responds to m_{12} with B instead of C, and to m_{34} with E instead of F; this response is optimal because he knows that t_1 and t_3 are not sending their equilibrium messages. Given this, we would add to the Sender's strategies the possibility that t_2 would in fact send the message $m(t_1)$ and t_4 would send $m(t_3)$. This, in turn, means that the Receiver might respond to $m(t_1)$ and $m(t_3)$ with actions C and F instead of B and E. Thus, the deviation correspondence will consist of strategies where types t_1 and t_2 send some combination of the messages $m(t_1)$ and m_{12} , and t_3 and t_4 will send some combination of messages $m(t_3)$ and m_{34} . In particular, the deviation correspondence contains the original partially pooling equilibrium (using either the original messages, or the neologisms). We do not add the fully pooling equilibrium, and hence (since there are no other equilibrium outcomes) the partially pooling equilibrium is quasi-focal.

constitutes credible deviations from an equilibrium, but is overly liberal in rejecting the partially pooling equilibrium because it is not focal.¹³

So far, all our examples of recurrent equilibria have been quasi-focal. In our final example, we illustrate that non-quasi-focal equilibria can be recurrent. Consider the game in Figure 4.¹⁴ There are three equilibrium outcomes that involve partial pooling (for example, type t_1 induces the action D, while the other two agents induce A) and a separating equilibrium. It is straightforward to check that all equilibria permit a mop deviation. Starting from the outcome in which the Receiver only plays actions A and D with positive probability, there is a deviation that induces C. The deviation theory must then include the action F (as the Receiver allows the possibility that types t_1 and t_2 will induce C and type t_3 will send another message). Consequently, the deviation correspondence must contain the outcome in which types t_1 and t_2 separate from the other type. In this way, one can verify that any minimal set for the game in Figure 4 must contain all three of the semi-pooling equilibria. In fact, while none of the equilibria are quasi-focal in this game, all four are recurrent.

We conclude with a more general result showing that recurrent mop can guarantee a minimal form of efficiency that other solution concepts do not

¹³ While our framework can be used to guarantee existence, we don't feel it should be applied only when there is non-existence using other solution concepts. There are instances where existing refinements clearly are over-selective in selecting equilibria even when existence is not a problem. In Figure 3, for instance, a recurrent version of Farrell's (1993) solution concept would conclude that both equilibria are possible, whereas Farrell (1993) allows only the fully pooling equilibrium. While Farrell (1993) may correctly identify the partially pooling equilibrium as somewhat unstable, his proposed deviations involve the players separating even more, so that nothing in the logic of his proposed deviations suggests that the players will deviate back to the pooling equilibrium.

¹⁴ This game is modeled after an example in Myerson (1989), which is itself a variation of a game introduced by Moulin and Vial (1978).

guarantee. In particular, we can show that all recurrent pools involve meaningful communication in games of *partial common interests*.¹⁵ In games of partial common interests, the Sender has a compelling interest to share with the Receiver some, but not necessarily all, of her private information. To define such games formally, we need to develop some initial definitions. For any non-empty subset of types L , let $A^*(L) \equiv \{\arg\max_{t \in L} u_R(t, a) \mid \mu \text{ is a probability distribution supported on } L\}$ and $A^*(L, \pi) \equiv \{\arg\max_{t \in L} u_R(t, a) \mid \pi(t)\}$. The set $A^*(L)$ contains all of the responses that an optimizing Receiver would consider assuming that the Sender's type is an element of L , while $A^*(L, \pi)$ requires the Receiver to derive the relative probabilities of the types in L from the prior π . Also let $\underline{u}_S(t; L) \equiv \min \{v_S(t, a_i) : a_i \in A^*(L)\}$ denote the lowest payoff that the type t would obtain if the Receiver believed that her type was in L , and let $v_S^D(t)$ denote the maximum payoff that a type t can obtain in a completely pooling equilibrium. (For cheap-talk games with generic payoffs, there will only be one pooling equilibrium payoff.)

Definition 8:

A cheap-talk game has *partial common interests* if there exists a partition $J_1, \dots, J_j$ of T such that:

- i) $\underline{u}_S(t; L) > \max \{u_S(t_i, a_k) : a_k \in A^*(J_k)\}$ for all $t_i \in J_i \neq J_k$,
- ii) for all $t_i \in J_i$, there exists $a_i \in A^*(J_i, \pi)$ such that $u_S(t_i, a_i) > v_S^D(t_i)$, and
- iii) if $L \cap J_k \neq \emptyset$ for at least two k , then for each $a \in A^*(L)$ there exists an i and $t_i \in J_i$ such that $\underline{u}_S(t_i; J_i) > u_S(t_i, a)$.

The definition is meant to capture the intuition that it is in the interest of both players for types in sets J_i to reveal (at least) that

¹⁵ Blume, Kim, and Sobel (1993) obtain a related result in their study of evolutionary stability in cheap-talk games, but to our knowledge, the result holds for no other cheap-talk solution concept.

information. Condition (i) is a strong condition that guarantees that types in J_i prefer to identify themselves as members of J_i rather than as members of any other element of the partition. Condition (ii) states that each type would prefer to identify herself as a member of the partition that contains it rather than be pooled.¹⁶ Condition (iii) requires that, relative to what is available by being treated as a member of the set it belongs to, at least one type loses when members of different elements of the partition pool. We use this condition to show that once the population arrives at a strategy that reveals the partition J , it will never drift to a less informative strategy profile. This condition follows from (i) whenever $A = \bigcup_{i=1}^J BR_R(J_i)$.

Proposition 3:

No pooling equilibrium is a recurrent mop in a game with partial common interests.

Proof:

We prove the proposition in two steps. First we show that in a game with partial common interests the deviation correspondence beginning from a pooling equilibrium must contain a partially revealing equilibrium in which types in different sets J_i send different messages. Second we show that the deviation correspondence that starts from an equilibrium in which types in different J_i send different messages contains only equilibria with the same property. Formally, let $\{J_i\}$, $i = 1, \dots, j$ be the partition in Definition 8, let γ^P be a pooling equilibrium and let $B = \{(\gamma_R, \gamma_S) | (\gamma_R, \gamma_S) \text{ is a pure-strategy Nash equilibrium and } \gamma_S(m, t_i) \neq \gamma_S(m, t_k) \text{ whenever } t_i \in J_i \text{ and } t_k \in J_k \text{ for } i \neq k\}$. Step 1 demonstrates that $ND(\gamma^P) \cap B \neq \emptyset$, while Step 2 demonstrates that $ND(B) \subset B$. It follows from the definitions that there is no recurrent mop containing γ^P . It follows from Definition 8 that B is non-empty (it must contain a partially pooling outcome in which types in the same J_i send the same message, but separate from types in other J_i ; to support the equilibrium the Receiver

¹⁶ Conditions (i) and (ii) therefore together guarantee that J is a self-signaling family of sets relative to any pooling equilibrium payoff.

can respond to any message with an action from one of the sets $A^*(J_i, \pi)$. It is apparent that the deviation correspondences starting from γ^P contains this outcome: Q_R and Q_S add the appropriate equilibrium strategies.

It remains to show that $ND(B) \subset B$. Start with any equilibrium in B . It follows from (iii) of Definition 7 that any self-signaling family that exists relative to an equilibrium in B must be a subpartition of $\{J_i\}$. It follows from (i) and the full-support assumption on conjectures that the expansion process only admits actions contained in $A^*(L)$ for $L \subset J_i$ for some i (and does not alter the Receiver's responses to messages that are not sent in the original equilibrium). Therefore, if type $t_i \in J_i$ sends message m_i with positive probability under $\gamma \in B$, then every strategy for the Receiver that is in $D(\gamma)$ must respond to m_i with an element of $A^*(J_i)$. It follows from (iii) that no element of $ND(\gamma)$ can pool types from different elements of the partition. This completes the proof. Q.E.D.

Proposition 3 states that in games with partial common interests, players will not babble uninformatively forever. Along with the examples of this Section, Proposition 3 demonstrates that we can construct solution concepts within our framework that are both more powerful and more realistic than existing concepts.

6. Discussion

We conclude with informal discussions of two areas of game-theoretic research to which the general principles outlined in this paper may, we feel, be usefully applied.

Bernheim, Peleg, and Whinston (1987) consider the frequent supposition in game theory that, in communication-rich environments, only Pareto-efficient equilibria will be played. They follow Aumann (1959) in pointing out a problem with this hypothesis in multi-person games: A Pareto-efficient Nash

equilibrium may be susceptible to a coalition of players renegotiating their behavior so as to yield them all higher payoffs, given that other players continue to play their equilibrium strategies. While this will leave some other players worse off, such renegotiation still seems likely if there is opportunity for private communication.

To deal with this issue, Aumann (1959) proposed the solution concept of *strong Nash equilibrium*, which rules out all Nash equilibria that are susceptible to any such deviating coalition. Bernheim, Peleg, and Whinston critique strong Nash equilibrium as ruling out too many equilibria, because it applies no test as to whether the outcome negotiated by the deviating coalition is *itself* susceptible to renegotiation. They define a solution concept, *coalition-proof Nash equilibrium*, in which equilibria are ruled out only if there exists some beneficial coalitional renegotiation that is itself free from further coalitional renegotiation.

Our model does not directly apply to this issue, partly because we have not formalized it for multi-player games, but mostly because it is hard to conceptualize common knowledge, etc., without an explicit model of how players communicate. Yet our framework suggests that both strong Nash equilibrium and coalition-proof Nash equilibrium may be misleading in focusing too much on whether an equilibrium is "stable", rather than whether it is "recurrent".

While Bernheim, Peleg, and Whinston may be correct in suggesting that many of the renegotiations allowed by strong Nash equilibrium need not lead to stable behavior, we question their inference that such renegotiations are unlikely. In our framework, we would instead allow all such renegotiations, and consider their the long-run implications. Doing so, we could obtain an equilibrium concept that incorporates intuitive notions of when equilibria are

subject to renegotiation, but which guarantees general existence.

While we have focused on equilibrium outcomes throughout this paper, the equilibrium hypothesis itself has come under attack in recent years (see, e.g., Bernheim (1984) and Pearce (1984)). We wish to conclude by discussing how our framework could perhaps be useful both in exploring the foundations of equilibrium analysis and in developing useful non-equilibrium theories.

Suppose that we defined deviation correspondences with respect to non-equilibrium outcomes as well as equilibrium outcomes, and said that a set of outcomes O is *minimal* with respect to a (self-contained) deviation correspondence $D(\cdot)$ if for all outcomes $\gamma \in O$, $D(\gamma) \subset O$. We could then define "recurrent outcomes" as those outcomes that are contained in some minimal set. While this definition maintains the hypothesis that the dynamic process is "unbiased" enough so that play eventually converges to a minimal set, it no longer leads us automatically to focus only on equilibrium outcomes.

Of course, because focal equilibria are themselves minimal sets, they would still be natural candidates as outcomes upon which play will settle. In fact, suppose that every deviation correspondence constructed from either the intuitive or NWBR criterion contains a focal equilibrium, even for deviation correspondences generated by non-equilibrium outcomes. Then our framework would suggest that, using either of these criteria, only focal equilibria will occur in the long run, even if we do not *a priori* focus only on equilibria.

As we showed in Section 5, however, there are games where no equilibrium is focal according to the mop deviation correspondence. Thus, the implied solution concept "recurrent mop rationalizability" would include non-equilibrium as well as equilibrium outcomes. It should nonetheless have

17 Neither Aumann (1959) nor Bernheim, Peleg, and Whinston (1987) guarantee existence.

some predictive power. In Example 3, for instance, although it would not uniquely predict the partially pooling equilibrium, recurrent mop rationalizability will include only outcomes in which there is partial separation of the differing types.

The above conjectures all suggest that a variant of our framework can be used not only to investigate which equilibria will occur in the long run, as we have done in this paper, but also to investigate whether play will converge to equilibrium and, if play does not necessarily converge, which non-equilibrium outcomes might recur infinitely often.

References

- Aumann, R. (1959), "Acceptable Points in General Cooperative N-Person Games," in Contributions to the Theory of Games IV, Princeton, N.J.: Princeton University Press.
- Banks, J. and Sobel, J. (1987), "Equilibrium Selection in Signaling Games," Econometrica 55, 647-662.
- Bernheim, B.D. (1984), "Rationalizable Strategic Behavior," Econometrica 52, 1007-1028, July.
- Bernheim, B.D., Peleg, B. and Whinston, M. (1987), "Coalition-Proof Nash Equilibria I: Concepts," Journal of Economic Theory 42, 1-12.
- Blume, A., Kim, Y.-G., and Sobel, J. (1993), "Evolutionary Stability in Games of Communication," Games and Economic Behavior, forthcoming.
- Blume, A. and Sobel, J. (1992), "Communication-Proof Equilibria in Cheap-Talk Games," mimeo, University of California--San Diego.
- Cho, I.-K. (1992), "Stationarity, Rationalizability and Bargaining," Department of Economics Working Paper 92-127, University of Chicago, May.
- Cho, I.-K. and Kreps, D. M. (1987), "Signaling Games and Stable Equilibria," Quarterly Journal of Economics 102 May, 179-221.
- Crawford, V. and Sobel, J. (1982), "Strategic Information Transmission," Econometrica 50, 1431-1451, November.
- Farrell, Joseph (1993), "Meaning and Credibility in Cheap Talk Games," Games and Economic Behavior, forthcoming.

- Foster, D. and Young, P. (1990), "Stochastic Evolutionary Game Dynamics," Theoretical Population Biology 38, 219-232.
- Grossman, S. and Perry, M. (1986), "Perfect Sequential Equilibrium," Journal of Economic Theory 39, 97-119.
- Kandori, M., Mailath, G., and Rob, R. (1992), "Learning, Mutation, and Long Run Equilibria in Games," manuscript, University of Pennsylvania.
- Kohlberg, E. and Mertens, J.F. (1986), "On the Strategic Stability of Equilibria," Econometrica 54, 1003-1038.
- Mailath, G. (1988), "A Reformulation of a Criticism of the Intuitive Criterion and Forward Induction," manuscript, University of Pennsylvania.
- Matthews, S., Okuno-Fujiwara, M. and Postlewaite, A. (1991), "Refining Cheap-Talk Equilibria," Journal of Economic Theory 55, 247-273.
- Moulin, H. and Vial, J.-P. (1978), "Strategically Zero-Sum Games: The Class Whose Completely Mixed Equilibria Cannot Be Improved Upon," International Journal of Game Theory 7, 201-221.
- Myerson, R. (1989), "Credible Negotiation Statements and Coherent Plans," Journal of Economic Theory 48, 264-291.
- Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," Econometrica 52, 1029-1050, July.
- Rabin, M. (1989), "Predictions and Solution Concepts In Non-Cooperative Games," PhD Dissertation, MIT.
- Rabin, M. (1990), "Communication Between Rational Agents," Journal of Economic Theory 51, 144-170.
- Rabin, M. (forthcoming(a)), "A Model of Pre-Game Communication," Journal of Economic Theory.
- Rabin, M. (forthcoming(b)), "Incorporating Behavioral Assumptions into Game Theory," in James Friedman (ed.), Problems of Coordination in Economic Activity, Norwell, MA: Kluwer Academic Publishers.
- Watson, J. (1992), "A 'Reputation' Refinement without Equilibrium," mimeo, Stanford University.
- Zapater, I. (1992), "Credible Proposals in Communication Games," mimeo, Brown University.

Examples

Figure 1

Figure 2

	A	B	C	D	E	F	G
t_1	9, 10	6, 7	7, 9	4, 0	4, 0	4, 0	3, 6
t_2	6, 7	5, 10	7, 9	4, 0	4, 0	4, 0	3, 6
t_3	4, 0	4, 0	4, 0	9, 10	6, 7	7, 9	3, 6
t_4	4, 0	4, 0	4, 0	6, 7	5, 10	7, 9	3, 6

Figure 3

	A	B	C	D	E	F
t_1	0, 0	5, 4	4, 5	2, 6	1, 0	1, 0
t_2	4, 5	0, 0	5, 4	1, 0	2, 6	1, 0
t_3	5, 4	4, 5	0, 0	1, 0	1, 0	2, 6

Figure 4

May 12, 1993

**Working Paper Series
Department of Economics
University of California, Berkeley**

Individual copies are available for \$3.50 within the USA and Canada; \$6.00 for Europe and South America; and \$7.00 for all other areas. Papers may be obtained from the Institute of Business and Economic Research: send requests to IBER, 156 Barrows Hall, University of California, Berkeley CA 94720. Prepayment is required. Make checks or money orders payable to "The Regents of the University of California."

- 91-175 "Market Socialism: A Case for Rejuvenation." Pranab Bardhan and John E. Roemer. July 1991.
- 91-176 "Designing A Central Bank For Europe: A Cautionary Tale from the Early Years of the Federal Reserve." Barry Eichengreen. July 1991.
- 91-177 "Restructuring Centrally-Planned Economies: The Case of China in the Long Term." John M. Letiche. September 1991.
- 91-178 "Willingness to Pay for the Quality and Intensity of Medical Care: Evidence from Low Income Households in Ghana." Victor Lavy and John M. Quigley. September 1991.
- 91-179 "Focal Points in Pre-Game Communication." Matthew Rabin. September 1991.
- 91-180 "Cognitive Dissonance and Social Change." Matthew Rabin. September 1991.
- 91-181 "European Monetary Unification and the Regional Unemployment Problem." Barry Eichengreen. October 1991.
- 91-182 "The Effects of Competition on Executive Behavior." Benjamin E. Hermalin. October 1991.
- 91-183 "The Use of an Agent in a Signalling Model." Bernard Caillaud and Benjamin Hermalin. October 1991.
- 91-184 "The Marshall Plan: History's Most Successful Structural Adjustment Program." J. Bradford De Long and Barry Eichengreen. November 1991.
- 91-185 "Incorporating Fairness into Game Theory." Matthew Rabin. December 1991.
- 91-186 "How Pervasive Is the Product Cycle? The Empirical Dynamics of American and Japanese Trade Flows." Joseph E. Gagnon and Andrew K. Rose. December 1991.
- 92-187 "Shocking Aspects of European Monetary Unification." Tamim Bayoumi and Barry Eichengreen. January 1992.
- 92-188 "Is There a Conflict Between EC Enlargement and European Monetary Unification?" Tamim Bayoumi and Barry Eichengreen. January 1992.
- 92-189 "The Marshall Plan: Economic Effects and Implications for Eastern Europe and the Soviet Union." Barry Eichengreen and Marc Uzan. January 1992.
- 92-190 "Exploring the Relationship between R&D and Productivity at the Firm Level in French Manufacturing." Bronwyn H. Hall and Jacques Mairesse. February 1992.
- 92-191 "Three Perspectives on the Bretton Woods System." Barry Eichengreen. February 1992.

- 92-192 "Are Futures Margins Adequate?" Roger Craine. April 1992.
- 92-193 "Heterogeneity in Organizational Form: Why Otherwise Identical Firms Choose Different Incentives for Their Managers." Benjamin E. Hermalin. May 1992.
- 92-194 "Investment and Research and Development at the Firm Level: Does the Source of Financing Matter?" Bronwyn H. Hall. May 1992.
- 92-195 "The Determinants of Efficiency and Solvency in Savings and Loans." Benjamin E. Hermalin and Nancy E. Wallace. May 1992.
- 92-196 "Economics of Development and the Development of Economics." Pranab Bardhan. June 1992.
- 92-197 "The Core and the Hedonic Core: Equivalence and Comparative Statics." Suzanne Scotchmer and Greg Engl. July 1992.
- 92-198 "Incorporating Behavioral Assumptions into Game Theory." Matthew Rabin. July 1992.
- 92-199 "Incorporating Fairness into Game Theory and Economics." Matthew Rabin. July 1992.
- 92-200 "A Consumer's Guide to EMU." Barry Eichengreen. July 1992.
- 92-201 "Macroeconomic Adjustment Under Bretton Woods and the Post-Bretton-Woods Float: An Impulse-Response Analysis." Tamim Bayoumi and Barry Eichengreen. August 1992.
- 92-202 "Oligopoly Structure and the Incidence of Cigarette Excise Taxes." Paul G. Barnett, Theodore E. Keeler, Teh-wei Hu. September 1992.
- 92-203 "The Impact of 1989 California Major Anti-Smoking Legislation on Cigarette Consumption: Three Years Later." Teh-wei Hu, Jushan Bai, Theodore E. Keeler, Paul G. Barnett. September 1992.
- 92-204 "A Dynamic Simultaneous-Equations Model for Cigarette Consumption in the Western States." Hai-Yen Sung, Teh-Wei Hu, and Theodore E. Keeler. November 1992.
- 92-205 "Maastricht: Prospect and Retrospect." John M. Letiche. December 1992.
- 92-206 "Nonparametric Methods to Measure Efficiency: A Comparison of Methods." Richard F. Garbaccio, Benjamin E. Hermalin, Nancy E. Wallace. December 1992.
- 93-207 "The Value of Intangible Corporate Assets: An Empirical Study of the Components of Tobin's Q." Bronwyn H. Hall. January 1993.
- 93-208 "R&D Tax Policy During the 1980s: Success or Failure?" Bronwyn H. Hall. January 1993.
- 93-209 "Hospital Costs and Excess Bed Capacity: A Statistical Analysis." Theodore E. Keeler and John S. Ying. April 1993.
- 93-210 "Homothetic Preferences, Homothetic Transformations, and the Law of Demand in Exchange Economies." John K.-H. Quah. April 1993.
- 93-211 "Deviations, Dynamics and Equilibrium Refinements." Matthew Rabin and Joel Sobel. May 1993.