Information Transmission and Preference Similarity*

Andreas Blume

May, 1996

Abstract

This paper examines sets of Nash equilibria in sender-receiver games that are stable against replacement by alternative Nash equilibria. Such stable sets exist. In *partial common interest games* they contain only informative equilibria. The stability requirement sharpens currently available predictions for such games by (1) weakening the partial common interest condition, (2) ruling out strictly dominated actions, (3) reflecting the informativeness of the sender's strategy in the receiver's reply, and (4) by ruling out pooling actions. This approach is then used as a step toward unifying the study of partial common interest games and of Crawford and Sobel's parametric model of preference similarity.

*I thank, without implicating, Doug DeJong and Joel Sobel for comments. I have benefitted from comments by seminar participants at the University of Indiana and the University of Iowa. Support by the National Science Foundation under award number SBR-9410588 is gratefully acknowledged.

1 Introduction

The goal of this paper is twofold. One is two provide strong and refutable predictions for sender-receiver games with imperfectly aligned interests, without having to rely on auxiliary assumptions like exogenous meanings of messages, nominal signaling costs or restrictions on the interpretation of zero-probability messages, which are common in the literature. The second objective is to begin a formal study of equilibrium selection in Crawford and Sobel's [1982] model of strategic information transmission. This dual purpose mirrors the two approaches in the literature to expressing imperfect incentive alignment in sender-receiver games formally, either through a *partial common interest* condition, Blume, Kim and Sobel [1993], Rabin and Sobel [1996], Blume [1996], or as in Crawford and Sobel's parametric model of preference similarity.

To this end this paper examines sets of Nash equilibria in sender-receiver games that are stable against replacement by alternative Nash equilibria. Such stable sets exist. In *partial common interest games* they contain only informative equilibria. The stability requirement sharpens currently available predictions for such games by (1) weakening the partial common interest condition, (2) ruling out strictly dominated actions, (3) reflecting the informativeness of the sender's strategy in the receiver's reply, and (4) by ruling out pooling actions. This approach is then used as a step toward unifying the study of partial common interest games and of Crawford and Sobel's parametric model.

In sender-receiver games a privately informed sender sends a message to a receiver who takes an action. Payoffs are determined jointly by the sender's private information and the receiver's type; messages do not affect payoffs directly. If the sender's and the receiver's incentives are sufficiently closely aligned, the sender wants to reveal at least some of his private information. However, there is always an uninformative "babbling" equilibrium, in which the sender's messages contain no information and the receiver responds to all messages with the same action.

If the sender and the receiver can both gain from sharing some information, we might expect them to do so if they are able to focus on particular interpretation of messages. In the refinement literature, Farrell [1993], Matthews, Okuno-Fujiwara and Postlewaite [1991] and Rabin [1990], a focal mechanism is provided through an existing commonly understood language. In this paper we take the view that focal points, as considered by Schelling [1960], may emerge endogenously in the process of repeated play.[1]

In an environment where population play has settled on a particular equilibrium, other strategies can emerge through either mistakes or experimentation. This paper takes the position that such strategies can become focal points for the reorganization of the environment. They have the best chance of doing so if (1) they are best replies to the current environment and (2) they are themselves equilibria. Under those circumstances no player incurs a loss by switching from the status quo to the alternative equilibrium, even if he is matched with a player who remains at the status quo. Furthermore, since the alternative is an equilibrium, there is no ex post regret if other players behave in the same manner, and switch. Finally, at least in two-player games, if the two conditions are satisfied for a candidate replacement equilibrium, they remain satisfied even if initially only a fraction of the population adopts the replacement strategy. Nonequilibrium mistakes or experiments lack this coordinating potential.

The notion that a solution should consist of equilibria and be stable against the replacement by alternative equilibria is formalized in the conditions for a set of strategies to be an *equilibrium entry resistant (EER)* set. An *EER* set is a minimal closed set of Nash equilibria that is closed under the inclusion of (replacement by) strategies that are best replies both against an element of the set, and against themselves. The solution concept proposed here combines ideas from the evolutionary literature and the refinement literature. From the evolutionary literature it takes the idea of endogeneity, from the refinement literature it borrows the idea of focal points.

EER sets exist in every finite game and consist solely of equilibria. This accords well with experimental results on the emergence of meaningful communication in sender-receiver games. Blume, DeJong, Kim and Sprinkle [1995] examine the endogenous emergence of meaning for *a priori* meaningless messages experimentally. They consider both *common interest games,* in which there is a unique payoff pair that maximizes both the sender's and the receiver's payoff, and *divergent interest games* in which the sender's and the receiver's preferences over equilibrium outcomes are opposed. Equilibria play a central role in classifying the data. This is even true in games of divergent interests where many set-valued evolutionary solutions either do not exist or contain nonequilibrium strategies.

[1] The endogenous assignment of meanings to messages is not only relevant for a priori meaningless messages but also for messages that are part of a commonly understood language and whose meanings are subject to deterioration and recoding in specific strategic environments.

They find that while both revealing and nonrevealing equilibria may emerge in these games, equilibrium behavior itself is robust.

EER sets are attractive for their conceptual simplicity. They make powerful predictions in sender-receiver games without relying on auxiliary assumptions like exogenously given meanings of messages, nominal signaling costs, limitations on drift, interpretation of zero-probability messages etc. that are commonly made in this literature, both in static solution concepts and in explicit dynamic models.

The simplicity of *EER* sets comes at a cost. We do not have a justification of *EER* sets as the long-run limits of some explicit dynamic that relies only on simple behavioral rules governed by local conditions. Equilibrium play is postulated rather than explained as the result of dynamic adjustment. This does not detract from the value of *EER* sets as a solution concept. In the laboratory setting we deal with subjects who are capable of some degree of strategic analysis of a game and who become more sophisticated strategists during repeated play. Their behavior is likely to be governed as much by a global analysis of the game as by myopic payoff considerations based on the current state of play. Undoubtedly, equilibria have a prominent place among the set of all strategies in a game. In this paper it is taken as given that players know an equilibrium when they see it. For such players an emergent equilibrium can become a rallying point if playing according to that equilibrium is to their advantage. Nonequilibrium experiments or mistakes lack this property. They cannot serve as organizing precedents because they lack the ability to coordinate players' expectations.

We bring the EER condition and a weaker variant of it to bear on the two approaches to imperfect incentive alignment in sender-receiver games that have been considered in the literature. One is via a *partial common interest condition,* the other is Crawford and Sobel's [1982] parametric model. The former has played a role in evolutionary and refinement analyses of such games. This theme is taken up in the first part of the paper. It is shown that *EER* sets reject pooling equilibria under a considerably weaker partial common interest condition than those proposed in the literature. In addition, the set of strategies included in a solution is in general much smaller and more tractable than if nonequilibrium strategies are admitted as part of the solution. This is a great advantage if we want to test the theory experimentally.

One can enhance tractability further by strengthening the partial common interest condition; the resulting condition is still weaker than those in the literature. The stronger condition permits a straightforward characterization of the equilibria that form the solution without the need for characterizing the entire set of equilibria.

In two-player games (such as sender-receiver games) there is a close relationship between *EER* sets and set-valued evolutionary solution concepts. Essentially, if one of the evolutionary solutions exists, it contains an *EER* set. Therefore, if a certain partition of the type set is identified in an evolutionary solution, there is also an *EER* set in which it is identified. On the other hand, it turns out that noninformative equilibria are just as easily destabilized under the *EER* condition as under the evolutionary solutions. As a result, *EER* sets make more definite predictions in sender-receiver games.

The second approach to imperfect incentive alignment has attracted less attention in the evolutionary and refinement literature. Farrell [1993] pointed out that his *neologism proofness* criterion tends to reject all equilibria in Crawford and Sobel's quadratic model. The evolutionary literature on sender-receiver games does not address the selection issue in Crawford and Sobel's model at all. In the present paper it is shown that by weakening the replacement condition in the definition of *EER* sets we can reject pooling equilibria both when incentives are closely aligned in Crawford and Sobel's model and in a class of partial common interest games.

The remainder of the paper is organized as follows. Section 2 introduces sender-receiver games and the solution concept. Section 3 establishes the important role played by unused messages. Section 4 characterizes stable outcomes in games satisfying a *partial common interest* condition. In section 5 the partial common interest condition is strengthened to permit a characterization of stable outcomes without having to identify the entire set of Nash equilibria. Section 6 relates *EER* sets to evolutionary solutions in two-player games. Section 7 explores ways of unifying the analysis of partial common interest games and of Crawford-Sobel games. The final section discusses the literature.

2 Preliminaries

This section describes sender-receiver games, defines and the solution concept, and proves existence of a solution in finite games.

In a sender-receiver game player 1, the sender, has private information and player 2, the receiver, takes an action. Before the receiver takes his action, the sender sends a message. Payoffs to both players depend solely on the sender's private information, his type, and on the receiver's action.

Let T be the finite set of types and $\pi(\cdot)$ the prior distribution of types. The sender's set of pure strategies is the set of mappings from the type set to the finite set of messages, M. The receiver's set of pure strategies is the set of mappings from $M = \{m_i\}_{i=1}^k$ to the finite set of actions A. Given type $t \in T$, message $m \in M$, and action $a \in A$, player i's payoff is $v_i(t, a),\ i = 1, 2$. Messages do not directly affect payoffs. For any finite set X, denote by $\Delta(X)$ the set of probability distributions over X. The payoff from a mixed action $\alpha \in \Delta(A)$ is $v_i(t, \alpha) = \sum_{a \in A} v_i(t, a)\alpha(a),\ i = 1, 2$.

$\sigma_1(m, t)$ is the probability of type t sending message m and $\sigma_2(a, m)$ stands for the probability that the receiver will choose action a in response to message m. $\sigma = (\sigma_1, \sigma_2)$ denotes a mixed strategy profile. A strategy profile σ gives rise to a payoff $V(\sigma) = (V_1(\sigma), V_2(\sigma))$, where

$$V_i(\sigma) = \sum_{t \in T} \sum_{m \in M} \sum_{a \in A} v_i(t, a)\sigma_1(m, t)\sigma_2(a, m)\pi(t),\ i = 1, 2.$$

Also, let

$$V_i(t, \sigma) = \sum_{m \in M} \sum_{a \in A} v_i(t, a)\sigma_1(m, t)\sigma_2(a, m)\ i = 1, 2.$$

σ is a Nash equilibrium if σ_1 and σ_2 are mutual best replies:

$$\text{if } \sigma_1(m, t) > 0, \text{ then } m \text{ solves } \max_{m' \in M} \sum_{a \in A} v_1(t, a)\sigma_2(a, m'),$$

and

$$\text{if } \sigma_2(a, m) > 0, \text{ then } a \text{ solves } \max_{a' \in A} \sum_{t \in T} v_2(t, a)\sigma_1(m, t)\pi(t).$$

The message space is assumed to be large to ensure the availability of sufficiently many unsent messages in any candidate for a solution. If $\#(X)$ denotes the cardinality of the set X, then

$$\#(M) > 2^{[\#(A)+\#(T)]} + \#(T).$$

For any finite game, let $\mathcal{N}$ be the set of Nash equilibria of the game.

We will define *EER* sets for a general n-player game (I, S, u) with player set I, the set of strategy profiles $S = \times_{i=1}^n S_i$ and payoff function $u = (u_i)_{i=1}^n$. Denote by $\Delta(S_i)$ the set of mixed strategies of player i and by $C(s)$ the carrier of s, the set of all pure strategies that have positive probability under s. Let $BR_i(\cdot)$ be the (pure strategy) best reply correspondence of player i, $i = 1, ..., n$. For $\sigma = (\sigma_1, ..., \sigma_n) \in \times_{i=1}^n \Delta(S_i)$ define $BR(\sigma) := (BR_1(\sigma_{-1}), ..., BR_n(\sigma_{-n}))$; define the mixed best reply correspondence $MBR(\sigma)$ analogously.

Definition 1 *A set $\Theta \subset \mathcal{N}$ is equilibrium entry resistant (EER) if it is minimal with respect to*

[R] $\forall \sigma \in \Theta$, if $C(\sigma') \subset BR(\sigma) \cap BR(\sigma')$, then $\sigma' \in \Theta$, and

[C] Θ is closed and nonempty.

Thus, an *EER* set is a closed and nonempty set of equilibrium strategy profiles that satisfies a *replacement condition,* [R] in the above definition. Only equilibrium strategies that are best replies to the status quo can replace the status quo. The focus on equilibrium strategies is deliberate. Cheap-talk refinements in the tradition of Farrell [1993] have for the most part considered equilibria as candidate solutions. Rabin [1990] is an exception but unlike in the present paper he is concerned with one-time play between rational players who base their decisions of which strategy to play only on common knowledge of rationality. The evolutionary approach to cheap-talk games, discussed in detail below, has for the most part also considered conditions on equilibria or on sets of equilibria. Even where there is no a priori restriction to equilibria, it is sometimes the case that solutions either consist of (self-confirming) equilibria, as in Nöldeke and Samuelson's [1992] discussion of cheap-talk games, or can be guaranteed to contain an equilibrium, Blume [1996].

Restricting successful replacements to be themselves equilibria is more novel. As outlined in the introduction this is intended to capture the intuition that because of their focal potential, equilibrium experiments may be more important than nonequilibrium experiments. Jointly these assumptions make it possible to insist on equilibria and yet to guarantee existence.

Existence of *EER* sets is proved next. First it is shown that closed sets of Nash equilibria that satisfy the replacement condition have a simple structure. Whenever one element of a closed connected component of Nash equilibria is part of such a set, all elements of the component must be included.

Lemma 1 *If $\sigma \in \Theta \cap K$, where Θ is a closed set of Nash equilibria satisfying the replacement condition and K a connected component of Nash equilibria, then $K \subset \Theta$.*

Proof: To derive a contradiction, suppose not. Then either (1) $\Theta \cap K$ contains a limit point of $K \setminus \Theta$ or (2) $K \setminus \Theta$ contains a limit point of $\Theta \cap K$. [Munkres, p.147] (2) is ruled out because Θ is closed. If (1) holds, then there exists a sequence of Nash equilibria $\sigma^n \rightarrow \sigma$, $\sigma^n \notin \Theta$, $\sigma \in \Theta$. For σ^n not to be in Θ, $\forall n$, it is necessary that σ^n not be a best reply to σ, $\forall n$. Thus for each n there exists a player i such that $V_i(\sigma_i, \sigma_{-i}) > V_i(\sigma_i^n, \sigma_{-i})$. Hence, there exists a subsequence indexed by k such that for one of the players, i, $V_i(\sigma_i, \sigma_{-i}) > V_i(\sigma_i^k, \sigma_{-i})$, $\forall\ k = 1, 2....$ Since there are finitely many pure strategies, there exists a pure strategy s_i that is in the support of σ_i^k infinitely often such that $V_i(\sigma_i, \sigma_{-i}) > V_i(s_i, \sigma_{-i})$. Without loss of generality reindex the subsequence such that s_i is in the support of σ_i^k for all k. Convergence of σ^k and continuity of $V(\cdot)$ imply that $V_i(s_i, \sigma_{-i}^k) \rightarrow V_i(s_i, \sigma_{-i})$ and $V_i(\sigma_i, \sigma_{-i}^k) \rightarrow V_i(\sigma_i, \sigma_{-i})$. Hence there exists an N such that for $k > N$ we have $V_i(\sigma_i, \sigma_{-i}^k) > V_i(s_i, \sigma_{-i}^k)$. This contradicts σ^k being Nash equilibria. $\square$

Lemma 1 establishes that in our search for a solution we can concentrate on unions of closed connected components of Nash equilibria. The lemma is used in the following proposition to establish existence of EER sets.

Proposition 1 *The strategy space of every finite game contains an EER set.*

Proof: The set of all Nash equilibria is closed and satisfies the replacement condition. By Lemma 1, any other set with these properties must be a union of connected components of Nash equilibria. Following Kohlberg and Mertens [1986], the set of Nash equilibria of a finite game consists of a finite union of closed connected components. Therefore we need to consider only finitely many possible such unions. Hence, a minimal set exists. $\square$

3 Unused Messages

We will show that in a class of sender-receiver games where the sender's and the receiver's preferences are partially aligned, *EER* sets contain only equilibria in which the sender reveals at least some information. This involves showing that partially informative equilibria form a set that is closed under the replacement condition and that less informative equilibria can be replaced by more informative ones.

Unused messages play an important role in the replacement of equilibria. They form natural entry points for candidate replacement equilibria. The following result shows that for any element of an *EER* set there exists an essentially equivalent equilibrium that is also an element of the EER set and does not use a certain number of messages.[2]

Lemma 2 *If Θ is an EER set of a sender-receiver game, then for any strategy $\sigma \in \Theta$, there exists a strategy $\sigma' \in \Theta$ such that (1) $\sigma'_2 = \sigma_2$, (2) $\sigma'_1(m,t) > 0 \Rightarrow \sigma_1(m,t) > 0$, and (3) and the sender assigns probability zero to at least $\#(T)$ messages under σ'.*

Proof: Let Θ be an *EER* set and $\sigma \in \Theta$. Call messages m_i and m_j equivalent if

$$\{t : \sigma_1(m_i,t) > 0\} = \{t : \sigma_1(m_j,t) > 0\},$$

and

$$\{a : \sigma_2(a,m_i) > 0\} = \{a : \sigma_2(a,m_j) > 0\}.$$

Call a message m_i redundant if $\sigma_1(m_i,t) = 0,\ \forall t$, or if there exists an equivalent message m_j with $j < i$. The cardinality assumption on M implies that there are at least $\#(T)$ redundant messages. From σ construct σ' by having the sender move all weight from redundant messages to the lowest index equivalent message; leave $\sigma'_2 = \sigma_2$. It is easily checked that σ' is a Nash equilibrium and $C(\sigma') \subset BR(\sigma)$. Thus $\sigma' \in \Theta$. □

Lemma 2 says that we can replace any element of an *EER* set by one that maintains the same separation of types, is payoff equivalent for the sender and leaves a large number of messages unused. The following result states that the resulting strategy can in turn be replaced by one that alters the receiver's responses after unused messages. The new response after an unused message can be any reply that supports the equilibrium.

The definition of *EER* sets implies

Lemma 3 *Let Θ be an* EER *set in a sender-receiver game, $\sigma \in \Theta$, and $\sigma_1(\overline{m},t) = 0\ \forall t$. For any $m' \neq \overline{m}$, the strategy σ' defined by $\sigma'_1 := \sigma_1$, $\sigma'_2(\cdot,m) := \sigma_2(\cdot,m)$ $\forall m \neq \overline{m}$, and $\sigma'_2(\cdot,\overline{m}) := \sigma_2(\cdot,m')$ is an element of Θ.*

If there is a sufficiently strong alignment of interests between senders and receivers, Lemma 2 and 3 can be used to construct an effective mechanism for the replacement of equilibria that do not utilize this alignment. First, using Lemma 2, the original equilibrium is replaced by one in which the receiver uses

[2] This and the following result are equivalent to results in Blume, Kim and Sobel [1993].

identical responses, the sender types who are separated in the original equilibrium remain separated and at least T messages are unused. Second, using Lemma 3 the receiver's responses after unused messages are changed to a reply to a message that some pooled set of types use in equilibrium. Third, suppose the alignment of interests is strong enough that additional separation of pooled types is possible without the need to undo any of the existing separation of types. Then, a replacement equilibrium exists in which the newly separated types use the "unused messages," the other types use the messages they used before, the receiver responds to the old positive probability messages as before and uses a best reply that reflects the additional separation of types after the newly activated messages.

4 Equilibrium Partial Common Interest

This section defines *equilibrium partial common interest (EPCI)* relative to an equilibrium and argues that if an equilibrium is EPCI dominated via a partition $J = \{J_i\},\ i = 1, ..., j$, of the type set T, then (1) it is not part of any *EER* set, and (2) there exists an *EER* set containing only equilibria in which members of different elements of the partition J send different messages.

EPCI is one way of formalizing the notion of partial alignment of interests between the sender and receiver. In keeping with the postulate of this paper of the primacy of equilibrium behavior, *EPCI* focuses on the benefits to the sender from being revealed as the member of some type set in some equilibrium. Roughly, other equilibria that undo the revelation either through misrepresentation (in terms of the status quo equilibrium) or through pooling formerly separated types must not be more attractive. The idea of *EPCI* is that if the receiver uses best replies to equilibrium strategies of the sender, then (1) types in J_i prefer being identified as members of J_i rather than as members of $J_l,\ i \neq l$ and that (2) for any set K containing types from multiple elements of J there is at least one type who prefers to be identified as a member of his own element rather than as a member of K.

In order to define equilibrium partial common interest, we need a few preliminary definitions. Define $\phi(J)$, the *set of separating strategies* relative to J, as the set of Nash equilibrium strategies $\sigma = (\sigma_1, \sigma_2)$ such that, if $\sigma_1(m, t_i) > 0$ for $t_i \in J_i$, then $\sigma_1(m, t) = 0,\ \forall\ t \notin J_i$. For $L \subset T$, let $\mathcal{N}(L)$ be the set of Nash equilibria σ such that there exists a partition J with $L \in J$ and $\sigma \in \phi(J)$. This is the set of equilibrium strategies in which the set of types L is identified. Let $P(L)$ be the set of Nash equilibria σ such that $\exists m : \sum_{s \in T} \sigma_1(m, s) > 0$ and $(\sigma(m, t) > 0 \Rightarrow t \in L)$.

This is the set of Nash equilibria in which there is a message that is exclusively sent by types in L.

Whenever $\mathcal{N}(L)$ is nonempty, we can define

$$\underline{v}_1^{eb}(t;L) := \min_{\sigma_2}\max_{\alpha}\{v_1(t,\alpha)|\theta \in \mathcal{N}(L),\ \sigma_1 = \theta_1,\ \sigma_2 \in MBR_2(\sigma_1),\ \exists s \in L,$$

$$m \in M, \sigma_1(m,s) > 0, \alpha = \sigma_2(\cdot, m)\}.$$

$\underline{v}_1^{eb}(t;L)$ is the payoff type t can guarantee for himself if the receiver best responds to an equilibrium strategy σ_1 in which the set of types L is separated and if the sender uses messages that have positive probability for types in L under σ_1. The value $\underline{v}_1^{eb}(t;L)$ is a lower bound on type t's payoff if he chooses to remain pooled with the type set L. The construction of this lower bound recognizes that in any candidate replacement equilibrium, following a message that is a positive probability messages under the status quo equilibrium, the receiver has *equilibrium beliefs,* i.e. beliefs dictated by the status quo equilibrium, and best responds to those beliefs.

With these preliminary definitions we can define equilibrium partial common interest relative to a reference equilibrium.

Definition 2 *A sender-receiver game has $(J;\sigma^0)$-equilibrium partial common interest relative to an equilibrium σ^0 ($(J;\sigma^0)$-EPCI) if there exists a partition $J = \{J_i\}_{i=1}^{j}$ of T such that*

[0] $\phi(J) \neq \emptyset$

[1] $\forall \sigma \in \phi(J),\ V_1(t_i,\sigma) > \sum_{a\in A} v_1(t_i,a)\sigma_2(a,m),\ \forall\ t_i \in J_i,\ \forall J_i,\ \forall m$ *such that* $\exists t_l \in J_l, i \neq l,$ *with* $\sigma_1(m,t_l) > 0.$

[2] $\exists \sigma \in \phi(J)$ *such that* $\forall J_i, \exists m_i : t_i \in J_i \Rightarrow \sigma_1(m_i,t_i) = 1$ *and* $V_1(t_i,\sigma) \geq V_1(t_i,\sigma^0)$,

[3] If $K \cap J_l \neq \emptyset$ *for at least two* l, *then* $\forall \sigma \in P(K), \exists i, t_i \in K \cap J_i : \underline{v}_1^{eb}(t_i;J_i) > V_1(t_i,\sigma)$, *and*

[4] $\forall i, \exists m_i$ *such that* $J_i \subset \{t : \sigma_1^0(m_i,t) > 0\}$.

This definition of equilibrium partial common interest is with respect to a reference equilibrium, σ^0. $(J;\sigma^0)$-EPCI requires that σ^0 be “dominated" by a class of equilibria that separate the partition J. Domination requires that types belonging to the same element of the partition send a common message in the dominated equilibrium, [4], and that there exists a dominating equilibrium that separates J in which types belonging to the same element of the partition send a common message that yields payoffs at least as large as σ^0, [2]. Jointly, these conditions make it possible that types belonging to a common element move from one common message that does not identify them to an identifying message without being penalized for doing so. Once J is separated, it is an optimal response for the receiver to accommodate the separation. And since each sender type's payoffs do not decrease by moving to the J-separating equilibrium the sender has no incentive to continue using the positive probability messages of the reference equilibrium. This is how J-separation can become established, starting with the reference equilibrium. It remains to ensure, that J-separation, once achieved will not become undone. This is accomplished by conditions [1] and [3]. The former guarantees that in any J-separating equilibrium types in one element of the partition strictly prefer not to mimic types in another element. Thus, a J-separating equilibrium cannot be replaced by one in which a type switches to a positive probability message of another partition element. Finally, condition [3] ensures that in any replacement equilibrium, types who start separated remain separated. If members from across different partition elements were to pool on an unused message in a candidate replacement equilibrium, the replacement condition would require the receiver to reply with a best response that is part of an equilibrium. Because the replacement condition requires the receiver to use a best reply against the status quo, the sender of type $t_i \in J_i$ can guarantee himself $\underline{v}_1^{eb}(t_i; J_i)$ by using one of the status quo messages. According to condition [3] there is at least one type in any set that pools types across partition elements, who prefers $\underline{v}_1^{eb}(t_i; J_i)$ to the payoff he would receive from following the candidate replacement equilibrium.

This discussion is summarized in the following proposition.

Proposition 2 *If G is a sender-receiver game with $(J; \sigma^0)$-EPCI, then*

[5] $\phi(J)$ contains an EER set, and

[6] σ^0 is not a member of any EER set.

Proof: To prove [5], consider $\sigma \in \phi(J)$. Let $R(\sigma)$ be the smallest closed set containing σ that satisfies [R]. Clearly, $R(\sigma)$ contains an *EER* set. Also, $R(\sigma) \subset$

$\phi(J)$. To see this, let $\sigma \in \phi(J)$ and let σ' be a replacement strategy. It will be the case that types in different elements of J use different message under σ' : If m is a positive probability message under σ for types in J_l, then under σ', m will not be used by types in J_i, $l \neq i$, because of [1] and because σ'_1 must be a best reply to σ_2. Suppose then that m is a message such that $\sigma_1(m,t) = 0\ \forall t$ and let $K \cap J_l \neq \emptyset$ for at least two l where $K = \{t : \sigma'_1(m,t) > 0\}$. [R] implies first that $\sigma'_2(\cdot, m)$ is an equilibrium reply to a positive probability message that induces beliefs concentrated on K, i.e. $\sigma' \in P(K)$, and second that σ'_2 is a best reply to σ_1, such that a type $t_l \in J_l$ sender can guarantee himself at least $\underline{v}_1^{eb}(t; J_i)$ by following σ_1, and will receive at most $V_1(t_l, \sigma')$ from following σ'. Hence (3) implies that there is at least one type in K who prefers following σ_1 to following σ'_1 against σ'_2. This inconsistent with σ' being an equilibrium, which contradicts [R]. Thus, $\sigma' \in \phi(J)$. Also, if $\overline{Q}$ is the closure of $Q \subset \phi(J)$, then $\overline{Q} \subset \phi(J)$. Claim (5) follows.

To prove [6], suppose that σ^0 is a member of an *EER* set. We will derive a contradiction by showing that then $R(\sigma^0) \cap \phi(J) \neq \emptyset$. Lemmata 2 and 3 imply that it is without loss of generality to assume that there are j unused messages $\tilde{m}_i$, $i = 1, ..., j$, under σ^0 such that $\sigma_2^0(\cdot, \tilde{m}_i) = \sigma_2^0(\cdot, m_i)$, $i = 1, ..., j$, where m_i, $i = 1, ..., j$ are the messages identified in [4]. [2] implies that there exists $\sigma' \in \phi(J)$ in which the sender signals only with the unused messages of σ^0, such that if $t_i \in J_i$, then $\sigma'_1(\tilde{m}_i, t_i) = 1$, and the receiver responds to the positive probability messages of σ^0 as he does under σ^0. These off-the-equilibrium path responses support the equilibrium σ' because of [2]. Thus $\sigma' \in \phi(J) \cap MBR(\sigma^0)$, which implies that $\sigma' \in R(\sigma^0) \cap \phi(J)$. □

The following game illustrates the equilibrium partial interest condition and demonstrates how *EER* sets strengthen predictions in games with partial alignment of interests. Let types be equally likely. The first entry in each cell denotes the sender's payoff, the second entry is the receiver's payoff.

	a_1	a_2	a_3	a_4	a_5
t_1	x,0	2,7	4,4	0,0	0,3
t_2	2,7	x,0	4,4	0,0	0,3
t_3	0,0	0,0	0,0	4,4	0,3

Figure 1: Game 1

Independent of the value of x, there exists an *EER* set in which types t_1 and t_2 separate from type t_3 in every strategy that is part of the set. Also, the pooling outcome is not part of any *EER* set. Other evolutionary solutions (discussed below) either do not exist or predict communication only if x is positive. Intuitively, if nonequilibrium strategies are permitted as part of the solution, then for $x < 0$ types t_1 and t_2 cannot be assured of positive payoffs, even if the set $\{t_1, t_2\}$ is separated *and* the receiver responds accordingly.

5 Partial Common Interest

The evaluation of the *EPCI* condition requires that the entire set of Nash equilibria be determined. However, often it is possible to make predictions on the basis of a simpler and more transparent condition.[3] In this section we propose a slightly stronger partial common interest condition *(PCI)* that avoids the need to determine the entire set of Nash equilibria. This concept is intermediate between *EPCI* and partial common interest definitions available in the literature.

PCI formalizes the requirement that if the sender uses a strategy in which the elements of the partition J are identified and if the receiver uses a best reply to the sender's strategy, then (1) $t_i \in J_i$ has no incentive to misrepresent himself as a member of J_l, and (2) a set of types K combining members from different elements of the partition cannot all gain if the receiver uses a best reply to beliefs concentrated on K.

For separation of the partition J not to unravel, once it is established, it is necessary that the minimum payoff from being identified as the member of an element J_i not be too low. In the previous sections these bounds where determined by the fact that players use equilibrium strategies. Here we will consider bounds that only use the fact (given the *EER* replacement condition) that the receiver uses best replies to a strategy of the sender.

If types separate according to the partition J, then there *always* exists a message that allows $t_i \in J_i$ not only to identify himself as a member of J_i but to rule out certain classes of beliefs over J_i. The following lemma verifies this claim. Let L_T^c denote the complement of L in T, and for $K \subset T$, let $\mu(K, m; \sigma_1)$ denote the posterior probability of the type set K given message m for strategy σ_1.

[3] This is convenient also for the purpose of parameterizing different incentive structures in game experiments.

Lemma 4 *Let $L \subset T$ and suppose that for $t_i \in L$, $\sigma_1(m, t_i) > 0$ implies $\sigma_1(m, t_j) = 0 \ \forall \ t_j \in L^c_T$ and for $t_i \in L^c_T$, $\sigma_1(m, t_i) > 0$ implies $\sigma_1(m, t_j) = 0 \ \forall \ t_j \in L$. Then, $\forall \ K \subset L$, $\exists \ m \ : \ \mu(K, m; \sigma_1) \geq \frac{\pi(K)}{\pi(L)}$.*

Proof: Suppose not. Let $\tilde{M} := \{m \in M : \sigma_1(m, t) > 0, \ t \in L\}$. Then $\forall m \in \tilde{M}$

$$\mu(K, m; \sigma_1) = \frac{\sum_{t \in K} \sigma_1(m,t)\pi(t)}{\sum_{t \in K} \sigma_1(m,t)\pi(t) + \sum_{t \in K^c_L} \sigma_1(m,t)\pi(t)} < \frac{\pi(K)}{\pi(L)} \Rightarrow$$

$$\sum_{t \in K} \sigma_1(m, t)\pi(t) < \left(\sum_{t \in K} \sigma_1(m, t)\pi(t) + \sum_{t \in K^c_L} \sigma_1(m, t)\pi(t)\right) \frac{\pi(K)}{\pi(L)}$$

$$\forall m \in \tilde{M} \ \Rightarrow$$

$$\sum_{m \in \tilde{M}} \sum_{t \in K} \sigma_1(m, t)\pi(t) < \sum_{m \in \tilde{M}} \left(\sum_{t \in K} \sigma_1(m, t)\pi(t) + \sum_{t \in K^c_L} \sigma_1(m, t)\pi(t)\right) \frac{\pi(K)}{\pi(L)} \ \Rightarrow$$

$$\sum_{t \in K} \sum_{m \in \tilde{M}} \sigma_1(m, t)\pi(t) <$$

$$\left(\sum_{t \in K} \sum_{m \in \tilde{M}} \sigma_1(m, t)\pi(t) + \sum_{t \in K^c_L} \sum_{m \in \tilde{M}} \sigma_1(m, t)\pi(t)\right) \frac{\pi(K)}{\pi(L)} \ \Rightarrow$$

$$\pi(K) < \pi(K).$$

Thus we have established a contradiction. □

In view of this result we can provide a lower bound on the payoff of a type who is identified as a member of J_i by a strategy profile σ_1 that separates J and in which the receiver uses a best reply.

For any $K \subset L \subset T$, let

$$BR_2(K, L) := \arg\max\Big\{\sum_{a \in A} \sum_{t \in L} v_2(t, a)\mu(t)\alpha(a) :$$

$$supp(\alpha) = A, supp(\mu) = L, and \ \mu(K) \geq \frac{\pi(K)}{\pi(L)}\Big\}.$$

$BR_2(K, L)$ is the set of best replies of the receiver against beliefs concentrated on L that put at least weight $\frac{\pi(K)}{\pi(L)}$ on the subset K of L. Let

$$\underline{v}_1(t; L) := \max_{K \subset L} \min_{\alpha} \{v_1(t, \alpha) : \alpha \in BR_2(K, L)\}.$$

From the lemma, if types in L are identified as such by their messages and the receiver uses a best reply, then for any $K \subset L$, the sender can always find a message that induces a reply in $BR_2(K, L)$. The definition of $\underline{v}_1(t; L)$ recognizes the fact that the sender can "choose" among K. If types in L are identified as such by their messages and the receiver uses a best reply, then $\underline{v}_1(t; L)$ is the payoff that a type t sender can guarantee himself by continuing to identify himself as belonging to L.

This motivates the following definition of partial common interest. For $J_i \subset T$, let $BR_2(J_i)$ be the set of best replies of the receiver against beliefs concentrated on J_i, and let $BR_2(J_i;\pi)$ be the set of best replies against prior beliefs restricted to J_i. Denote the corresponding mixed best replies by $MBR_2(\cdot)$.

Definition 3 *A sender-receiver game G has (J,σ^0)-partial common interest ((J,σ^0)-PCI) relative to an equilibrium σ^0 if there exists a partition $J = \{J_i\},\ i = 1,...,j$, of T such that*

[7] $\underline{v}_1(t_i;J_i) > \max\{v_1(t_i,a_l) : a_l \in BR_2(J_l)\}$ *for all* $t_i \in J_i$ *and* $i \neq l$.

[8] for all $i = 1,...,j$, *there exists* $a_i \in BR_2(J_i;\pi)$ *such that* $v_1(t_i,a_i) \geq V_1(t_i,\sigma^0)$ *for all* $t_i \in J_i$,

[9] if $K \cap J_l \neq \emptyset$ *for at least two l, then for each* $\alpha \in MBR_2(K)$ *there exists i and* $t_i \in K \cap J_i$ *such that* $\underline{v}_1(t_i;J_i) > v_1(t_i,\alpha)$, *and*

[10] $\forall i,\ \exists m:\ J_i \subset \{t : \sigma_1(m,t) > 0\}$.

Comparing Definitions 2 and 3, it is easily verified that (1) $\underline{v}_1(t_i;J_i) \leq V_1(t_i,\sigma)\forall\sigma \in \phi(J)\forall t_i \in J_i$ (2) $\sum_{a\in A} v_1(t_i,a)\sigma_2(a,m) \leq \max\{v_1(t_i,a_l) : a_l \in BR_2(J_l)\}, \forall t_i \in J_i, \forall J_i, \forall m$ such that $\exists t_l \in J_l, i \neq l$, with $\sigma_1(m,t_l) > 0$, $\forall\sigma \in \phi(J)$, and thus [7] implies [1], and [7] and [8] imply [2]. (3) [9] implies [3] because (1) the set of best replies against beliefs concentrated on a set K is a superset of the set of equilibrium best replies against such beliefs and (2) $\forall K,\sigma$ we have $\min_\alpha\{v_1(t,\alpha)|\alpha \in BR_2(K,L) \leq \max_\alpha\{v_1(t,\alpha)|\theta \in \mathcal{N}(L),\ \sigma_1 = \theta_1,\ \sigma_2 \in MBR_2(\sigma_1),\ \exists s \in L,\ m \in M, \sigma_1(m,s) > 0, \alpha = \sigma_2(\cdot,m)\}\}$. Thus we have the following corollary to Proposition 2.

Corollary 1 *If G is a sender-receiver game with $(J;\sigma^0)$-PCI, then*

[11] $\phi(J)$ *contains an EER set, and*

[12] σ^0 *is not a member of any EER set.*

6 Entry as Replacement

EER sets resemble static set-valued solution concepts from evolutionary game theory. This section examines the relationship in greater detail. It is shown that in two-player games, as for example in sender-receiver games, the essential difference is in turning the entry condition of the evolutionary solution concepts into a replacement condition.

Recall the following definition of *equilibrium evolutionarily stable (EES)* sets, due to Swinkels [1992].

Definition 4 *A set* $\Theta \subset \times_{i=1}^{n}\Delta(S_i)$ *is* equilibrium evolutionarily stable (EES), *if it is minimal with respect to*

[13] There exists $\epsilon' \in (0,1)$ *such that for all* $\epsilon \in (0,\epsilon')$ *and for all* $\sigma \in \Theta$, *if* $C(\sigma') \subset BR((1-\epsilon)\sigma + \epsilon\sigma')$, *then* $(1-\epsilon)\sigma + \epsilon\sigma' \in \Theta$.

[14] Θ *is closed and nonempty.*

[15] $\Theta \subset \mathcal{N}$

An *EES* set is a closed set of Nash equilibria that meets an entry condition, [13]. Blume, Kim and Sobel (BKS) [1993] showed that *EES* sets predict communication in sender-receiver games of common interest, i.e. when there exists a unique Pareto-efficient point. *EES* sets need not exist in every game and fail to exist in large classes of partial common interest games. To address the existence issue BKS propose to drop condition [15]. They call sets satisfying the remaining requirements *ER* sets. *ER* sets are fairly permissive with regard to receiver's actions; even if senders are locked into separation according to a partition *J,* the senders' actions after positive probability messages need not fully reflect this separation.

Consider Game 2 below, with types being equally likely. The first entry in each cell denotes the sender's payoff, the second entry is the receiver's payoff.

	a_1	a_2	a_3	a_4
t_1	2,4	4,4	0,0	0,3
t_2	0,0	0,0	3,4	0,3

Figure 2: Game 2

This game illustrates both the nonexistence problem of *EES* sets and that in an *ER* set sender separation need not imply that the receiver uses only separating replies.

The pooling component of Nash equilibria can be invaded by a separating equilibrium and thus cannot be part of an *EES* set. The only other equilibria in this game are separating equilibria. In such an equilibrium the receiver is indifferent among actions a_1 and a_2 following any message that identifies type t_1. A separating equilibrium σ can invade where the receiver uses only action a_1 after positive probability messages that identify t_1 and action a_4 after zero probability messages. This invasion can continue until the entire population plays according to σ. Our assumption on message space size ensures that it is without loss of generality to assume that there is at least one unused message, m, under σ. σ can be invaded by a strategy σ' that is identical to σ, except that following m, the receiver responds with an equal probability mixture over a_2 and a_4. The invasion can continue until σ' represents population play. σ' can be invaded by a strategy σ'' in which t_2 uses the same messages as before, t_1 uses only m, the receiver responds to all messages other than m as before, and responds to m with action a_2. However, $(1-\epsilon)\sigma' + \epsilon\sigma''$ is not a Nash equilibrium for any $\epsilon \in (0,1)$. This contradicts the requirement that *EES* sets consist only of Nash equilibria.

One can show that there exists an *ER* set in this game that contains a (separating) Nash equilibrium, σ. From there one can argue as above, except that now the strategy $(1-\epsilon)\sigma' + \epsilon\sigma''$ becomes part of the *ER* set for some $\epsilon \in [0,1]$. Because there is positive weight on σ'', m is used with positive probability, and because there is positive weight on σ', the message m induces action a_4 with positive probability.

The approach to preserving existence in this paper is different. The requirement that solutions must consist of Nash equilibria is maintained. What changes, is the entry condition [13]. It changes in two ways; (1) entering strategies must be Nash equilibria and (2) a strategy that passes the entry requirement replaces the present population. This is best seen in the following version of the definition of *EER* sets for two-player games. The remainder of this section is concerned with two-player games.

Definition 5 *A set* $\Theta \subset \Delta(S_1) \times \Delta(S_2)$ *is* equilibrium entry resistant (EER), *if it is minimal with respect to*

[16] There exists $\epsilon' \in (0,1)$ such that for all $\epsilon \in (0,\epsilon')$ and for all $\sigma \in \Theta$, if $\sigma' \in \mathcal{N}$ and $C(\sigma') \subset BR((1-\epsilon)\sigma + \epsilon\sigma')$, then $\sigma' \in \Theta$.

[17] Θ is closed and nonempty.

[18] $\Theta \subset \mathcal{N}$

In two-player games, Definitions 1 and 5 of *EER* sets are equivalent. To show this we have to establish the equivalence of [16] and [R]. To see that [R] implies [16], let $\sigma \in \Theta$ and let σ' satisfy $\sigma' \in \mathcal{N}$ and $C(\sigma') \subset BR((1-\epsilon)\sigma + \epsilon\sigma')$ for ϵ sufficiently small. Then it must be that $C(\sigma') \subset BR(\sigma) \cap BR(\sigma')$. Thus, if Θ satisfies [R] it satisfies [16] as well. The converse follows from the following lemma.

Lemma 5 *Let $C(\sigma') \subset BR(\sigma) \cap BR(\sigma')$. Then $C(\sigma') \in BR(\lambda\sigma + (1-\lambda)\sigma'), \forall \lambda \in [0,1]$.*

Proof: Suppose not. Then there exists σ'', $i \in \{1,2\}$ such that $V_i(\sigma''_i, \lambda\sigma_{-i} + (1-\lambda)\sigma'_{-i}) = \lambda V_i(\sigma''_i, \sigma_{-i}) + (1-\lambda)V_i(\sigma''_i, \sigma'_{-i}) > \lambda V_i(\sigma'_i, \sigma_{-i}) + (1-\lambda)V_i(\sigma'_i, \sigma'_{-i})$, which implies that we have either $V_i(\sigma''_i, \sigma_{-i}) > V_i(\sigma'_i, \sigma_{-i})$ or we have $V_i(\sigma''_i, \sigma'_{-i}) > V_i(\sigma'_i, \sigma'_{-i})$. Either inequality contradicts $C(\sigma') \subset BR(\sigma) \cap BR(\sigma')$. $\square$

According to Definition 5 an *EER* set is a closed set of Nash equilibria that cannot be replaced by entrants that (1) are themselves equilibria and (2) are best replies to a post-entry population consisting mainly of the original population and a small fraction of entrants. This is close to the definition of *weakly equilibrium evolutionarily stable (WEES)* sets of Kim and Sobel [1992]. WEES replaces condition [16] by:

[19] There exists $\epsilon' \in (0,1)$ such that for all $\epsilon \in (0,\epsilon')$ and for all $\sigma \in \Theta$, if $\sigma' \in \mathcal{N}$ and $C(\sigma') \subset BR((1-\epsilon)\sigma + \epsilon\sigma')$, then $(1-\epsilon)\sigma + \epsilon\sigma' \in \Theta$.

Thus *EER* sets differ from *WEES* sets only in that the entry condition takes the form of a replacement condition. Both concepts have the potential drawback that the sets need not contain a proper equilibrium. *WEES* sets need not exist in general whereas *EER* sets do.

Proposition 3 *Every EES set and every WEES set contain an EER set. If an ER set contains a Nash equilibrium, the ER set contains an EER set.*[4]

[4] Lemma 5 and Proposition 3 are only valid in two-player games. In the appendix an example is given in which an *EES* set does not contain an *EER* set as defined in Definition 1.

The proof cannibalizes a result in Kim and Sobel [1992].

Proof: Let $\sigma \in \mathcal{N} \cap \Theta$ where Θ is either an EES, a WEES or an ER set. If $C(\sigma') \in BR(\sigma) \cap BR(\sigma')$, then $\sigma' \in \Theta$. This can be seen as follows. By Lemma 5, $C(\sigma') \subset BR((1-\lambda)\sigma + \lambda\sigma')\ \forall\lambda$. Let $\epsilon'' = \sup\{\epsilon' : (1-\epsilon)\sigma + \epsilon\sigma' \in \Theta\ \forall\epsilon \in (0, \epsilon')\}$. ϵ'' could depend on whether Θ is an EES, a WEES or an ER set. By [13], for EES and ER sets, or [19] for WEES sets, $\epsilon'' > 0$. Closedness of Θ implies that $\sigma'' := (1-\epsilon'')\sigma + \epsilon''\sigma' \in \Theta$. By Lemma 5, $C(\sigma') \subset BR((1-\lambda)\sigma'' + \lambda\sigma')\ \forall\lambda \in [0, 1]$. Therefore [13], for EES and ER sets, and [19] for WEES sets, implies that $\epsilon'' = 1$. Hence $\sigma' \in \Theta$. Now consider $R(\sigma)$, the smallest closed set containing σ that satisfies [R]. The foregoing showed that $R(\sigma) \subset \Theta$. Combining this with the fact that $R(\sigma)$ contains an EER set establishes the result. $\square$

7 Selection in Crawford-Sobel Games

Crawford and Sobel (CS) [1982] launched the literature on cheap talk games with a parametric model of incentive alignment. Variations of their model have played an important role in applied work on cheap talk in accounting, Newman and Sansing [1993] and Gigler [1994], political science, Austen-Smith [1990], and economics, Stein [1989]. Equilibria in CS's model have a simple structure and an intuitive comparative statics result holds. In every equilibrium, the sender's strategy partitions the private information, the unit interval, into finitely many intervals. The more closely incentives are aligned, the finer is the partition under the most influential equilibrium[5], the equilibrium that induces the maximum number of distinct responses, and in that sense the more information is revealed in equilibrium.

The applied literature has primarily focussed on the most influential equilibria. This is sometimes justified by appealing to the ex ante efficiency of these equilibria. No other argument, of either the refinement sort, or an evolutionary one has ever been given for why the most influential equilibria ought to be selected or even for why the ever present pooling equilibria ought to be rejected.

The CS game has been notoriously impervious to any refinement or evolutionary analysis. Farrell [1993], for example pointed out that in CS's quadratic example (discussed below), whenever a nontrivial partition equilibrium exists, none of the equilibria in the game passes the neologism proofness test. EES sets

[5] This terminology is adapted from Austen-Smith and Banks [1995]

typically also do not exist in versions of the game with a plausible discretization of the type space. ER sets are plainly nontractable and, as we saw above, do not lead to plausible restrictions of receiver behavior. This makes it interesting to ask whether it is possible to say something about the CS-game within our framework. A second reason for bringing up CS-games is that the notion of partial alignment of interests in those games is quite different from the various *PCI* conditions used here and elsewhere in the literature. The alignment conditions in the evolutionary/refinement literature are stronger than in the CS-game. This ought to make it harder to reject uninformative outcomes in CS-games using the approach taken in this paper.

It is desirable to have a unified approach to both *PCI* and CS-games. This section shows that a modification of our solution concept indeed rejects pooling equilibria in *both* CS-games and under a strengthened *EPCI* requirement. The modified solution concept, *asymmetric entry resistant (AER)* set, relaxes the replacement condition in the definition of EER sets. It permits the transition to a new equilibrium to be initiated also by a subset of the player set rather than only by the entire set.

For any strategy profile σ, and player set $K \subset I$, denote by σ_K the partial profile which contains only the strategies of players in the set K. The complementary partial profile is indicated by σ_{-K}. Similarly, we write BR_K for partial best response correspondences.

Definition 6 *A set $\Theta \subset \mathcal{N}$ is m-asymmetric entry resistant (m-AER) for $m \in \{0, ..., n\}$, if it is minimal with respect to*

[20] $\forall \sigma \in \Theta$, if $\exists\ K \subset I$, $\#(K) \geq m$ such that $C(\sigma'_K) \subset BR_K(\sigma) \cap BR_K(\sigma')$, and $C(\sigma'_{-K}) \subset BR_{-K}(\sigma')$, then $\sigma' \in \Theta$, and

[21] Θ is closed and nonempty.

For $m = n$, *m-AER* sets are the *EER* sets defined earlier, for $m = 0$ they coincide with the entire set of Nash equilibria, and with increasing m the replacement condition becomes increasingly restrictive. For the consideration of sender-receiver games we will let $m = 1$ and simply refer to *AER* sets. This leads to a simple distinction between *sender-led replacements* of σ by σ', where $C(\sigma'_1) \subset BR_1(\sigma_2)$ (S- replacements), and *receiver-led replacements,* where $C(\sigma'_2) \subset BR_2(\sigma_1)$ (R-replacements).

Existence of *AER* sets for finite sender-receiver game is established in the same manner as for *EER* sets. It is also straightforward to show that the results that guarantee availability of unused messages continue to hold. This can be used to prove a result on *EPCI*-games if one strengthens the *EPCI*-requirement. Modify the *EPCI*-definition by replacing condition [3] by

[3'] if $K \cap J_l \neq \emptyset$ for at least two l, then for all $\alpha \in \Delta(A)$, $\exists i,\ t_i \in K \cap J_i$: $\underline{v}_1^{eb}(t_i, J_i) > v_1(t_i, \alpha)$.

While this is a stringent condition, it is satisfies in many interesting classes of games. Game 1 is one example.

Proposition 4 *If G is a sender-receiver game with $(J; \sigma^0)$-EPCI, then*

[22] $\phi(J)$ contains an AER set, and

[23] σ^0 is not a member of any AER set.

Proof: The proof of [23] is identical to the proof of [6]. To verify [22] it is necessary to deal with both sender-led replacements and receiver-led replacements. To derive a contradiction, let σ' be a candidate replacement for $\sigma \in \phi(J)$ under which types from different elements of J send a common message.

Under sender-led replacements, $C(\sigma_1') \subset BR_1(\sigma_2)$. Against σ_2, each type t_i can guarantee $V_1(t_i, \sigma) \geq \underline{v}_1^{eb}(t_i, J_i)$ by following σ_1. Condition [3'] ensures that for no strategy of the receiver, including σ_2, does there exists a message m that would at least achieve a payoff of $\underline{v}_1^{eb}(t_i, J_i)$ for all types in a set $K = \{t : \sigma_1'(m, t) > 0\}$ such that $K \cap J_l \neq \emptyset$ for at least two l.

Under receiver-led replacements $C(\sigma_2') \subset BR_2(\sigma_1)$. Thus, type $t_i \in J_i$ can guarantee a payoff of at least $\underline{v}_1^{eb}(t_i, J_i)$ against σ_2'. According to condition [3'], there does not exist a message $\tilde{m}$ and set $K = \{t : \sigma'(\tilde{m}, t) > 0\}$ such that $K \cap J_l \neq \emptyset$ for at least two l and all types t_i in K would at least achieve a payoff of $\underline{v}_1^{eb}(t_i, J_i)$. This contradicts σ' being an equilibrium. □

In the remainder of this section we will show that AER sets also have predictive power in CS games and thus get us closer to the desired unified treatment of *PCI* and CS games. In a sender-receiver game of the Crawford-Sobel type, the sender's type t is drawn from a cumulative distribution $F(\cdot)$, with density $f(\cdot)$ and support $[0, 1]$. The receiver's action is a real number $a \in \Re$. The sender's payoff function

$v^1(a,t,b)$ depends on a parameter $b \in \Re$ that measures the degree of alignment of interests between senders and receivers. Both the sender's and the receiver's payoff function, $v^2(a,t)$, are twice continuously differentiable. Indicating partial derivatives by subscripts, assume that for all t and $i = 1, 2$, $v^i_1(a,t) = 0$ for some action a, $v^i_{11} < 0$ to guarantee a unique maximum in a for any (t,b), and $v^i_{12} > 0$, a sorting condition that ensures that for either player under full information the optimal action is a strictly increasing function of the sender's type. Assume that the message space M is a superset of $[0,1]$.

In a *partition equilibrium,* the type set can be divided into intervals (which we represent via their interiors) (t_j, t_{j+1}), $j = 0, 1, ..., N-1$, such that $t_0 = 0$, $t_N = 1$, types in (t_j, t_{j+1}) send a common message m_j unique to that type set and the receiver responds to each m_j with the best reply to prior beliefs restricted to (t_j, t_{j+1}). Since there always exists a pooling equilibrium in which all types send the same message, the existence of a partition equilibrium is guaranteed.

Let $a^1(t,b) := \arg\max v^1(a,t,b)$ denote the maximizer of the sender's utility if his type is t and the value of the alignment parameter equals b. Similarly, for the receiver, $a^2(t) := \arg\max v^2(a,t)$. Crawford and Sobel show that if $a^1(t,b) \neq a^2(t)$ for all t, then (1) in every equilibrium the relationship between type and action is as in a partition equilibrium, (2) there exists a partition equilibrium that maximizes the number of elements of the partition, and (3) if the maximal partition equilibrium has N elements, then there exist a partition equilibrium with n elements for all $n \in \{1, ..., N\}$.

To address the equilibrium selection issue, Crawford and Sobel impose additional structure on their model. They examine a quadratic example (that has become the basis for most of the applications of cheap talk games). In that example $F(\cdot)$ is uniform on $[0,1]$, $v^1(a,t,b) := -(a-(t+b))^2$, $v^2(a,t) := -(a-t)^2$, and $b > 0$. This example satisfies all of the above assumptions implying that all equilibria are partition equilibria. In addition (1) the equilibrium with the maximal number of partition elements (the "finest" equilibrium) is ex ante efficient among the set of equilibria, and (2) as $b \to 0$ the finest equilibrium becomes finer approaching full revelation in the limit. Crawford and Sobel argue in favor of the finest partition equilibrium. It is one of the two extreme and therefore distinct equilibria, and they find the coarsest equilibrium not to be plausible. Secondly, the finest equilibrium is ex ante efficient.

While this rule of selecting the finest partition equilibria has been adopted widely in the applied literature, no argument besides distinctness and efficiency has been advanced. We will show below that with quasi-evolutionary arguments we can at least reject the completely uninformative equilibria in a version of Crawford and Sobel's model that generalizes their quadratic model.

To generalize the lessons learned from the quadratic model, Crawford and Sobel add an additional assumption on the payoff functions. Let $v^1(a,t,0) \equiv v^2(a,t)$ and $v^1_{13} > 0$ to guarantee that for all $b > 0$ the sender's preferred response strictly exceeds the receiver's full information action. Our next proposition will be proved under this assumption, and for $b > 0$.

In addition, we assume that receivers use pure replies off the equilibrium path and we make a technical modification in the definition of AER sets. We drop the requirement that AER sets be closed to avoid having to introduce topologies on strategy spaces. The restriction on receiver strategies is for tractability. The modification of AER sets is purely cosmetic. In the case of finite type, message and action sets closedness helps in characterizing and proving existence of EER and AER sets. Fortunately, for CS games, given our assumptions, characterizing AER sets and proving existence is immediate without appealing to closedness of AER sets. We know from Crawford and Sobel that there are only finitely many partition equilibria and that every equilibrium induces a type-action association from a corresponding partition equilibrium. The set of all equilibria is invariant under S- and R-replacement. Clearly, if some equilibrium corresponding to a partition equilibrium belongs to an AER set, so does every other equilibrium corresponding to that AER set. Since there are only finitely many partition equilibria, there is a minimal set that is closed under S- and R-replacement.

Following Austen-Smith and Banks [1995], we will refer to any equilibrium that induces multiple actions as *influential*. Define a pooling equilibrium as any equilibrium that induces only the pooling action.

We will show that pooling equilibria can be replaced by influential equilibria and that the reverse replacement is impossible. In addition there are severe restrictions on replacements among influential equilibria. Essentially an influential equilibrium σ^0 can be replaced by another influential equilibrium σ^1 only if the partition of the latter refines the partition of the former. This is established in the following three lemmata.

To state and prove these results it is convenient to adopt the following conventions. Represent any partition by the collection of interiors of its elements. For an equilibrium σ, denote the partition corresponding to that equilibrium by $J(\sigma)$. A partition $J(\sigma^1)$ refines $J(\sigma^0)$ if the interiors of elements of $J(\sigma^1)$ are subsets of interiors of elements of $J(\sigma^0)$. If (s,t) is a nonempty interval of types, let $a(s,t)$ stand for the receiver's best reply to beliefs restricted to $(s,t.)$ Also, given an equilibrium

σ, let $a_h(\sigma)$ denote the highest action taken after a positive probability message of σ.

The first of our three preliminary results establishes limits on sender-led replacement.

Lemma 6 *If the equilibrium σ^1 S-replaces the equilibrium σ^0, then $J(\sigma^1)$ refines $J(\sigma^0)$.*

Proof: In order to derive a contradiction, suppose not. Then there are two adjacent elements of $J(\sigma^0)$, say (t_{j-1}, t_j) and (t_j, t_{j+1}), and types $t' \in (t_{j-1}, t_j)$, $t'' \in (t_j, t_{j+1})$ who send a common message under σ^1. Under sender-led replacement this means that there exists a message, m, that is unused in σ^0 and in that equilibrium induces a reply a that makes both types t' and t'' indifferent between their σ^0 messages and m. Note that for t'' to be willing to send m we need $a > a(t_{j-1}, t_j)$. Similarly, we need $a < a(t_j, t_{j+1})$. Since the optimal action from the sender's perspective is an increasing function of the sender's type, there exists a type $t^* \in (t', t'')$ who strictly prefers the action a to any of the receiver's equilibrium replies under σ^0. This is incompatible with σ^0 being an equilibrium. $\square$

The next two preliminary results establish analogous limits on receiver-led replacement.

Lemma 7 *If the equilibrium σ^1 R-replaces the equilibrium σ^0, then the maximal element of $J(\sigma^0)$ contains the maximal element of $J(\sigma^1)$.*

Proof: Suppose not. Then the maximal actions under the two equilibria satisfy $a_h(\sigma^1) < a_h(\sigma^0)$. From $a^2(t) = a^1(t, 0)$, and $v^1_{13} > 0$ we have $a^1(t, b) \geq a^2(t)$. Therefore there is a positive measure of types who prefer $a_h(\sigma^0)$ to $a_h(\sigma^1)$. Since under receiver-led replacement the receiver's replacement strategy σ^1_2 is a best reply to the sender's status quo σ^0_1, the status quo action $a_h(\sigma^0)$ must be attainable for the sender under σ^1. This contradicts σ^1 being an equilibrium. $\square$

This result shows that there is a sense in which an R-replacement must induce a finer partition. A complementary result demonstrates that just having a greater number of elements in the candidate replacement partition is not enough.

Lemma 8 *If the equilibrium σ^1 R-replaces the equilibrium σ^0, then there do not exist actions $a^1_1 < a^0 < a^1_2$ such that a^1_1 and a^1_2 are adjacent equilibrium actions under σ^1, and a^0 is an equilibrium action under σ^0.*

Proof: Strict concavity in actions of payoff functions implies that the marginal type who is indifferent between a_1^1 and a_2^1 strictly prefers a^0. Under receiver-led replacements, the action a^0 is available to senders in the candidate replacement equilibrium σ^1. This breaks the equilibrium σ^1. □

In summary, influential equilibria are hard to replace. Sender-led replacement requires nestedness and receiver-led replacement requires analogous strong conditions. Fortunately this does not prevent pooling equilibria from being replaced. Therefore we have the following proposition.

Proposition 5 *If G is a CS game in which the set of influential equilibria is nonempty, then*

[24] the set of influential equilibria contains an AER set, and

[25] no pooling equilibrium is a member of an AER set.

Proof: The first claim of the proposition, [24], follows directly from Lemma 6 and Lemma 7.

For the second statement in the proposition, [25], let σ^p be a pooling equilibrium. We will show that σ^p can be replaced by an influential equilibrium. First, σ^p can be replaced by σ^1 where $\sigma_1^1 = \sigma_1^p$ and under σ_2^1 the receiver responds to all message with the pooling response a_p. Second, note that, by assumption, there exists an influential equilibrium σ^2 and without loss of generality types belonging to the same partition element use the same message under σ^2. Observe that σ_1^2 is a best reply to σ_2^1. Thus the pooling equilibrium σ^1 can be replaced by the influential equilibrium σ^2. □

This result cannot be substantially strengthened because in general partitions will not be nested. As an example consider Crawford and Sobel's quadratic example with an alignment parameter $b = 1/20$. Crawford and Sobel show that in that case there are three partition equilibria with partitions $\{(0, 1)]\}$, $\{(0, \frac{2}{5}), (\frac{2}{5}, 1)\}$, and $\{(0, \frac{2}{15}), (\frac{2}{15}, \frac{7}{15}), (\frac{7}{15}, 1)\}$. The two influential partitions are not nested; therefore by Lemma 6 they cannot be attained one from the other via sender-led replacement. The finer partition among the influential ones cannot be attained from the coarser one via receiver-led replacement because of Lemma 8. The reverse replacement is impossible because of Lemma 7. Thus, in this game there are two AER sets, each corresponding to one of the influential partitions.

The type of incentive alignment considered by Crawford and Sobel does play an important role in the rejection of pooling equilibria by *AER* sets. The following example shows that outside the class of Crawford-Sobel games pooling equilibria need not be rejected even in the presence of partition equilibria. Let $F(\cdot)$ be uniform on $[0,1]$, $v^1(a,t,b) := -(a-t)^2 - b(a-\frac{1}{2})^2$, $v^2(a,t) := -(a-t)^2$, and $b > 0$. Then, regardless of the value of b there always exists a partition equilibrium with a two-element partition in which types $t \in [0, \frac{1}{2})$ send a common message distinct from the message sent by types in $[\frac{1}{2}, 1]$. For sufficiently large b these are the only influential partition equilibria. It is easily checked that in that case any *AER*-set contains pooling equilibria. This is so because for large b all types strictly prefer the pooling action $\frac{1}{2}$ to either of the two separating actions $\frac{1}{4}$ and $\frac{3}{4}$.

8 Relation to the Literature

Thus far equilibrium selection in sender-receiver games satisfying a partial common interest condition has been approached with a number of different quasi-dynamic solution concepts. Here, we pursue a similar strategy. It yields the sharpest predictions to date, guarantees existence, predicts tractable sets of equilibria and does not rely on commonly used auxiliary assumptions like nominal message costs, exogenously given meanings of messages, limitations on drift, limitations on the interpretation of zero-probability messages, etc.. Of course, our paper owes a considerable debt to these other works. We will conclude with a brief discussion of how predictions differ across the various approaches.

In two-player games, *EER* sets resemble static evolutionary solutions. In both instances a strategy or a set of strategies is a solution if it satisfies an entry (or a replacement) condition. However, there are substantial differences. Unlike *evolutionarily stable strategies,* Maynard Smith and Price [1973], *EER* sets always exist. Unlike *neutrally stable strategies*, Maynard Smith and Price [1973], *EER* sets predict efficient outcomes in all common interest games. As for set-valued solution concepts, again *ES* sets, Thomas [1985 a,b], and *EES* sets, Swinkels [1992], Blume, Kim and Sobel [1993], do not exist in general. However, whenever an *ES* set or an *EES* set does exists, it contains an *EER* set.

Blume, Kim and Sobel [1993] propose to consider *ER* sets to address the existence problem of *EES* sets. *ER* sets do exist and do not include pooling equilibria in partial common interest games. However, in these games, they do not rule out pooling actions, in general. *ER* sets are closely related to the *cyclically* *stable sets (CSS)* of Gilboa and Matsui [1991] and Matsui [1992]. *CSSs* can be given an explicit dynamic interpretation. Like *ER* sets, *CSS* sets can be shown not to rule out pooling actions in partial common interest games, and not to rule out dominated actions in general. The latter point is illustrated in the following game, Game 3.

	a_1	a_2	a_3
t_1	7,7	2,2	9,0
t_2	2,2	7,7	0,0

Figure 3: Game 3

In this game no *EES* set exists. There exists an *ER* set and a *CSS* in which the two types of the sender separate. Either set contains a strategy in which the strictly dominated action a_3 is used with positive probability.

Blume's [1994] *perturbed message persistence (PMP)* does eliminate dominated actions. It also eliminates pooling actions in a class of partial common interest games that includes Game 1 with $x > 0$. PMP ensures that the receiver's responses reflect the separation of types in partial common interest games. The partial common interest condition in Blume [1994] is more restrictive, in part, because the retracts that form the solution will in general include nonequilibrium strategies. *PMP* retracts exist in every game.

Rabin and Sobel [1994] present yet another approach to partial common interest games. Their solutions, *recurrent MOPs*, are sets of equilibria and exist in every game. Unlike *EER* sets, neither recurrent MOPs nor PMP sets predict communication in Game 4; again, assume that types are equally likely.

	a_1	a_2	a_3	a_4	a_5
t_1	5,5	6,0	4,4	0,0	3,3
t_2	.5,-1	-1,5	4,4	0,0	3,3
t_3	0,0	0,0	0,0	4,4	3,3

Figure 4: Game 4

In the present paper it is postulated that solutions consist of sets of equilibria, and that these sets must be stable only against other equilibria. Both ideas are familiar. Farrell [1993, p.523] discusses the possibility of requiring a neologism to be itself part of an equilibrium. Myerson's [1988] core mechanism examines this idea in an environment that permits correlation. As for the requirement that solutions be sets of equilibria, Swinkels [1992] argues that his requirement on entrants to use best replies against the post-entry population is more plausible if this behavior does not lead outside the set of Nash equilibria. There are also explicit dynamic models that derive equilibrium behavior from simple behavioral adjustment rules. For example, in the class of sender-receiver games they examine, the limiting outcomes of Nöldeke and Samuelson's [1992] dynamic contain only (self-confirming) equilibria.

None of these solution concepts has been used to distinguish influential from pooling equilibria in Crawford-Sobel games. In CS's quadratic example typically no neologism-proof equilibrium exists, for natural discretizations of the game no EES set exists, and ER sets are not tractable. The present paper is the first one to integrate the analysis of Crawford-Sobel games with that of partial common interest games.

A Appendix

This appendix demonstrates that the equivalence of definitions 1 and 5 of EER sets does not extend beyond the class of two-player games. Consider the following three-player game in which the row player chooses row a_1 or a_2, the column player chooses column b_1 or b_2 and the matrix player chooses matrix c_1 or c_2.

c_1:

	b_1	b_2
a_1	3,3,3	2,2,0
a_2	2,2,0	2,2,0

c_2:

	b_1	b_2
a_1	0,0,0	1,1,100
a_2	1,1,100	2,2,0

Both the equilibrium (a_1, b_1, c_1) and the component $(a_2, b_2, (\gamma, 1-\gamma))$, $\gamma \in [0, 1]$ are EER sets according to Definition 5 (naturally extended to n-player games). Both of these sets are also *EES* sets. However, only the former is an EER set according to Definition 1.

References

Austen-Smith, D. [1990], "Information Transmission in Debate," *American Journal of Political Science*, **34**, 124-152.

Austen-Smith, D. and J. S. Banks [1995], "Cheap Talk and Burned Money," University of Rochester Working Paper.

Blume, A. [1996], "Neighborhood Stability in Sender-Receiver Games," *Games and Economic Behavior*, **13**, 2-25.

Blume, A., D. DeJong, Y.-G. Kim, and G. Sprinkle [1995], "Evolution of the Meaning of Messages in Sender-Receiver Games: An Experiment," University of Iowa Working Paper.

Blume, A., Y.-G. Kim and J. Sobel [1993], "Evolutionary Stability in Games of Communication," *Games and Economic Behavior*, **5**, 547-575.

Crawford, V. and J. Sobel [1982], "Strategic Information Transmission," *Econometrica*, **50**, 1431-1451.

Farrell, J. [1993], "Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior*, **5**, 514-531.

Gigler, F. [1994], "Self-Enforcing Voluntary Disclosures," *Journal of Accounting Research*, **32**, 224-240.

Gilboa, I. and A. Matsui [1991], "Social Stability and Equilibrium," *Econometrica*, **59**, 859-867.

Kim, Yong-Gwan and Joel Sobel [1992], "An Evolutionary Approach to Pre-Play Communication," UCSD Working Paper.

Kohlberg, E. and J.-F. Mertens, "On the Strategic Stability of Equilibrium," *Econometrica*, **50**, 1003-1038.

Matsui, A. [1992], "Best Response Dynamics and Socially Stable Strategy," *Journal of Economic Theory*, **57**, 343-362.

Matthews, S., M. Okuno-Fujiwara and A. Postlewaite [1991], “Refining Cheap-Talk Equilibria," *Journal of Economic Theory*, **55**, 247-273.

Maynard Smith, J. and Price, J. [1973], “The Logic of Animal Conflict," *Nature*, **246**, 15-18.

Munkres, J. [1975], *Topology: A First Course,* Englewood Cliffs, Prentice-Hall.

Myerson, R. [1988], “Credible Negotiation Statements and Coherent Plans," *Journal of Economic Theory*, **48**, 264-291.

Newman, P. and R. Sansing [1993], “Disclosure Policies with Multiple Users," *Journal of Accounting Research*, **31**, 92-112.

Nöldeke, G. and L. Samuelson [1992], “The Evolutionary Foundations of Backward and Forward Induction," University of Wisconsin.

Rabin, M. [1990], “Communication Between Rational Agents," *Journal of Economic Theory*, **51** 144- 70.

Rabin, M. and J. Sobel [1996], “Deviations, Dynamics and Equilibrium Refinements," *Journal of Economic Theory,* **68**, 1-25.

Schelling, Thomas C. [1960], *The Strategy of Conflict,* Cambridge, Harvard University Press.

Stein, J. [1989], “Cheap Talk and the Fed: A Theory of Imprecise Policy Announcements," *American Economic Review,* **79**, 32-42.

Swinkels, J. [1992], “Evolutionary Stability with Equilibrium Entrants," *Journal of Economic Theory,* **57**, 306-332.

Thomas, B. [1985a], “On Evolutionarily Stable Sets," *Journal of Mathematical Biology,* **22**, 105-115.

Thomas, B. [1985b], “Evolutionarily Stable Sets in Mixed-Strategist Models," *Theoretical Population Biology,* **28**, 332-341.