Common Knowledge of Language and Iterative Admissibility in a Sender-Receiver Game

Pei-yu Lo*†

This version: January, 2006

Abstract

This paper investigates the implications of common knowledge of language on cheap talk games. A general framework is proposed where language is modeled as a direct restriction on players' strategies, and the predictions under iterative admissibility (IA) are characterized. We apply this framework to sender-receiver games a la Crawford and Sobel (1982) (CS), where the Receiver takes a one-dimensional action. We incorporate two observations about natural language into the language assumption: 1) there always exists a natural expression to induce a certain action, if that action is indeed inducible by some message, 2) messages that are more different from each other induce actions that are weakly more different. It is assumed to be common knowledge that the Receiver plays only strategies that belong to language. Typically, there is a severe multiplicity issue in CS games. This procedure, on the other hand, eliminates outcomes where only a small amount of information is transmitted. Under certain regularity conditions, all equilibrium outcomes are eliminated except the most informative one. However, with an example, we point out that the normal form procedure does not take care of sequential rationality. To address this issue, we propose an extensive form procedure and characterize the solution.

---

*I feel indebted to Stephen Morris, Dino Gerardi and Benjamin Polak for invaluable guidance and support. I am also grateful to Dirk Bergemann and Itzhak Gilboa for useful suggestions. Finally I would like to thank Daniel Monte, Rebecca Sawyer, Amalavoyal Chari, Christopher Ksoll, Ulrich Wagner, Siddharth Sharma, Dmitry Shapiro for their help and support throughout the process.

†Adress: Department of Economics, Yale University, P.O.Box 208268, New Haven, CT 06520-8268. Tel: (203) 809-8006. E-mail: pei-yu.lo@yale.edu.

1. Introduction

Common sense suggests that speaking the same language helps with cooperation and efficiency, as long as there is room for cooperation. However, this phenomenon is not quite captured in economic analyses of communication. Game theoretic predictions do not depend on whether or not the players speak the same language. This is not surprising, since the notion of language is absent from standard models of cheap talk games. In the standard cheap talk analysis, all messages are treated symmetrically, in that the exact labeling does not matter. That is, two messages can have their names swapped with each other without changing the strategy set or the equilibrium outcome. However, if players speak the same language, convention offers a way of interpreting messages, suggesting that labeling does matter in reality. For example, suppose a man and a woman are both native English speakers and they have to choose between going to the opera or to a boxing match. The woman can send a message beforehand. It is natural that two messages, "Opera" and "Boxing", uttered in a serious manner, are either taken literally or ignored for strategic reasons. It is counter-intuitive that the message "Opera" would indicate going to the boxing match while the message "Boxing" would indicate going to the opera.

Language manifests itself in the asymmetry among messages. This paper attempts to formalize the notion of language in terms of players' strategy sets. We propose the following general framework to incorporate language. First, we model language as a direct restriction on players' strategies. The restriction does not by itself shrink the set of communication outcomes. It eliminates only strategies that are replicas of other strategies up to the name change. We call this new game "the game with language". Second, we characterize the prediction of the game with language under iterative admissibility, i.e., iterative deletion of weakly dominated strategies. Applying language or iterative admissibility alone does not shrink the set of outcomes, but the combination can give a sharp prediction.

We make a first attempt to apply this general framework to a classic sender-receiver game as in Crawford and Sobel (1982) (CS). The simple structure of CS games provides a straightforward implication for the language assumption, which we will describe below. In a CS game, the Sender (she) is the only player with private information, which is called the Sender's type and is assumed to be one-dimensional. The Receiver (he), upon receiving the message, takes a one-dimensional action, which affects the utility of both. The Sender always prefers a different action from the Receiver. Since the Sender communicates in an attempt to influence the behavior of the Receiver, messages can be mapped to recommendations. Equating message space with action space allows us to linearly order messages because the action space is on the real line. Two observations of natural language usage are imposed as assumptions: (i) there always exists a natural expression to induce a certain action, if that action is indeed inducible by some message; (ii) messages that are more different from each other induce actions that are weakly more different, i.e., if two messages induce the same action, any message in between the two will induce the same action. The second assumption exploits the linear order on the action space. It gives more structure to language and is important for our characterizations.

We first take the normal form approach to this multi-stage game described above. It seems natural as language is a normal form restriction, and sequential rationality is not an issue in standard cheap talk games, since all messages can get used with positive probability. We find that if the players' interests are sufficiently aligned, this procedure eliminates outcomes where only a small amount of information is transmitted. Under certain regularity conditions, all equilibrium outcomes but the most informative one are eliminated.

However, we find that the normal form approach might even eliminate the most informative equilibrium of the game without language. We show an example where our procedure yields a unique informative outcome, in contrast with the game without language where babbling is the unique equilibrium—thus the most informative equilibrium in this game. This example illustrates how normal form procedure might allow the Receiver to take a sub-optimal action after receiving some messages, though it requires strategies to be ex ante optimal for the Receiver with respect to his belief. This is because modeling language as a direct restriction on the strategy sets gives language the highest priority, overriding rationality at times. We then illustrate the tension between language, iterative deletion of weakly dominated strategies and sequential rationality. To resolve this tension, we propose a notion of weak sequential rationality with language and an extensive form iterative procedure. The key observation motivating our definition of weak sequential rationality is that the Sender does not distinguish between messages that induce the same action, and hence the Receiver does not either. We view messages as a coordination device to achieve a mapping from types to actions, which is called an outcome. We decompose a strategy profile into the usage of messages and the induced outcome. In order to capture the idea that language takes care of the usage of messages, while rationality concerns determine the set of possible outcomes, we define a concept of sequential rationality in terms of the outcome induced by a strategy profile, instead of the strategy profile itself. This extensive form iterative procedure always yields a nonempty limiting set. When the original game has multiple equilibria, it eliminates some of the less informative outcomes like the normal form procedure. When babbling is the unique equilibrium in the original game, it also yields babbling as the unique prediction.

Our approach falls into the tradition of trying to incorporate literal meanings into cheap talk games. Farrell (1993), Rabin (1990) and Zapater (1997) share the assumption that the literal meaning of a message is believed if it is credible, but they propose different credibility criteria. Farrell (1993) uses credible literal meaning to restrict off-equilibrium-path beliefs held by the Receiver and proposes neologism-proofness as an approach to equilibrium refinement, which, however, suffers from Stiglitz Critique, because in establishing credibility, the Sender is assumed to be guaranteed her equilibrium payoff, even if the equilibrium in question is not stable. Additionally, it might make an empty prediction. In particular, no equilibrium in a nontrivial CS game is neologism-proof.

Rabin (1990) and Zapater (1997) both use rationalizability to establish credibility. To get the unraveling going in rationalizability, they restrict the Sender's strategies and ask whether that restriction is consistent with rationality and common knowledge of the restriction. In making the restriction, certain communication outcomes are ruled out a priori. Credibility assures that if it is common knowledge that this information is going to be transmitted, the eliminated outcomes will not be realized. Rabin's "credible message rationalizability" represents the minimal amount of information the Sender can credibly transmit, while a "credible proposal" as defined by Zapater represents the maximal amount. Credible message rationalizability always yields a non-empty (if sometimes weak) prediction, while a credible proposal is not guaranteed to exist. In particular, every CS equilibrium is credible message rationalizable.

Our approach is closely related to Rabin's and Zapater's, since in a two-player setting, rationalizability is equivalent to iterative deletion of dominated strategies. Our approach differs from the literature in two key aspects. First, we make restrictions on the Receiver's strategy set instead of on the sender's strategy set, and hence avoid ruling out babbling or any equilibrium outcome a priori. Our definition of language applies without modification to the entire class of CS games, in contrast to Rabin and Zapater's definitions which are not independent of the specifics of the game, such as the utility functions and the prior, since restrictions have to be credible and credibility differs with games. Second, looking at any message in isolation, we make no assumption about the actions the Receiver will take. Instead, all our assumptions concern the relation between messages in terms of the induced actions. On the other hand, Rabin and Zapater assume that the Receiver believes credible messages and carries out credible recommendations, while the relation between messages is roughly determined by the model. We argue that, in reality, messages have relative meanings in addition to absolute meanings. For example, when the audience says "good job," they might sincerely mean that they appreciate the performance, but they might just be polite. However, for the receiver of the comment, it is probably weakly better than if they say "horrible." We share with the literature the view on absolute meanings, but stress the asymmetry among messages as an important implication of language. In addition to equilibrium selection, our prediction then reflects the effect of the properties of language.

The rest of the paper is structured as follows. Section 2 provides a simple example to motivate our approach. Section 3 discusses the solution concept in use and the language assumptions. Section 4 outlines the setup of the game. Section 5 presents the results using normal form concept and highlights the conflict with sequential rationality. Section 6 motivates a weak concept of sequential rationality for this game with language, defines an extensive form procedure and presents the results. Section 7 concludes.

2. Motivating Example

Consider a two-player game with one-sided pre-play communication. Rob the pirate is planning to set sail for the treasure island. He does not know whether it is on the West sea or the East sea. He only knows that with probability $\frac{2}{3}$, the treasure island is on the West sea. The prior is common knowledge. Sally the witch, however, knows where the treasure island is. Rob asks Sally in which direction he should go and commits to giving Sally a commission if he finds the treasure. Their payoff matrix is as in table 2.1. The row indicates whether the treasure island is on the West Sea or East Sea. The column indicates the direction Rob chooses. W stands for west and E stands for east. The number on the left is Sally's payoff and the number on the right is Rob's payoff. The game goes like this: Sally tells Rob which direction to take, either west or east, and Rob chooses one direction and sets sail. If he finds the treasure, he has to give Sally a payoff of 2. If he does not, neither of them loses anything.

		a	
location of treasure		W	E
	West	2,1	0,0
	East	0,0	2,1

Table 2.1: Treasure Hunt Game

	"west"	*"east"*
Stubborn W	W	W
Stubborn E	E	E
Literal	W	E
Opposite	E	W

Table 2.2: Receiver Strategies in the Treasure Hunt Game

Given the true location of the treasure island, t, Sally chooses a message $s^S(t)$: either "west" or "east." Her strategy is therefore $s^S = \left(s^S(West), s^S(East)\right)$. A strategy for Rob, denoted by s^R, is a function from the message space M to the set of actions $A = \{W, E\}$. Table 2.2 lists all of Rob's possible strategies. Both the *Stubborn W* and the *Stubborn E* strategies completely ignore Sally's recommendation. *Literal* strategy and *Opposite* strategy are essentially the same strategy up to relabeling. This is because they both react to one message with the action W and the other with the action E.

This game has two equilibrium outcomes. One is the so-called "babbling" equilibrium, in which Rob always chooses W and Sally "babbles". The other equilibrium is what we call the informative equilibrium, in which Rob's decision changes with Sally's recommendation and Sally's recommendation depends non-trivially on the true state. There is an innocuous multiplicity here in terms of relabeling the two messages. Actually, if we relabel the messages, we will end up with the same strategy set. The symmetry between messages suggests that language does not play a role in standard analysis. Game theoretic predictions for an English speaking Rob and an English speaking Sally would be the same as the predictions for an English speaking Rob and an alien Sally.

However, suppose Rob and Sally do share a common first tongue, say English. In the language English, "west" means the direction where the sun falls and "east" means the direction where the sun rises. It seems absurd that Rob and Sally would coordinate in such a way that the message "west" induces Rob to go east and the message "east" induces Rob to go west, if they are going to play the informative equilibrium. The issue is not credibility: if Rob does not believe that Sally's recommendation conveys information, Rob would ignore the message and take the same action regardless. If Rob's action depends nontrivially on Sally's message, then it seems more natural that he would go west upon hearing the suggestion "west" and he would go east upon hearing the suggestion "east."

Suppose it is common knowledge that Rob follows the convention of language and does not use the *Opposite* strategy. That is to say, in the game with language G_L, the set of strategies for Rob is $S_L^R \equiv \{Stubborn\ W,\ Stubborn\ E,\ Literal\}$. Then when the true state is $West$, for Sally, sending the message "east" is weakly dominated by sending the message "west". To see this, notice that both messages yield the same payoff if Rob plays either the *Stubborn W* strategy or the *Stubborn E* strategy. Sally's choice of message matters only if Rob plays the *Literal* strategy. In that case, message "west" induces the action W, which is strictly preferred by Sally when the true state is $West$. Similarly when the true state is *East*, the message "west" is weakly dominated for Sally. In conclusion, if Sally does not play weakly dominated strategies, then she says "west" when the true state is $West$ and "east" when the true state is *East*.

If the Receiver knows that Sally does not play weakly dominated strategies, then when he receives the recommendation "west", he knows that the true location must be $West$, and when he receives the recommendation "east", he knows that the true location must be *East*. The optimal strategy then is to follow Sally's advice and play the strategy *Literal*. We therefore end up with a unique prediction that Rob and Sally play the informative equilibrium outcome, which is what we would "expect".

Eliminating the *Opposite* strategy by way of the language assumption is a key step in getting the unique prediction. In the game *without* language, both strategies *Literal* and *Opposite* belong to Rob's strategy set. When the true state is $West$ (*East*), sending message "*east*" ("*west*") performs better for Sally than sending message "*west*" ("*east*") if Rob plays the strategy *Opposite*, while sending "*west*" ("*east*") performs better if Rob plays the strategy *Literal*. In short, none of Sally's strategies are weakly dominated. Eliminating the strategy *Opposite* from Rob's strategy set gets the unraveling process going.

However, language alone does not do the trick. It is language combined with iterative deletion of weakly dominated strategies that sharpens the predictions. Table 2.3 summarizes the predictions under different combinations of solution concepts and language assumption.

IA means iterative admissibility, i.e., iterative deletion of weakly dominated

	Equilibrium	**IA**	**ID**
No Language	babbling,informative	everything	everything
Language	babbling,informative	informative	everything

Table 2.3: Predictions of Treasure Hunt Game

strategies. ID stands for iterative deletion of strictly dominated strategies. "Everything" means every pair of strategies in the game except those where Rob plays *Stubborn E*. As long as Rob is rational, he will not play *Stubborn E* because going east blindly is worse ex ante than going west blindly. Here the language assumption alone does not change the set of equilibrium outcomes. It only eliminates the innocuous equilibrium multiplicity where meanings are reversed. Comparing the prediction using IA and ID suggests that weak dominance is key in getting rid of the babbling outcome. This is not surprising since messages are costless, and therefore Sally does not have a strict preference for any message if she believes that Rob will ignore it.

3. General Framework

The example of section 2 suggests modeling language as a direct restriction on players' strategies. Let Γ denote a cheap talk game where a one-shot game is preceded by a communication stage. Let I denote the set of players, and T^i denote the set of types for player i. A strategy for player i, denoted by $s^i \in S^i$, is a mapping from player i's type space T^i to his action plans. Write player i's ex ante expected utility function as $U^i : (S^i)_{i \in E} \to R$. That is, U^i is a mapping from the set of strategy profiles to the real line. We can represent Γ in the strategic form $G = \left(I, (S^i)_{i \in I}, (U^i)_{i \in I}\right)$. Language transforms the game into $G_L = \left(I, (S^i_L)_{i \in I}, (U^i)_{i \in I}\right)$, which we call "the game with language". To make predictions about cheap talk games with language, we need to know two things: (1) the implications of "language," that is, which strategies belong to S^i_L for each $i \in I$, and (2) given $(S^i_L)_{i \in I}$, the solution to the game G_L. This is a clean way to incorporate language since all assumptions about language are embodied in $(S^i_L)_{i \in I}$. By altering the assumptions, we can understand the implications of specific properties of language. This section first discusses the solution concept employed and then motivates a specific way to model language.

3.1. Solution Concept

The solution concept employed here is iterative admissibility (IA) when the normal form is used and a variation when the extensive form is used. The discussion of the variation for the extensive form analysis is deferred to section 6. Here we recall the definition of iterative admissibility and discuss the choice of this solution concept over others.

The definitions below follow Brandenburger et al (2004).

Definition 1. *Fix $(X^j)_{j\in I} \subseteq (S^j)_{j\in I}$. A strategy s^i is weakly dominated with respect to X^{-i} if there exists $\hat{\sigma}^i \in \Delta X^i$ such that $U^i\left(\hat{\sigma}^i, s^{-i}\right) \geq U^i\left(s^i, s^{-i}\right)$ for every $s^{-i} \in X^{-i}$ and that $U^i\left(\hat{\sigma}^i, \hat{s}^{-i}\right) > U^i\left(s^i, \hat{s}^{-i}\right)$ for some $\hat{s}^{-i} \in X^{-i}$. Otherwise, say that s^i is admissible with respect to $(X^j)_{j\in I}$. If s^i is admissible w.r.t. $(S^j)_{j\in I}$, simply say that s^i is admissible.*

Definition 2. *Set $S^i(0) = S^i$ for $i \in I$ and iteratively define*

$$S^i(k+1) = \left\{ s^i \in S^i(k) : s^i \text{ is not weakly dominated with respect to } \left(S^i(k)\right)_{i\in I} \right\}.$$

Write $\cap_{k=0}^{\infty} S^i(k) = S^i(\infty)$ and $\cap_{k=0}^{\infty} S(k) = S(\infty)$. A strategy $s^i \in S^i(\infty)$ is called iteratively admissible.

Denote by ΔX the set of probability distribution on X, and by $\Delta^+ X$ the set of probability distribution which puts positive weight on every element of X.

Brandenburger et al (2004) show that if there are only two players, say player S and player R, a strategy is weakly dominated if and only if it is never a best response to a totally mixed strategy. For completeness of arguments, this equivalence result is restated as Lemma 1 below. Note that this result does not hold if there are more than two players unless players can play correlated strategies.

Lemma 1 (Brandenburger et al (2004)). *A strategy $\hat{s}^R \in X^R$ is admissible with respect to $X^S \times X^R$ if and only if there exists $\hat{\sigma}^S \in \Delta^+ S^S$ such that $U^R\left(\hat{\sigma}^S, \hat{s}^R\right) \geq U^R\left(\hat{\sigma}^S, s^R\right)$ for every $s^R \in X^R$.*

As our analysis of the Treasure Hunt game revealed, weak dominance is crucial. This is not surprising since messages are costless, and therefore senders are indifferent between messages per se. It is this indifference that causes severe multiplicity. In the evolutionary approach, it is important that any strategy that is weakly better than the current strategies gets used with strictly positive probability and gets taken into account by opponents. This corresponds to weak dominance in the iterative procedure instead of strong dominance, since a strategy that survives weak dominance is a best response to a belief that puts strictly positive weight on every surviving opponent strategy.

One reason we choose iterative admissibility over other non-equilibrium concepts employing weak dominance is its epistemic foundation. Brandenburger et al (2004) provide a sufficient epistemic condition under which the predicted strategy profiles are characterized by IA. More specifically, they show that if there is rationality and n-th order assumption of rationality, where n is higher than the number of iterations needed to arrive at IA, then players play strategies in IA. However, we are not incorporating language into the epistemic framework provided by Brandenburger et al (2004). We simply take the solution concept as given and apply it directly to the game transformed by our language assumption.

By representing the original game as $G = (I, (S^i)_{i \in E}, U)$, we have implicitly assumed that players make their decisions at the initial node before nature makes her move. This ex ante interpretation implies that each player believes that different types of his opponent hold the same belief about his behavior. Alternatively, we can think of different types as representing different "individuals", chosen to appear by nature, and thus assume that players make their decisions after nature makes her move[1]. This interim interpretation implies that each player believes that different types of his opponent may hold different beliefs about his behavior. Let's rewrite the set of players, I, as $I^m \equiv \cup_{i \in I} T^i$. Every player q in I^m can then be written as $t^i \in T^i$ for some $i \in I$. Let $\tilde{S}^q = S^{t^i}$ be the set of action plans available to type t^i of player i. Define $\tilde{U}^q \equiv U^i \forall q \in I^m$. Then iterative admissibility in the game $(I, (S^i_L)_{i \in I}, U)$ under the interim interpretation is equivalent to iterative admissibility in the game $(I^m, \left(\tilde{S}^q_L\right)_{q \in I^m}, \tilde{U})$.

In equilibrium concepts, it does not matter whether players make their decisions before or after nature makes her move, because in equilibrium, every type of player i holds the correct belief about the behavior of the opponents, and thus every type of player i holds the same belief. However, the two interpretations make a difference in nonequilibrium solution concepts, since players are not assumed to hold the "correct" belief about the behavior of the opponents. The interim interpretation is more appealing if we think of private information as some hard-wired characteristics of the players. However, the ex ante interpreta-

[1] See p.226 in [5].

tion is more closely related to the equilibrium concept in that it is as if players of different types hold the same belief about the opponents. Analysis is conducted under both interpretations. In general, it is easier to include strategies under the interim interpretation , while it is easier to exclude strategies under the ex ante interpretation.

Lastly, Lemma 2 shows that the equivalence between weak dominance and never best response to a totally mixed belief holds under the interim interpretation with only two players. This characterization, instead of weak dominance, is directly used in operation. To simplify the analysis, we assume that there is only one-sided incomplete information. Player S holds private information while player R does not. It should easily generalize to cases with two-sided incomplete information. Under the interim interpretation, each type of S is considered an individual player, so Lemma 1 does not directly apply. Let $X^S \equiv \Pi_t X^S(t)$. The proof for equivalence is similar to that in Pearce (1984).

Lemma 2. *s^R is weakly dominated w.r.t. $\left(\Pi_t X^S(t)\right) \times X^R$ if and only if there does not exist a $\sigma^S(t) \in \Delta^+ X^S(t)$ for every t such that*

$$s^R \in \arg\max_{s' \in X^R} U^R\left(\left(\sigma^S(t)\right)_{t \in T^S}, s'\right).$$

3.2. Incorporating Language

Recall that language here is simply a subset of players' strategies resembling conventional language usage. This paper focuses on sender-receiver games where only the Sender (S) possesses private information and only the Receiver (R) has a non-trivial one-dimensional action space A. (In arbitrary communication games, our notion of language remains valid , although a different set of restrictions may be appropriate.) The relative simplicity of the communication protocol and the complete linear order on the action space A give language more structure and generate the assumptions discussed below.

Before talking about the implication of language for players' strategy sets, we need to discuss the message space. It will be assumed throughout the paper that the message space M has the same number of elements as the action space A. With that assumption in the background, we argue that (1) language should restrict only the Receiver's strategy set, (2) the message space M can be identified with the action space A, i.e., $M = A$, and (3) every Receiver strategy in language should satisfy the *literal meaning* condition and the *convexity* condition. Lastly, we discuss the implications of these restrictions.

It is desirable that our definition of language itself does not restrict the set of communication outcomes in a given game, while eliminating the "innocuous" multiplicities in terms of how messages are used. An outcome of a sender-receiver game dictates which action (or probability distribution over actions) each type of sender induces. Messages are only means to implement a possibly nontrivial outcome since they are costless. As pointed out in the example in section 2, relabeling of messages might produce the same outcome. Language is a restriction only insofar as these relabelings are concerned.

In particular, our definition of language itself should not rule out the babbling outcome. This would imply that language does not place any restriction on the Sender's strategy set. To see this, notice that in the babbling equilibrium, the Receiver ignores all messages and the Sender sends every message with equal probability[2]. In other words, language should not force the Sender to convey information nor should it force the Receiver to react differently to different messages. Thus, language includes any Receiver strategy that is a constant on the message space. This implies that when looking at every message in isolation, any action is possible. For a given type of Sender, every message belongs to her message space if she can play the strategy that puts equal probability on every message. If we do not take it as a literal assumption that there is one Sender at the initial node before nature decides on the true state, the message sent by a given type should not be physically linked to the message sent by another type. Therefore, the set of pure Sender strategies in language as mappings from the type space to the message space should be the product space of M. We conclude that language does not place any restriction on the Sender's strategies.

To justify the simplification that the message space M is equivalent to the Receiver's action space A, notice that the sender talks in an attempt to induce a certain behavior from the Receiver. Say that an action a can be induced in language S_L^R if there exists a Receiver strategy $s^R \in S_L^R$ and a message $m \in M$ such that $s^R(m) = a$. Every action $a \in A$ can be induced in language since language should contain all constant s^R. If the language is rich enough, there is usually a conventional way to express the literal meaning of a. For example, in the Treasure Hunt game, if Sally can get Rob to go east in some way, she can successfully do so simply by saying "Go east!"

Formally, this implies that for every action $\hat{a}$, there is at least one message $\hat{m}$ that invariably induces $\hat{a}$ whenever the Receiver is going to take $\hat{a}$ after some

[2]We focus on the babbling equilibrium strategy profile where every message is used with strictly positive probability, so that Bayesian update can be performed on every message.

message. Call such $\hat{m}$ a message with literal "$\hat{a}$"-meaning. Whether the Receiver is going to take action $\hat{a}$ after any message is up to strategic considerations, but there will be no ambiguity about the literal meanings of messages. If the language is rich enough, there exists a canonical message for every action, that is, for every action a in A, there exists a message m in M with literal "a" $-$ *meaning*. It is easy to show that a message cannot have different literal meanings. Given the assumption that the number of messages in M is the same as the number of actions in A, we can label the message with literal "a" $-$ *meaning* by "a." Therefore, we can simply assume that $M = A$, given the assumption that $|M| = |A|$.

We can then compare messages since $M = A$. It is intuitive that "similar" messages should induce "similar" actions. For example, when a friend tells you that restaurant A is "faaabulous" instead of telling you that it is "so-so", it would appear that she means that restaurant A is drastically better than average. If you know that your friend have a tendency to exaggerate, and you wouldn't go to restaurant A even if she told you it was "faaabulous", then it is unlikely that you would go to restaurant B if she told you it was "good". Messages that lie on the two extremes should convey weakly more information than messages that lie in between them.

The preceding discussion leads us naturally to define language as follows.

Definition 3. *$s^R : M \to A$ belongs to S_L^R if and only if*

1. *(literal meaning) $s^R(\hat{a}) = \hat{a}$ if $\exists \hat{m} \in M$ such that $s^R(\hat{m}) = \hat{a}$;*
2. *(convexity) If $s^R(m_1) = s^R(m_2)$ where $m_1 < m_2$, then $s^R(m_1) = s^R(m)$ for all m such that $m_1 \leq m \leq m_2$.*

Definition 4. *Say that s^R belongs to language if $s^R \in S_L^R$.*

Lemma 3 (Property of strategies in language). *If s^R belongs to language, then*

1. *(relative meaning) s^R is weakly increasing on M. That is, $\forall m_1 < m_2$, $s^R(m_1) \leq s^R(m_2)$,*
2. *(absolute meaning) if $m_1 < m_2$ and $s^R(m_1) \neq s^R(m_2)$, then $s^R(m_1) < m_2$ and $s^R(m_2) > m_1$.*

The first property reflects the relative difference in messages: a higher message induces a weakly higher action. A deadline of tomorrow signals a more urgent deadline than a deadline 10 days later, if they do convey any information at all. The second property reflects the absolute difference in messages: if "excellent" means something different from "good," then "excellent" means something at least as good as the absolute quality of being good.

Given s^R and $Q \subset M$, define

$$s^R(Q) \equiv \left\{a | \exists m \in Q \text{ s.t. } s^R(m) = a\right\}.$$

That is, $s^R(Q)$ is the set of actions induced by a message m in Q under the Receiver strategy s^R.

We began the discussion by arguing that it is desirable that a definition of language does not a priori rule out any outcomes in the original game. Lemma 4 stated below confirms that the specific way of modeling language given by definition 3 satisfies this condition.

Lemma 4 (Completeness of Language). *For all $B \subset A$, there exists a $s^R \in S_L^R$ such that $s^R(M) = B$.*

Proof To see this, we simply need to construct a Receiver strategy, s^R, taking exactly the actions in a given $B \subset A$. We can linearly order the elements in B and write $B = \{a_1, a_2, ..., a_n\}$ where $a_j < a_{j+1}$ for every j. We can construct the Receiver strategy s^R by defining

$$\hat{s}^R(m) \equiv \begin{cases} a_1 & m \in [0, a_1] \\ a_j & m \in (a_{j-1}, a_j], j = 2, ..., n \\ a_n & m \in [a_n, 1] \end{cases}.$$

It is easy to check that $\hat{s}^R$ satisfies definition 3 and $\hat{s}^R(M) = B$.

Corollary 1. *Every equilibrium outcome in the game without language is also an equilibrium outcome in the game with language.*

4. The Setup

We apply this general framework to a discretized version of sender-receiver games as in Crawford and Sobel (1982). There are two players, a Sender (S) and a

Receiver (R). Only the Sender has private information, represented by her type $t \in T$. The common prior on T is $\pi \in \Delta T$. The Sender sends a message $m \in M$, and the Receiver takes an action a in A after receiving the message m. It is helpful to think of $T = A = M = \{0, \Delta, 2\Delta, ..., 1\}$, though all we need is that they are all finite spaces, and that $A = M$. Both players have Von Neumann-Morgenstern utility function $u^i(t,a)$, $i = S, R$. Though the type space and the action space are both discrete, we assume that u^i can be extended to a function from $[0,1] \times [0,1]$ to the real line. It is assumed that u^i is twice continuously differentiable.

As in Crawford and Sobel (1984), it is assumed throughout the paper that $\frac{\partial^2}{\partial a^2} u^i < 0$ and $\frac{\partial^2}{\partial t \partial a} u^i > 0$ for $i = S, R$. Define

$$y^i(t) := \arg\max_{a \in A} u^i(t,a).$$

From the conditions on u^i, $y^i(t)$ is weakly increasing in t for both $i = S, R$. Since A is discretized, $\arg\max_{a \in A} u_i(t,a)$ might not be a singleton. For simplicity, assume that $y^i(t_S)$ is a singleton for all t and both $i = S, R$. The bias is represented by

$$b := \min_{t \in T} \left\{ y^S(t) - y^R(t) \right\}.$$

To simplify the analysis, we also assume that $y^R(t) = t$. Let $E([t_1, t_2])$ denote the optimal action for the Receiver if he only knows that the Sender's type lies in the interval $[t_1, t_2]$. That is, for any $t_1 < t_2$,

$$E([t_1, t_2]) \equiv \arg\max_a \sum_{\substack{t \in T, \\ t_1 \le t \le t_2}} u^R(t,a)\, \pi(t).$$

A pure strategy of the Receiver (s^R) is a function from the message space M to the action space A which belongs to the language, that is, $s^R \in S_L^R$. Denote by σ^R a mixed strategy of the Receiver. Under the interim interpretation, a pure strategy of the type t Sender, $s^S(t)$, is an element in the message space M. Write $s^S \equiv \left(s^S(t)\right)_{t \in T}$. Let $\sigma^S(t) \in \Delta M$ denote a mixed strategy of type T Sender. In ex ante interpretation, a pure strategy of the Sender, s^S, is a function from the type space T to the message space M. Denote a pure Sender strategy by s^S and a mixed Sender strategy by σ^S. With some abuse of notation, write $\left(\sigma^S(t)\right)_{t \in T}$ as σ^S.

For ease of exposition, we restate the related CS results here. In their paper, both the type space and the action space are the unit interval. That is, $T = A =$

$[0,1]$. They showed that every equilibrium is characterized by a finite partition of the type space, $\{t_0, t_1, ..., t_N\}$, where $t_0 = 0$, $t_N = 1$, and type t_i is indifferent between being pooled with the immediately lower step and getting the action $E([t_{i-1}, t_i])$ and being pooled with the immediately higher step and getting the action $E([t_i, t_{i+1}])$. They proved that there exists a finite upper bound $N(b)$ on the maximum number of steps in an equilibrium, and that for every $1 \leq n \leq N(b)$, there exists an equilibrium with n steps.

They used a monotonicity condition to conduct comparative statics. Call a sequence $\tau \equiv \{\tau_0; \tau_1; ...; \tau_N\}$ a forward solution if type τ_i is indifferent between action $E([\tau_{i-1}, \tau_i])$ and action $E([\tau_i, \tau_{i+1}])$ for $i = 1, ..., N-1$. Call N the size of the forward solution τ. Say that τ is a size-N forward solution on $[\tau_0, \tau_N]$ and that $[\underline{\tau}, \bar{\tau}]$ has a forward solution of size N if there exists a forward solution $\{\tau_0; \tau_1; ...; \tau_N\}$ where $\tau_0 = \underline{\tau}$ and $\tau_N = \bar{\tau}$. With abuse of notation, we define

$$t_j^N([[\underline{\tau}, \bar{\tau}]]) \equiv \tau_j, \ j = 1, ..., N-1$$

where $\{\tau_0; \tau_1; ...; \tau_N\}$ is a forward solution on $[\underline{\tau}, \bar{\tau}]$. Write $\alpha_j^N([\underline{\tau}, \bar{\tau}]) \equiv E([\tau_{j-1}, \tau_j])$.

(M) If $\hat{\tau}$ and $\tilde{\tau}$ are two forward solutions with $\hat{\tau}_0 = \tilde{\tau}_0$ and $\hat{\tau}_1 > \tilde{\tau}_1$, then $\hat{\tau}_i > \tilde{\tau}_i$ for all $i \geq 2$.

CS proved that condition (M) implies that ex ante, the Receiver always prefers an equilibrium with more steps. Therefore, the most informative equilibrium, i.e. the equilibrium with the largest number of steps, gives the Receiver the highest ex ante utility. This condition will play an important role in some of our results.

5. Normal Form Iterative Admissibility

Section 5.1 characterizes the solution to $NIAL$, which is simply iterative admissibility of the game with language. Section 5.2 compares $NIAL$ with equilibria of the game without language and discusses the caveats of $NIAL$.

5.1. Characterizations

The notation here implies the use of the interim interpretation. However, the main results hold under both interpretations. Recall that $NIAL$ is simply iterative admissibility in the game with language. By the equivalence of weak dominance and never best response to a totally mixed belief in two player incomplete information games, we rewrite the procedure of $NIAL$ as the following:

Definition 5. $S^R(0) = S^R_L$. $S^S(0;t) = M \ \forall t$. *Defined iteratively:*

$$S^R(k+1) = \left\{ \begin{array}{c} s^R \in S^R(k) \mid \\ \text{there exists } \sigma^S(t) \in \Delta^+ S^S(k;t) \text{ for every } t \text{ such that} \\ U^R\left(\left(\sigma^S(t)\right)_{t\in T}, s^R\right) \geq U^R\left(\left(\sigma^S(t)\right)_{t\in T}, s'\right) \text{ for all } s' \in S^R(k) \end{array} \right\}$$

and

$$s^S(k+1;t) = \left\{ \begin{array}{c} m \in S^S(k;t) \mid \\ \text{there exists } \sigma^R \in \Delta^+ S^R(k) \text{ such that} \\ u^S\left(t, \sigma^R(m)\right) \geq u^S\left(t, \sigma^R(m')\right) \text{ for all } m' \in S^S(k;t) \end{array} \right\}$$

where $u^S\left(t, \sigma^R(m)\right) \equiv \sum_{s^R \in S^R} \sigma^R\left(s^R\right) u^S\left(t, s^R(m)\right)$. *Write* $\cap_{k=0}^{\infty} S^R(k) = S^R(\infty)$ *and* $\cap_{k=0}^{\infty} S^S(k;t) = S^S(\infty;t)$. *That is,* $S^R(\infty)$ *and* $S^S(\infty)$ *are the limiting set of strategies for the Receiver and the Sender respectively under this normal form iterative procedure.*

We need some more notations here.

Notation
1. $l(k;t) \equiv \min S^S(k;t)$;
2. $g(k;t) \equiv \max S^S(k;t)$;
3. $l^{-1}(k;m) \equiv \max\{t | l(k;t) \leq m\}$;
4. $g^{-1}(k;t) \equiv \min\{t | g(k;t) \geq m\}$.

$l(k;t)$ and $g(k;t)$ are respectively the smallest and the largest message that a type t Sender might send in round k. $l^{-1}(k;m)$ is the highest type t that might send a message smaller than or equal to m in round k, while $g^{-1}(k;m)$ represents the lowest type t that might send a message greater than or equal to m in round k. Given k, if $l(k;t)$ and $g(k;t)$ as functions from T to M are bijective when the range is restricted to $l(k;T)$ and $g(k;T)$ respectively, then $l(k;t)$ and $l^{-1}(k;m)$ are inverse functions to each other, while $g(k;t)$ and $g^{-1}(k;m)$ are inverse functions to each other.

Before characterizing the solution to cases where $b > 0$, let's look at the benchmark case where players' interests are aligned, that is, where $y^S(t) = y^R(t)$ for all t. Proposition 1 characterizes the $NIAL$ solution. It confirms conventional wisdom that players should be able to coordinate on the efficient outcome if the interests are aligned and they can communicate before playing the game.

We need the following observation for the proof. It says that if a message m is used sometimes in round k, and if it can only come from types smaller than or equal to itself, then any Receiver strategy that takes an action greater than the message m is weakly dominated.

Observation If $l^{-1}(k;m) \leq m$, and $m \in M(k)$, then $s^R(m) \leq m$ for all $s^R \in S^R(k+1)$.

Proof Suppose $\hat{s}^R(\hat{m}) > \hat{m}$ and $l^{-1}(k;\hat{m}) \leq \hat{m}$. Let m_1 be the smallest message m such that $\hat{s}^R(m) = \hat{s}^R(\hat{m})$. From the assumption that $l^{-1}(k;\hat{m}) \leq \hat{m}$, according to any belief on $\Delta S^S(k)$, any message in $[m_1, \hat{m}]$ can only come from types smaller than or equal to $\hat{m}$. Then $\hat{s}^R$ can be improved upon by lowering the value on $[m_1, \hat{m}]$ down to $\hat{m}$. This does not violate the language conditions. So $\hat{s}^R$ is weakly dominated w.r.t. $S(k)$ and does not belong to $S^R(k+1)$.

Proposition 1. *If $y^S(t_S) = y^R(t_S)\,\forall t_S$, then there is full communication in $S(\infty)$, i.e. $S(\infty) = \{s^R_{id}\}$ where $s^R_{id}(a) = a$ for all a.*

Proof Type 0 prefers a lower action to a higher one. In addition, she prefers action 0 the most. Recall that from the discretization, Δ is the second lowest message. From the relative property of language (see Lemma 3 in Section 3.2), a lower message always induces a weakly lower action. Therefore, in the first round, every message $m \geq \Delta$ is weakly dominated by message 0 with respect to $S(0)$ for type 0. Similarly, every message $m \leq 1-\Delta$ is weakly dominated by message 1 with respect to $S(0)$ for the highest type Sender. Knowing this, in the second round, the Receiver knows that any message between Δ and $1-\Delta$ can only come from types in $[\Delta, 1-\Delta]$. Then after receiving any message $m \in [\Delta, 1-\Delta]$, the Receiver will not take either action 0 or action 1. Suppose to the contrary, there exists $\hat{s}^R \in S^R(2)$ such that $\hat{s}^R(\hat{m}) = 0$ for some $\hat{m} > 0$ and $\hat{s}^R(\hat{m}+\Delta) \neq 0$. From the supermodularity condition of u^R, action Δ is better than action 0 for the Receiver whatever belief he has. If we change $\hat{s}^R$ by changing the action taken on $[\Delta, \hat{m}]$ from 0 to Δ, we strictly improve the Receiver's utility with respect to any belief in $\Delta S^S(1)$. The case for action 1 is similar.

Given $S^R(2)$, the lowest action that message Δ can induce is action Δ, and the highest action message $1-\Delta$ can induce is action $1-\Delta$. Since on the interval $[\Delta, 1-\Delta]$, type Δ prefers a lower action to a higher one, every message m greater than Δ is weakly dominated for type Δ by message Δ with respect to $S(2)$ because message Δ induces a weakly lower action in $[\Delta, 1-\Delta]$. The same holds here for type $1-\Delta$. We have used the same logic to get $S^S(1)$. Likewise for all s^R in $S^R(4)$, s^R takes on actions between $[2\Delta, 1-2\Delta]$ for messages in $[2\Delta, 1-\Delta]$. Repeating the process iteratively, we get that $s^R(m) = m$ for all $s^R \in S^R(\infty)$ and $s^S(t) = t$ for all type $t \in T$.

The finiteness assumption imposed on the type space T is crucial to the proof above. In the iterative process, we first showed that the lowest type of the Sender does not send any message other than the lowest one, because she prefers the lowest action (action 0) to a higher action, and message 0 induces a weakly lower action than any other message greater than 0. In response to that, the Receiver does not take the lowest action unless he receives the lowest message (message 0). Therefore, the lowest action a Sender will get by sending a message higher than 0 is the action preferred by the second lowest type of the Sender. Hence, the second lowest type Sender does not send any message higher than her most preferred action, which is equal to her type. However, if T is dense, the second lowest type does not exist. Therefore, this argument does not carry through. Nonetheless, the full communication result itself does not necessarily rely on the finiteness assumption. As a corollary to Proposition 4 to be stated later, when the monotonicity condition (M) holds (defined in section 4), full communication is the unique outcome in the limiting set even without the finiteness assumption. However, without the finiteness assumption, we do not know at this stage about convergence.

From now on, it is assumed that $b > 0$. It implies that $b \geq \Delta$ since y^i is defined on the discretized space. The case that $b < 0$ is done in the same way.

$NIAL$ gives a nontrivial upper bound and lower bound on the amount of information transmitted in a given game. Before stating the results, we need to define how we measure the amount of information transmitted. Let Q be a subset of the message space M. Say that an action a is inducible on Q under a Receiver strategy s^R if there exists a message m in Q such that $s^R(m) = a$. For the Receiver to be willing to use a strategy s^R that has many different inducible actions, he has to believe that the Sender can credibly transmit significant amounts of fine-tuned information. Let s^R be in $S^R(\infty)$. Define $M(\infty)$ to be the set of messages used by some type t under some strategy in $S^S(\infty; t)$. Say that a is inducible under s^R if a is inducible on $M(\infty)$ under s^R. When measuring the number of different inducible actions taken by s^R, we confine the attention to the message subset $M(\infty)$ since messages outside of $M(\infty)$ are never used by any type of the Sender, and hence actions taken by the Receiver outside of $M(\infty)$ are irrelevant. Proposition 2 stated below implies a limited number of different values for s^R. This is intuitive because as the interests of the Sender and the Receiver diverge, it becomes more difficult to transmit fine-tuned information credibly. Proposition 3 stated below says that if given a bias b sufficiently small, the number of inducible actions on $M(\infty)$ under s^R is at least $L \geq 2$, where L varies with the bias. In addition, L does not depend on how finely we discretize the action space, as long as it is not greater than the bias. These two results hold whether the interim interpretation or the ex ante interpretation is used in this incomplete information game.

Lemma 5 is the building block for these two results. It states that no type of the Sender ever recommends an action that is smaller than what is most preferred by the Receiver. We say that two messages are equivalent if they receive the same action under any s^R in $S^R(\infty)$. Lemma 5 sets forth, more precisely, that the Sender always gives a recommendation at least as high as an equivalent recommendation of her most preferred action.

Lemma 5. *$l(\infty;t) \geq t$ for all $t \in T$. Moreover, either $l(\infty;t) \geq y^S(t)$ or $s^R(l(\infty;t)) = s^R\left(y^S(t)\right)$ for all $s^R \in S^R(\infty)$.*

Proof (Sketch) Every type of the Sender prefers an action higher than her type t. Though the assumption of language implies that a higher message induces a weakly higher action, a type t Sender might want to send a recommendation smaller than her most preferred action for fear of being pooled with types that are too high. By definition, the distance between each type's most preferred action and her type is at least b. In particular, every type greater than or equal to $1-b$ prefers a higher action to a lower action and will send only message 1 after step 1. Any type smaller than $1-b$ will not be pooled with extremely high types if she sends a message no greater than $1-b$. Thus, on the message interval $[0, 1-b]$, type $1-b$ becomes the extreme high type. Using the same logic, we can show that the extreme high types in $[0, 1-b]$ will not send any message smaller than $1-b$. After iteration, we find that every type gives a recommendation at least as high as her most preferred action. Therefore, it is never optimal for the Receiver to take an action higher than the recommendation.

A further complication of this proof arises because one message cannot be weakly dominated by another if they always receive the same action. Because the iterative process depends on high types eliminating low messages, we must rule out as many situations as we can in which messages always receive the same action.

We need some notations here. Write $M(k) = \cup_{t \in T} S^S(k;t)$. $M(k)$ is the set of messages used by some type up to round k. Then $M(\infty) = \cup_{t \in T} S^S(\infty;t)$. Define

$$\left(s^R\right)^{-1}(a) \equiv \left\{m \in M | s^R(m) = a\right\}.$$

It is the set of messages in M that induces the action a under the Receiver strategy s^R.

Now we state the two main results for $NIAL$.

Proposition 2 (Coarseness). *Given any $s^R \in S^R(\infty)$, suppose $a_1 < a_2 < a_3$ are adjacent actions in the range of s^R, and $\left(s^R\right)^{-1}(a_2) \cap M(\infty) \neq \emptyset$, that is, a_2 might be received by some type t Sender under some Sender strategy in $S^S(\infty;t)$. Then the following inequality holds:*

$$a_3 - a_1 > a_3 - y_S^{-1}(a_3) \geq b,$$

where $y_3^{-1}(a_3)$ is the type that prefers action a_3, or the lowest type that prefers a_3 to any action lower than a_3.

Proof From the definition of S_L^R, $s^R(a_j) = a_j, j = 1, 2, 3$. Suppose to the contrary that there exists $\hat{s}^R \in S^R(\infty)$ where $a_1 < a_2 < a_3$ are adjacent actions taken in $\hat{s}^R$ and $a_3 - a_1 \leq b$. Let $[m_2, \bar{m}_2]$ be the interval on which $\hat{s}^R(m) = a_2$. That is, $\left(s^R\right)^{-1}(a_2) = [m_2, \bar{m}_2]$. From Lemma 5, types that prefer action a_3 to any lower one will send a message no smaller than a message which is equivalent to message a_3. (Recall the definition that a message m' is equivalent to message m if $s^R(m) = s^R(m') \forall s^R \in S^R(\infty)$.) $\bar{m}_2$ is not equivalent to message a_3 because $\hat{s}^R(a_3) \neq \hat{s}^R(\bar{m}_2)$ and $\hat{s}^R$ belongs to $S^R(\infty)$. From the definition of b, every type no smaller than $a_3 - b$ prefers action a_3 to any lower one. Therefore, messages in $[m_2, \bar{m}_2]$ can only come from types smaller than $a_3 - b$. By assumption, $a_3 - b \leq a_1$. Then if $[m_2, \bar{m}_2] \cap M(\infty) \neq \emptyset$, $\hat{s}^R$ can be improved upon by changing the action taken on $[m_2, \bar{m}_2]$ from action a_2 to action a_1. Hence $\hat{s}^R$ is weakly dominated and should not belong to $S^R(\infty)$. We need the qualifier that $\left(\hat{s}^R\right)^{-1}(a_2)$ is used by at least one type under some strategy because otherwise the Receiver would not care what he does on the interval $\left(\hat{s}^R\right)^{-1}(a_2)$.

Remark 1. *Proposition 2 shows that communication cannot be perfectly informative as long as the bias is greater than* 2Δ. *If the bias is large, we can be sure that very little information will be transmitted. Proposition 2 is parallel to Lemma 1 in Crawford and Sobel (1984), stating that there exists* $\varepsilon > 0$ *such that any two actions induced on the equilibrium path differ by at least* ε.

Remark 2. *The number of inducible actions under Receiver strategies in* $S^R(\infty)$ *is less than or equal to* $\frac{2}{b}$.

Proposition 3. *There exists* $L > 0$ *such that the number of inducible actions on* $M(\infty)$ *under any* $s^R \in S^R(\infty)$ *is at least* L. L *increases as the bias* b *decreases.*

Proof Lemma 5 shows that $l(\infty; t) \geq t$ for all t. From observation 5.1, $s^R(m) \leq m$ for all $m \in M(\infty)$. We can also show that the maximum action taken by a strategy $s^R \in S^R(\infty)$ must be greater than or equal to $E([0,1])$. If b is small enough, then $g(\infty; 0) < E([0,1])$, which implies that there will exist some types that will never elicit the highest action because they always send lower messages. Therefore, every s^R in $S^R(\infty)$ must partition $M(\infty)$ into at least 2 subintervals. Let m_q be the lowest message that takes on $\max_m \hat{s}^R(m)$ where $\hat{s}^R \in S^R(\infty)$. Then $E([0,1]) \leq \hat{s}^R(m_q) = m_q$. Let m_{q-1} be the smallest message that takes on $\max_{m<m_q} \hat{s}^R(m)$, then by the same argument, $m_{q-1} = \hat{s}^R(m_{q-1}) \geq E([0, g^{-1}(\infty; m_q)])$. If b is small, then $g^{-1}(\infty; E([0, g^{-1}(\infty; m_q)])) > 0$, that is, there will be types sending only low messages which never elicit an action higher than the second highest one. If we stop after L steps, then we know that every Receiver strategy s^R in $S^R(\infty)$ partitions $M(\infty)$ into at least L intervals.

CS showed that under the monotonicity condition (M) restated here in section 4, the Receiver prefers the most informative equilibrium. They argued that focusing on the most informative equilibrium would be natural. It is natural to ask whether $NIAL$ provides grounds for doing so. To relate $NIAL$ to the equilibrium concept, we will use ex ante interpretation, which is equivalent to assuming that different Sender types hold the same belief about the behavior of the Receiver. Proposition 4 states that every equilibrium which is not as informative as the largest equilibrium will be eliminated.

Discretization compels us to make certain assumptions. When $T = [0, 1]$, continuity insures that boundary types, which are indifferent between two equilibrium actions, are of measure zero. We assume that this condition holds in the discrete case.

Assumption Given any $\underline{\tau} < \bar{\tau} \in T$, every equilibrium in the game restricted to the subset $[\underline{\tau}, \bar{\tau}] \cap T$ is such that no boundary types are indifferent between two equilibrium actions.

This assumption implies that every forward solution $\{\tau_0; \tau_1; ...; \tau_n\}$ is such that type $\tau_i - \Delta$ prefers action $E([\tau_{i-1}, \tau_i - \Delta])$ to action $E([\tau_i, \tau_{i+1} - \Delta])$, while type τ_i prefers action $E([\tau_i, \tau_{i+1} - \Delta])$ to action $E([\tau_{i-1}, \tau_i - \Delta])$. This assumption will be carried throughout the paper.

Proposition 4. *Under condition (M), every Receiver strategy satisfying $NIAL$ takes at least as many different actions on $M(\infty)$ as the most informative equilibrium in the game without language. That is, $L \geq N(b)$.*

The proposition follows immediately from the following claim.

Claim For any $\hat{\alpha} \in \hat{s}^R(M(\infty))$ where $\hat{s}^R \in S^R(\infty)$ and $\hat{a} \neq \min \hat{s}^R(M(\infty))$, $\hat{s}^R$ takes at least q different actions on $M(\infty) \cap [0, \hat{a} - \Delta]$ if $[0, g^{-1}(\infty; \hat{a}) - \Delta]$ has a forward solution of size-q. Let $\tilde{a} \equiv \max \hat{s}^R([0, \hat{a} - \Delta] \cap M(\infty))$. Then $\tilde{a}$ is greater than or equal to the largest action on the size-q forward solution on $[0, g^{-1}(\infty; \hat{a}) - \Delta]$ and $g^{-1}(\infty; a_q) \geq t^q_{q-1}([0, g^{-1}(\infty; \hat{a}) - \Delta])$ where a_q is the smallest message equivalent to $\tilde{a}$.

Proof (Sketch) Prove by induction. The claim holds by the construction of $\hat{s}^R$ and $\hat{a}$ for $q = 1$. Suppose the claim holds for any such $\hat{s}^R \in S^R(\infty)$ and any $\hat{a} \in \hat{s}^R(M(\infty))$ for $q = 1, 2, ..., \bar{q}$. To abuse notation, write $\bar{q}$ as q. Recall that for every $s^R \in S^R(\infty)$, $s^R(m) \leq m$ for all $m \in M(\infty)$. To show that it holds for $q+1$, it suffices to show that $[0, g^{-1}(\infty; \tilde{a}) - \Delta]$ must have a size-q forward solution. It is easy to see that $\tilde{a} \geq E([g^{-1}(\infty; \tilde{a}), g^{-1}(\infty; \hat{a}) - \Delta])$. Then we can show that $\hat{s}^R$ must partition $[0, g^{-1}(\hat{a}) - \Delta] \cap M(\infty)$ into at least two intervals. That is, roughly, $\hat{s}^R(\tilde{a} - \Delta) \neq \hat{s}^R(\tilde{a})$ and $g^{-1}(\infty; \tilde{a}) > 0$. Suppose $[0, g^{-1}(\infty; \tilde{a}) - \Delta]$ has a maximum of size-j forward solution where $j < q$. Notice that by the definition of $g^{-1}(\infty; a)$, type $g^{-1}(\infty; \tilde{a})$ has to prefer the action $\tilde{a}$ to action $\hat{s}^R(\tilde{a} - \Delta)$. By $\frac{\partial^2}{\partial a^2} u^S < 0$, type $g^{-1}(\infty; \tilde{a})$

has to prefer action $E\left([g^{-1}(\infty;\tilde{a}), g^{-1}(\infty;\hat{a})-\Delta]\right)$ to action $\hat{s}^R(\tilde{a}-\Delta)$. Let $\bar{t}$ be such that type $g^{-1}(\infty;\tilde{a})$ prefers action $E\left([g^{-1}(\infty;\tilde{a}),\bar{t}]\right)$ to action $\alpha_j^j\left([0, g^{-1}(\infty;\tilde{a})-\Delta]\right)$, and type $g^{-1}(\infty;\tilde{a})-\Delta$ prefers action

$$\alpha_j^j\left([0, g^{-1}(\infty;\tilde{a})-\Delta]\right)$$

to action $E\left([g^{-1}(\infty;\tilde{a}),\bar{t}]\right)$. Then we form a size-$(j+1)$ forward solution by adding $\bar{t}$ to the size-j forward solution $t^j\left([0, g^{-1}(\infty;\tilde{a})-\Delta]\right)$. Since the largest forward solution on $[0, g^{-1}(\infty;\tilde{a})-\Delta]$ is of size j, the largest forward solution on $[0,\bar{t}]$ is of size $j+1$ by condition (M). But $[0, g^{-1}(\infty;\tilde{a})-\Delta]$ has a size-$(q+1)$ forward solution, so $\bar{t} < g^{-1}(\infty;\hat{a})-\Delta$ and therefore $E\left([g^{-1}(\infty;\tilde{a}),\bar{t}]\right) < E\left([g^{-1}(\infty;\tilde{a}), g^{-1}(\infty;\hat{a})]\right)$. Since $\bar{t}$ is defined such that type $g^{-1}(\infty;\tilde{a})$ is roughly indifferent between $E\left([g^{-1}(\infty;\tilde{a}),\bar{t}]\right)$ and $\alpha_j^j\left([0, g^{-1}(\infty;\tilde{a})-\Delta]\right)$, type $g^{-1}(\infty;\tilde{a})$ prefers action $\alpha_j^j\left([0, g^{-1}(\infty;\tilde{a})-\Delta]\right)$ to action $E\left([g^{-1}(\infty;\tilde{a}), g^{-1}(\infty;\hat{a})]\right)$ because $E\left([g^{-1}(\infty;\tilde{a}), g^{-1}(\infty;\hat{a})]\right) > E\left([g^{-1}(\infty;\tilde{a}),\bar{t}]\right)$. By assumption, $s^R(\tilde{a}-\Delta) \geq \alpha_j^j\left([0, g^{-1}(\infty;\tilde{a})-\Delta]\right)$ for any $s^R \in S^R(\infty)$ such that $\tilde{a} \in s^R(M(\infty))$. By the property of $S^R(\infty)$, $s^R(\tilde{a}) = \tilde{a} > s^R(\tilde{a}-\Delta)$ if $s^R \in S^R(\infty)$ and $\tilde{a} \in s^R(M(\infty))$. We have shown that $\tilde{a} \geq E\left([g^{-1}(\infty;\tilde{a}), g^{-1}(\infty;\hat{a})-\Delta]\right)$. Then type $g^{-1}(\infty;\tilde{a})$ prefers action $s^R(\tilde{a}-\Delta)$ to action $\tilde{a} = s^R(\tilde{a})$ for all $s^R \in S^R(\infty)$ where $s^R(\tilde{a}) \neq s^R(\tilde{a}-\Delta)$. Such s^R exists by construction, indicating that it is weakly dominated for type $g^{-1}(\infty;\tilde{a})$ to send message $\tilde{a}$. But this violates the definition of $g^{-1}(\infty;\tilde{a})$. We hence arrive at a contradiction. Therefore $j \geq q$. But then we can show similarly that $g^{-1}(\infty;\tilde{a}) \geq t_q^{q+1}\left([0, g^{-1}(\infty;\hat{a})]\right)$.

Remark 3. *We prove it by showing that it is necessary for the limiting set. The arguments do not depend on the finiteness assumption. As a corollary, this is also a necessary condition for the limiting set under $NIAL$ even if $T = A = M = [0,1]$. In fact, we do not need the assumptions we impose in the discrete case.*

5.2. Relating $NIAL$ to Equilibria in the Game without Language

Denote by $EQ(G)$ the set of equilibria in G, where G represents the game without language. Recall that $NIAL$ is iterative admissibility in G_L, the game WITH language. In general, there is no containment between $NIAL$ and $EQ(G)$. As a non-equilibrium concept, $NIAL$ naturally gives rise to non-equilibrium outcome being contained in $NIAL$. Proposition 3 implies that $NIAL$ may eliminate some

types of S sending message m in $S^S(1)$	$\emptyset$	0	$\frac{1}{2}, 1$		
s^R **in language \ message** m	"0"	"$\frac{1}{2}$"	"1"	in $S^R(1)$	in $S^R(2)$
	0	0	0		
	0	$\frac{1}{2}$	$\frac{1}{2}$	v	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	v	
	1	1	1		
	0	1	1	v	
	$\frac{1}{2}$	$\frac{1}{2}$	1	v	
	0	$\frac{1}{2}$	1	v	
s^R_{nice}	0	0	1	v	v

Table 5.1: Receiver Strategy Set in Language

of the less informative equilibria. However, in this section, we present an example demonstrating that $NIAL$ can be disjoint from $EQ(G)$. In our example, the unique equilibrium in G is the babbling equilibrium, while the unique prediction given by $NIAL$ is partially informative.

Example 1. *There are three types: type* $0, \frac{1}{2}$ *and* 1. *The common prior is such that* $\pi(0) = \frac{1}{3}$; $\pi\left(\frac{1}{2}\right) = \frac{4}{9}$ *and* $\pi(1) = \frac{2}{9}$. *Both the Sender and the Receiver have quadratic loss function:* $u^R(t,a) = -(t-a)^2$ *and* $u^S(t,a) = -\left(t + \frac{1}{2} - a\right)^2$.

The unique equilibrium in this game without language is babbling. Because both type $\frac{1}{2}$ and type 1 Senders prefer a higher action to a lower one, it is impossible to separate these two types in any equilibrium. To show that there is no informative equilibria, let's suppose to the contrary that there is an equilibrium in which type 0 separates from type $\frac{1}{2}$ and 1. The best action against pooling of type $\frac{1}{2}$ and type 1 is action $\frac{1}{2}$, while the best action against type 0 is action 0. However, this cannot be an equilibrium because type 0 prefers action $\frac{1}{2}$ to action 0, and therefore would have an incentive to imitate type $\frac{1}{2}$ and type 1. Thus, in an equilibrium, type 0 cannot separate from type $\frac{1}{2}$ and type 1. Since the best action with respect to the prior is $\frac{1}{2}$, the unique equilibrium in the game without language is babbling, where all types pool and the Receiver takes action $\frac{1}{2}$ after receiving every message.

Now let's derive the solution to this game under $NIAL$. Write s^R as a 3-tuple of actions taken after messages $0, \frac{1}{2}$ and 1 respectively, i.e., $s^R = \left(s^R(0), s^R\left(\frac{1}{2}\right), s^R(1)\right)$. The bottom part of table 5.1 shows all the Receiver strategies in S_L^R. In the first round of deletion, $(0,0,0)$ and $(1,1,1)$ are eliminated since the unique best response to pooling of all types is $\frac{1}{2}$. For every other strategy $\hat{s}^R$, a totally mixed Sender strategy $\hat{\sigma}^S$ can be constructed such that $\hat{s}^R$ is a best response to $\hat{\sigma}^S$. Thus no further Receiver strategies can be eliminated in the first round.

Now we must determine the set of Sender strategies that survive the first round. Both type 1 and type $\frac{1}{2}$ Sender prefer a higher action to a lower one. For these two types, both message 0 and message $\frac{1}{2}$ are weakly dominated by message 1 because Receiver strategies in S_L^R are weakly increasing, and hence message 1 induces the weakly highest action. A type 0 Sender prefers action $\frac{1}{2}$ the most, and is indifferent between action 0 and action 1. Recall that by the absolute meaning property of language, s^R satisfies the following inequalities if $s^R(0) \neq s^R\left(\frac{1}{2}\right)$:

$$\begin{cases} s^R\left(\frac{1}{2}\right) > s^R(0) \\ s^R\left(\frac{1}{2}\right) > 0 \\ s^R(0) < \frac{1}{2} \end{cases}.$$

This implies that $s^R(0) = 0$ and $s^R\left(\frac{1}{2}\right) \geq \frac{1}{2}$ if $s^R(0) \neq s^R\left(\frac{1}{2}\right)$. That is, whenever message 0 and message $\frac{1}{2}$ induce different actions, message 0 induces action 0 while message $\frac{1}{2}$ induce either action $\frac{1}{2}$ or action 1. As the type 0 Sender weakly prefer both action $\frac{1}{2}$ and action 1 over action 0, and strictly prefers action $\frac{1}{2}$ over action 0, message 0 is weakly dominated by message $\frac{1}{2}$ for the type 0 Sender. Similarly, message 1 is weakly dominated by message $\frac{1}{2}$ for a type 0 Sender among $S^S(0)$. In conclusion, after the first round of deletion, a type 0 Sender will send only message $\frac{1}{2}$, and both type $\frac{1}{2}$ and type 1 Sender will send only message 1. This Sender strategy, called s_{nice}^S, is shown in the first row of table 5.1.

Now we show that $S^R(2) = \left\{s_{nice}^R\right\}$ where $s_{nice}^R = (0,0,1)$. In the second round, the only conjecture the Receiver can hold about the Sender's strategy is s_{nice}^S. Under s_{nice}^S, no type of Sender ever sends message 0. Therefore the Receiver's predetermined response to message 0 is irrelevant. Hence, the relevant difference among Receiver strategies $\left(\frac{1}{2},\frac{1}{2},1\right), \left(0,\frac{1}{2},1\right), (0,1,1)$ and $(0,0,1)$ lies only in their responses at message $\frac{1}{2}$. When receiving message $\frac{1}{2}$, action 0 is the best because only type 0 sends this message. Therefore, Receiver strategy $(0,0,1)$ yields a higher utility than either strategy $\left(\frac{1}{2},\frac{1}{2},1\right), \left(0,\frac{1}{2},1\right)$ or $(0,1,1)$. Then we need only compare s_{nice}^R with the strategy $\left(0,\frac{1}{2},\frac{1}{2}\right)$. Simple calculation of ex ante utility shows that $U^R\left(s_{nice}^S,(0,0,1)\right) > U^R\left(s_{nice}^S,\left(0,\frac{1}{2},\frac{1}{2}\right)\right)$. So $S^R(2) = \left\{s_{nice}^R\right\}$. The process then stops and $S(\infty) = \left(\left\{s_{nice}^S\right\},\left\{s_{nice}^R\right\}\right)$. Call this strategy profile s_{nice}. $NIAL$ predicts that type 0 Sender receives action 0 and both type $\frac{1}{2}$ and

	Equilibrium	**IA**	ID
No Language	s_{babble}	everything	everything
Language	s_{babble}, s_{nice}	s_{nice}	everything

Table 5.2: Comparison of Predictions

types of S sending message m **in** s^S_{nice}	$\emptyset$	0	$\frac{1}{2}, 1$	
$s^R \setminus$ **message** m	"0"	"$\frac{1}{2}$"	"1"	in language
s^R_{nice}	0	0	1	yes
s^R_{cheat}	0	0	$\frac{1}{2}$	no
s^R_{ignore}	0	$\frac{1}{2}$	$\frac{1}{2}$	yes

Table 5.3:

type 1 Sender receive action 1.

Table 5.2 summarizes this game's predictions under different combinations of language restriction and solution concepts. s_{nice} emerges as an equilibrium in the game with language, though babbling is the unique equilibrium in the game without language. In arriving at $S(\infty)$, we previously showed that s^R_{nice} is optimal with respect to s^S_{nice} among S^R_L. s^S_{nice} is optimal among S^S with respect to s^R_{nice} because every other Sender strategy is weakly dominated by s^S_{nice} with respect to S^R_L. It follows that s_{nice} is an equilibrium in G_L.

To understand why s_{nice} is not an equilibrium in G, note that according to s^S_{nice}, message 1 is transmitted by either type $\frac{1}{2}$ or type 1. The best response against pooling of these two types is action $\frac{1}{2}$, not action 1. Therefore, the strategy $\left(0, 0, \frac{1}{2}\right)$ yields a higher utility than $(0, 0, 1)$ with respect to s^S_{nice}. It then follows that s_{nice} is not an equilibrium in G. Table 5.3 illustrates all the relevant strategies. To see why s_{nice} is an equilibrium in G_L but not an equilibrium in G, note that $\left(0, 0, \frac{1}{2}\right)$ does not satisfy the language assumptions, and therefore does not belong to S^R_L. Recall that by the literal meaning assumption of language (see definition 3), if action $\frac{1}{2}$ belongs to the range of a strategy, action $\frac{1}{2}$ must be taken in response to message $\frac{1}{2}$. Thus, if the Receiver wants to take action $\frac{1}{2}$ after receiving message 1, he must choose strategy $\left(0, \frac{1}{2}, \frac{1}{2}\right)$. Though strategy $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ yields a higher interim utility than strategy $(0, 0, 1)$ when message 1 is received, it yields a lower interim utility when message $\frac{1}{2}$ is received, because only type 0 sends message $\frac{1}{2}$, and action 0 is the best against type 0. When deriving the solution to $NIAL$, we have shown that strategy $(0, 0, 1)$ gives a higher ex ante payoff than strategy $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ against the Sender strategy s^S_{nice}. Therefore,

$(0,0,1)$ is optimal among S_L^R with respect to s_{nice}^S, though it is not optimal among the unrestricted strategy set.

This example points out that in a game with language, ex ante utility maximization does not necessarily imply interim utility maximization on every information set on the equilibrium path. We showed above that s_{nice}^R is ex ante optimal against s_{nice}^S. However, s_{nice}^R takes a suboptimal action against the posterior generated by s_{nice}^S in response to message 1, which is reached with positive probability by the profile s_{nice}. Thus, in the game with language, s_{nice}^R is not interim optimal with respect to s_{nice}^S even on the equilibrium path.

The break down of the link between ex ante optimality and interim optimality on the equilibrium path results from the non-separability of the second-stage-action space created by the language restriction. In the first stage of the two-stage sender-receiver game, the Sender decides to send a message based on her private information. In the second stage, the Receiver takes an action in response to the message from the Sender. In a standard game, the action space available to a player at a given information set (a particular message in a sender-receiver game) does not depend on the action the player plans to take at any other information set. We can conceive of the second stage action space as "separable". The language restriction breaks the separability: the set of actions available to the Receiver upon receiving a message depends on which actions he plans to take in response to other messages. Although for any single message taken in isolation, language does not restrict the Receiver to a strict subset of his action space, when holding fixed the Receiver's responses to other messages, language assumption does often impose restrictions on the available action space. In Example 1, if the Receiver wants to respond optimally to message 1 with respect to s_{nice}^S, he should take action $\frac{1}{2}$. However, if he takes action $\frac{1}{2}$ after receiving message 1, he must then take action $\frac{1}{2}$ at message $\frac{1}{2}$ by the literal meaning condition of language. Note that our assumption prohibits the Receiver from any strategy which violates language. Therefore, the set of actions available to the Receiver at message $\frac{1}{2}$ is $\left\{\frac{1}{2}\right\}$ given that he takes action $\frac{1}{2}$ after receiving message 1. Thus it is often the case that when the Receiver decides to take an optimal action in response to a message based on a conjecture he holds about the Sender's behavior, he will be forced to take a suboptimal action in response to another message reached with positive probability. The Receiver gauges the gains and losses ex ante and chooses one that maximizes his ex ante utility.

Though the language restriction does not limit strategic contents, in that $EQ(G) \subset EQ(G_L)$ for every sender-receiver game G, it does provide an artificial commitment device that may make $EQ(G)$ strictly contained in $EQ(G_L)$. As is often the case, commitment makes the Receiver weakly better off. For example, s_{nice} gives the Receiver a higher ex ante payoff than the babbling outcome. However, the Receiver does not really have a commitment device. Incorporating language with a normal form approach fails to take into account sequential rationality, since in the game with language, interim optimality on the equilibrium path is no longer implied by ex ante optimality. This prompts us to develop an extensive form version incorporating language, iterative admissibility and sequential rationality.

6. Extensive Form Iterative Admissibility with Language

To address the issue presented in example 1, an extensive form version of iterative admissibility with language ($EIAL$) is proposed. Section 6.1 motivates and defines weak sequential rationality and the procedure for $EIAL$. Section 6.2 characterizes the solution to $EIAL$.

6.1. Weak Sequential Rationality and the Extensive Form Procedure

Sequential rationality is not a novel issue, and a natural first step is to add the requirement into the iterative procedure. Recall the standard definition of sequential rationality. A Receiver strategy σ^R is sequentially rational with respect to a belief σ^S if and only if $\sigma^R(m)$ is optimal at every message m according to the Bayesian update of σ^S. However, with a simple opposing-interest example, we show that this definition may clash with language combined with iterative admissibility. We argue that a weaker notion of sequential rationality, in terms of the induced outcome instead of the strategy profile, can better capture the idea of language, because messages serve only as coordination device. We then develop the extensive form procedure $EIAL$. Example 1 is revisited to show the predictions of IA combined with different sequential rationality notion. It is shown that the limiting set of $EIAL$ is nonempty.

6.1.1. The Opposing-interest Game

Let's look at the game in figure 6.1 where the Sender and the Receiver have opposing interest. When the true state is $West$, the Receiver wants to take action W while the Sender wants the Receiver to take action E and vice versa

		a	
		West	East
t	West	0,1	2,0
	East	2,0	0,1

Table 6.1: Opposing Interest Game

when the true state is $East$. The probability that the true state is $West$ is $\frac{2}{3}$ and the probability that the true state is $East$ is $\frac{1}{3}$. If the players cannot communicate before the Receiver takes an action, it's optimal for the Receiver to take action W. This game has a unique babbling equilibrium. $NIAL$ gives a unique solution where the Receiver takes action W to both messages, which is the same as in the babbling equilibrium.

We derive the solution to $NIAL$ in this game as follows. The bottom part of table 6.2 shows all the Receiver strategies in G_L, the game with language. In the first round of deletion, the strategy *Stubborn E* is eliminated because it is strongly dominated by *Stubborn W* since taking action W is optimal without communication. Nothing else can further be eliminated for the Receiver. For type $West$ Sender, sending message "*west*" is weakly dominated by sending message "*east*," because type $West$ prefers action E to action W, and either both messages lead to the same action, or message "*west*" leads to action W and message "*east*" leads action E. Similarly, for type $East$ Sender, sending message "*east*" is weakly dominated by sending message "*west*." In summary, the only strategy that survives the first round of deletion for each type of Sender is to utter the desired action. Call it s^S_{prefer}. That is, $s^S_{prefer}(West) =$ "*east*" and $s^S_{prefer}(East) =$ "*west*". In the second round of deletion, the only conjecture the Receiver can have about the Sender's behavior is s^S_{prefer}. *Stubborn W* strategy strictly dominates *Literal* strategy with respect to s^S_{prefer}. The two strategies differ only on the actions taken after receiving message "*east*." At round 2, message "*east*" can only come from a type $West$ Sender, and *Stubborn W* strategy takes action W there, which is better against type $West$ than action E, the action taken by *Literal* strategy. Therefore, we end up with a unique prediction $S(\infty) = \left\{s^S_{prefer}, \textit{Stubborn } W\right\}$, which gives the babbling outcome.

However, *Stubborn W* is not interim optimal with respect to s^S_{prefer} when the Receiver receives message "*west*", even though $NIAL$ prediction is equal to the unique equilibrium outcome in the original game. The top row in table 6.2 illustrates the correspondence between messages and types under the sender

types of Sender sending m in s^S_{prefer}	$East$	$West$
s^R**message**	**"West"**	**"East"**
Stubborn W	W	W
Stubborn E	E	E
Literal	W	E

Table 6.2: Language in Opposing Interest Game

strategy s^S_{prefer}. As is shown, message "*west*" can only come from type $East$. But when the true state is $East$, the optimal action is action E , not action W which is taken by the *Stubborn W* strategy. The unique strategy which is sequentially rational with respect to s^S_{prefer} is the *Opposite* strategy (E, W). But (E, W) does not belong to language, and therefore is physically unavailable. Since s^S_{prefer} is the only conjecture the Receiver can have in the second round, none of the strategies in language satisfies standard sequential rationality in the second round. Therefore, imposing standard sequential rationality in the iterative procedure would yield an empty set.

6.1.2. Weak Sequential Rationality

To see what drives this result and how to tackle it, it might be worthwhile to look at sequential rationality by its components. In this two-stage game, sequential rationality can be broken down into ex ante rationality and interim rationality. Ex ante rationality means utility maximization at the hypothetical initial node, before the receiver receives the message. Interim rationality means utility maximization at every information set in the second stage, after receiving the message. Interim rationality implies ex ante rationality. In the game without language, ex ante rationality implies interim rationality. We showed that the latter does not hold in the game with language. While ex ante rationality is taken care of by normal form analysis, the problem lies in interim rationality. The above discussion shows that no Receiver strategy in language is sequentially rational with respect to s^S_{prefer}. This is because no Receiver strategy in language is interim rational with respect to s^S_{prefer}. Since every outcome can be achieved in the game with language by some strategy profile, and thus any information can be successfully transmitted by some message usage specified by language, we wonder what s^S_{prefer} represents in the game with language: is it meant to convey information?

This project focuses on the set of outcomes: language specifies how messages are used to achieve a given set of outcomes. But standard interim rationality is defined in terms of strategy profiles, not outcomes. In addition, it is sensitive to the number of messages employed to convey the given information. Consider the opposing-interest game without language. We want to find the smallest set of strategy profiles which satisfies the following two properties: 1) it contains all babbling strategies; 2) it contains all pure strategies getting positive weight in the set; 2) it is closed under standard interim rationality. Suppose the message space is trivial and contains only one message. Then the smallest such set contains only babbling outcome. But if there are two messages in the message space, then one babbling strategy for each type of the sender is to randomize over the two messages. Then to contain all supporting pure strategies, this set needs to contain the two sender strategies where type $West$ sends one message and type $East$ sends the other message. One such strategy is s^S_{prefer}, where type $West$ utters "*east*" and type $East$ says "*west*." Let's call the other strategy s^S_{honest}. Then to contain Receiver strategies that are interim rational with respect to these two sender strategies, this set needs to contain both *Literal* and *Opposite*, where the Receiver takes different actions after receiving different messages. This set then has to contain two separating outcomes. However, if we look at interim rationality in terms of outcomes, ignoring altogether how messages are used, we'll avoid this dependence. Since we use language to take care of how messages are used, it might be natural to look for a notion of sequential rationality that deals only with the outcomes.

Given a Receiver strategy s^R and a belief σ^S. Typically we say that the profile $\left(s^R, \sigma^S\right)$ gives rise to an outcome which is a mapping from the type space to distributions over the action space. From the Receiver's point of view, however, the profile $\left(\sigma^S, s^R\right)$ gives rise to an association between actions in the range of s^R and distributions over the type space. Let $\beta_{(\sigma^S, s^R)}$ denote the association induced by the profile $\left(\sigma^S, s^R\right)$ and $s^R(M)$ denote the range of s^R. Then, $\beta_{(\sigma^S, s^R)}$: $s^R(M) \rightarrow \Delta T$ is defined by

$$\beta_{(\sigma^S, s^R)}(a)(t) = \frac{\sum_{m \in (s^R)^{-1}(a)} \sigma^S(m; t)}{\sum_{t'} \sum_{m \in (s^R)^{-1}(a)} \sigma^S(m; t')},$$

which is simply a Bayesian update. For each action a that the Receiver takes in response to some message $m \in M$, $\beta_{(\sigma^S, s^R)}$ associates with it a probability distribution on the type space T, which represents the distribution of the types of the Sender that might receive this action a under the profile $\left(\sigma^S, s^R\right)$. Standard interim rationality looks at $\left(\sigma^S, s^R\right)$. We propose checking interim rationality from the point of view of $\beta_{(\sigma^S, s^R)}$. It is formally stated as follows:

Definition 6 (Outcome Interim Rationality). *Let B denote a subset of A. Say that $\beta : B \to \Delta T$ is outcome interim rational if and only if*

$$a \in \arg\max_{a' \in A} \sum_t \beta(a)(t)\, u^R(t, a')$$

for all $a \in B$.

Definition 7 (Weak Interim Rationality). *Say that s^R is weakly interim rational with respect to σ^S if and only if $\beta_{(\sigma^S, s^R)}$ is outcome interim rational.*

It is easy to see that *Stubborn W* is weakly interim rational with respect to s^S_{prefer}. Actually, *Stubborn W* is weakly interim rational with respect to every $\sigma^S \in \Delta S^S$. In general, for every $\sigma^S \in \Delta S^S$, there exists a s^R in language that is weakly interim rational with respect to σ^S. Using outcome interim rationality, we avoid the problem that there might exist some conjectures σ^S with respect to which no Receiver strategy in language is interim rational. However, unlike standard interim rationality, weak interim rationality does not necessarily imply ex ante rationality. Given a belief σ^S, there are typically many Receiver strategies that are weakly interim rational with respect to σ^S and can be Pareto ranked. The idea of sequential rationality is that strategies that are not "rational" at the interim stage are not credible. This motivates a weaker notion of ex ante rationality: compare ex ante payoff among only "credible" Receiver strategies. More precisely, given σ^S, only strategies that are weakly interim rational with respect to σ^S are credible. Ex ante, the Receiver picks among these "credible" strategies one that gives him the highest ex ante payoff. We combine the weaker notion of ex ante rationality and weak interim rationality analogously to define weak sequential rationality. Breaking standard sequential rationality into the two parts and putting them back this way does not alter the implication, since every Receiver strategy that is interim rational with respect to a conjecture σ^S is ex ante payoff equivalent to each other.

Definition 8 (Weak Sequential Rationality). *Let $X^R \subset S^R$. Say that s^R is weakly sequentially rational among X^R with respect to σ^S if and only if s^R is ex* *ante optimal with respect to σ^S among Receiver strategies that are weakly interim rational with respect to σ^S. That is,*

$$s^R \in \arg \max_{\substack{s^{R\prime} \in X^R \\ s^{R\prime} \text{ is weakly interim optimal} \\ \text{w.r.t. } \sigma^S}} U^R \left(\sigma^S, s^{R\prime}\right).$$

Call $\left(\sigma^S, \sigma^R\right)$ a weak sequential equilibrium if and only if σ^S is sequentially rational with respect to σ^R and s^R is weakly sequentially rational with respect to σ^S for every s^R in the support of σ^R. Let $WSEQ(G)$ denote the set of weak sequential equilibrium in the game without language and $WSEQ(G_L)$ denote the set of weak sequential equilibrium in the game with language. Then $WSEQ(G) = WSEQ(G_L)$. Recall that in Example 1, where $NIAL$ selects a unique informative outcome while the unique equilibrium in the original game is babbling, the set of equilibrium outcomes in the game with language strictly contain the set of equilibrium outcomes in the game without language. That is, $EQ(G) \subsetneqq EQ(G_L)$. It is then not that surprising that $NIAL$, being iterative admissibility on G_L, does not select any equilibrium in $EQ(G)$. Imposing weak sequential rationality restores the equilibrium outcomes in G_L to the equilibrium outcomes in G. This gives us hope that this definition might work.

The motivation for outcome interim rationality is that, instead of truly conveying information, s^S_{prefer} might simply be a supporting pure strategy of the mixed babbling sender strategy. But in the second round, s^S_{prefer} is the only conjecture the Receiver can hold. If s^S_{prefer} represents only a supporting pure strategy of the mixed babbling strategy, the other supporting pure strategy should also be contained as a possible conjecture held by the Receiver.

We now explain how the combination of language and weak dominance selects $\left\{s^S_{prefer}\right\}$ as the unique conjecture the Receiver can hold in the second round and why it is more properly viewed as a pure strategy supporting the mixed babbling sender strategy. In the first round, we eliminated all sender strategies except s^S_{prefer}. The elimination takes place because the sender takes into account the possibility of the strategy *Literal* being used. Suppose instead that the sender believes that *Literal* is not going to be used in the game with language, and therefore the Receiver always ignores messages. Then the two messages have exactly the same implication to the Sender, and therefore she might as well randomize. No message is weakly dominated for either type, and all sender strategies are possible. This points to the well-known force of weak dominance: the reason for eliminating one strategy might later be eliminated. To show the role language plays, consider the game without language. If the sender takes into account all four strategies, she is not sure which one induces her preferred action more often. The two messages again look the same to her, though she might prefer one message under some conjecture, while another under another conjecture. We'll end up with everything in the prediction, which is not clear whether it represents no information transmission, or simply no predicting power.

Language gives a bite by specifying the asymmetry. Though it does not rule out any outcome, weak dominance forces the sender to take into account all communication outcomes. Babbling is present in every cheap talk game because it is self-fulfilling: if the receiver always takes the same action, and the sender wholeheartedly believes that, then the sender sees the two messages as the same and might very well randomize between the two. This in turn makes it optimal for the Receiver to treat the two messages equally and therefore always take the same action. Language and weak dominance breaks out of this by making the sender take into account all outcomes. But the danger lies in going to the other extreme and taking into account outcomes that cannot happen in the game in question. This gives rise to selecting $\left\{s^S_{prefer}\right\}$ as the unique sender strategy profile in a babbling outcome, because the sender takes into account even outcomes that are not possible in the situation.

6.1.3. The Procedure for the Extensive Iterative Admissibility with Language (EIAL)

The idea is to let language take care of how messages are used and use weak sequential rationality to take care of rationality. That $WSEQ\left(G_L\right) = WSEQ\left(G\right)$ is encouraging. We then define the iterative procedure in an analogous way. We call this procedure extensive form iterative admissibility with language ($EIAL$).

Let $ES^S\left(0;t\right) = M$, $\forall t$ and $ES^R\left(0\right) = S^R_L$.

Procedure 1. $s^R \in ES^R\left(k+1\right)$ iff

   a. $s^R \in ES^R\left(k\right)$

   b. there exists a totally mixed conjecture $\sigma^S \in \Pi_{t\in T}\left(\Delta^+ ES^S\left(k;t\right)\right)$ such that s^R is weakly sequentially rational with respect to σ^S.

2. $s^S\left(t\right) \in ES^S\left(k+1;t\right)$ iff

   a. $s^S\left(t\right) \in ES^S\left(k;t\right)$

   b. there exists a totally mixed conjecture $\sigma^R \in \Delta^+ ES^R\left(k\right)$ such that $s^S\left(t\right)$ is a best response among $ES^S\left(k+1;t\right)$ with respect to σ^R.

Definition 9. *Write* $\cap_{k=0}^{\infty} ES^i(k) = ES^i(\infty)$ *and* $ES(k) = \left(ES^S(k), ES^R(k)\right)$.

It is easy to see that $EIAL$ gives the same prediction as $NIAL$ in the opposing interest game. Now let's look at the prediction of $EIAL$ on the game in example 1 in section 5.2.

Example 1 Revisited There are three types: type $0, \frac{1}{2}$ and 1. The common prior is such that $\pi(0) = \frac{1}{3}$; $\pi\left(\frac{1}{2}\right) = \frac{4}{9}$ and $\pi(1) = \frac{2}{9}$. Both the Sender and the Receiver have quadratic loss function: $u^R(t,a) = -(t-a)^2$ and $u^S(t,a) = -\left(t + \frac{1}{2} - a\right)^2$.

Recall that the unique equilibrium in the game without language is babbling, while $NIAL$ selects a unique informative strategy $\left(s^S_{nice}, s^R_{nice}\right)$. Table 5.1 shows all the Receiver strategies in language. The first round of deletion is the same as in the normal form procedure: for the Receiver, every strategy in language except $(0,0,0)$ and $(1,1,1)$ are retained; for each type of the Sender, only $s^S_{nice}(t)$ is retained. The second round of the extensive form procedure is different from that of $NIAL$. Suppose s^R survives the second round of deletion. Then it is necessary for s^R to be weakly interim rational with respect to s^S_{nice}, which is the only conjecture the Receiver can hold at the second round. It is then necessary that s^R takes the same action at both message $\frac{1}{2}$ and message 1. Suppose to the contrary that $s^R\left(\frac{1}{2}\right) \neq s^R(1)$, then by the assumption of language, $s^R(1) = 1$. From the first two rows in table 5.1, we can see that both type $\frac{1}{2}$ and type 1 senders send only message 1. Since $s^R(1) \neq s^R\left(\frac{1}{2}\right)$, action 1 is associated with a posterior belief that puts probability $\frac{2}{3}$ on type $\frac{1}{2}$ and probability $\frac{1}{3}$ on type 1. The best action given this distribution is action $\frac{1}{2}$, not action 1. So s^R is not weakly interim rational with respect to s^S_{nice}. We've then shown that to be weakly interim rational with respect to s^S_{nice}, it is necessary to take the same action at both message $\frac{1}{2}$ and message 1. Then every type receives the same action since all types of the Sender send either message $\frac{1}{2}$ or message 1. Thus $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ and $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ are both weakly sequentially rational with respect to s^S_{nice}. The unique outcome predicted by $EIAL$ is that all types of the Sender receive action $\frac{1}{2}$, which is the same as the babbling outcome.

Table 6.1.3 summarizes the predictions of the game in example 6.1 under different procedures.

We now establish nonemptiness of the limit. We need one notation here. Define $\varepsilon\left(\sigma^S, \sigma^{S\prime}\right) : T \to \Delta M$ by $\varepsilon\left(\sigma^S, \sigma^{S\prime}\right)(t) \equiv (1-\varepsilon)\,\sigma^S(t) + \varepsilon\sigma^{S\prime}(t)$.

	Language	**No Language**
EQ	s^R_{babble}, s^R_{nice}	s^R_{babble}
WSEQ	s^R_{babble}	s^R_{babble}
IA	s^R_{nice}	everything
IA+Standard Sequential Rationality	empty	everything
Weak IA+Weak Sequential Rationality	s^R_{babble}	everything

Lemma 6. *$ES(\infty)$ is nonempty.*

Proof Since $S_L = \left(S^S, S^R_L\right)$ is finite, the elimination process must stop after finite steps. It suffices to show that $ES(k+1)$ is nonempty if $ES(k)$ is nonempty. It is obvious from the iterative procedure that $ES^S(k+1)$ is nonempty, since S_L is finite and for every $\sigma^R \in \Delta^+ ES^R(k)$, there exists $s^S \in ES^S(k)$ which attains the maximum payoff among Sender strategies in $ES^S(k)$. To show that $ES^R(k+1)$ is nonempty, it suffices to show that there exists a totally mixed conjecture on $ES^S(k)$, i.e. $\sigma^S \in \Delta^+ ES^S(k)$, such that there exists s^R in $ES^R(k)$ which is weakly interim rational with respect to σ^S. It is obvious that for every $\sigma^S \in \Delta^+ ES^S(0) = \Delta^+ S^S$, there exists $s^R \in ES^R(0) = S^R$ which is weakly interim rational with respect to σ^S, since a constant s^R which plays the best action against the prior is weakly interim rational with respect to σ^S. Let us pick any $\sigma^S_{k-1} \in \Delta^+ ES^S(k-1)$. Since $ES^S(k) \subset ES^S(k-1)$, $\varepsilon\left(\sigma^S_k, \sigma^S_{k-1}\right) \in \Delta^+ ES^S(k-1)$ for any σ^S_k in $ES^S(k)$. To shorten the notation, write $\varepsilon\left(\sigma^S_k, \sigma^S_{k-1}\right)$ simply as $\sigma^S_{k,\varepsilon}$. Lemma 7 implies that for every ε small enough,

$$\begin{aligned} & \left\{s^R \in ES^R(k-1) : s^R \text{ is weakly interim rational with respect to } \sigma^S_{k,\varepsilon}\right\} \\ \subset\ & \left\{s^R \in ES^R(k) : s^R \text{ is weakly interim rational with respect to } \sigma^S_k\right\}. \end{aligned}$$

By hypothesis, the set

$$\left\{s^R \in ES^R(k-1) : s^R \text{ is weakly interim rational with respect to } \sigma^S_{k,\varepsilon}\right\}$$

is nonempty. Since $ES^R(k-1)$ is finite, there exists a Receiver strategy $s^R \in ES^R(k-1)$ which attains the maximum expected utility among all strategies that are weakly interim rational with respect to $\sigma^S_{k,\varepsilon}$. Therefore, the set

$$\left\{s^R \in ES^R(k) : s^R \text{ is weakly sequentially rational with respect to } \sigma^S_k\right\}$$

is nonempty. The proof is then completed by induction.

Lemma 7. *Given any $\sigma^S \in \Delta S^S$, there exists $\bar{\varepsilon} > 0$ such that for all $\sigma^{S\prime} \in \Delta S^S$ and $\varepsilon < \bar{\varepsilon}$,*

$$\begin{aligned} & \left\{ s_R \in S_R | s_R \text{ is weakly sequentially rational with respect to } \varepsilon\left(\sigma^S, \sigma^{S\prime}\right) \right\} \\ \subset \; & \left\{ s_R \in S_R | s_R \text{ is weakly sequentially rational with respect to } \sigma^S \right\}. \end{aligned}$$

The proof is left to the Appendix.

6.2. Characterization

The extensive form procedure ($EIAL$) is motivated by the example illustrating that $NIAL$ might not select any equilibrium outcome of the original game. We showed that $EIAL$ restores babbling as the unique prediction in that example. In this section, we show that this result is general, i.e., $EIAL$ selects babbling as the unique outcome when babbling is the unique equilibrium in the original game. However, we are able to show that $EIAL$ contains at least one equilibrium outcome in the original game only under monotonicity condition (M) (defined in section 4) and with the interim interpretation. On the other hand, with $EIAL$, we are able to show the lower bound on the amount of information transmission only under ex ante interpretation. We do not have a tight characterization when the monotonicity condition (M) is satisfied. Showing inclusion of strategies under interim representation is easier, while showing exclusion of strategies under ex ante representation is easier. Therefore, our current results under $EIAL$ depend on whether interim representation or ex ante representation is employed.

Proposition 5. *If babbling is the unique equilibrium, then babbling is the only outcome under $EIAL$.*

Proof (Idea) Say that $X \subset S$ contains an informative outcome if there exists $\left(s^S, s^R\right) \in X$ such that there are two different types $t_1 \neq t_2$ where $s^R\left(s^S\left(t_1\right)\right) \neq s^R\left(s^S\left(t_2\right)\right)$. We show that if $ES\left(k\right)$ contains an informative outcome, then the iterative process does not stop, i.e., $ES\left(k+1\right) \subsetneq ES\left(k\right)$. Since we have shown that the limiting set is nonempty, it is necessary that $ES\left(\infty\right)$ contains no informative outcomes. Thus $EIAL$ predicts that every type receives the same action. Since the strictly best constant strategy is to play the best action against the prior, we get the babbling outcome.

To see why the iterative process does not stop when $ES(k)$ contains an informative outcome, note that for babbling to be the unique equilibrium in the original game, it has to be the case that every type t prefers to be pooled with all higher types than with all lower types. Whenever it is not the case that all messages induce the same action, some type t will want to discard a message which always receives the lowest action.

The details of the proof is left to the Appendix.

Proposition 6. *If condition (M) holds, then the most informative equilibrium outcome is contained in $EIAL$ under the interim interpretation.*

We prove it by showing that if every type exaggerates the most they want and sends the highest message they might use in $ES^S(\infty)$, the best the Receiver can do without violating either language or weak interim rationality is to play the most informative equilibrium strategy. The details are left in the Appendix.

When ex ante interpretation is employed, however, we need to make sure that there exists one single Receiver strategy with respect to which every type of the Sender wants to exaggerate the most. Therefore, the proof of proposition 6 does not carry through directly. Under the interim interpretation, different types are allowed to hold different beliefs about the behavior of the Receiver. Therefore, it is easier to construct a sender strategy profile in the limiting set, and therefore easier to show that a Receiver strategy belongs to the limiting set.

Now we state the result of the lower bound on the amount of information transmission. Proposition 7 says that every Receiver strategy in the limit partitions the set of messages used in the limit into at least L intervals.

Proposition 7. *With ex ante interpretation, under $EIAL$, there exists a non-trivial lower bound on the number of different actions taken on $ES^R(\infty)$. Specifically, if the game admits a non-babbling equilibrium, then the number of different actions taken in $ES^R(\infty)$ is at least 2.*

Proof A Receiver strategy s^R partitions the message space. If s^R is weakly interim rational with respect to σ^S, then a partition determines s^R. A finer partition is unambiguously better. However, a finer partition might violate the language restriction. We show that a Receiver strategy cannot have a step that is too wide, because otherwise there exists a finer partition that satisfies language and is weakly interim rational. The same logic is used to attain a lower bound.

We need the following two observations to proceed with the proof. The first claim relates the minimum action the Receiver might take at message $m-\Delta$ to the value m. It holds only under ex ante interpretation. The second claim gives a relation between the lowest type that might send messages in $[m,1]$ to the value m.

Claim $s^R(m-\Delta) \geq E([0,\rho(m)])$ for all $s^R \in ES^R(1)$ such that $s^R(m) \neq s^R(m-\Delta)$.

Claim Type $g^{-1}(\infty; m)$ prefers action $a = m$ to action $E([0, g^{-1}(\infty; m) - \Delta])$.

In particular, $g^{-1}(\infty; E([t_1^2, 1])) \geq t_1^2$. Given any $\sigma^S \in \Delta S^S$, define

$$s^R_{sep2}(m) \equiv \begin{cases} \arg\max_a U^R|_{[0,E([t_1^2,1])-\Delta]}\left(\sigma^S, a\right) & m \in [0, E([t_1^2,1]) - \Delta] \\ \arg\max_a U^R|_{[E([t_1^2,1]),1]}\left(\sigma^S, a\right) & m \in [E([t_1^2,1]), 1] \end{cases}$$

It is obvious that s^R_{sep2} is weakly interim rational with respect to σ^S. To show that s^R_{sep2} belongs to language, we need to show that

$$\arg\max_a U^R|_{[0,E([t_1^2,1])-\Delta]}\left(\sigma^S, a\right) \leq E\left(\left[t_1^2, 1\right]\right) - \Delta \tag{6.1}$$

$$\arg\max_a U^R|_{[E([t_1^2,1]),1]}\left(\sigma^S, a\right) \geq E\left(\left[t_1^2, 1\right]\right). \tag{6.2}$$

Ex ante interpretation implies that every pure Sender strategy $s^S \in ES^S(1)$ is weakly increasing in t. Therefore, $\arg\max_a U^R|_{[0,m]}\left(s^S, a\right) \leq E([0,1])$ for every m and every $s^S \in ES^S(1)$. It follows that $\arg\max_a U^R|_{[0,m]}\left(\sigma^S, a\right) \leq E([0,1])$ for every m and every $\sigma^S \in \Delta ES^S(1)$. Thus,

$$\begin{aligned} \arg\max_a U^R|_{[0,E([t_1^2,1])-\Delta]}\left(\sigma^S, a\right) &\leq E([0,1]) \\ &< E\left(\left[t_1^2, 1\right]\right). \end{aligned}$$

This gives us inequality 6.1. We showed that the smallest type that can send any message higher than or equal to message $E([t_1^2, 1])$ is greater than $t_1^2 - \Delta$. Therefore, $E([g^{-1}(\infty; E([t_1^2,1])), 1]) \geq E([t_1^2, 1])$. Inequality 6.2 then follows. So a constant Receiver strategy cannot be weakly sequentially rational with respect to σ^S, because s^R_{sep2} is weakly interim rational with respect to σ^S and gives a higher ex ante payoff than a constant Receiver strategy.

Therefore, s^R must partition $M(\infty)$ into at least two intervals. Let $\{a_1, ..., a_q\}$ be the set of actions taken by s^R on $M(\infty)$, where $a_j < a_{j+1}$. Let m_j be the smallest message on which s^R takes the value a_j. Let

$$\hat{m}_q \equiv \max\left\{m \in M | E\left(\left[g^{-1}(\infty; m), 1\right]\right) \leq m\right\}.$$

It follows that $\hat{m}_q \geq E([t_1^2, 1])$.

Claim $a_q \geq \hat{m}_q$ or a_q is such that

$$g(\infty; \rho(a_k)) \geq E\left(\left[\rho(a_k), g^{-1}(\infty; \hat{m}_q) - \Delta\right]\right). \tag{6.3}$$

To show this, define

$$s^R_{sep2}(m) \equiv \begin{cases} \arg\max_a U^R|_{\left[m_q, E\left(\left[t_1^2, 1\right]\right) - \Delta\right]}\left(\sigma^S, a\right) & m \in [m_q, \hat{m}_q - \Delta] \\ \arg\max_a U^R|_{\left[E\left(\left[t_1^2, 1\right]\right), 1\right]}\left(\sigma^S, a\right) & . \quad m \in [\hat{m}_q, 1] \\ s^R(m) & \text{otherwise} \end{cases}.$$

For s^R to be weakly interim rational with respect to σ^S, it has to be the case that $g\left(\infty; \rho_{[0,1]}(a_q)\right) \geq m_q$. If $g(\infty; \rho(a_k)) < E([\rho(a_k), g^{-1}(\infty; \hat{m}_q) - \Delta])$, then

$$\begin{aligned} m_q &\leq g(\infty; \rho(a_k)) \\ &< E\left(\left[\rho(a_k), g^{-1}(\infty; \hat{m}_q) - \Delta\right]\right) \\ &\leq \arg\max_a U^R|_{\left[m_q, E\left(\left[t_1^2, 1\right]\right) - \Delta\right]}\left(\sigma^S, a\right). \end{aligned}$$

Therefore s^R_{sep2} satisfies language. Since s^R_{sep2} is weakly interim rational w.r.t. σ^S by construction, we have thus reached a contradiction.

The above claim gives a lower bound on a_q. This in turn gives a lower bound on $\arg\max_a U^R|_{[0, m_q - \Delta]}\left(\sigma^S, a\right)$. Look at σ^S restricted on the interval $[0, m_q - \Delta]$. We can then apply the same argument and get a lower bound on a_{q-1}.

Define $\psi(\tau_1)$ to be the longest forward solution with an initial condition τ_1. That is, $\psi(\tau_1) \equiv \{0, \tau_1, \psi_2, \psi_3, ..., \psi_n\}$ is a forward solution on $[0, \psi_n]$ where $\psi_n \leq 1$, and there does not exist a forward solution $\{0, \tau_1, \psi'_2, ..., \psi'_{n'}\}$ where $\psi'_{n'} \leq 1$ and $n' > n$. Define $\lambda(\tau_1) \equiv n$ where n is the size of the forward solution $\psi(\tau_1)$. A necessary condition is that either $\lambda(\rho(a_k)) = 1$

or $\psi_2(\rho(a_q)) \geq t_1^2([0,1])$. So when there is a size-2 forward solution on $[0, t_1^2([0,1])]$, $q-1 \geq 2$. A lower bound of the lower bound on a_{q-1} can be interpreted this way by restricting types to the subset $[0, \rho(a_k) - \Delta]$. So if there is a size-2 forward solution on $[0, t_1^2([0, t_1^2([0,1])])]$, then $q-2 \geq 2$. Define $f_1 \equiv t_1^2([0,1])$, and $f_{j+1} \equiv t_1^2([0, f_j])$ whenever $[0, f_j]$ has a size-2 forward solution. The process ends when we reach $[0, f_l]$ where there is no size-2 forward solutions on it.

7. Conclusion

This paper is an exercise to demonstrate the power of incorporating the asymmetry implied by language, when language is regarded as one way to transmit a given amount of information. Taking a literal approach, we model common knowledge of language by directly restricting players' strategy sets without a priori ruling out any outcome. We then characterize the solution to this new game under iterative admissibility. Applying the general framework to sender-receiver games a la Crawford and Sobel (1982), we assume that strategies satisfy "language" if and only if they satisfy the literal meaning condition and the convexity condition. Using normal form iterative admissibility, under a regularity condition, we show that all outcomes are at least as informative (in terms of number of distinct actions possibly received by the sender) as the most informative equilibrium.

However, we illustrate through an example that this procedure may eliminate even the most informative equilibrium, and we point out the tension among language, iterative admissibility and sequential rationality. These conflicts arise because modeling language through physically restricting a player's strategy set gives language the highest priority. Therefore, language always overrides interim optimality, although normal form iterative admissibility takes care of ex ante optimality. We show that simply adding standard sequential rationality does not resolve these conflicts. In order to capture and reconcile the ideas of language, iterative weak dominance and sequential rationality, we develop a weaker version of sequential rationality which seems more suited to cheap talk games, and we propose an extensive form procedure. However, we do not have a complete characterization of the predictions under our extensive form iterative procedure, and there exist alternative methods for analyzing language while giving rationality a comparable weight. A worthwhile future investigation might involve capturing the idea of language through restricting beliefs, instead of strategy sets, and finding a way to analyze the interaction between beliefs about language and beliefs about rationality.

In sum, we point out the asymmetry among messages as the driving force for language to advance coordination and efficiency. We propose a framework which provides a systematic and structural approach to understanding the implications of different properties of language. We leave open the questions of how to incorporate language with rationality while giving the two concepts comparable weights and the epistemic story of common knowledge of language.

8. Appendix

8.1. Proofs for Section 3

Proof for Lemma 2 We first establish the equivalence between strong dominance and best response. Lemma 1 then follows using the same method in the proof of lemma 4 in the Appendix of Pearce (1984). For the completeness of the argument, we restate the proof below.

Suppose that $\hat{s}^R$ is not a best response to any $\sigma^S \in \Pi_t \left(\Delta^+ X^S(t)\right)$. Define

$$A = \left\{\sigma^R \in \Delta X^R : U^R\left(\sigma^S, \sigma^R\right) = U^R\left(\sigma^S, \hat{s}^R\right) \forall \sigma^S \in \Pi_t \left(\Delta X^S(t)\right)\right\}.$$

Let k_t be the number of pure strategies in $X^S(t)$ and let $k = \Pi_{t\in T} k_t$, and κ be the open interval $\left(0, \frac{1}{k}\right)$. Define

$$\begin{aligned}
\delta_\varepsilon^t &= \left\{\sigma^S \in \Delta X^S(t) : \sigma_i^S \geq \varepsilon \forall i = 1, 2, ..., k_t\right\}, \\
\delta_\varepsilon &= \Pi_t \kappa_\varepsilon^t \\
B_\varepsilon &= \left\{\sigma^R \in \Delta X^R : U^R\left(\sigma^R, \sigma^S\right) > U^R\left(\hat{s}^R, \sigma^S\right) \forall \sigma^S \in \kappa_\varepsilon\right\}, \\
W_\varepsilon &= \left\{\sigma^R \in \Delta X^R : U^R\left(\sigma^R, \sigma^S\right) \geq U^R\left(\hat{s}^R, \sigma^S\right) \forall \sigma^S \in \kappa_\varepsilon\right\}.
\end{aligned}$$

$\hat{s}^R$ is not a best response to any $\sigma^S \in \Pi_t \left(\Delta^+ X^S(t)\right)$, so for each $\varepsilon \in \kappa$, $\hat{s}^R$ is not a best response to any $\sigma^S \in \delta_\varepsilon$. If we view δ_ε^t as set of strategies for type t Sender, then the equivalence between strong dominance and never best response establishes that B_ε is nonempty. Since W_ε is closed and nonempty, for each $\varepsilon \in \kappa$ we can choose $s_\varepsilon^R \in \Delta X^R$ that is a best response in W_ε to $\sigma_\varepsilon^S \in \delta_\varepsilon$, where $\sigma_\varepsilon^S(t)$ puts probability $\frac{1}{k_t}$ on every pure strategy in $X^S(t)$. Notice that s_ε^R yields player R strictly higher utility against σ_ε^S than $\hat{s}^R$, since $B_\varepsilon \sqsubseteq W_\varepsilon$. Choose a sequence of $\varepsilon_i' s$ in T converging to 0, such that $\left\{\sigma_{\varepsilon_i}^R\right\}$ converges. Let σ_*^R be the limit of the sequence $\left\{\sigma_{\varepsilon_i}^R\right\}$. We will show that σ_*^R weakly dominates $\hat{s}^R$.

Continuity of U^R guarantees that σ_*^R is at least as good for player R as $\hat{s}^R$ against all $\sigma^S \in \Pi\left(\Delta X^S(t)\right)$. It remains only to show that $\sigma_*^R \notin A$. If $\exists \sigma'^R \in A$ with $\sigma'^R = \sigma_*^R$, then for all sufficiently small ε_i, $\sigma_{\varepsilon_i}^R$ gives positive weight to very pure strategy given positive weight by σ'^R. Then $\lambda > 0$ can be chosen sufficiently small so that all components of

$$\bar{\sigma}_{\varepsilon_i}^R = \left(\sigma_{\varepsilon_i} - \lambda \sigma'^R\right) \frac{1}{1-\lambda}$$

are nonnegative. For any $\sigma^R \in \delta_{\varepsilon_i}$,

$$U^R\left(\sigma^S, \bar{\sigma}_{\varepsilon_i}^R\right) - U^R\left(\sigma^S, \sigma_{\varepsilon_i}^R\right) = \frac{\lambda}{1-\lambda}\left[U^R\left(\sigma_{\varepsilon_i}^R, \sigma^S\right) - U^R\left(\hat{s}^R, \sigma^S\right)\right] \geq 0$$

because $\sigma_{\varepsilon_i}^R \in W_{\varepsilon_i}$. Moreover, the inequality is strict when σ^S is such that, for every t, $\sigma^S(t)$ puts probability $\frac{1}{k_t}$ on every pure strategy in $X^S(t)$ (denote it by $\tilde{\sigma}^S$). Thus $\bar{\sigma}_{\varepsilon_i}^R$ is in W_{ε_i} and yields player R higher utility than $\sigma_{\varepsilon_i}^R$ against $\tilde{\sigma}^S$, a contradiction.

Claim s^R is strongly dominated w.r.t. $\left(\Pi_t X^S(t)\right) \times X^R$ if and only if there does not exist a $\sigma^S(t) \in \Delta X^S(t)$ for every t such that $s^R \in \arg\max_{s' \in X^R} U^R\left(\left(\sigma^S(t)\right)_t, s'\right)$.

Proof The "only-if" part is trivial. To show the "if" part, suppose to the contrary that $\hat{s}^R$ is not a best response to any $\sigma^S \in \Pi_t\left(\Delta X^S(t)\right)$. Then there exists a function $b : \Pi_t\left(\Delta X^S(t)\right) \to X^R$ with $U^R\left(\sigma^S, b\left(\sigma^S\right)\right) > U^R\left(\sigma^S, \hat{s}^R\right) \forall \sigma^S$. Consider the zero-sum game

$$\bar{G} = \left(T, \Pi_t X^S(t), X^R, \bar{U}^S(;t), \bar{U}^R\right)$$

where $\bar{U}^R\left(s^S, s^R\right) = U^R\left(s^S, s^R\right) - U^R\left(s^S, \hat{s}^R\right)$ and $\bar{U}^S\left(s^S, s^R; t\right) = -U^R\left(s^S, s^R\right) \forall t$. Let $\left(\Pi_t \sigma_*^S(t), \sigma_*^R\right)$ be a Bayesian Nash equilibrium of $\bar{G}$. Since the interim interpretation results in the same equilibria as the ex ante interpretation,

$$\bar{U}^S\left(\Pi_t \sigma_*^S(t), \sigma_*^R\right) \geq \bar{U}^S\left(\Pi_t \sigma^S(t), \sigma_*^R\right)$$

for any $\Pi_t \sigma^S(t) \neq \Pi_t \sigma_*^S(t)$. For any $\sigma^S \in \Delta X^S$,

$$\begin{aligned} \bar{U}^R\left(\Pi_t \sigma^S(t), \sigma_*^R\right) &\geq \bar{U}^R\left(\Pi_t \sigma_*^S(t), \sigma_*^R\right) \\ &\geq \bar{U}^R\left(\Pi_t \sigma_*^S(t), b\left(\Pi_t \sigma_*^S(t)\right)\right) \\ = U^R\left(\Pi_t \sigma_*^S(t), b\left(\Pi_t \sigma_*^S(t)\right)\right) &- U^R\left(\Pi_t \sigma_*^S(t), \hat{s}^R\right) \\ &> 0 \end{aligned}$$

So

$$U^R\left(\Pi_t\sigma^S(t),\sigma^R_*\right) > U^R\left(\Pi_t\sigma^S(t),\hat{s}^R\right)\ \forall\Pi_t\sigma^S(t)\in\Pi_t\Delta X^S(t)$$

So $\hat{s}^R$ is strongly dominated by σ^R_*.

8.2. NIAL Results under the Interim Interpretation

To show lemma 5 strictly under the interim interpretation, we need the following lemmas.

Lemma 8. [3]*If $y^S(t)\geq t+b\forall t$, then $S(k)$ satisfies the following properties for all k:*

1. *Given any messages m_0, m_1 where $m_0 < m_1 \leq y^S(t)$ and $m_0\in S^S(k;t)$, then m_1 belongs to $S^S(k;t)$.*
2. *If $l(k;t) < y^S(t)$, then $[l(k;t), y_S(t)]\in S^S(k;t)$;*
3. *$\forall m_0 < m_1$ such that $m_1\in M(k)$ and $\exists s^R\in S^R(k)$ such that $s^R(m_0)\neq s^R(m_1)$, there exists $\hat{s}^R\in S^R(k)$ such that $\hat{s}^R(m_0) < \hat{s}^R(m_1)\leq m_1$.*
4. *For all message $\hat{m}\in M(k)$ such that there exists a strategy $s^R_1\in S^R(k)$ where $s^R_1(\hat{m}-\Delta) < \hat{m} < s^R_1(\hat{m})$, there exists another strategy $s^R_2\in S^R(k)$ where $s^R_2(\hat{m}+\Delta) = s^R_1(\hat{m}+\Delta) > s^R_2(\hat{m}) \geq s^R_1(\hat{m}-\Delta)$.*

Proof. Prove by induction. It is obvious that properties 1 through 4 hold for $k=0$. Suppose they hold for $j=0,...,k$. Property 2 is a re-phrasing of property 1. Property 1 follows from property 4. To show property 4 holds for $j=k+1$, suppose $\hat{s}^R\in S^R(k+1)$ is such that $\hat{s}^R(\hat{m}-\Delta)\neq\hat{s}^R(\hat{m})$. If $l(k;t) > \hat{m}$ for all $t>\hat{m}$, then message $\hat{m}$ can only come from types smaller or equal to $\hat{m}$. $\hat{m}\in M(k)$, so it can be shown that s^R such that $s^R(\hat{m}) > \hat{m}$ cannot be a best response to any $\sigma^S\in\Pi_{t\in T}\left(\Delta^+S^S(k;t)\right)$. So if $\hat{s}^R(\hat{m})\neq\hat{s}^R(\hat{m}-\Delta)$ and $\hat{s}^R\in S^R(k+1)$, then $\hat{s}^R(\hat{m})=\hat{m}$. Property 3 is thus shown to hold and property 4 holds automatically since there does not exist $s^R\in S^R(k)$ such that $s^R(\hat{m}) > \hat{m}$. Now discuss the case where $l(k;t)\leq\hat{m}$ for some $t>\hat{m}$. According to the procedure, there exists $\hat{\sigma}^S\in\Pi_{t\in T}\left(\Delta^+S^S(k;t)\right)$ to which $\hat{s}^R$ is a best

[3] I have a proof for property 3 for ex ante representation. The proof for property 4 relies on interim representation. Do not know whether it holds under ex ante representation.

response. Recall that $\sigma^S(.;t) \in \Delta M$. For type $t > \hat{m}$, construct $\hat{\sigma}_2^S(.;t)$ to be such that the weight on message $\hat{m}$ is moved to message $\hat{m}+\Delta$, i.e.

$$\begin{aligned} \hat{\sigma}_2^S(\hat{m}+\Delta;t) &= \hat{\sigma}^S(\hat{m};t) + \hat{\sigma}^S(\hat{m}+\Delta;t) \\ \hat{\sigma}_2^S(\hat{m};t) &= 0 \\ \hat{\sigma}_2^S(m;t) &= \hat{\sigma}^S(m;t) \ \forall m < \hat{m} \end{aligned}$$

Define $\hat{\sigma}_2^S(.;t) = \hat{\sigma}^S(.;t)$ for every type $t \leq \hat{m}$. Message $\hat{m}+\Delta$ belongs to $S^S(k;t)$ $\forall t > \hat{m}$ if $\hat{\sigma}^S(\hat{m};t) > 0$ because $y^S(t) > t + b \geq \hat{m} + b \geq \hat{m} + \Delta$ and property 2 implies that $\left[\hat{m}, y^S(t)\right] \subset S^S(k;t)$ if $\hat{m} \in S^S(k;t)$. Therefore $\hat{\sigma}_2^S(.;t) \in \Delta S^S(k;t)$ $\forall t$. Define $\hat{\sigma}_\alpha^S(.;t) \equiv (\alpha)\hat{\sigma}_2^S(.;t) + (1-\alpha)\hat{\sigma}^S(.;t)$. Because $\hat{s}^R$ is a best response to $\hat{\sigma}^S$ and $\hat{s}^R(\hat{m}) > \hat{m}$, it must be the case that $\arg\max_a U^R|_{\{\hat{m}\}}\left(\hat{\sigma}^S, a\right) > \hat{m}$. Since types that send message $\hat{m}$ under $\hat{\sigma}_2^S$ must be smaller or equal to $\hat{m}$, $\arg\max_a U^R|_{\{\hat{m}\}}\left(\hat{\sigma}_2^S, a\right) \leq \hat{m}$. So there exists $\hat{\alpha} \in (0,1)$ such that $\arg\max_a U^R|_{\{\hat{m}\}}\left(\hat{\sigma}_{\hat{\alpha}}^S, a\right) = \hat{m}$. This comes from the condition that $\frac{\partial^2 u}{\partial a^2} < 0$ and can be shown by mean value theorem.

Let $s_{\hat{\alpha}}^R$ be a best response to $\hat{\sigma}_{\hat{\alpha}}^S$. Since $\hat{\sigma}_{\hat{\alpha}}^S(t) \in \Delta^+ S^S(k;t)$ $\forall t$, it follows that $s_{\hat{\alpha}}^R \in S^R(k+1)$. It remains to show that $s_{\hat{\alpha}}^R(\hat{m}) = \hat{m}$ and $s_{\hat{\alpha}}^R(\hat{m}+\Delta) \geq \hat{s}^R(\hat{m}+\Delta)$ for property 4 to hold for $j = k+1$. If $s_{\hat{\alpha}}^R(\hat{m}-\Delta) \leq \hat{m}-\Delta$, then property 3 is shown to hold for $j = k+1$. Otherwise, define $\hat{\sigma}_3^S$ to be such that all types smaller than $\hat{m}$ send messages smaller than $\hat{m}$ and all types geater than or equal to $\hat{m}$ send messages greater than or equal to $\hat{m}$. Then if $\tilde{s}^R$ is a best response to σ^S close to $\hat{\sigma}_3^S$, it must be the case that $\tilde{s}^R(\hat{m}-\Delta) \leq \hat{m}-\Delta$ and $\tilde{s}^R(\hat{m}) \geq \hat{m}$. If $\tilde{s}^R(\hat{m}) = \hat{m}$ then property 3 is shown to hold. Otherwise, $\tilde{s}^R(\hat{m}) > \hat{m} > \tilde{s}^R(\hat{m}-\Delta)$ and we can apply the technique for $\hat{s}^R$ again to show that there exists $\tilde{s}_{\tilde{\alpha}}^R \in S^R(k+1)$ such that $\tilde{s}_{\tilde{\alpha}}^R(\hat{m}) = \hat{m}$ and $\tilde{s}_{\tilde{\alpha}}^R(\hat{m}-\Delta) \leq \hat{m}-\Delta$.

Now show that $s_{\hat{\alpha}}^R(\hat{m}) = \hat{m}$ and $s_{\hat{\alpha}}^R(\hat{m}+\Delta) \geq \hat{s}^R(\hat{m}+\Delta)$. If s^R is such that $s^R(\hat{m}) = s^R(\hat{m}+\Delta)$, then $U^R\left(\hat{\sigma}_2^S, s^R\right) = U^R\left(\hat{\sigma}^S, s^R\right)$ and therefore $U^R\left(\hat{\sigma}_{\hat{\alpha}}^S, s^R\right) = U^R\left(\hat{\sigma}^S, s^R\right)$. Construct a strategy $\phi_-\left(\hat{s}^R, \hat{m}\right)$ which is equal to $\hat{s}^R$ except on $\hat{m}$ and $\phi_-\left(\hat{s}^R, \hat{m}\right)(\hat{m}) = \hat{m}$. The strategy $\phi_-\left(\hat{s}^R, \hat{m}\right) \in S_L^R$ because $\hat{s}^R(\hat{m}) > \hat{m} > \hat{s}^R(\hat{m}-\Delta)$ by construction. It's easy to show that

$$U^R\left(\hat{\sigma}_{\hat{\alpha}}^S, \phi_-\left(\hat{s}^R, \hat{m}\right)\right) > U^R\left(\hat{\sigma}_{\hat{\alpha}}^S, \hat{s}^R\right)$$

because $\arg\max_a U^R|_{\{\hat{m}\}}\left(\hat{\sigma}^S_{\hat{\alpha}}, a\right) = \hat{m}$. From the construction that $\hat{s}^R$ is a best response to $\hat{\sigma}^S_{\hat{\alpha}}$,

$$\begin{aligned} U^R\left(\hat{\sigma}^S_{\hat{\alpha}}, \phi_-\left(\hat{s}^R, \hat{m}\right)\right) &> U^R\left(\hat{\sigma}^S_{\hat{\alpha}}, \hat{s}^R\right) \\ &= U^R\left(\hat{\sigma}^S, \hat{s}^R\right) \\ &\geq U^R\left(\hat{\sigma}^S, s^R\right) \end{aligned}$$

for all s^R such that $s^R(\hat{m}) = s^R(\hat{m}+\Delta)$. Therefore, being a best response to $\hat{\sigma}^S_{\hat{\alpha}}$ by construction, $s^R_{\hat{\alpha}}(\hat{m}) \neq s^R_{\hat{\alpha}}(\hat{m}+\Delta)$ and hence $s^R_{\hat{\alpha}}(\hat{m}) \leq \hat{m}$. A similar argument can be used to show that $s^R_{\hat{\alpha}}(\hat{m}) = \hat{m}$ because $\arg\max_a U^R|_{\{\hat{m}\}}\left(\hat{\sigma}^S_{\hat{\alpha}}, a\right) = \hat{m}$. Now construct a strategy $\phi\left(s^R_{\hat{\alpha}}, \hat{m}, \hat{s}^R\right)$ which is equal to $s^R_{\hat{\alpha}}$ for $m \leq \hat{m}$ and is equal to $\hat{s}^R$ for $m > \hat{m}$. This new strategy $\phi\left(s^R_{\hat{\alpha}}, \hat{m}, \hat{s}^R\right)$ belongs to the language S^R_L because $s^R_{\hat{\alpha}}(\hat{m}) \leq \hat{m} < \hat{s}^R(\hat{m}+\Delta)$. For $s^R_{\hat{\alpha}}$ to be a best response to $\hat{\sigma}^S_{\hat{\alpha}}$, it has to be the case that $U^R\left(\hat{\sigma}^S_{\hat{\alpha}}, s^R_{\hat{\alpha}}\right) \geq U^R\left(\hat{\sigma}^S_{\hat{\alpha}}, \phi\left(s^R_{\hat{\alpha}}, \hat{m}, \hat{s}^R\right)\right)$. Let $\phi_+\left(s^R_{\hat{\alpha}}, \hat{m}\right)$ be a strategy which is equal to $s^R_{\hat{\alpha}}$ except on message $\hat{m}$ and $\phi_+\left(s^R_{\hat{\alpha}}, \hat{m}\right)(\hat{m}) = s^R_{\hat{\alpha}}(\hat{m}+\Delta)$. Since $s^R_{\hat{\alpha}}(\hat{m}+\Delta) > \hat{m}$, the new strategy $\phi_+\left(s^R_{\hat{\alpha}}, \hat{m}\right)$ belongs to S^R_L. Therefore,

$$\begin{aligned} 0 &\leq U^R\left(\hat{\sigma}^S_{\hat{\alpha}}, s^R_{\hat{\alpha}}\right) - U^R\left(\hat{\sigma}^S_{\hat{\alpha}}, \phi\left(s^R_{\hat{\alpha}}, \hat{m}, \hat{s}^R\right)\right) \\ &= U^R|_{[\hat{m}+\Delta,1]}\left(\hat{\sigma}^S_{\hat{\alpha}}, s^R_{\hat{\alpha}}\right) - U^R|_{[\hat{m}+\Delta,1]}\left(\hat{\sigma}^S_{\hat{\alpha}}, \hat{s}^R\right) \\ &= \sum_{t \geq \hat{m}+\Delta} \left(\hat{\sigma}^S(\hat{m}+\Delta; t) + \alpha \times \hat{\sigma}^S(\hat{m}; t)\right)\left[u^R\left(t, s^R_{\hat{\alpha}}(\hat{m}+\Delta)\right) - u^R\left(t, \hat{s}^R(\hat{m}+\Delta)\right)\right] \\ &\quad + \sum_{t \leq \hat{m}} \left(\hat{\sigma}^S(\hat{m}+\Delta; t)\right)\left[u^R\left(t, s^R_{\hat{\alpha}}(\hat{m}+\Delta)\right) - u^R\left(t, \hat{s}^R(\hat{m}+\Delta)\right)\right] \\ &\quad + \sum_{m \geq \hat{m}+2\Delta} \sum_t \hat{\sigma}^S(m; t)\left[u^R\left(t, s^R_{\hat{\alpha}}(m)\right) - u^R\left(t, \hat{s}^R(m)\right)\right] \\ &= \sum_t \left(\hat{\sigma}^S(\hat{m}+\Delta; t) + \hat{\sigma}^S(\hat{m}; t)\right)\left[u^R\left(t, s^R_{\hat{\alpha}}(\hat{m}+\Delta)\right) - u^R\left(t, \hat{s}^R(\hat{m}+\Delta)\right)\right] \\ &\quad + \sum_{m \geq \hat{m}+2\Delta} \sum_t \hat{\sigma}^S(m; t)\left[u^R\left(t, s^R_{\hat{\alpha}}(m)\right) - u^R\left(t, \hat{s}^R(m)\right)\right] \\ &\quad - \left\{ \begin{array}{c} \sum_{t \leq \hat{m}} \hat{\sigma}^S(\hat{m}; t)\left[u^R\left(t, s^R_{\hat{\alpha}}(\hat{m}+\Delta)\right) - u^R\left(t, \hat{s}^R(\hat{m}+\Delta)\right)\right] \\ + (1-\alpha)\sum_{t \geq \hat{m}+\Delta} \hat{\sigma}^S(\hat{m}; t)\left[u^R\left(t, s^R_{\hat{\alpha}}(\hat{m}+\Delta)\right) - u^R\left(t, \hat{s}^R(\hat{m}+\Delta)\right)\right] \end{array} \right\} \\ &= U^R\left(\hat{\sigma}^S, \phi\left(\hat{s}^R, \hat{m}-\Delta, \phi_+\left(s^R_{\hat{\alpha}}, \hat{m}\right)\right)\right) - U^R\left(\hat{\sigma}^S, \hat{s}^R\right) \\ &\quad - \left(U^R|_{\{\hat{m}\}}\left(\hat{\sigma}^S_{\hat{\alpha}}, s^R_{\hat{\alpha}}(\hat{m}+\Delta)\right) - U^R|_{\{\hat{m}\}}\left(\hat{\sigma}^S_{\hat{\alpha}}, \hat{s}^R(\hat{m}+\Delta)\right)\right) \end{aligned}$$

Thus,

$$\begin{aligned} & U^R|_{\{\hat{m}\}}\left(\hat{\sigma}_{\hat{\alpha}}^S, s_{\hat{\alpha}}^R(\hat{m}+\Delta)\right) - U^R|_{\{\hat{m}\}}\left(\hat{\sigma}_{\hat{\alpha}}^S, \hat{s}^R(\hat{m}+\Delta)\right) \\ \leq\ & U^R\left(\hat{\sigma}^S, \phi\left(\hat{s}^R, \hat{m}-\Delta, \phi_+\left(s_{\hat{\alpha}}^R, \hat{m}\right)\right)\right) - U^R\left(\hat{\sigma}^S, \hat{s}^R\right) \\ \leq\ & 0 \end{aligned}$$

because $\hat{s}^R$ is a best response to $\hat{\sigma}^S$. Since

$$\arg\max_{a\in A} U^R|_{\{\hat{m}\}}\left(\hat{\sigma}_{\hat{\alpha}}^S, a\right) = \hat{m} < \hat{s}^R(\hat{m}+\Delta)$$

and $U^R|_{\{\hat{m}\}}\left(\hat{\sigma}_{\hat{\alpha}}^S, a\right)$ as a function of a inherits the concavity from u^R, we get

$$s_{\hat{\alpha}}^R(\hat{m}+\Delta) \geq \hat{s}^R(\hat{m}+\Delta)$$

It has now been shown that $s_{\hat{\alpha}}^R(\hat{m}+\Delta) \geq \hat{s}^R(\hat{m}+\Delta)$ and $s_{\hat{\alpha}}^R(\hat{m}) = \hat{m} \geq \hat{s}^R(\hat{m}-\Delta)$. ■

Define η_{k+1} iteratively to be the largest type $\hat{t} < \eta_k$ such that $l(k;t) \leq t$. That is, define

$$\eta_{k+1} \equiv \max\{t < \eta_k | l(k;t) \leq t\}$$

. Define

$$l^{-1}(k;m) = \max\{t | l(k;t) \leq m\}$$

. Then by definition, $l^{-1}(k;\eta_k) = \eta_k$ and $l^{-1}(k;m) < m$ for all $m > \eta_k$.

Lemma 9. *There exists $s^R \in S^R(k+1)$ such that $s^R(\eta_k) \neq s^R(\eta_k - \Delta)$, for any k.*

Proof. This can be done by showing that there exists $\hat{s}^S \in S^S(k)$ such that $\hat{s}^S(\eta_k) = \eta_k$ and $\hat{s}^S(t) \leq \eta_k - \Delta$ for all $t \leq \eta_k - \Delta$. Then show that $s^R(m) \leq m$ for all $m \geq \eta_k$ given any $s^R \in C_*^R(k+1)$. Then $\hat{s}^R$ where $\hat{s}^R(\eta_k) = \eta_k$ and $\hat{s}^R(\eta_k - \Delta) \leq \eta_k - \Delta$ does strictly better w.r.t. $\hat{\sigma}^S$ close to $\hat{s}^S$ than any other s^R. So there exists $\hat{s}^R \in C_*^R(k+1)$ where $\hat{s}^R(\eta_k) = \eta_k$ and $\hat{s}^R(\eta_k - \Delta) \leq \eta_k - \Delta \neq \hat{s}^R(\eta_k)$. ■

Now lemma 5 follows.

Lemma 10. *$l(\infty;t) \geq t$ for all $t \in T$. Moreover, either $l(\infty;t) \geq y^S(t)$ or $s^R(l(\infty;t)) = s^R\left(y^S(t)\right)$ for all $s^R \in S^R(\infty)$.*

Proof. Suppose given k, there exists a type t such that $l(k;t) \leq t$. Then η_k is well defined. So for all $s^R \in S^R(k+1)$, $s^R(\eta_k) \leq \eta_k$ and there exists $\hat{s}^R \in S^R(k+1)$ where $\eta_k = \hat{s}^R(\eta_k) \neq \hat{s}^R(\eta_k - \Delta)$. So for every type t where $y_S(t) \geq \eta_k$, message $\eta_k - \Delta$ is weakly dominated by message η_k, because they prefer action η_k to any smaller action. It then follows that every message $m \leq \eta_k - \Delta$ is also weakly dominated by message η_k. So $l(k+2;t) \geq \eta_k$ for all t where $y_S(t) \geq \eta_k$. So $l(k+2;t) \geq \eta_k$ for all $t \geq \eta_k - b$. It follows that $\eta_{k+2} \leq \eta_k - b$ and thus the process does not stop at round k. So when the process stops, it has to be the case that $l(\infty;t) \geq t$ for all type t. Furthermore, either $l(\infty;t) \geq y_S(t)$ or $s^R(l(\infty;t)) = s^R\left(y^S(t)\right)$ for all $s^R \in S^R(\infty)$ because otherwise , message $l(\infty;t)$ is weakly dominated by message $y^S(t)$ since $s^R\left(y^S(t)\right) \leq y^S(t)$ for all $s^R \in S^R(\infty)$ and thus type t always prefers the action induced by message $y^S(t)$ to that induced by message $l(\infty;t)$, which contradicts the definition of $l(\infty;t)$. ■

8.3. NIAL Results under Ex Ante Interpretation

Ex ante interpretation is equivalent to assuming that different types of Sender hold the same belief about the behavior of the Receiver. Therefore, $S^S(k)$ is no longer a product space of $S^S(k;t)$, and the proof under the interim interpretation does not apply. We'll make an assumption which is satisfied when neither type space nor action space is discretized. We make use of the assumption in the proof when type space and action space are discretized.

Assumption For any $a_1 < a_2 < a_3$ where $a_j \in A$ for $j = 1,2,3$ and $a_1 \geq y^S(0)$, there exists $\hat{t} \in T$ such that $a_2 = \arg\max_{j=1,2,3} u^S\left(\hat{t}, a_j\right)$.

Proof for Lemma 5 Suppose to the contrary that there exists type $\hat{t}$ such that $l\left(\infty;\hat{t}\right) < \hat{t}$ and $l(\infty;t) \geq t$ for all $t > \hat{t}$. From the definition of b, we know that every type greater or equal to $1-b$ prefers a higher action to a lower one. Thus, $l(\infty;t) = 1$ for all $t \geq 1-b$ because any message smaller than 1 induces a weakly smaller action. So $\hat{t} < 1-b$. $l(\infty;t) \geq t$ for all $t > \hat{t}$ implies that $l^{-1}(\infty;m) \leq m$ for all $m \geq \hat{t}$. From observation 5.1, we know that $s^R(m) \leq m$ for all $s^R \in S^R(\infty)$ and $m \in M(\infty) \cap \left[\hat{t}, 1\right]$. In particular, $s^R\left(\hat{t}\right) \leq \hat{t}$ for all $s^R \in S^R(\infty)$. If there exists $\tilde{s}^R \in S^R(\infty)$ such that $\tilde{s}^R\left(\hat{t}\right) \neq \tilde{s}^R\left(\hat{t}-\Delta\right)$, then $\tilde{s}^R\left(\hat{t}\right) = \hat{t} > \tilde{s}^R\left(\hat{t}-\Delta\right)$. Since $b \geq \Delta$, $y^S\left(\hat{t}\right) \geq \hat{t}+\Delta$. Therefore, any message smaller than $\hat{t}$ is weakly dominated for type $\hat{t}$ by the message $\hat{t}$. This contradicts the assumption that $l\left(\infty;\hat{t}\right) <$

$\hat{t}$. Likewise, if there exists $m \in \left[\hat{t}, y^S\left(\hat{t}\right)\right]$ such that there exists $s^R \in S^R(\infty)$ where $s^R(m) \neq s^R\left(l\left(\infty;\hat{t}\right)\right)$, then the message $l\left(\infty;\hat{t}\right)$ is weakly dominated by the message m. Since $l\left(\infty;\hat{t}\right) < \hat{t}$, there exists $\hat{m} \geq y^S\left(\hat{t}\right)$ such that $s^R(m) = s^R\left(l\left(\infty;\hat{t}\right)\right)$ for all $m \in \left[l\left(\infty;\hat{t}\right), \hat{m}\right]$.

Moreover, $s^R\left(l\left(\infty;\hat{t}-\Delta\right)\right) = s^R\left(l\left(\infty;\hat{t}\right)\right)$ for all $s^R \in S^R(\infty)$ because type $\hat{t}-\Delta$ prefers action $\hat{t}$ to any smaller action, due to the assumption that $b \geq \Delta$. So $s^R(m) = s^R\left(l\left(\infty;\hat{t}-\Delta\right)\right)$ for all $m \in \left[l\left(\infty;\hat{t}-\Delta\right), \hat{m}\right]$.

Now we need to construct $s^S_* \in S^S(\infty)$ such that the Receiver's ex ante best response to s^S_* must be a non-constant on the interval $\left[l\left(\infty;\hat{t}-\Delta\right), \hat{m}\right]$, which would contradict the construction of $\hat{m}$ and we would be done. We can find such sender strategy in $S^S(1)$. We'll suppose we can find one in $S(k)$, and then show that we can find one in $S(k+1)$. Then by induction, we an find one in $S(\infty)$.

Definition 10. *Say that property * holds for j if there exists $m \in S^S\left(j;\hat{t}\right) \cap \left[\hat{t}, \hat{m}-\Delta\right]$ and $\sigma^R \in \Delta S^R(j-1)$ such that $u^S\left(\hat{t}, \sigma^R(m)\right) > u^S\left(\hat{t}, \sigma^R\left(\hat{t}-\Delta\right)\right)$.*

Suppose k is such that property * holds for $j = 1, 2, ..., k$. Let $m_1^k \equiv \min\left\{\begin{array}{c} m \in S^S\left(k;\hat{t}\right) \cap \left[\hat{t}, \hat{m}-\Delta\right] \mid \\ \exists \sigma^R \in \Delta S^R(k-1) \text{ s.t.} \\ u^S\left(\hat{t}, \sigma^R(m)\right) > u^S\left(\hat{t}, \sigma^R\left(\hat{t}-\Delta\right)\right) \end{array}\right\}$.

Step 1 Show that there exists $s^S_{*k} \in S^S(k)$ such that $s^S_{*k}\left(\hat{t}\right) = m_1^k$ (or a message m such that $s^R(m) = s^R\left(m_1^k\right)$ for all $s^R \in S^R(k-1)$) and $s^S_{*k}\left(\hat{t}-\Delta\right) \in \left[s^S_0\left(\hat{t}-\Delta\right), \hat{t}-\Delta\right]$. Let $S^R_{*k} \equiv \left\{s^R \in S^R(k) \mid s^R \text{ is a best response among } S^R(k-1) \text{ to } s^S_{*k}\right\}$.

Proof We want to construct $\sigma^R_{*k} \in \Delta^+ S^R(k-1)$ such that

$$\begin{aligned} m_1^k &\in \arg\max_m u^S\left(\hat{t}, \sigma(m)\right) \\ u^S\left(\hat{t}, \sigma^R_{*k}\left(m_1^k\right)\right) &> u^S\left(\hat{t}, \sigma^R_{*k}\left(\hat{t}-\Delta\right)\right) \\ u^S\left(\hat{t}-\Delta, \sigma^R_{1k}(m')\right) &> u^S\left(\hat{t}-\Delta, \sigma^R_{1k}(m)\right) \\ \text{for some } m' &\in \left[l\left(\infty;\hat{t}-\Delta\right), \hat{t}-\Delta\right] \\ \text{and all } m &\geq \hat{t} \end{aligned}$$

. From the construction of m_1^k, there exists σ_{1k}^R such that

$$\begin{aligned} m_1^k &\in \arg\max_m u^S\left(\hat{t}, \sigma_{1k}^R(m)\right) \\ u^S\left(\hat{t}, \sigma_{1k}^R\left(m_1^k\right)\right) &> u^S\left(\hat{t}, \sigma_{1k}^R\left(\hat{t}-\Delta\right)\right) \\ &\geq u^S\left(\hat{t}, \sigma_{1k}^R(m)\right) \text{ for all } m \leq \hat{t}-\Delta. \end{aligned}$$

W.l.o.g. assume $u^S\left(\hat{t}-\Delta, \sigma_{1k}^R(m)\right)$ is weakly increasing in m on the set $S^S\left(k; \hat{t}\right) \cap \left[0, m_1^k\right]$. If there exists $m' \in \left[l\left(\infty; \hat{t}-\Delta\right), \hat{t}-\Delta\right]$ such that

$$\begin{aligned} u^S\left(\hat{t}-\Delta, \sigma_{1k}^R(m')\right) &> u^S\left(\hat{t}-\Delta, \sigma_{1k}^R(m)\right) \\ \text{for all } m &\geq \hat{t} \end{aligned}$$

, then let $\sigma_{*k}^R \equiv \sigma_{1k}^R$. Otherwise, let

$$m^+ \equiv \min\left\{ \begin{array}{c} m \geq \hat{t} | \\ u^S\left(\hat{t}-\Delta, \sigma_1^k(m)\right) \geq u^S\left(\hat{t}-\Delta, \sigma_1^k\left(\hat{t}-\Delta\right)\right); \\ \exists s^R \in S^R(k-1) \text{ s.t.} \\ s^R(m) \neq s^R\left(\hat{t}-\Delta\right) \end{array} \right\}$$

if $\hat{t}-\Delta \in S^S\left(k; \hat{t}-\Delta\right)$. If not, then choose the highest message smaller than $\hat{t}-\Delta$ that belongs to $S^S\left(k; \hat{t}-\Delta\right)$. Then $m^+ \in \left[\hat{t}, m_1^k\right]$ and $s^R(m) = s^R\left(\hat{t}-\Delta\right)$ for all $m \in \left[\hat{t}, m^+ - \Delta\right]$. By construction of $\hat{s}_0^S$, $\hat{s}_0^S\left(\hat{t}-\Delta\right) \in M(\infty) \cap \left[l\left(\infty; \hat{t}-\Delta\right), \hat{t}-\Delta\right]$. We want to show that $m^+$is not equivalent to message $\hat{s}_0^S\left(\hat{t}-\Delta\right)$ for type $\hat{t}-\Delta$ w.r.t. $S^R(k-1)$. Suppose true. Then m^+ is equivalent with message $\hat{s}_0^S\left(\hat{t}\right)$ for type $\hat{t}$ w.r.t. $S^R(k-1)$ because $\hat{s}_0^S\left(\hat{t}\right) \in \left[\hat{s}_0^S\left(\hat{t}-\Delta\right), m^+\right]$. Then message m^+weakly dominates message $\hat{s}_0^S\left(\hat{t}\right)$ for type $\hat{t}$ w.r.t. $S^R(k-1)$ by single crossing condition. Thus $\hat{s}_0^S\left(\hat{t}\right) \notin S^S\left(\infty; \hat{t}\right)$, which contradicts the construction of $\hat{s}_0^S$. So $m^+$is not equivalent with message $\hat{s}_0^S\left(\hat{t}-\Delta\right)$ for type $\hat{t}-\Delta$ w.r.t. $S^R(k-1)$. So $u^S\left(\hat{t}, \sigma_{1k}^R(m^+)\right) > u^S\left(\hat{t}, \sigma_{1k}^R\left(\hat{t}-\Delta\right)\right)$. In addition, there exists $\sigma_{0k}^R \in \Delta^+ S^R(k-1)$ such that

$$\begin{aligned} \hat{s}_0^S\left(\hat{t}-\Delta\right) &\in \arg\max_m u^S\left(\hat{t}-\Delta, \sigma_{0k}^R(m)\right) \\ u^S\left(\hat{t}-\Delta, \sigma_{0k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) &> u^S\left(\hat{t}-\Delta, \sigma_{0k}^R\left(m^+\right)\right) \end{aligned}$$

. If $u^S\left(\hat{t},\sigma_{0k}^R\left(m_1^k\right)\right)\geq u^S\left(\hat{t},\sigma_{0k}^R\left(m\right)\right)$ for all $m\in\left[l\left(\infty;\hat{t}-\Delta\right),\hat{t}-\Delta\right]$, then let $\sigma_{*k}^R\equiv\left(1-\varepsilon\right)\sigma_{0k}^R+\varepsilon\sigma_{1k}^R$, and then we get the property we want. Otherwise, let

$$m^-\equiv\max\left\{\begin{array}{c} m\leq\hat{t}-\Delta| \\ u^S\left(\hat{t},\sigma_{0k}^R\left(m_1^k\right)\right)<u^S\left(\hat{t},\sigma_{0k}^R\left(m\right)\right)\end{array}\right\}$$

Define $\mu\left(\alpha;\sigma_{1k}^R,\sigma_{0k}^R\right)\equiv\alpha\sigma_{1k}^R+\left(1-\alpha\right)\sigma_{0k}^R$. W.l.o.g., let $u^S\left(\hat{t},\sigma_{0k}^R\left(m\right)\right)$ be decreasing in m on $S^S\left(k;\hat{t}\right)\cap\left[m_1^k,1\right]$. Summarizing what we have, we get that

$$\begin{array}{rcl} u^S\left(\hat{t}-\Delta,\sigma_{0k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) & > & u^S\left(\hat{t}-\Delta,\sigma_{0k}^R\left(m^+\right)\right) \\ u^S\left(\hat{t},\sigma_{0k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) & > & u^S\left(\hat{t},\sigma_{0k}^R\left(m^+\right)\right) \\ u^S\left(\hat{t}-\Delta,\sigma_{1k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) & \leq & u^S\left(\hat{t}-\Delta,\sigma_{1k}^R\left(m^+\right)\right) \\ u^S\left(\hat{t},\sigma_{1k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) & < & u^S\left(\hat{t},\sigma_{1k}^R\left(m^+\right)\right) \end{array}$$

. Suppressing σ_{1k}^R and σ_{0k}^R in the notation and write $\mu\left(\alpha;\sigma_{1k}^R,\sigma_{0k}^R\right)$ as $\mu\left(\alpha\right)$. Then we get

$$\begin{array}{rl} & u^S\left(\hat{t},\mu\left(1\right)\left(m^+\right)\right)-u^S\left(\hat{t},\mu\left(1\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ > & u^S\left(\hat{t}-\Delta,\mu\left(1\right)\left(m^+\right)\right)-u^S\left(\hat{t}-\Delta,\mu\left(1\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ \geq & 0 \end{array}$$

and

$$\begin{array}{rl} & 0 \\ > & u^S\left(\hat{t},\mu\left(0\right)\left(m^+\right)\right)-u^S\left(\hat{t},\mu\left(0\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ > & u^S\left(\hat{t}-\Delta,\mu\left(0\right)\left(m^+\right)\right)-u^S\left(\hat{t}-\Delta,\mu\left(0\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \end{array}$$

and

$$\begin{array}{rl} & u^S\left(\hat{t},\mu\left(\alpha\right)\left(m^+\right)\right)-u^S\left(\hat{t},\mu\left(\alpha\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ > & u^S\left(\hat{t}-\Delta,\mu\left(\alpha\right)\left(m^+\right)\right)-u^S\left(\hat{t}-\Delta,\mu\left(\alpha\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \end{array} \quad (8.1)$$

for all α. μ is continuous with α. Then from intermediate value theorm, we know that there exists $\hat{\alpha}$ such that

$$u^S\left(\hat{t}-\Delta,\mu\left(\hat{\alpha}\right)\left(m^+\right)\right)-u^S\left(\hat{t}-\Delta,\mu\left(\hat{\alpha}\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right)=0.$$

Combined with inequality 8.1, we get that

$$\begin{aligned} & u^S\left(\hat{t}, \mu(\hat{\alpha})\left(m^+\right)\right) - u^S\left(\hat{t}, \mu(\hat{\alpha})\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ > \ & 0. \end{aligned}$$

We then get

$$\begin{aligned} & u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(m^+\right)\right) - u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ > \ & 0 \\ > \ & u^S\left(\hat{t}-\Delta, \mu(\hat{\alpha}-\varepsilon)\left(m^+\right)\right) - u^S\left(\hat{t}-\Delta, \mu(\hat{\alpha}-\varepsilon)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \end{aligned}$$

for ε small enough. Since by assumption, $u^S\left(\hat{t}, \sigma_{1k}^R(m)\right)$ is weakly increasing in m on $S^S\left(k;\hat{t}\right)\cap\left[0, m_1^k\right]$, and by construction, $u^S\left(\hat{t}-\Delta, \sigma_{0k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \geq u^S\left(\hat{t}-\Delta, \sigma_{0k}^R(m)\right)$ for all $m \leq \hat{s}_0^S\left(\hat{t}-\Delta\right)$, and thus $u^S\left(\hat{t}, \sigma_{0k}^R\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \geq u^S\left(\hat{t}, \sigma_{0k}^R(m)\right)$ for all $m \leq \hat{s}_0^S\left(\hat{t}-\Delta\right)$, we know that

$$\begin{aligned} u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)(m)\right) & \leq u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ & < u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(m^+\right)\right) \end{aligned}$$

for all $m \leq \hat{s}_0^S\left(\hat{t}-\Delta\right)$. Suppose $\exists m' \in \left[\hat{s}_0^S\left(\hat{t}-\Delta\right)+\Delta, \hat{t}-\Delta\right]$ such that

$$u^S\left(\hat{t}, \mu(\hat{\alpha})\left(m'\right)\right) \geq u^S\left(\hat{t}, \mu(\hat{\alpha})\left(m^+\right)\right),$$

then

$$u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(m'\right)\right) > u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(m^+\right)\right).$$

Let m' be the smallest such message in $S^S\left(k;\hat{t}\right)$. Then

$$\begin{aligned} u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)(m)\right) & \leq u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(m^+\right)\right) \\ & < u^S\left(\hat{t}, \mu(\hat{\alpha}-\varepsilon)\left(m'\right)\right) \end{aligned}$$

for all $m < m'$. We can redo the analysis with $\mu(\hat{\alpha}-\varepsilon)$ as σ_{0k}^R and m' as $\hat{s}_0^S\left(\hat{t}-\Delta\right)$. Eventually we will run out of messages in between $m+\Delta$ and $\hat{t}-\Delta$. Without loss of generality, we can assume that

$$u^S\left(\hat{t}, \mu(\hat{\alpha})\left(m'\right)\right) < u^S\left(\hat{t}, \mu(\hat{\alpha})\left(m^+\right)\right)$$

for all $m' \in \left[\hat{s}_0^S\left(\hat{t}-\Delta\right), \hat{t}-\Delta\right]$. Define $\sigma_{*k}^R \equiv \mu(\hat{\alpha}-\varepsilon)$. So $u^S\left(\hat{t}, \mu(\alpha-\varepsilon)\left(m'\right)\right) < u^S\left(\hat{t}, \mu(\alpha-\varepsilon)\left(m^+\right)\right)$ for all $m' \leq \hat{t}-\Delta$. From the construction of σ_{1k}^R and σ_{0k}^R, we know that $u^S\left(\hat{t},\sigma_{1k}^R\left(m_1^k\right)\right) \geq u^S\left(\hat{t},\sigma_{1k}^R\left(m\right)\right)$ and $u^S\left(\hat{t},\sigma_{0k}^R\left(m_1^k\right)\right) \geq u^S\left(\hat{t},\sigma_{0k}^R\left(m\right)\right)$ for all $m \in S^S\left(k;\hat{t}\right) \cap \left[m_1^k, 1\right]$ and that strict inequality holds for any m which is not equivalent to m_1^k for type $\hat{t}$ w.r.t. $S^R\left(k-1\right)$. Let

$$\bar{m}_1^k \equiv \max\left\{\begin{array}{c} m \in S^S\left(k;\hat{t}\right) \mid \\ u^S\left(\hat{t},s^R\left(m\right)\right) = u^S\left(\hat{t},s^R\left(m_1^k\right)\right) \text{ for all } s^R \in S^R\left(k-1\right) \end{array}\right\}.$$

Then $\arg\max u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) \subset \left[\hat{t},m_1^k\right]$. By definition, if $m < m_1^k$, either $s^R\left(m\right) = s^R\left(\hat{t}-\Delta\right)$ for all $s^R \in S^R\left(k-1\right)$, and thus $u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) = u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(\hat{t}-\Delta\right)\right) < u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m^+\right)\right)$, or $m \notin S^S\left(k;\hat{t}\right)$, and thus $\arg\max_{m''} u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m''\right)\right) \setminus \{m\}$ is not empty. So $\left\{m_1^k\right\} = \arg\max u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) \cap S^S\left(k;\hat{t}\right)$. If $\arg\max_m u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) \in \left[\hat{t},m^+-\Delta\right]$, then by construction of m^+, $\hat{t}-\Delta \in \arg\max_m u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m\right)\right)$, because $s^R\left(\hat{t}-\Delta\right) = s^R\left(m\right)$ for all $m \in \left[\hat{t},m^+-\Delta\right]$ and all $s^R \in S^R\left(k-1\right)$. Since $\hat{t}-\Delta \in S^S\left(k;\hat{t}-\Delta\right)$ by construction, from **??** and that $\left\{m_1^k\right\} = \arg\max u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) \cap S^S\left(k;\hat{t}\right)$, there exists $s_*^S \in S^S\left(k\right)$ such that $s_*^S\left(\hat{t}-\Delta\right) = \hat{t}-\Delta$ and $s_*^S\left(\hat{t}\right) = m_1^k$, where s_*^S is a best response to $\mu\left(\alpha-\varepsilon\right)$. If there exists $m' > m^+$ such that $u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m'\right)\right) > u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right)$, then let m' be the smallest such m' (so $m_1^k \geq m' > m^+$), then we get

$$\begin{aligned} u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m'\right)\right) &> u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(m'\right)\right) &> u^S\left(\hat{t},\mu\left(\alpha-\varepsilon\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \end{aligned}$$

and

$$\begin{aligned} u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m'\right)\right) &> u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) \\ \text{for all } m &\in \left[\hat{t},m'-\Delta\right] \end{aligned}$$

because

$$\begin{aligned} u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m'\right)\right) &> u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(\hat{s}_0^S\left(\hat{t}-\Delta\right)\right)\right) \\ &\geq u^S\left(\hat{t}-\Delta,\mu\left(\alpha-\varepsilon\right)\left(m\right)\right) \\ &\quad \text{by the definition of } m'. \end{aligned}$$

Then we can perform the process using m' as m^+ and $\mu(\alpha-\varepsilon)$ as σ^R_{1k}. We'll eventually reach $m^+ = m^k_1$ and $\left[m^+ + \Delta, m^k_1\right] = \emptyset$, so the process ends after finite steps. So we can assume w.l.o.g. that $u^S\left(\hat{t}-\Delta, \mu(\alpha-\varepsilon)(m')\right) \leq u^S\left(\hat{t}-\Delta, \mu(\alpha-\varepsilon)\left(\hat{s}^S_0\left(\hat{t}-\Delta\right)\right)\right)$ for all $m' > m^+$. So let $\mu(\alpha-\varepsilon) \equiv \sigma^R_{*k}$. Then $\sigma^R_{*k} \in \Delta^+ S^R(k-1)$ and $\arg\max_m u^S\left(\hat{t}, \sigma^R_{*k}(m)\right) = \left\{m^k_1\right\}$ and $u^S\left(\hat{t}-\Delta, \sigma^R_{*k}\left(\hat{s}^S_0\left(\hat{t}-\Delta\right)\right)\right) \geq u^S\left(\hat{t}-\Delta, \sigma^R_{*k}(m)\right)$ for all $m \geq \hat{t}$, where strict inequality holds if m is not equivalent to message $\hat{t}-\Delta$ for type $\hat{t}-\Delta$ w.r.t. $S^R(k-1)$. Then there exists $s^S_{*k} \in S^S(k)$ which is a best response to $(1-\varepsilon)\sigma^R_0 + \varepsilon\sigma^R_{*k} \in \Delta^+ S^R(k-1)$, where $s^S_{*k}\left(\hat{t}\right) = m^k_1$ and $s^S_{*k}\left(\hat{t}-\Delta\right) \in \left[\hat{s}^S_0\left(\hat{t}-\Delta\right), \hat{t}-\Delta\right]$ and $u^S\left(t, \sigma^R_0\left(s^S_{*k}(t)\right)\right) = u^S\left(t, \sigma^R_0\left(\hat{s}^S_0(t)\right)\right)$ for all $t \geq \hat{t}+\Delta$. So $s^S_{*k}(t) \geq \hat{m}+\Delta$ for all $t \geq \hat{t}+\Delta$.

Step 2 Show that $\exists s^R_{*k} \in S^R_{*k}$ such that $s^R_{*k}\left(\hat{t}-\Delta\right) \leq \hat{t}-\Delta$.

Proof This does not hold only if $\exists \tilde{k} \leq k$ and $\tilde{m} \leq \hat{t}-\Delta$ such that the set of types sending messages in $\tilde{m}$ is nonempty and every type sending messages in $\left[\tilde{m}, \hat{t}-\Delta\right]$ under $S^S\left(\tilde{k}-1\right)$is no smaller than $\hat{t}$. Otherwise, we can use $\tilde{m}$ instead of $\hat{s}^S_0\left(\hat{t}-\Delta\right)$ in step one and then $s^S_{*\tilde{k}}\left(\hat{t}-\Delta\right) \in \left[\hat{s}^S_0\left(\hat{t}\right), \hat{t}-\Delta\right] = \left[\tilde{m}, \hat{t}-\Delta\right]$ and so $s^R_{*\tilde{k}}(\tilde{m}) \leq \hat{t}-\Delta$. Now suppose $\tilde{k}$ exists and is the smallest one. Let $\bar{m} \equiv \max_t \max S^S\left(\tilde{k}-1; t\right) \cap \left[0, \hat{t}-\Delta\right]$. Then $s^R_{*\tilde{k}}(\bar{m}) \leq \hat{t}-\Delta < \hat{t} \leq s^R_{*\tilde{k}}(\tilde{m})$. W.l.o.g. assume $\tilde{m} = \min\left(s^R_{*\tilde{k}}\right)^{-1}\left(s^R_{*\tilde{k}}(\tilde{m})\right)$. Assume that there exists $t \in T$ such that $u^S(t, \hat{a}) = \max_{a \in \hat{A}} u^S(t, a)$ for any pair $\left(\hat{a}, \hat{A}\right)$ where $\hat{A}$ is a subset of A and there exists $a_l, a_h \in \hat{A}$ where $a_l \geq y^S(0)$ and $\hat{a} \in (a_l, a_h)$. Then there exists type $\tilde{t} \leq \hat{t}-\Delta$ such that $\arg\max_m u^S\left(\tilde{t}, s^R_{*\tilde{k}}(m)\right) \geq \tilde{m}$. But by assumption, $\tilde{m} \notin S^S\left(\tilde{k}-1; \tilde{t}\right)$. So $\tilde{m}$ must be weakly dominated for type $\tilde{t}$ w.r.t. $S^R(k')$ for some $k' \leq \tilde{k}-2$. But then $\tilde{m}$ is weakly dominated for every type $t \geq \tilde{t}$ w.r.t. $S^R(k')$. We then arrive at a contradiction.

Step 3 Show that it cannot be the case that $s^R_{*k}(\hat{m}+\Delta) = s^R_{*k}\left(m^k_1\right)$ and type $\hat{t}$ prefers action $s^R_{*k}(\tilde{m})$ to $s^R_{*k}\left(m^k_1\right)$ for some $\tilde{m} \leq \hat{t}-\Delta$

Proof Suppose to the contrary that $s^R_{*k}(\hat{m}+\Delta) = s^R_{*k}\left(m^k_1\right)$ and type $\hat{t}$ prefers $s^R_{*k}(\tilde{m})$ to $s^R_{*k}\left(m^k_1\right)$ for some $\tilde{m} \leq \hat{t}-\Delta$. Then it has to be the case that $s^R_{*k}\left(m^k_1\right) \leq m^k_1$, because otherwise, it is better off to take action $\hat{t}$ at message m^k_1 with respect to s^S_{*k}. It also has to be the case that $s^R_{*k}\left(m^k_1\right) > y^S\left(\hat{t}\right)$. Let $\tilde{k}$ be the largest $k' < k$ such that $m^{k'}_1 < s^R_{*k}\left(m^k_1\right)$. Then $m^{\tilde{k}}_1 < s^R_{*k}\left(m^k_1\right)$ and $m^{\tilde{k}+1}_1 \geq s^R_{*k}\left(m^k_1\right)$. Then we have $s^S_{*\tilde{k}}\left(\hat{t}\right) = m^{\tilde{k}}_1$ and $s^S_{*\tilde{k}}\left(\hat{t}-\Delta\right) \in \left[s^S_{*k}\left(\hat{t}-\Delta\right), \hat{t}-\Delta\right]$. Then there exists $\hat{s}^R_{*\tilde{k}} \in BR^R\left((1-\varepsilon)\, s^S_{*k} + \varepsilon s^S_{*\tilde{k}}\right) \cap S^R\left(\tilde{k}\right)$ such that

$$\begin{aligned} \hat{s}^R_{*\tilde{k}}\left(m^k_1\right) &= s^R_{*k}\left(m^k_1\right) \\ \hat{s}^R_{*\tilde{k}}\left(m^{\tilde{k}}_1\right) &= \hat{t} \\ \hat{s}^R_{*\tilde{k}}\left(\hat{t}-\Delta\right) &\leq \hat{t}-\Delta \end{aligned}$$

. If $s^R_{*k}(\tilde{m}) \leq \hat{t}-\Delta$, then type $\hat{t}$ must prefer action $\hat{t}$ to action $s^R_{*k}(\tilde{m})$ to action $s^R_{*k}\left(m^k_1\right)$. So $u^S\left(\hat{t}, \hat{s}^R_{*\tilde{k}}\left(m^{\tilde{k}}_1\right)\right) > u^S\left(\hat{t}, \hat{s}^R_{*\tilde{k}}\left(m^k_1\right)\right)$ and $u^S\left(\hat{t}, \hat{s}^R_{*\tilde{k}}\left(m^{\tilde{k}}_1\right)\right) > u^S\left(\hat{t}, \hat{s}^R_{*\tilde{k}}\left(\hat{t}-\Delta\right)\right)$ and $\arg\max_m u^S\left(\hat{t}, \hat{s}^R_{*\tilde{k}}(m)\right) \subset \left[\hat{t}, s^R_{*k}\left(m^k_1\right)-\Delta\right]$. Therefore, $m^{\tilde{k}+1}_1 < s^R_{*k}\left(m^k_1\right)$, which contradicts the construction of $\tilde{k}$.

Step 4 From step 3, we know that either

a. $s^R_{*k}(m_2) \geq \hat{m}+\Delta$ and therefore $s^R_{*k}\left(m^k_1\right) = \hat{t}$. Thus, $\arg\max_m u^S\left(\hat{t}, \left((1-\varepsilon)\,\sigma^R_0 + \varepsilon s^R_{*k}\right)(m)\right)$ $\left[\hat{t}, \hat{m}\right]$ because σ^R_0 is constant on $\left[l\left(\infty; \hat{t}-\Delta\right), \hat{m}\right]$ while $s^R_{*k}\left(\hat{t}-\Delta\right) \leq \hat{t}-\Delta < \hat{t} = s^R_{*k}\left(\hat{t}\right)$.

b. or $s^R_{*k}(m_2) \leq \hat{m}$. We know that type $\hat{t}$ necessarily prefers action $s^R_{*k}(m_2)$ to action $s^R_{*k}(m)$ for any $m > m_2$ and $s^R_{*k}(m) \neq s^R_{*k}(m_2)$ because $s^R_{*k}(m) > m_2 > s^R_{*k}(m_2) \geq \hat{t}$ and type $\hat{t}$ prefers action $\hat{t}$ to action m_2. But there exists $m^* \in \left[\hat{t}, \hat{m}\right]$ such that type $\hat{t}$ prefers $s^R_{*k}(m^*)$ to $s^R_{*k}\left(\hat{t}-\Delta\right)$, because from step 3, we know that type $\hat{t}$ must prefer action $s^R_{*k}\left(m^k_1\right)$ to action $s^R_{*k}\left(\hat{t}-\Delta\right)$. So $\arg\max_m u^S\left(\hat{t}, \left((1-\varepsilon)\,\sigma^R_0 + \varepsilon s^R_{*k}\right)(m)\right) \subset \left[\hat{t}, \hat{m}\right]$.

Step 5 We conclude from step 4 that $\arg\max_m u^S\left(\hat{t}, \left((1-\varepsilon)\,\sigma^R_0 + \varepsilon s^R_{*k}\right)(m)\right) \subset \left[\hat{t}, \hat{m}\right]$. Since $(1-\varepsilon)\,\sigma^R_0 + \varepsilon s^R_{*k} \in \Delta S^R(k)$, we've shown that property * holds for $k+1$.

Step 6 By induction, property * holds for ∞.

Proof for Claim 5.1 in section 5.1 Suppose it's true for $1, ..., q$ and any $\hat{s}^R$ and $\hat{a}$, want to show that it is true for $q+1$ and any $\hat{s}^R$ and any $\hat{a}$. Suppose that $\hat{\alpha} \in \hat{s}^R(M(\infty))$ where $\hat{s}^R \in S^R(\infty)$ and $[0, g^{-1}(\infty; \hat{a}) - \Delta]$ admits a size-$(q+1)$ equilibrium Let $\tilde{a} \equiv \max \hat{s}^R([0, \hat{a} - \Delta] \cap M(\infty))$. Then $\tilde{a} \geq E([g^{-1}(\infty; \tilde{a}), g^{-1}(\infty; \hat{a}) - \Delta])$ because the posterior maximizer is greater than or equal to $E([g^{-1}(\infty; \tilde{a}), g^{-1}(\infty; \hat{a}) - \Delta])$ and smaller than $\hat{a} - \Delta$. Therefore if $\tilde{a} < E([g^{-1}(\infty; \tilde{a}), g^{-1}(\infty; \hat{a}) - \Delta])$, can improve upon $\hat{s}^R$ against any $s^S \in S^S(\infty)$ by increasing it to $E([g^{-1}(\infty; \tilde{a}), g^{-1}(\infty; \hat{a}) - \Delta])$. Suppose that $\tilde{a} < \alpha_{q+1}^{q+1}([0, g^{-1}(\infty; \hat{a}) - \Delta])$. Since $[0, g^{-1}(\infty; \hat{a}) - \Delta]$ admits a size-$(q+1)$ equilibrium and $q+1 \geq 2$, $g^{-1}(\infty; \tilde{a}) > 0$ because $\tilde{a} \geq E([0, g^{-1}(\infty; \hat{a}) - \Delta])$ and type 0 prefers 0 to $E([0, g^{-1}(\infty; \hat{a}) - \Delta])$. So there exists $m < \tilde{a}$, $m \in M(\infty)$ and $s^R \in S^R(\infty)$ such that $s^R(m) \neq s^R(\tilde{a})$. So $g^{-1}(\infty; \tilde{a})$ has to prefer action $\tilde{a}$ to $s^R(\tilde{a} - \Delta)$ for some $s^R \in S^R(\infty)$. $s^R(\tilde{a} - \Delta) \geq E([0, g^{-1}(\infty; \tilde{a}) - \Delta])$. So type $g^{-1}(\infty; \tilde{a})$ has to prefer action $\tilde{a}$ to action $E([0, g^{-1}(\infty; \tilde{a}) - \Delta])$. So $g^{-1}(\infty; \tilde{a}) \geq \tau_1^2(\tilde{a})$. If $\tilde{a} < \alpha_2^2([0, g^{-1}(\infty; \hat{a}) - \Delta])$, then the system $\{0, \tau_1^2(\tilde{a}), \tau_2^2(\tilde{a})\}$ (this is a system such that $E([\tau_1^2(\tilde{\alpha}) + \Delta, \tau_2^2(\tilde{\alpha})]) = \tilde{\alpha})$ is such that $\tau_2^2(\tilde{a}) < g^{-1}(\infty; \tilde{a}) - \Delta$ and hence $E([g^{-1}(\infty, \tilde{a}), g^{-1}(\infty; \hat{a}) - \Delta]) > \tilde{a}$, which is a contradiction. So $\tilde{a} \geq \alpha_2^2([0, g^{-1}(\infty; \hat{a}) - \Delta])$. But then $[0, g^{-1}(\infty; \tilde{a}) - \Delta]$ admits a size-2 equilibrium. Suppose $[0, g^{-1}(\infty; \tilde{a}) - \Delta]$ admits a size-j equilibrium where $j \leq q$. Then by assumption, $\max s^R([0, \tilde{a} - \Delta] \cap M(\infty)) \geq \alpha_j^j([0, g^{-1}(\infty; \tilde{a}) - \Delta])$ for all $s^R \in S^R(\infty)$ such that $\tilde{a} \in s^R(M(\infty))$. So it has to be the case that type $g^{-1}(\infty; \tilde{a})$ prefers action $\tilde{a}$ to action $\alpha_j^j([0, g^{-1}(\infty; \tilde{a}) - \Delta])$. Let

$$\left\{0, t_1^j\left(\left[0, g^{-1}(\infty; \tilde{a}) - \Delta\right]\right), t_2^j\left(\left[0, g^{-1}(\infty; \tilde{a}) - \Delta\right]\right), ..., g^{-1}(\infty; \tilde{a}) - \Delta, \bar{t}\right\}$$

be a system where $E([g^{-1}(\infty; \tilde{a}), \bar{t}]) = \tilde{a}$. If $j < q$, then this system is of size $j+1 < q+1$. Since $[0, g^{-1}(\infty; \hat{a}) - \Delta]$ admits a size-$(q+1)$ equilibrium, $\bar{t} < g^{-1}(\infty; \hat{a}) - \Delta$. Therefore

$$\begin{aligned} \tilde{a} &\geq E\left(\left[g^{-1}(\infty; \tilde{a}), g^{-1}(\infty; \hat{a})\right]\right) \\ &> E\left(\left[g^{-1}(\infty; \tilde{a}), \bar{t}\right]\right) = \tilde{a}, \end{aligned}$$

which is a contradiction. So $j = q$, and by assumption,

$$\left|\hat{s}^R([0, \tilde{a} - \Delta] \cap M(\infty))\right| \geq q.$$

Thus type $g^{-1}(\infty;\tilde{a})$ prefers $\tilde{a}$ to $\alpha_q^q([0,g^{-1}(\infty;\tilde{a})-\Delta])$. Moreover, type $g^{-1}(\infty;\tilde{a})$ prefers action $E([g^{-1}(\infty;\tilde{a}),g^{-1}(\infty;\hat{a})-\Delta])$ to action

$$\alpha_q^q\left(\left[0,g^{-1}(\infty;\tilde{a})-\Delta\right]\right)$$

since

$$\tilde{a}\geq E\left(\left[g^{-1}(\infty;\tilde{a}),g^{-1}(\infty;\hat{a})-\Delta\right]\right).$$

By condition (M),

$$g^{-1}(\infty;\tilde{a})\geq t_q^{q+1}\left(\left[g^{-1}(\infty;\tilde{a}),g^{-1}(\infty;\hat{a})-\Delta\right]\right),$$

and hence $\tilde{a}\geq\alpha_{q+1}^{q+1}([g^{-1}(\infty;\tilde{a}),g^{-1}(\infty;\hat{a})-\Delta])$. Therefore,

$$\left|\hat{s}^R([0,\hat{a}-\Delta]\cap M(\infty))\right|\geq q+1.$$

We are then done with the induction. So if $[0,1]$ admits a maximum of size-$N(b)$ equilibrium, then $\left|s^R(M(\infty))\right|\geq N(b)$ and $\max s^R(M(\infty))\geq\alpha_{N(b)}^{N(b)}$.

8.4. EIAL Results

8.4.1. Proof for Lemma 7

Proof. The idea is that, if s_2^R is not σ^S- *compatible*, then there must exist an action a_2 taken by s_2^R exactly on some interval I_2 such that a_2 does not maximize expected utility conditional on I_2. If ε is small enough, $U^R|_{I_2}\left(\varepsilon\left(\sigma^S,\sigma^{S\prime}\right),a\right)$ is very close to $U^R|_{I_2}\left(\sigma^S,a\right)$ as a function of a, then a_2 cannot maximize expected utility conditional on I_2, hence s_2^R is not $\varepsilon\left(\sigma^S,\sigma^{S\prime}\right)-$ *compatible* either.

u^R is bounded, so

$$\bar{D}^R\equiv\max_{(t,a),(t',a')\in T\times A}\left|u^R(t,a)-u^R(t',a')\right|$$

is well-defined. Then for any $\left(\sigma_1^S,\sigma_1^R\right),\left(\sigma_2^S,\sigma_2^R\right)\in\Delta S_L$,

$$\begin{aligned}&\left|U^R\left(\sigma_1^S,\sigma_1^R\right)-U^R\left(\sigma_2^S,\sigma_2^R\right)\right|\\ \leq\ &\bar{D}^R\end{aligned}$$

Given an interval $I\subset M$, let $U^R|_I$ be the expected Receiver utility conditional on receiving a message in I. Let a denote both action $a\in A$ and the constant strategy which reacts to every message with action a. Then

$$\begin{aligned}
& \left|U^R|_I\left(\sigma^S, a\right) - U^R|_I\left(\sigma^S_\varepsilon\left(\sigma^{S\prime}\right), a\right)\right| \\
= & \left|U^R|_I\left(\sigma^S, a\right) - (1-\varepsilon)\, U^R|_I\left(\sigma^S, a\right) - \varepsilon U^R|_I\left(\sigma^{S\prime}, a\right)\right| \\
= & \ \varepsilon\left|U^R|_I\left(\sigma^S, a\right) - U^R|_I\left(\sigma^{S\prime}, a\right)\right| \\
\leq & \ \varepsilon \bar{D}^R
\end{aligned}$$

The bound is does not depend on $\sigma^S, \sigma^{S\prime}, a$ or I. A^R is finite, so a best response $a \in A$ to any conjecture σ^S gives a strictly higher expected utility than any non-best response a'. Let d_{I,σ^S} denote the difference in expected utility conditional on I against conjecture σ^S between the best action and the second best action. Formally, define

$$d_{I,\sigma^S} \equiv \min_{a_2 \notin \arg\max_{a'} U^R|_I(\sigma^S, a')} \left(\left(\max_{a''} U^R|_I\left(\sigma^S, a''\right)\right) - U^R|_I\left(\sigma^S, a_2\right)\right)$$

Then $d_{I,\sigma^S} > 0$.

For all $\varepsilon < \frac{1}{2}\frac{d_{I,\sigma^S}}{\bar{D}^R}$, $a \notin \arg\max_{a'} U^R|_I\left(\sigma^S, a'\right)$ and $a^* \in \arg\max_{a'} U^R|_I\left(\sigma^S, a'\right)$,

$$\begin{aligned}
& U^R|_I\left(\varepsilon\left(\sigma^S, \sigma^{S\prime}\right), a\right) - U^R|_I\left(\varepsilon\left(\sigma^S, \sigma^{S\prime}\right), a^*\right) \\
= & \ U^R|_I\left(\varepsilon\left(\sigma^S, \sigma^{S\prime}\right), a\right) - U^R|_I\left(\sigma^S, a\right) + U^R|_I\left(\sigma^S, a\right) - U^R|_I\left(\sigma^S, a^*\right) \\
& + U^R|_I\left(\sigma^S, a^*\right) - U^R|_I\left(\varepsilon\left(\sigma^S, \sigma^{S\prime}\right), a^*\right) \\
\leq & \ \varepsilon \bar{D}^R - d_{I,\sigma^S} + \varepsilon \bar{D}^R \\
< & \ \frac{d_{I,\sigma^S}}{2} - d_{I,\sigma^S} + \frac{d_{I,\sigma^S}}{2} = 0
\end{aligned}$$

So

$$\arg\max_a U^R|_I\left(\varepsilon\left(\sigma^S, \sigma^{S\prime}\right), a\right) \subset \arg\max_a U^R|_I\left(\sigma^S, a\right) \tag{8.2}$$

Define

$$\bar{\varepsilon}_{\sigma^S} \equiv \min_{I \subset M} d_{I,\sigma^S}$$

Since M is finite, $\bar{\varepsilon}_{\sigma^S}$ is well defined. So the containment relation 8.2 holds for any $\varepsilon < \bar{\varepsilon}_{\sigma^S}$, and for any $\sigma^{S\prime} \in \Delta S^S$. If s_1^R is $\varepsilon\left(\sigma^S, \sigma^{S\prime}\right)$ for some $\varepsilon < \bar{\varepsilon}_{\sigma^S}$, then for any $\hat{m}$ which is sent by some type with strictly positive probability given the conjecture $\varepsilon\left(\sigma^S, \sigma^{S\prime}\right)$, and for the interval $I_{\hat{m}}$ on which s_1^R takes the same value as $s_1^R(\hat{m})$,

$$\begin{aligned}
s_1^R(\hat{m}) & \in \arg\max_a U^R|_{I_{\hat{m}}}\left(\varepsilon\left(\sigma^S, \sigma^{S\prime}\right), a\right) \\
& \subset \arg\max_a U^R|_{I_{\hat{m}}}\left(\sigma^S, a\right)
\end{aligned}$$

Since $\varepsilon\left(\sigma^{S},\sigma^{S\prime}\right)(t)=(1-\varepsilon)\,\sigma^{S}(t)+\varepsilon\sigma^{S\prime}(t)$ for all t, any message that receives positive probability given the conjecture $\sigma^{S}(t)$ also receives positive probability under $\varepsilon\left(\sigma^{S},\sigma^{S\prime}\right)(t)$, it is just shown that s_1^R is also σ^S- *compatible*. ■

8.4.2. Proof for Proposition 5

Proof. Assume to the contrary there exist $m_1<m_2\in M(k-1)$ which receive different reactions under some $s^R\in S^R(k)$, i.e. $s^R(m_1)\neq s^R(m_2)$. Consider $\hat{m}_k$ being such that $\hat{m}_k$ always attains the minimum on $M(k-1)$ for any $s^R\in S^R(k)$ and that there exists $s^R\in S^R(k)$ such that $s^R(\hat{m}_k)\neq s^R(\hat{m}_k)$. Suppose $\hat{t}$ is the highest type that sends messages smaller or equal to $\hat{m}_k$. Then since $\hat{m}_k$ always takes on the minimum of s^R for any s^R in $C_R^*(k)$, the highest values $\hat{m}_k$ and $\hat{m}_k+\Delta$ can take on when $s^R(\hat{m}_k)\neq s^R(\hat{m}_k+\Delta)$ would be $E\left(\left[0,\hat{t}_k\right]\right)$ and $E\left(\left[\hat{t}_k+\Delta,1\right]\right)$ respectively. But since there is only babbling equilibria, for every Sender type t, she prefers being thought of as pooling with all higher types than pooling with all lower types. So $\hat{t}_k$ would prefer $E\left(\left[\hat{t}_k+\Delta,1\right]\right)$ to $E\left(\left[0,\hat{t}_k\right]\right)$, where $E\left(\left[0,\hat{t}_k\right]\right)$ is the best $\hat{t}_k$ can hope for from sending message $\hat{m}_k$ (because $E\left(\left[0,\hat{t}_k\right]\right)\leq\hat{t}_k$ is on the increasing part of $\hat{t}_k$'s utility curve) and $E\left(\left[\hat{t}_k+\Delta,1\right]\right)$ is the worst $\hat{t}_k$ would anticipate from sending message $\hat{m}_k+\Delta$ when $\hat{m}_k$ induces a different action from $\hat{m}_k+\Delta$. So sending message $\hat{m}_k$ is weakly dominated by sending message $\hat{m}_k+\Delta$ for type $\hat{t}_k$. Hence in $\Pi_{t\in T}S^S(k+1;t)$, the highest type that sends messages smaller or equal to $\hat{m}_k$ would be strictly smaller than $\hat{t}_k$ and thus $\Pi_{t\in T}S^S(k+1;t)\subsetneqq\Pi_{t\in T}S^S(k-1;t)$ (in particular, $S^S\left(k+1;\hat{t}\right)\subsetneqq S^S\left(k-1;\hat{t}\right)$) and the process does not stop at round k.

Formally, define

$$\hat{m}_k:=\min\left\{\begin{array}{c} m\in M(k-1)\,| \\ \exists s^R\in S^R(k)\ \text{ such that} \\ m\in\arg\min_{m'\in M(k)}s^R(m') \\ \text{and } s^R(m)\neq s^R(m+\Delta)\end{array}\right\}$$

From the definition, $s^R(\hat{m}_k)=\min_{m'\in M(k-1)}s^R(m')$ for all $s^R\in S^R(k)$ and there exists $\hat{s}^R\in S^R(k)$ such that $\hat{s}^R(\hat{m}_k)\neq\hat{s}^R(\hat{m}_k+\Delta)$. From weak monotonicity of $\hat{s}^R$ and the construction,

$$\begin{aligned}\hat{s}^R(\min M(k-1)) &= \min_{m'\in M(k-1)}\hat{s}^R(m') \\ &= \hat{s}^R(\hat{m}_k) \\ &\neq \hat{s}^R(\hat{m}_k)\end{aligned}$$

That is, the interval that $\hat{s}^R$ takes on the same value as on $\hat{m}_k$ is $[\min M^* (k-1), \hat{m}_k]$. Let the interval that $\hat{s}^R$ takes on the sme value as on $\hat{m}_k + \Delta$ be $[\hat{m}_k + \Delta, \bar{m}_k]$. By the procedure, there exists $\hat{\sigma}^S \in \Pi_{t \in T} \left(\Delta^+ S^S (k-1; t)\right)$ to which $\hat{s}^R$ is $\hat{\sigma}^S -$ *compatible*. It then follows that

$$\hat{s}(\hat{m}_k) \in \arg\max_{a \in A} U^R|_{[\min M(k-1), \hat{m}_k]} \left(\hat{\sigma}^S, a\right)$$

$$\begin{aligned} & U^R|_{[\min M(k-1), \hat{m}_k]} \left(\hat{\sigma}^S, a\right) \\ = & \sum_{s^S \in \Pi_{t \in T} S^S(k-1;t)} \hat{\sigma}^S \left(s^S\right) \sum_{\substack{t \in T: \\ s^S(t) \in [\min M(k-1), \hat{m}_k]}} \pi(t) u^R(t, a) \\ = & \sum_{s^S \in \Pi_{t \in T} S^S(k-1;t)} \sum_{\substack{t \leq \\ s^S(t) \in [\min M^*(k-1), \hat{m}_k]}} \hat{\sigma}^S \left(s^S\right) \end{aligned}$$

■

8.4.3. Proof for Proposition 6

Proposition 6 follows immediately from the following claim.

Claim For all k, there exists $s^R \in S^R$ such that.

1. $s^R \in ES^R(k)$, and $s^R(M(k)) = \left\{\alpha_1, ..., a_{N(b)}\right\}$ where $\alpha_i = E([t_{i-1}, t_i - \Delta])$;
2. $\forall m \in [\alpha_i, \alpha_{i+1} - \Delta]$, either there exists $m' < m$ such that $u^S\left(t_i - \Delta, s^R(m)\right) \leq u^S\left(t_i - \Delta, s^R(m')\right)$ for all $s^R \in ES^R(k)$, or $s^R(m) = \alpha_{i-1}$.

Proof Show by induction. Suppose they hold for k. Then there exists $\hat{s}^R \in ES^R(k)$ satisfying condition 1 and 2. From the definition that $\left\{t_0, ..., t_{N(b)}\right\}$ is a forward solution and that $\alpha_i = E([t_{i-1}, t_i - \Delta])\ \forall i = 1, ..., N(b)$, every type $t \in [t_{i-1}, t_i - \Delta]$ strictly prefers action α_i the most in the range of $\hat{s}^R$. Therefore, there exists one message m such that $\hat{s}^R(m) = \alpha_i$ and $m \in ES^S(k+1; t)$. Since $\hat{s}^R(\alpha_{i+1}) = \alpha_{i+1} > \alpha_i$, such message must be smaller than $\alpha_{i+1} - \Delta$. Thus $l(k+1; t) \leq \alpha_{i+1} - \Delta$ for all $t \in [t_{i-1}, t_i - \Delta]$. Therefore, we can define

$$s^S_{big}(t) \equiv \max \left\{ \begin{array}{c} m \in ES^S(k+1; t), m \leq \alpha_{i+1} - \Delta \\ \text{where } i \text{ is such that } t \in [t_{i-1}, t_i - \Delta] \end{array} \right\} \forall t.$$

By definition, $s^S_{big} \in ES^S(k+1)$, and thus $s^S_{big} \in ES^S(k)$.

Claim s_{big}^S is increasing in t.

Proof Given $\hat{t}$. Let i be such that $\hat{t} + \Delta \in [t_{i-1}, t_i - \Delta]$. To show that $s_{big}^S\left(\hat{t} + \Delta\right) \geq s_{big}^S\left(\hat{t}\right)$, it suffices to show that $\left[s_{big}^S\left(\hat{t}\right), \alpha_{i+1} - \Delta\right] \cap ES^S\left(\hat{t} + \Delta; k\right) \neq \emptyset$. We break the discussion into two cases.

Case 1 $\hat{s}^R\left(s_{big}^S\left(\hat{t}\right)\right) \leq \alpha_i - \Delta$. Then $\left(\hat{s}^R\right)^{-1}(\alpha_i) \subset \left[s_{big}^S\left(\hat{t}\right) + \Delta, \alpha_{i+1} - \Delta\right]$ by the construction of $\hat{s}^R$ and the assumption that $\hat{s}^R\left(s_{big}^S\left(\hat{t}\right)\right) \leq \alpha_i - \Delta$. Since

$$\left(\hat{s}^R\right)^{-1}(\alpha_i) \cap ES^S\left(k; \hat{t} + \Delta\right) \neq \emptyset,$$

we know that

$$\left[s_{big}^S\left(\hat{t}\right) + \Delta, \alpha_{i+1} - \Delta\right] \cap ES^S\left(k; \hat{t} + \Delta\right) \neq \emptyset.$$

Case 2 $\hat{s}^R\left(s_{big}^S\left(\hat{t}\right)\right) \geq \alpha_i$. Then $\hat{s}^R\left(s_{big}^S\left(\hat{t}\right)\right) = \alpha_i$, because $\hat{t} \leq t_i - \Delta$ and by construction of s_{big}^S, $s_{big}^S\left(\hat{t}\right) \leq \alpha_{i+1} - \Delta$, and the assumption that $\hat{s}^R$ satisfies condition 2. Let $\tilde{\sigma}^R \in \Delta^+ ES^R(k-1)$ such that

$$s_{big}^S\left(\hat{t}\right) \in \arg\max_m U^S\left(\hat{t}, \tilde{\sigma}^R(m)\right).$$

(Existence is guaranteed by the definition of $s_{big}^S\left(\hat{t}\right)$) Then by super modularity of U^S and weak monotonicity of s^R in the support of $\tilde{\sigma}^R$, $U^S\left(\hat{t} + \Delta, \tilde{\sigma}^R\left(s_{big}^S\left(\hat{t}\right)\right)\right) > U^S\left(\hat{t} + \Delta, \tilde{\sigma}^R(m)\right)$ for all $m < s_{big}^S\left(\hat{t}\right)$. Since $\hat{s}^R\left(s_{big}^S\left(\hat{t}\right)\right) = \alpha_i$, we know that $s_{big}^S\left(\hat{t}\right) \in \arg\max_m U^S\left(\hat{t} + \Delta, \hat{s}^R(m)\right)$. Therefore, for ε very small,

$$\arg\max_m U^S\left(\hat{t} + \Delta, \left((1-\varepsilon)\hat{s}^R + \varepsilon\tilde{\sigma}^R\right)(m)\right) \subset \left[s_{big}^S\left(\hat{t}\right), \alpha_{i+1} - \Delta\right].$$

Since $(1-\varepsilon)\hat{s}^R + \varepsilon\tilde{\sigma}^R \in \Delta^+ ES^R(k-1)$,

$$\arg\max_m U^S\left(\hat{t} + \Delta, \left((1-\varepsilon)\hat{s}^R + \varepsilon\tilde{\sigma}^R\right)(m)\right) \cap ES^S\left(k; \hat{t} + \Delta\right) \neq \emptyset$$

and thus

$$\left[s_{big}^S\left(\hat{t}\right), \alpha_{i+1} - \Delta\right] \cap ES^S\left(k; \hat{t} + \Delta\right) \neq \emptyset.$$

Lemma 7 implies that $ES^R(k+1)$ must contain one Receiver strategy that is weakly sequentially rational with respect to s^S_{big}. Suppose $s^R_{big} \in \arg\max_{\substack{s^R \in S^R \\ s^R \text{ is interim rational} \\ \text{w.r.t. } s^S_{big}}} U^R\left(s^S_{big}, s^R\right)$, then $s^R_{big}\left(s^S_{big}(t)\right)$ is increasing in t because s^R_{big} is increasing and $s^S_{big}(t)$ is increasing in t. Therefore, s^R_{big} partitions the type space into $\{\tau_0, ..., \tau_n\}$ where $\tau_0 = 0$ and $\tau_n = 1$. By definition, $s^R_{big}\left(s^S_{big}(t)\right) = s^R_{big}\left(s^S_{big}(t')\right)$ if and only if t and t' both belong to the same step $[\tau_{i-1}, \tau_i - \Delta]$ for some i and $s^R_{big}\left(s^S_{big}(\tau_{i-1})\right) = E([\tau_{i-1}, \tau_i - \Delta])$ for $i = 1, ..., n$.

Claim $[0, \tau_{i+1} - \Delta]$ has a forward solution of size $i+1$ and

$$\tau_i \le t^{i+1}_i([0, \tau_{i+1} - \Delta]) \, for i = 1, ..., n.$$

Proof Show by induction. Suppose $[0, \tau_{j+1} - \Delta]$ has a forward solution of size $j+1$ and $\tau_j \le t^{j+1}_j([0, \tau_{j+1} - \Delta])$ for all $j = 1, ..., i-1$. Condition (M) implies that $\tau_{j+1} > t_j$ for all $j = 1, ..., i-1$ because $\left\{t_0, ..., t_{N(b)}\right\}$ is the largest forward solution on $[0, 1]$. First we want to show that type $\tau_i - \Delta$ must weakly prefer action $E([\tau_{i-1}, \tau_i - \Delta])$ to action $E([\tau_i, \tau_{i+1} - \Delta])$.

Case 1 $\tau_i \ne t_q$ for any q.

Therefore, there exists q such that $\tau_i - \Delta, \tau_i \in [t_{q-1}, t_q - \Delta]$. By construction, $s^S_{big}(\tau_i) < \alpha_{q+1} - \Delta$. By the construction of $\hat{s}^R$, $\hat{s}^R\left(s^S_{big}(\tau_i)\right) = \alpha_q$. Suppose to the contrary that type $\tau_i - \Delta$ prefers action $E([\tau_i, \tau_{i+1} - \Delta])$ to action $E([\tau_{i-1}, \tau_i - \Delta])$. By the definition of $\{\tau_0, \tau_1, ...\tau_n\}$ and the construction that s^R_{big} is sequentially rational w.r.t. s^S_{big}, we know that $s^R_{big}\left(s^S_{big}(\tau_i)\right) = E([\tau_i, \tau_{i+1} - \Delta])$ and $s^R_{big}\left(s^S_{big}(\tau_i - \Delta)\right) = E([\tau_{i-1}, \tau_i - \Delta])$. Therefore, given the Receiver strategy s^R_{big}, type $\tau_i - \Delta$ prefers message $s^S_{big}(\tau_i)$ to message s^S_{big}, and

$$\arg\max_m U^S\left(\tau_i - \Delta, \left((1-\varepsilon)\hat{s}^R + \varepsilon s^R_{big}\right)(m)\right) \subset \left[s^S_{big}(\tau_i - \Delta) + \Delta, \alpha_{i+1} - \Delta\right].$$

Since $s^R_{big} \in ES^R(k+1) \subset ES^R(k)$,

$$\left[s^S_{big}(\tau_i - \Delta) + \Delta, \alpha_{i+1} - \Delta\right] \cap ES^S(k+1; \tau_i - \Delta) \ne \emptyset.$$

But this contradicts the construction of $s^S_{big}(\tau_i - \Delta)$.

Case 2 $\tau_i = t_q$ for some q.

We've shown that $\tau_i > t_{i-1}$. So $q \geq i$. Suppose $q > i$. But then s^R_{big} can be improved upon by partitioning $[0, \tau_i]$ as $\{0, t_1, ..., t_q\}$ by the monotonicity condition (M), and there exists a Receiver strategy that is interim rational w.r.t. s^S_{big} which does that partition. So $q = i$. But then by the same argument, $\tau_j = t_j$ for all $j < i$. In particular, $\tau_{i-1} = t_{i-1}$. Suppose to the contrary, type $t_i - \Delta$ prefers action $E([t_i, \tau_{i+1} - \Delta])$ to action $E([t_{i-1}, t_i - \Delta])$. Then it has to be the case that $E([t_i, \tau_{i+1} - \Delta]) < E([t_i, t_{i+1} - \Delta]) = \alpha_{i+1}$. By the literal condition of the language assumption, $s^R_{big}(E([t_i, \tau_{i+1} - \Delta])) = E([t_i, \tau_{i+1} - \Delta])$. Therefore, given s^R_{big}, type $t_i - \Delta$ prefers message $E([t_i, \tau_{i+1} - \Delta])$ to message $s^S_{big}(t_i - \Delta)$. So there exists some message $m \geq E([t_i, \tau_{i+1} - \Delta])$ such that $m \in ES^S(k; t_i - \Delta)$. Since $E([t_i, \tau_{i+1} - \Delta]) \leq \alpha_{i+1} - \Delta$, from the assumption that condition 2 holds for k, $\hat{s}^R(E([t_i, \tau_{i+1} - \Delta])) = \alpha_i$. So

$$\arg\max_m U^S\left(t_i - \Delta, \left((1-\varepsilon)\hat{s}^R + \varepsilon s^R_{big}\right)(m)\right) \subset \left[s^S_{big}(t_i - \Delta) + \Delta, \alpha_{i+1} - \Delta\right].$$

And it follows that

$$\left[s^S_{big}(t_i - \Delta) + \Delta, \alpha_{i+1} - \Delta\right] \cap ES^S(k+1; \tau_i - \Delta) \neq \emptyset.$$

A contradiction.

By assumption, $\tau_{i-1} \leq t^i_{i-1}([0, \tau_i - \Delta])$. So $E([\tau_{i-1}, \tau_i - \Delta]) \leq \alpha^i_i([0, \tau_i - \Delta])$. We have just shown that type $\tau_i - \Delta$ prefers $E([\tau_{i-1}, \tau_i - \Delta])$ to $E([\tau_i, \tau_{i+1} - \Delta])$. Thus, type $\tau_i - \Delta$ must prefer action $\alpha^i_i([0, \tau_i - \Delta])$ to action $E([\tau_i, \tau_{i+1} - \Delta])$ because

$$E([\tau_{i-1}, \tau_i - \Delta]) \leq \alpha^i_i([0, \tau_i - \Delta]) < \tau_i \leq E([\tau_i, \tau_{i+1} - \Delta]).$$

So there exists $\bar{t} \in [\tau_i, \tau_{i+1} - \Delta]$ such that type $\tau_i - \Delta$ prefers action $\alpha^i_i([0, \tau_i - \Delta])$ to action $E([\tau_i, \bar{t}])$ and type τ_i prefers action $E([\tau_i, \bar{t}])$ to action $\alpha^i_i([0, \tau_i - \Delta])$. By definition of $\bar{t}$, $\tau_i = t^{i+1}_i([0, \bar{t}])$. By the monotonicity condition (M), $t^{i+1}_i([0, \bar{t}]) \leq t^{i+1}_i([0, \tau_{i+1} - \Delta])$ because

$\bar{t} \leq \tau_{i+1} - \Delta$. It follows that $\tau_i \leq t_i^{i+1}([0, \tau_{i+1} - \Delta])$. Moreover, $[0, \bar{t}]$ has a forward solution of size $i+1$, so $[0, \tau_{i+1}]$ has a forward solution of size $i+1$.

Claim 8.4.3 implies that $\tau_{n-1} \leq t_{n-1}^{n}([0, \tau_n - \Delta]) = t_{n-1}^{n}([0,1])$ and that $[0,1]$ has a forward solution of size n. Since $N(b)$ is the maximum of the size of a forward solution on $[0,1]$, $n \leq N(b)$. So $\tau_i \leq t_i$ for all $i = 1, ..., n$. Condition (M) implies that $U^R\left(s_{big}^S, s_{big}^R\right) \leq U^R\left(s_{big}^S, \hat{s}^R\right)$ because $\hat{s}^R\left(s_{big}^S\right)$ partitions the type space into $\left\{0, t_1, ..., t_{N(b)-1}, 1\right\}$ and this is a better partition. But by assumption, $\hat{s}^R \in ES^R(k)$ where $\hat{s}^R(M(k)) = \left\{\alpha_1, ..., a_{N(b)}\right\}$ and $\alpha_i = E([t_{i-1}, t_i - \Delta])$ $\forall i = 1, ..., N(b)$. $\hat{s}^R$ is weakly interim rational with respect to s_{right}^S, so

$$\max_{\substack{s^R \in ES^R(k); \\ s^R \text{ is weakly interim rational} \\ \text{w.r.t. } s_{big}^S}} U^R\left(s_{big}^S, s^R\right) \geq U^R\left(s_{big}^S, \hat{s}^R\right).$$

Therefore, equality holds and for any $\tilde{s}^R \in ES^R(k+1)$ such that $\tilde{s}^R$ is weakly sequentially rational w.r.t. s_{big}^S, $\tilde{s}^R$ partitions the type space into $\left\{0, t_1, ..., t_{N(b)-1}, 1\right\}$.

Suppose $m \in [\alpha_i, \alpha_{i+1} - \Delta]$ is such that there exists $\tilde{s}^R \in ES^R(k+1)$ where $u^S\left(t_i - \Delta, \tilde{s}^R(m)\right) > u^S\left(t_i - \Delta, \tilde{s}^R(m - \Delta)\right)$. Since statement 2 holds for k, $\hat{s}^R(m) = \alpha_i$. So

$$\arg\max_m U^S\left(t_i - \Delta, \left((1-\varepsilon)\hat{s}^R + \tilde{s}^R\right)(m)\right) \subset [m, \alpha_{i+1} - \Delta].$$

Therefore, $ES^S(k+1; t_i - \Delta) \cap [m, \alpha_{i+1} - \Delta] \neq \emptyset$. It follows that $s_{big}^S(t_i - \Delta) \geq m$ and therefore $s_{big}^R(m) \leq s_{big}^R\left(s_{big}^S(t_i - \Delta)\right) = \alpha_i$. We have thus shown that statement 2 holds for $k+1$.

References

[1] Brandenburger, A., Friedenberg, A., and H.J. Keisler., "Admissibility in Games," mimeo, 2004. Available at www.stern.nyu.edu/~abranden.

[2] Brandenburger, A., and H.J. Keisler, "An Impossibility Theorem on Beliefs in Games," mimeo, 2004. Available at www.stern.nyu.edu/~abranden.

[3] Crawford, V., and J. Sobel, "Strategic Information Transmission," Econometrica, 50(6), 1982, 1431-1451.

[4] Farrell, J. "Meaning and Credibility in Cheap-Talk Games," Games and Economic Behavior, 5(4), 1993, 514-531.

[5] Fudenberg, D. and J. Tirole, *Game Theory*, The MIT Press, 1991.

[6] Kim, Y.-G. and J. Sobel, "An Evolutionary Approach to Pre-play Communication," Econometrica, 63(5), 1995, 1181-1193.

[7] Kreps, D., and R. Wilson, "Sequential Equilibria," Econometrica, 50(4), 1982, 863-894.

[8] Pearce, D.G. "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52(4), 1984, 1029-1050.

[9] Rabin, M. "Communication between Rational Agents," Journal of Economic Theory, 51, 1990, 144-179.

[10] Warneryd, K. "Cheap Talk, Coordination, and Evolutionary Stability," Games and Economic Behavior, 5, 1993, 532-546.

[11] Zapater, I. "Credible Proposals in Communication Games," Journal of Economic Theory, 72, 1997, 173-197.