Effective Communication in Cheap-Talk Games*

Sidartha Gordon,[†] Navin Kartik,[‡] Melody Pei-yu Lo,[§]
Wojciech Olszewski,[¶] and Joel Sobel[‖]

October 1, 2023

Abstract

This paper presents arguments based on weak dominance and learning for selecting informative equilibria in a model of cheap-talk communication where players must use monotonic strategies. Under a standard regularity condition, only one equilibrium survives iterated deletion of dominated strategies. Under the same condition, we establish that best-response dynamics converges to this outcome.

Journal of Economic Literature Classification Numbers: C72, D83. Keywords: Communication, Learning, Equilibrium Selection.

*This paper combines results from three projects: Gordon [16], Kartik and Sobel [20], and Lo and Olszewski [23]. We are grateful to many seminar audiences and to Pierpaolo Battigalli, Andreas Blume, Richard Brady, Vincent Crawford, Françoise Forges, Sjaak Hurkens, Philip Neary, Alexsandr Levkun, Jeffrey Mensch, Olivier Tercieux, Yuehui Wang, and Yangfan Zhou for useful comments.

†LEDA, Université Paris-Dauphine, email: sidartha.gordon@dauphine.psl.eu
‡Department of Economics, Columbia University, email: nkartik@gmail.com
§Department of Economics, National Taiwan University, email: peiyulo2006@gmail.com
¶Department of Economics, Northwestern University, email: wo@kellogg.northwestern.edu
‖Department of Economics, University of California, San Diego, email: jsobel@ucsd.edu

1 Introduction

Talk is a useful way to communicate private information in strategic situations, as formalized by Crawford and Sobel [10] and Green and Stokey [17]. These early papers recognized, however, that their equilibrium analysis is generally indeterminate. Models of cheap-talk communication typically have multiple equilibrium outcomes, including an uninformative one in which no information is transmitted. A central concern in the literature has been finding conditions under which communication is effective, that is in which the predicted outcome involves non-trivial information transmission. We hope this paper advances the literature on effective communication.

We develop arguments that lead to equilibrium selection in the Crawford and Sobel [10] (hereafter CS) model. We establish that iterated deletion of interim weakly dominated strategies selects an outcome with effective communication when such an outcome exists, for example, when the regularity condition of CS holds. We also show that best-response dynamics lead to the same selection. These arguments are subtle because they require a reformulation of the strategic situation in order to work. The paper presents two different ways to think about the reformulation: one reformulation studies a game in which the players are restricted to monotonic strategies; the other looks at learning processes. In the first case, our solution concept involves iterated deletion of weakly dominated strategies. In the second case, it involves best-response dynamics.

The CS model that underlies our analysis has an informed Sender sending a message to an uninformed Receiver. The Receiver responds to the message by making a decision that is payoff relevant to both players. Talk is cheap because the payoffs of the players do not depend directly on the Sender's message. CS characterize the set of equilibrium outcomes in a one-dimensional environment with an "upward bias" conflict of interest: the Sender prefers higher decisions that the Receiver. CS demonstrate that there is a finite upper bound, N^*, to the number of distinct actions that the Receiver takes in equilibrium, and that for each $N = 1, \ldots, N^*$, there is an equilibrium in which the Receiver takes N actions. In addition, when a technical regularity condition holds, CS demonstrate that for all $N = 1, \ldots, N^*$, there is a unique equilibrium outcome in which the Receiver takes N distinct actions, and the ex-ante expected payoff for both Sender and Receiver is strictly increasing in N. The outcome with N^* actions is often selected in applications.

The multiple-equilibria problem arises in three different ways in cheap-talk games. Typically, some messages are not used in equilibrium. There will often be multiple ways to specify the Receiver's behavior off the path of play. This first kind of multiplicity, off-path indeterminacy, is familiar in games with incomplete information and is generally not essential. The second kind of multiplicity, message indeterminacy, is that the meaning of messages is arbitrary. Given any equilibrium, one can generate another equilibrium by changing the use and interpretation of messages. This kind of problem identifies a way in which language is arbitrary. The word used to describe the color of a white house in Paris is blanche and in Warsaw is biały. Predictions are still possible with this kind of indeterminacy when the different equilibria induce the same relationship between types and actions. What matters is that French speakers and Polish speakers classify the same set of houses as "white" (and their audiences understand that) rather than the particular word they use to describe the color. The third type of multiplicity, type-action indeterminacy, is fundamental. Cheap-talk games typically have an uninformative equilibrium[1] and may have qualitatively different equilibria in which the Receiver takes at least two different actions with positive probability. It is this type of multiplicity that we wish to examine, but our approach shows how eliminating the problem of message indeterminacy can resolve type-action indeterminacy.

Our approach relies on a restriction to monotonic strategies. We assume that there is an exogenous order on messages and restrict players to strategies that are monotonic with respect to this order. We view this as a way to incorporate "exogenous meaning" into communication: players enter the strategic setting with a shared ordering of messages and it is common knowledge that they will behave in a way that is consistent with this ordering. The resulting monotonic cheap-talk game has all three kinds of multiplicity, but monotonic strategies eliminate some message indeterminacy. Our main result is that combining our monotonicity assumption with an equilibrium refinement solves the problem of type-action indeterminacy. Under the CS regularity condition, we select the outcome with the maximum number of actions. As a bonus, we also obtain a selection of the messages used in equilibrium. That is, our approach also eliminates message indeterminacy. Only the highest messages are used in the selected equilibrium. We find it intuitive that the Sender's upward bias leads to exaggeration.

Key to our analysis is identifying two sequences of strategy profiles. The sequences are defined by specifying an initial condition and then iterating best responses. One sequence starts with the highest strategy profile; the other starts with the lowest.[2] We show that these sequences are monotonic in a suitable sense and converge to equilibria. Furthermore, under the CS regularity condition, the two sequences have a common limit whose outcome is the CS equilibrium outcome with the maximum number of actions. Our results on limits of best-response dynamics follow because any sequence of best responses must be sandwiched between the highest and lowest sequence. Our results on iterated deletion follow because we can show that strategies larger than the higher limit or lower that the lower limit must eventually be deleted.

This paper collects the main ideas common to earlier working papers by various subsets of the authors, namely, Gordon [15], Gordon [16], Kartik and Sobel [20], Lo and Olszewski [23], Lo [24], and Lo [25]. The working papers contain analyses for other structures of conflict of interest, and other results that are more general, or simply different, in certain directions than those presented here. This paper concentrates on what we view as the most significant findings from the working papers.

The paper proceeds as follows. Section 2 introduces the basic cheap-talk game. Section 3 contains a simple example that illustrates how the combination of a restriction to monotonic strategies and removal of weakly dominated strategies has the power to select an equilibrium. Section 4 contains some preliminary results for our general analysis, and Section 5 the main results. Section 6 presents an example that illustrates the proof technique. Section 7 interprets the main result and connects it to the literature. The proofs

[1]To be precise, they have many uninformative equilibria when one takes into account the first two kinds of multiplicity.

[2]Strategies can be partially ordered under our monotonic strategies restriction.

are in the appendices.[3]

2 The Basic Cheap-Talk Model

2.1 The cheap-talk game

We study the following basic cheap-talk game, as in CS. There are two players: the Sender (S) and the Receiver (R). They respectively have utility functions $u^S(a,t)$ and $u^R(a,t)$, where $a \in \mathbb{R}$ is an action taken by the Receiver and $t \in [0,1]$ is the Sender's type. We assume that for $j = S, R$, $u^j : \mathbb{R} \times [0,1] \to \mathbb{R}$ is twice continuously differentiable, strictly concave in its first argument, and has strictly positive mixed partial derivative. The Sender's type is drawn from a strictly positive, continuous density f on $[0,1]$.

For $j = S, R$, we assume that $\max_{a\in\mathbb{R}} u^j(a,t)$ exists and let $a^j(t)$ be the unique maximizer. For $t' < t''$, let

$$a^R(t', t'') = \arg\max_{a\in\mathbb{R}} \int_{t'}^{t''} u^R(a,t) f(t) dt$$

and let $a^R(t,t) = a^R(t)$. We assume that there is an upward bias: $a^R(t) < a^S(\ t)$ for all t. We normalize preferences so that $a^R(0) = 0$ and $a^R(1) = 1$. With this normalization, actions outside of the unit interval are dominated for R and would not be part of any equilibrium outcome. We therefore restrict the action set to be $[0,1]$.

The timing is that the Sender privately learns the type t, sends the Receiver a message m, and the Receiver then takes an action a. More precisely, let M be a finite set of available messages. A (pure) strategy for S is a mapping $s : [0,1] \to M$ that associates with every type t the message $s(t)$. A pure strategy for R is a mapping $a : M \to \mathbb{R}$ that associates with every message m an action $a(m)$. Given a strategy profile (s,a) the payoff of player j is $\int_0^1 u^j(a(s(t)),t)f(t)dt$. We denote R's (pure) strategy set by $\mathcal{A}$ and S's (pure) strategy set by $\mathcal{S}$.

We impose the following monotonicity restriction. Let N be the number of available messages, and let the messages be ordered as $m_1 < \cdots < m_N$. Assume M is linearly ordered, and denote the order by $\leq$. A strategy $a(\cdot)$ for R is *monotonic* if $m < m'$ implies that $a(m) \leq a(m')$, and a strategy $s(\cdot)$ for S is *monotonic* if $t < t'$ implies that $s(t) \leq s(t')$. The monotonicity restriction permits us to describe strategies in a convenient way. Every monotonic strategy of the Receiver can be uniquely represented by actions $a_1 \leq \cdots \leq a_N$ such that a_i for $i = 1, \ldots, N$ is induced by message m_i. We will identify a monotonic strategy of the Sender with cutoffs $t_0 \leq t_1 \leq \cdots \leq t_N$ such that $t_0 = 0$ and $t_N = 1$, with the interpretation that types in the interval (t_{i-1}, t_i) send the message m_i.

There is a natural partial order on monotonic strategies. For R, $(a_1, \ldots, a_N) \leq (a'_1, \ldots, a'_N)$ if and only if $a_i \leq a'_i$ for all i. For S, when strategies are represented by

[3] Appendix A contains details of the main argument. Appendix B presents an example that demonstrates that hthe limit of best-response dynamics may depend on the initial condition if our regularity condition fails. Supplementary Appendix C (not for publication) contains a spreadsheet associated with an example in Section 6.

cutoffs, $(t_0, \ldots, t_N) \leq (t'_0, \ldots, t'_N)$ if and only if $t_i \leq t'_i$ for all i. Although these are partial orders, there exists a largest and smallest strategy for both S and R.

Given a strategy profile (s, a), the associated type-action mapping γ is defined by $\gamma(t) = a(s(t))$. Every monotonic strategy profile uniquely determines the associated type-action mapping up to the actions induced by the cutoffs.

A *monotonic cheap-talk game* is derived from the basic cheap-talk game by assuming that players are allowed to play only monotonic strategies. We denote R's (pure) monotonic strategy set by $\mathcal{A}_0$ and S's (pure) monotonic strategy set by $\mathcal{S}_0$.

The monotonicity assumption is restrictive in the trivial sense that non-monotonic strategies exist. It is also restrictive in the stronger sense that non-monotonic strategies may be best responses even if the opponent is restricted to playing monotonic strategies. Finally, if we allowed for playing mixed strategies, then there would also exist mixtures of monotonic pure strategies with only non-monotonic best responses. However, the Sender's best response to any strictly monotonic strategy of the Receiver must be monotonic, and the Sender always has a monotonic best response to a monotonic strategy of the Receiver. Similarly, if the Sender plays a monotonic strategy, then the Receiver has a monotonic best response. In addition, any best response of the Receiver must be monotonic on the path, but the Receiver may also have a non-monotonic best response off the path.

Any equilibrium type-action mapping for the original game can be derived from monotonic strategies. To see this, use the result (described in Section 2.2) that any equilibrium involves a finite partition of the Sender's types into adjacent intervals. Construct a strategy for the Sender in which the types from higher partition elements send higher messages. Any best response to this strategy will be monotonic on the equilibrium path. One can define specific off-the-path actions to preserve monotonicity and support the equilibrium.

2.2 The structure of equilibria

CS demonstrate that there exists a positive integer N^* such that for every integer $1 \leq N \leq N^*$, there exists at least one equilibrium in which there are N induced actions, and moreover, there is no equilibrium that induces strictly more than N^* actions. Any equilibrium can be characterized by cutoffs $0 = t_0 < t_1 < \cdots < t_N = 1$, and actions $a_1 \leq \cdots \leq a_N$ such that

$$u^S(a_{i+1}, t_i) - u^S(a_i, t_i) = 0 \tag{1}$$

for $i = 1, \ldots, N-1$, and

$$a_i = a^R(t_{i-1}, t_i) \tag{2}$$

for $i = 1, \ldots, N$. In such an equilibrium, adjacent types pool and send a common message. Condition (1) states that the cutoff types are indifferent between pooling with types immediately below or immediately above. Condition (2) states that R best responds to the information in S's message. Except for the specification of the Sender's behavior at cutpoints, Conditions (1) and (2) uniquely describe an equilibrium relationship between types and actions.

CS introduce a regularity assumption that permits them to strengthen the characterization of equilibria. For $t_{i-1} \leq t_i \leq t_{i+1}$, let

$$V(t_{i-1}, t_i, t_{i+1}) = u^S(a^R(t_i, t_{i+1}), t_i) - u^S(a^R(t_{i-1}, t_i), t_i).$$

A (forward) solution to (1) of length L is a sequence $t_0, \ldots, t_L$ such that $V(t_{i-1}, t_i, t_{i+1}) = 0$ for $0 < i < L$ and $t_0 < t_1$.

Definition 1 (Regularity Condition)**.** *The cheap-talk game satisfies the Regularity Condition (RC) if for any two solutions to (1) of length L, $(t_0, \ldots, t_L)$ and $(t'_0, \ldots, t'_L)$ with $t_0 = t'_0$ and $t_1 < t'_1$, we have that $t_i < t'_i$ for all $i \geq 2$.*

(RC) is satisfied by the leading "uniform-quadratic" example in CS, which has been the focus of many applications. CS prove that if (RC) holds, then there is exactly one equilibrium type-action mapping (up to the Sender's cutoffs) for each $N = 1, \ldots, N^*$, and the ex-ante equilibrium expected utility for both S and R is increasing in N. These results provide an argument for the salience of the N^* equilibrium.

Another argument in support of this equilibrium outcome requires a definition.

Definition 2 (NITS)**.** *An equilibrium (a^*, s^*) satisfies the No Incentive to Separate (NITS) Condition if $u^S(a^*(m_1), 0) \geq u^S(a^R(0), 0)$.*

NITS states that the lowest type of the Sender prefers her equilibrium payoff to the payoff she would receive if the Receiver knew her type (and responded optimally). Chen, Kartik, and Sobel [8] show that only the (essentially unique) equilibrium type-action mapping with N^* actions induced satisfies NITS when (RC) holds. Uniqueness fails only because type t_i is indifferent between messages i and $i + 1$ for $i = 1, \ldots, N - 1$. From now on, we drop the word "essentially," and call such mappings unique.

3 Coordination Game Example

Although our main results concern the basic cheap-talk model, our first example is a coordination game. This will enable us to present the main ideas in the simplest manner. Examples in Section 6 are conducted within the basic cheap-talk model, but are more involved.

There are two types (high and low), two actions (high and low), and two messages (high and low). The types are equally likely. The Sender and Receiver have common interests. The players receive a payoff of two if the action matches the type and a payoff of zero otherwise. The strategic form of the game is given in the following table, where rows (columns) correspond to Sender (Receiver) strategies.

	(H,H)	(L,H)	(H,L)	(L,L)
(h,h)	$1,1$	$1,1$	$1,1$	$1,1$
(h,l)	$1,1$	$0,0$	$2,2$	$1,1$
(l,h)	$1,1$	$2,2$	$0,0$	$1,1$
(l,l)	$1,1$	$1,1$	$1,1$	$1,1$

A strategy for the Sender is a pair (i, j) where the Sender sends message i when her type is low and j when her type is high. Similarly, the first component in the Receiver's strategy is his response to a low message and the second is his response to a high messages. Hence (l, h) is the Sender strategy that reports a message that matches the state, (h, l) is the strategy that sends the high message when the state is low and the low message when the state is high. Similarly, (H, L) is the strategy of the Receiver that responds to the low message with the high action and the high message with the low action.

The game has an uninformative and ex-ante Pareto inefficient equilibrium in which S mixes equally between (h, l) and (l, h) and R mixes equally between (H, L) and (L, H). (There are also inefficient pure-strategy equilibria and other inefficient mixed equilibria.) There are also two efficient equilibria in which the Sender distinguishes between the states and the Receiver correctly interprets this information. The mixed-strategy equilibrium satisfies standard refinements from perfection to strategic stability, although it may seem intuitively implausible. Our approach is to replace the original game by a game in which non-monotonic strategies (with respect to the natural ordering on messages, types, and actions) are not available. The strategic form of the monotonic game is given in the following table.

	(H,H)	(L,H)	(L,L)
(h,h)	$1, 1$	$1, 1$	$1, 1$
(l,h)	$1, 1$	$2, 2$	$1, 1$
(l,l)	$1, 1$	$1, 1$	$1, 1$

Deleting non-monotonic strategies eliminates some inefficient equilibria, but it does not eliminate any equilibrium payoffs. However, weak dominance in the monotonic game selects the $(l, h), (L, H)$ equilibrium, which is efficient. We emphasize that the example demonstrates why our approach requires both a restriction to monotonic strategies and an equilibrium refinement. On one hand, weak dominance arguments in the original game have no power. On the other hand, the monotonic game still has multiple (Nash) equilibria.

Our subsequent analysis of the basic cheap-talk model has nuances in how weak dominance is applied, but the current example captures an essential idea. Namely, restricting attention to monotonic strategies eliminates some coordination problems. Without the restriction, every informative equilibrium type-action distribution can be supported in multiple ways by permuting the assignment of types to messages. Imposing an order on messages removes this indeterminacy.

Turning to dynamics, we note that the best-response dynamic does not converge to the efficient equilibrium for the initial conditions $(h, h), (H, H)$ or $(l, l), (L, L)$. However, the best-response dynamic converges to the efficient equilibrium for any initial conditions if we restrict attention to what we call robust best responses: a player's strategy is a robust best response to a strategy of the opponent if the player has close best responses to close strategies of the opponent.

Balkenborg, Hofbauer, and Kuzmics [2, Section 6] and Myerson and Weibull [28, Example 6] use this coordination example to illustrate the power of refinement arguments. Both of these papers select an efficient outcome. Balkenborg, Hofbauer, and Kuzmics eliminate the Sender's strategies (h, h) and (l, l) as "weakly inferior" (they are equivalent to a equal mixture of the other two strategies) and point out the efficient outcomes are the only locally stable equilibria of a refined best-reply dynamic that avoids these strategies. Myerson and Weibull show that only the efficient equilibria are settled.[4]

4 Concepts

Our main results are in Section 5. The first result, Theorem 1, describes the limiting behavior of certain sequences of best responses. It enables us to provide conditions under which these sequences converge to an equilibrium in which communication is effective. The second result, Theorem 2, describes the strategies that survive a procedure of iterated deletion of weakly dominated strategies. This section develops ideas needed to state and prove these results.

There are two subtleties with how we apply weak dominance compared to the example in Section 3. First, we need an iterative procedure, which is described in Section 4.2. Second, at times we require a modification of weak dominance that we impose at the interim stage. Section 4.3 introduces the concept of interim weak dominance.

If we start at an equilibrium, then the sequence of best replies remains constant. We argue, however, that best-reply sequences can be used to select equilibria. To do so, we must refine the best-reply correspondence. Section 4.1 introduces a refinement that we use in the proof of Theorem 1. The proof of our dynamic result relies on the observation that a sequence of strategies generated by taking "refined best replies" is bounded by a pair of sequences of strategies that we can describe explicitly. We will show that the same sequence provides bounds for strategies that survive an iterative process of deleting weakly dominated strategies. Section 4.4 describes the bounding sequences.

4.1 Robust best responses

In our discussion on learning, we will refer to the concept of *robust best response*. A strategy s of the Sender is a robust best response to a strategy a of the Receiver if for every strategy a' close to a, there is a best response s' to a' which is close to s. More precisely, for every $\varepsilon > 0$ there is a $\delta > 0$ such that if a' is such that $|a'_i - a_i| < \delta$ for $i = 1, ..., N$, then there is a best response s' to a' such that $|s'_i - s_i| < \varepsilon$ for $i = 0, ..., N$. A robust best response a of the Receiver to a strategy s of the Sender is defined similarly.

For example, there are many best responses to the Sender's strategy $(0, 1/2, 1/2, 1)$, because message m_2 is used with probability 0, and the best response to this message is not uniquely determined. However, the unique robust best response must prescribe action $a^R(1/2)$ in response to m_2, because the best responses to strategies $(0, t_1, t_2, 1)$ for $t_1 < t_2$ close to $1/2$ must prescribe actions close to $a^R(1/2)$. More generally, there is a unique best response of the Receiver to any strategy $t_0 < t_1 < ... < t_N$. Because the best response varies continuously with t, there is a unique robust best response of the Receiver to any strategy of the Sender.

[4] Settled equilibria are a refinement of the set of (Kalai-Samet [18]) proper equilibria.

Similarly, any strategy of the Sender is a best response to the Receiver's strategy $(1/2, 1/2, 1/2)$. However, the unique robust best response prescribes message m_1 for types from $[0, t_1)$, message m_2 for type from t_1, and message m_3 for types from $(t_1, 1]$, where t_1 is the type of the Sender for whom $1/2$ is the most preferred action. More generally, there is a unique best response of the Sender to any strategy $a_1 < a_2 < \cdots < a_N$. Because the best response varies continuously with a, there is a unique robust best response of the Sender to any strategy of the Receiver.

4.2 Iterated deletion of weakly dominated strategies

Let $U^j(s, a)$ for j =S, R be the payoff of player j given the strategy profile (s, a). Consider a game with strategy sets $\tilde{\mathcal{S}} \subset \mathcal{S}$ for S and $\tilde{\mathcal{A}} \subset \mathcal{A}$ for R. A strategy $s \in \tilde{\mathcal{S}}$ *weakly dominates* a strategy $s' \in \tilde{\mathcal{S}}$ if $U^S(s, a) \geq U^S(s', a)$ for every strategy $a \in \tilde{\mathcal{A}}$ and $U^S(s, \tilde{a}) > U^S(s', \tilde{a})$ for some strategy $\tilde{a} \in \tilde{\mathcal{A}}$; a strategy $a \in \tilde{\mathcal{A}}$ *weakly dominates* a strategy $a' \in \tilde{\mathcal{A}}$ if $U^R(s, a) \geq U^R(s, a')$ for every strategy $s \in \tilde{\mathcal{S}}$ and $U^R(\tilde{s}, a) > U^R(\tilde{s}, a')$ for some strategy $\tilde{s} \in \tilde{\mathcal{S}}$. A general procedure of iterated deletion of weakly dominated strategies (IDWDS) produces a sequence of sets $\widetilde{\mathcal{S}}_k$ and $\widetilde{\mathcal{A}}_k$ such that:

1. $\widetilde{\mathcal{S}}_0 = \mathcal{S}_0$, $\widetilde{\mathcal{A}}_0 = \mathcal{A}_0$;
2. $\widetilde{\mathcal{S}}_k$ is a subset of $\widetilde{\mathcal{S}}_{k-1}$ obtained by deleting a (possibly empty) subset of S's weakly dominated strategies in the game with strategy sets $(\widetilde{\mathcal{S}}_{k-1}, \widetilde{\mathcal{A}}_{k-1})$;
3. $\widetilde{\mathcal{A}}_k$ is a subset of $\widetilde{\mathcal{A}}_{k-1}$ obtained by deleting a (possibly empty) subset of R's weakly dominated strategies in the game with strategy sets $(\widetilde{\mathcal{S}}_{k-1}, \widetilde{\mathcal{A}}_{k-1})$;
4. The sets $\widetilde{\mathcal{S}}^* = \bigcap_{k=0}^{\infty} \widetilde{\mathcal{S}}_k$ and $\widetilde{\mathcal{A}}^* = \bigcap_{k=0}^{\infty} \widetilde{\mathcal{A}}_k$ are non-empty.
5. There are no weakly dominated strategies in either $\widetilde{\mathcal{S}}^*$ or $\widetilde{\mathcal{A}}^*$ in the game with strategy sets $(\widetilde{\mathcal{S}}^*, \widetilde{\mathcal{A}}^*)$.

The second and third conditions permit the deletion of only some weakly dominated strategies (and for deletions to be simultaneous). The fourth condition guarantees that the limit of the process exists. The fifth condition guarantees that the process continues as long as weakly dominated strategies remain. In general games (beyond our setting), there may exist no procedure satisfying our conditions (see Lipman [22]), and $(\widetilde{\mathcal{S}}^*, \widetilde{\mathcal{A}}^*)$ may depend on the order of deletion (see Dufwenberg and Stegeman [12]).

Our extension of IDWDS to games with infinite sets of strategies coincides with that of Dufwenberg and Stegeman [12]. Alternatively, one may consider procedures that produce transfinite sequences of sets $\widetilde{\mathcal{S}}_\kappa$ and $\widetilde{\mathcal{A}}_\kappa$ (κ stands here for an ordinal number), or define sets $\widetilde{\mathcal{S}}^*$ or $\widetilde{\mathcal{A}}^*$ in terms of stable sets (see Chen, Long, and Luo [6]).

4.3 Interim dominance

Our results on iterative deletion of dominated strategies can be formulated by referring to a different notion of dominance.

Consider a game with strategy sets $\tilde{\mathcal{S}} \subset \mathcal{S}$ for S and $\tilde{\mathcal{A}} \subset \mathcal{A}$ for R. A strategy $a \in \tilde{\mathcal{A}}$ *interim weakly dominates* a strategy $a' \in \tilde{\mathcal{A}}$ if for every strategy $s \in \tilde{\mathcal{S}}$ and message m used by s with positive probability, strategy a yields a weakly higher payoff than that of strategy a', both contingent on the Sender's types sending message m; in addition, the payoff of a is strictly higher than that of a' for at least one strategy of the Sender in $\tilde{\mathcal{S}}$ and at least one message m. For the Sender, a strategy s interim dominates a strategy s' if every type of the Sender weakly prefers s to s' for every strategy of the Receiver, and some type strictly prefers s to s' for some strategy of the Receiver.

For the Receiver interim dominance implies dominance. For the Sender, a strategy may interim dominate another strategy without weakly dominating the other strategy. This possibility arises if the interim domination is the result of strictly better performance of types whose probability is zero. A strategy may weakly dominate another strategy without dominating it in the interim sense. This possibility arises if the weakly dominating strategy fails to respond optimally on a set of types with probability zero. In addition, weak dominance (for both players) allows for compensating for an inferior response to one message or of some types by a superior response to another message or of some other types.

We can modify the definition of IDWDS by replacing weakly dominated by interim weakly dominated to obtain the concept of IDIWDS.

4.4 Equilibrium Bounds

Our construction uses two sequences of strategy profiles. The sequences consist of best responses to an initial specification of strategies for S and R. One specification begins with the lowest possible strategies; the other specification begins with the highest possible strategies. Given initial conditions, the best response property does not uniquely define the sequences because best responses need not be unique. We will show, however, that there is a way to select best responses that guarantee that the lower sequence is increasing and converges to an equilibrium and the upper sequence is decreasing and converges to an equilibrium. Furthermore, when (RC) holds, the limit of the lower sequence is equal to the limit of the upper sequence. To show convergence of the best-response dynamic we show that any sequence of strategy profiles generated by interim best responses remains sandwiched between the lower and upper sequence. Hence, whenever the lower and upper sequences have a common limit, any sequence converges to the same limit. To show that iterated deletion of strategies selects a particular outcome, we show that strategies that are less than the lower limit and strategies that are greater than the upper limit are eventually deleted.

We now describe the lower and upper sequences consistent with the best-response dynamic. Let

$$0 = \underline{t}_0^0 = \cdots = \underline{t}_{N-1}^0 < \underline{t}_N^0 = 1 \text{ and } 0 = \overline{t}_0^0 < \overline{t}_1^0 = \cdots = \overline{t}_N^0 = 1,$$

$$\underline{a}_1^0 = \cdots = \underline{a}_N^0 = 0 \text{ and } \overline{a}_1^0 = \cdots = \overline{a}_N^0 = 1.$$

By induction, let $(\underline{a}_1^{k+1}, \ldots, \underline{a}_N^{k+1})$ be defined as the interim best response of the Receiver to strategy $(\underline{t}_0^k, \underline{t}_1^k, \ldots, \underline{t}_N^k)$ of the Sender, and let $(\overline{a}_1^{k+1}, \ldots, \overline{a}_N^{k+1})$ be the interim best response of the Receiver to strategy $(\overline{t}_0^k, \overline{t}_1^k, \ldots, \overline{t}_N^k)$ of the Sender. These best responses specify optimal actions contingent on all messages, even the messages that correspond to degenerate intervals of the Sender's strategy. Because the interim best responses of the Receiver are unique, actions $(\underline{a}_1^{k+1}, \ldots, \underline{a}_N^{k+1})$ and $(\overline{a}_1^{k+1}, \ldots, \overline{a}_N^{k+1})$ are uniquely defined.

Let $(\underline{t}_0^{k+1}, \underline{t}_1^{k+1}, \ldots, \underline{t}_N^{k+1})$ be a best response of the Sender to strategy $(\underline{a}_1^k, \ldots, \underline{a}_N^k)$ of the Receiver, and let $(\overline{t}_0^{k+1}, \overline{t}_1^{k+1}, \ldots, \overline{t}_N^{k+1})$ be a best response of the Sender to strategy $(\overline{a}_1^k, \ldots, \overline{a}_N^k)$ of the Receiver. Because the Sender has more than one best response to any strategy of the Receiver such that $a_i = a_{i+1}$ for some i, we must pick among them. For $(\underline{t}_0^{k+1}, \underline{t}_1^{k+1}, \ldots, \underline{t}_N^{k+1})$, we pick the smallest best response to $(\underline{a}_1^k, \ldots, \underline{a}_N^k)$, that is, the best response such that if $(t_0, t_1, \ldots, t_N)$ is another best response of the Sender to $(\underline{a}_1^k, \ldots, \underline{a}_N^k)$, then $t_i \geq \underline{t}_i^{k+1}$ for $i = 0, 1, \ldots, N$. The smallest best response exists. We pick for $\underline{t}_i^{k+1}$ the lowest type that weakly prefers a_i to all strictly lower actions in the profile $(\underline{a}_1^k, \ldots, \underline{a}_N^k)$. For $(\overline{t}_0^{k+1}, \overline{t}_1^{k+1}, \ldots, \overline{t}_N^{k+1})$, we pick greatest best response, that is, the best response such that if $(t_0, t_1, \ldots, t_N)$ is another best response of the Sender to $(\overline{a}_1^k, \ldots, \overline{a}_N^k)$, then $t_i \leq \overline{t}_i^{k+1}$ for $i = 0, \ldots, N$, except for all i such that $a_i = a_{i+1} = 0$ when we pick $t_i = 0$. This requires picking for $\overline{t}_i^{k+1}$ the highest type that weakly prefers a_i to all strictly higher actions in the profile $(\overline{a}_1^k, \ldots, \overline{a}_N^k)$.

We now describe the construction of the sequence of upper bounds in more detail. First, notice that $\overline{a}_1^1 < 1$ and $\overline{a}_2^1 = \cdots = \overline{a}_N^1 = 1$; in turn,

$$\overline{t}_i^1 = \overline{t}_i^0 = 1 \tag{3}$$

for $i = 1, \ldots, N$ and therefore

$$\overline{a}_i^2 = \overline{a}_i^1 \tag{4}$$

for $i = 1, \ldots, N$. Continuing, $\overline{t}_1^2$ is the type that is indifferent between actions $\overline{a}_1^1$ and $\overline{a}_2^1$ ($\overline{t}_1^2 = 0$ if no such type exists), and $\overline{t}_2^2 = \cdots = \overline{t}_N^2 = 1$. If all types prefer $\overline{a}_2^1$ to $\overline{a}_1^1$, then $\overline{t}_1^2 = 0$. Further, $\overline{a}_1^3 < \overline{a}_2^3 < \overline{a}_3^3 = \cdots = \overline{a}_N^3 = 1$, and $\overline{t}_1^4$ is the type that is indifferent between actions $\overline{a}_1^3$ and $\overline{a}_2^3$, $\overline{t}_2^4$ is the type that is indifferent between actions $\overline{a}_2^3$ and $\overline{a}_3^3$, and $\overline{t}_3^4 = \cdots = \overline{t}_N^4 = 1$. If all types prefer $\overline{a}_2^3$ to $\overline{a}_1^3$, then $\overline{t}_1^4 = 0$. And if all types prefer $\overline{a}_3^3$ to $\overline{a}_2^3$, then also $\overline{t}_2^4 = 0$. We continue in this fashion. Note that equations (3) and (4) imply that $\overline{t}_i^{k+1} = \overline{t}_i^k$ when k is even and $\overline{a}_i^{k+1} = \overline{a}_i^k$ when k is odd. After $N-1$ changes to R's strategy, we reach a stage k^* such that the strategy $(\overline{a}_1^{k^*}, \ldots, \overline{a}_N^{k^*})$ has all positive actions different, and from that moment (that is, for $k \geq k^* = 2N-3$), the Sender's best responses to $(\overline{a}_1^k, \ldots, \overline{a}_N^k)$ with the property that $t_i = 0$ when $a_i = a_{i+1} = 0$ are unique.[5]

Thus, sequences $(\underline{t}_i^k)_{k=1}^\infty$, $(\overline{t}_i^k)_{k=1}^\infty$, $(\underline{a}_i^k)_{k=1}^\infty$ and $(\overline{a}_i^k)_{k=1}^\infty$ are well-defined by induction. Recall now two properties of the basic cheap-talk model:

(i) the Receiver's optimal action $a^R(t_l, t_h)$ given the belief that the Sender's types belong to an interval $[t_l, t_h]$ strictly increases in t_l and in t_h;

[5]Because we assume upward bias, $\underline{t}_{N-1}^k < \underline{t}_N^k = 1$ and $\overline{t}_{N-1}^k < \overline{t}_N^k = 1$ for $k > k^*$. Therefore, positive actions will be distinct.

(ii) for each pair of actions $a_l \leq a_h$, if there is a type that is indifferent between actions a_l and a_h,[6] then this type strictly increases in a_l and in a_h.

An inductive proof referring to (i) and (ii) shows that sequences $(\underline{t}_i^k)_{k=1}^{\infty}$, $(\overline{t}_i^k)_{k=1}^{\infty}$, $(\underline{a}_i^k)_{k=1}^{\infty}$ and $(\overline{a}_i^k)_{k=1}^{\infty}$ are monotonic:

$$\underline{t}_i^k \leq \underline{t}_i^{k+1} \text{ and } \overline{t}_i^k \geq \overline{t}_i^{k+1}$$

and

$$\underline{a}_i^k \leq \underline{a}_i^{k+1} \text{ and } \overline{a}_i^k \geq \overline{a}_i^{k+1}$$

for all i and k.[7] Monotonicity implies convergence, so $\underline{t}_i^k \to_k \underline{t}_i^*$, $\overline{t}_i^k \to_k \overline{t}_i^*$, $\underline{a}_i^k \to_k \underline{a}_i^*$ and $\overline{a}_i^k \to_k \overline{a}_i^*$. The limit actions and cutoff strategies are an equilibrium.

5 Results

We now state the two main results of our paper.

5.1 Convergence of best-response sequences

Definition 3. *Sequences of strategies $(t^k)_{k=0}^{\infty}$ and $(a^k)_{k=0}^{\infty}$, where $t^k = (t_0^k, \ldots, t_N^k)$ and $a^k = (a_1^k, \ldots, a_N^k)$, are called (interim, robust) best-response sequences if for $k = 0, 1, \ldots$, strategy a^{k+1} is the Receiver's (interim, robust) best response to the strategy t^k of the Sender, and strategy t^{k+1} is the Sender's (interim, robust) best response, to the strategy a^k of the Receiver.*

The sequence of robust best responses $(t_i^k)_{k=0}^{\infty}$ and $(a_i^k)_{k=0}^{\infty}$ starting from any initial conditions $(t_1^0, \ldots, t_N^0)$ and $(a_1^0, \ldots, a_N^0)$ is sandwiched between $(\underline{t}_i^k)_{k=0}^{\infty}$ and $(\overline{t}_i^k)_{k=0}^{\infty}$, and $(\underline{a}_i^k)_{k=0}^{\infty}$ and $(\overline{a}_i^k)_{k=0}^{\infty}$, respectively. That is,

$$\underline{t}_i^k \leq t_i^k \leq \overline{t}_i^k \text{ and } \underline{a}_i^k \leq a_i^k \leq \overline{a}_i^k$$

for all k and i. These inequalities follow by induction from the monotonicity of the greatest best responses and the smallest best responses with respect to the opponent's strategy. This yields the following result.

Theorem 1. *For any robust best-response sequences $(t^k)_{k=0}^{\infty}$ and $(a^k)_{k=0}^{\infty}$ we have that*

$$\underline{t}_i^* \leq \liminf_k t_i^k \leq \limsup_k t_i^k \leq \overline{t}_i^*$$

for $i = 1, \ldots, N-1$, and

$$\underline{a}_i^* \leq \liminf_k a_i^k \leq \limsup_k a_i^k \leq \overline{a}_i^*$$

for $i = 1, \ldots, N$.

[6] For $a_l = a_h$, we take this to be the type for whom $a_l = a_h$ is the most-preferred action.

[7] Actually, they are eventually strictly monotonic, except at the lower end, but we will not need this property.

In Appendix A, we show that Theorem 1 implies the following corollary.

Corollary 1. *Assume $N \geq N^*$. If there exists a unique equilibrium type-action mapping that satisfies NITS, then any robust best-response sequence converges to an equilibrium with this type-action mapping.*

In follows from Claim 2 in Appendix A that the limit equilibrium uses robust best responses. In addition, any equilibrium using robust best responses in a monotonic cheap-talk game must satisfy NITS if there are at least N^* messages. (See Appendix B for the proof.)

Appendix B also contains an example in which there exist two equilibrium type-action mappings that satisfy NITS. The example illustrates a general property. When there are multiple type-action distributions that satisfy NITS, then the limit of a robust best-response sequence will depend on the initial condition. In particular, if the initial conditions specify that the highest messages induce the on-path actions of a NITS equilibrium and all lower messages induce the action 0, then the sequence of robust best responses is constant.

5.2 Characterization of Iterated Undominated Strategies

We present results parallel to those from the previous section for iterated deletion of weakly dominated strategies.

Theorem 2. *There exists a procedure of iterated deletion of interim weakly dominated strategies such that the sets $\widetilde{\mathcal{S}}^*$ and $\widetilde{\mathcal{A}}^*$ consist of the strategies $(t_0, \ldots, t_N)$ and $(a_1, \ldots, a_N)$ such that*

$$\underline{t}_i^* \leq t_i \leq \overline{t}_i^* \text{ for } i = 0, \ldots, N \text{ and } \underline{a}_i^* \leq a_i \leq \overline{a}_i^* \text{ for } i = 1, \ldots, N. \quad (5)$$

In every round of deletion in the procedure we use to prove Theorem 2 we eliminate strategies that are weakly dominated and interim weakly dominated. However, there may be weakly dominated strategies that satisfy condition (5). Therefore, Theorem 2 does not imply that there exists a procedure of deleting weakly dominated strategies in which $(\tilde{\mathcal{S}}^*, \tilde{\mathcal{A}}^*)$ is (5). Therefore, we do not offer an analogue of Theorem 2 for weakly dominated strategies. However, the following corollary holds true for both interim dominance and dominance if there exists a unique equilibrium type-action mapping that satisfies NITS. The corollary follows from Theorem 2 because the uniqueness of NITS type-action mapping implies that $\underline{t}_i^* = \overline{t}_i^*$ for $i = 0, \ldots, N$ and $\underline{a}_i^* = \overline{a}_i^*$ for $i = 1, \ldots, N$.

Corollary 2. *Assume $N \geq N^*$. If there exists a unique equilibrium type-action mapping that satisfies NITS, then there is a procedure of iterated deletion of weakly dominated and interim weakly dominated strategies that retains only this type-action mapping. Furthermore, the surviving strategy uses only the highest N^* messages with positive probability.*

Corollary 2 states that IDWDS selects a unique type-action mapping when there is a unique equilibrium type-action mapping that satisfies NITS. In this mapping, the Sender "exaggerates" by using only the highest messages. Consequently the combination of restriction to the monotonic cheap-talk game and elimination of weakly dominated strategies resolves message indeterminacy in addition to type-action indeterminacy.

We conjecture, but we have not managed to prove that a result similar to Corollary 2 holds for general IDWDS procedures. It is difficult to show that an arbitrary procedure eliminates some strategies such that some actions coincide or such that some cutoffs coincide. An inspection of the proof of Corollary 2 shows that the cases in which agents have multiple best responses are the only obstacles to obtaining an analogous result for general IDWDS.

6 Examples

6.1 Convergence of Best-Response Dynamic

The proof of Theorem 1 and Corollary 1 involves constructing sequences of best responses starting from two extreme initial conditions. If the initial condition is high, then the resulting sequence monotonically decreases to an outcome that satisfies NITS. If the initial condition is low, then the resulting sequence monotonically increases to an outcome that satisfies NITS. In the example, there is only one equilibrium that satisfies NITS. Hence the two sequences have a common limit. Our result follows because the sequence of robust best responses starting from an arbitrary initial condition is sandwiched between the two extreme sequences. However, to illustrate our results we directly analyze in this subsection the sequences of robust best responses in an example.

Suppose that the Sender's type is distributed uniformly on interval $[0, 1]$, and the players' utilities are: $u^S(a, t) = -(a - t - b)^2$ and $u^R(a, t) = -(a - t)^2$, where $b > 0$. In this case, there is an $N^* \geq 1$ such that for every $N \leq N^*$ there exists a unique equilibrium type-action mapping with N partition intervals (i.e., with N equilibrium actions). There exist no other equilibrium type-action mapping. For $b = 0.05$, we have that $N^* = 3$, i.e., the game has three equilibrium type-action mappings. In the largest of them, the types from $[t_0^*, t_1^*) = [0, 4/30)$ induce action $a_1^* = 2/30$, the types from $(t_1^*, t_2^*) = (4/30, 14/30)$ induce action $a_2^* = 9/30$, and the types from $(t_2^*, t_3^*] = (14/30, 1]$ induce action $a_3^* = 22/30$.

Assume the message space consists of three messages, $m_1 < m_2 < m_3$. Let $0 = t_0^0 \leq t_1^0 \leq t_2^0 \leq t_3^0 = 1$ denote the cutoffs of a strategy of the Sender in period 0. That is, the types from interval (t_{i-1}^0, t_i^0) send message m_i in period 0. Let (a_1^0, a_2^0, a_3^0) denote a strategy of the Receiver in period 0. Suppose that the Sender's cutoffs and the Receiver's actions in period $k+1$ are determined by the following equations:

$$t_i^{k+1} + 0.05 - a_i^k = a_{i+1}^k - t_i^{k+1} - 0.05, \text{ i.e., } t_i^{k+1} = \frac{a_{i+1}^k + a_i^k}{2} - 0.05, \tag{6}$$

for $i = 1, 2,$[8] and

$$a_i^{k+1} = \frac{t_{i-1}^k + t_i^k}{2}, \tag{7}$$

for $i = 1, 2, 3$. Recall that we always fix t_0^k at 0 and t_3^k at 1.

[8] If $a_{i+1}^k + a_i^k < .1$ so that the equation defining t_i^{k+1} has no solution in $[0, 1]$, we set $t_i^{k+1} = 0$, $i = 1, 2$.

Formulas (6)–(7) define best-responses of the players to the strategies of their opponents from period k. The best responses of the Sender are defined by imposing a specific tie-breaking rule. This guarantees that they are robust. For example, if $a_1^k = a_2^k = a_3^k = 0.5$, then any strategy of the Sender is a best response. Our formulas imply that the Sender chooses the strategy $t_1^{k+1} = t_2^{k+1} = 0.45$, which is the robust rest-response. In particular, the type-action mapping of babbling equilibria is the limit of a constant best-response sequence, but this sequence is not a robust best-responses sequence.

We show that $(t_1^k)_{k=0}^{\infty}$ and $(t_2^k)_{k=0}^{\infty}$ converge by considering two specific initial conditions: (i) $t_1^0 = t_2^0 = 0$ and $a_1^0 = a_2^0 = a_3^0 = 0$; (ii) $t_1^0 = t_2^0 = 1$ and $a_1^0 = a_2^0 = a_3^0 = 1$. Note these are not the bounding sequences we defined in Section 4.4. The bounding sequences must be defined more carefully, taking into account the multiplicity of the Sender's best responses when the Receiver's responses to some messages coincide. However, to obtain convergence of sequences defined by formulas (6)–(7), it is enough to consider bounding sequences that are also defined by these formulas. In case (i), since t_i^0 and a_i^0 take the lowest possible values, so $t_i^0 \le t_i^1$ for $i = 1, 2$ and $a_i^0 \le a_i^1$ for $i = 1, 2, 3$. This implies that the sequences $(t_1^k)_{k=0}^{\infty}$ and $(t_2^k)_{k=0}^{\infty}$ and $(a_i^k)_{k=0}^{\infty}$ for $i = 1, 2, 3$ are increasing. We obtain this from (6)–(7) by induction. So, $(t_1^k)_{k=0}^{\infty}$ and $(t_2^k)_{k=0}^{\infty}$ must converge to some t_1^* and t_2^*, and the sequences $(a_i^k)_{k=0}^{\infty}$, $i = 1, 2, 3$, must converge to some a_i^*. In addition, $t_1^{k+1} = t_1^k = t_1^*$, $t_2^{k+1} = t_2^k = t_2^*$, and $a_i^{k+1} = a_i^k = a_i^*$, $i = 1, 2, 3$, must satisfy (6)–(7). It follows that $t_1^* = 4/30$, $t_2^* = 14/30$, $a_1^* = 2/30$, $a_2^* = 9/30$, and $a_3^* = 22/30$.

In case (ii), the sequences $(t_1^k)_{k=0}^{\infty}$ and $(t_2^k)_{k=0}^{\infty}$ and $(a_i^k)_{k=0}^{\infty}$ for $i = 1, 2, 3$ are decreasing, but they also converge to $t_1^* = 4/30$, $t_2^* = 14/30$, $a_1^* = 2/30$, $a_2^* = 9/30$, and $a_3^* = 22/30$. Therefore, $(t_1^k)_{k=0}^{\infty}$ and $(t_2^k)_{k=0}^{\infty}$ for arbitrary initial conditions (t_1^0, t_2^0) and (a_1^0, a_2^0, a_3^0) converge to $t_1^* = 4/30$, $t_2^* = 14/30$, $a_1^* = 2/30$, $a_2^* = 9/30$, and $a_3^* = 22/30$, because they are "sandwiched" between the sequences from case (i) and case (ii). The sequences $(t_1^k)_{k=0}^{\infty}$ and $(t_2^k)_{k=0}^{\infty}$ may not be monotonic in general. For example, if $t_1^0 = 0.25$ and $t_2^0 = 0.75$, and (a_1^0, a_2^0, a_3^0) is the Receiver's best response to this strategy of the Sender, then $t_1^0 < t_1^1$ but $t_1^1 > t_1^2 > t_1^3 > \cdots$, while $t_2^0 > t_2^1 > t_2^2 > t_2^3 > \cdots$. Therefore, showing their convergence requires our slightly more subtle argument.

For the general case with possibly type-dependent biases, sequences $(t_1^k)_{k=0}^{\infty}$, $(t_2^k)_{k=0}^{\infty}$, and $(a_i^k)_{k=0}^{\infty}$ defined as in cases (i) and (ii) are monotonic, and their limits induce equilibrium type-action mappings. Also, the sandwich argument applies. This does not guarantee the convergence of sequences $(t_1^k)_{k=0}^{\infty}$, $(t_2^k)_{k=0}^{\infty}$, and $(a_i^k)_{k=0}^{\infty}$ for arbitrary initial conditions, because the two limit equilibrium type-action mappings: that from case (i) and that from case (ii), may not coincide.[9] However, if the number of available messages is N^* or higher, and there is only one equilibrium type-action mapping that satisfies NITS (e.g., (RC) is satisfied),[10] then these limit mappings must coincide, and so $(t_1^k)_{k=0}^{\infty}$, $(t_2^k)_{k=0}^{\infty}$, and $(a_i^k)_{k=0}^{\infty}$ for arbitrary initial conditions converge to the cutoffs and actions of

[9] Olszewski [30] shows that $(t_1^k)_{k=0}^{\infty}$, $(t_2^k)_{k=0}^{\infty}$, and $(a_i^k)_{k=0}^{\infty}$ actually converge for an arbitrary set of initial conditions. However, the initial conditions may affect the limit equilibrium type-action mapping.

[10] These conditions are satisfied in the present example.

this equilibrium.[11,12]

6.2 Iterated Deletion of Weakly Dominated Strategies

We will now use the uniform-quadratic example to illustrate the proofs of Theorem 2 and Corollary 2. The central idea is that the strategies such that the cutoff t_i is smaller than $\underline{t}_i^k$ or such that the cutoff t_i is greater than $\overline{t}_i^k$ defined in Section 4.4 are weakly dominated and are eventually deleted. So are the strategies such that some actions are smaller than $\underline{a}_i^k$ or such that some actions are greater than $\overline{a}_i^k$. In the example, all strategies that we delete will be interim weakly dominated and weakly dominated. For simplicity, we will refer to them as dominated.

The cutoffs $\underline{t}_0^0 = \underline{t}_1^0 = \underline{t}_2^0 = 0$ and $\underline{t}_3^0 = 1$ determine a monotonic strategy of the Sender. The strategy that responds with actions $\underline{a}_1^1 = \underline{a}_2^1 = 0$ and $\underline{a}_3^1 = 0.5$ is a best response of the Receiver to this strategy of the Sender. Any strategy such that $a_3 < 0.5 = \underline{a}_3^1$ is weakly dominated by the strategy $(a_1,\, a_2,\, \underline{a}_3^1)$. This follows because no matter what the strategy (t_0, t_1, t_2, t_3) of the Sender, the Receiver weakly prefers playing action 0.5 to playing any action $a < 0.5$ in response to message m_3 (i.e., the message sent by the types from $(t_2, t_3] = (t_2, 1])$, and she strictly prefers playing 0.5 to any $a < 0.5$ if $t_2 < 1$. Denote the set of dominated strategies described in this paragraph as $\underline{D}_1^R$.

The cutoffs $\overline{t}_1^0 = 0$ and $\overline{t}_1^0 = \overline{t}_2^0 = \overline{t}_3^0 = 1$ also determine a monotonic strategy of the Sender. The strategy that responds with actions $\overline{a}_1^1 = 0.5$ and $\overline{a}_2^1 = \overline{a}_3^1 = 1$ is a best response of the Receiver to this strategy of the Sender. Any strategy such that $a_1 > 0.5 = \overline{a}_1^1$ is weakly dominated by the strategy $(\overline{a}_1^1,\, a_2,\, a_3)$. Indeed, no matter what the strategy (t_0, t_1, t_2, t_3) of the Sender, the Receiver weakly prefers playing action 0.5 to playing any action $a > 0.5$ in response to message m_1, and she strictly prefers playing 0.5 to any $a > 0.5$ if $t_1 > 0$. Denote the set of dominated strategies described in this paragraph as $\overline{D}_1^R$. Note that there may exist other strategies of the Receiver (i.e., not belonging to $\underline{D}_1^R \cup \overline{D}_1^R$) that are weakly dominated, about which we make no claim.[13]

At this point, we have illustrated how IDWDS may remove strategies of the Receiver. We have established a non-trivial lower bound $(\underline{a}_1^1,\, \underline{a}_2^1,\, \underline{a}_3^1)$ and a non-trivial upper bound $(\overline{a}_1^1,\, \overline{a}_2^1,\, \overline{a}_3^1)$ on the retained strategies of the Receiver. It is plausible to conjecture that iterating the process will eliminate more strategies. In fact, we use the fact that we have deleted some of the Receiver's strategies to impose non-trivial lower and upper bounds on the Sender's strategies. Specifically, we can delete (as weakly dominated) all strategies of the Sender with some coordinate lower than the corresponding coordinate of $(\underline{t}_0^2,\, \underline{t}_1^2,\, \underline{t}_2^2,\, \underline{t}_3^2)$ or some coordinate higher than the corresponding coordinate of $(\overline{t}_0^2,\, \overline{t}_1^2,\, \overline{t}_2^2,\, \overline{t}_3^2)$. That is, we can delete any strategy of the Sender with a coordinate that is lower (respectively, higher) that the corresponding coordinate of the best response to the lower (respectively, upper) bound on the Receiver's strategies. These bounds on the Sender's

[11] In Appendix B, we give an example in which two equilibria satisfy NITS, and our result no longer holds true.

[12] By Proposition 1 in Chen et al. (2008), all equilibria with N^*-interval partition satisfy NITS, but there may exist more than one such equilibrium.

[13] Indeed, it can be checked that $a_1 = a_2 = 0$, $a_3 = 1$ is dominated by $a_1 = a_2 = 0.1$, $a_3 = 0.9$. However, the argument is not as obvious as for the strategies such that $a_3 < 0.5$ or $a_1 > 0.5$.

strategies make tighter bounds on the Receiver's strategies. We can delete (as weakly dominated) all strategies of the Receiver with a coordinate that specifies an action less than the best response to ($\underline{t}_0^2$, $\underline{t}_1^2$, $\underline{t}_2^2$, $\underline{t}_3^2$) or with a coordinate that specifies an action greater than the best response to ($\overline{t}_0^2$, $\overline{t}_1^2$, $\overline{t}_2^2$, $\overline{t}_3^2$). And we can continue in this fashion to obtain tighter and tighter bounds.

Formally, $\underline{t}_1^2 = 0$ and $\underline{t}_2^2 = 0.2$ (with $\underline{t}_0^2 = 0$ and $\underline{t}_3^2 = 1$) is a best response of the Sender to $\underline{a}_1^1 = \underline{a}_2^1 = 0$ and $\underline{a}_3^1 = 0.5$, and if the Receiver plays only strategies (a_1, a_2, a_3) from the complement of $\underline{D}_1^R$, any strategy (t_0, t_1, t_2, t_3) such that $t_2 < \underline{t}_2^2$ is weakly dominated by the strategy $(t_0, t_1, \underline{t}_2^2, t_3)$. This follows because types $t < 0.2$ weakly prefer actions a_2 to $a_3 \geq \max\{0.5, a_2\}$, and strictly so if $a_2 < 0.5$. Thus, types $t < 0.2$ will never induce action a_3. Denote the set of dominated strategies described in this paragraph as $\underline{D}_2^S$.

Similarly, $\overline{t}_1^2 = 0.7$ and $\overline{t}_2^2 = 1$ is a best response of the Sender to $\overline{a}_1^1 = 0.5$ and $\overline{a}_2^1 = \overline{a}_3^1 = 1$, and if the Receiver plays only strategies (a_1, a_2, a_3) from the complement of $\overline{D}_1^R$, any strategy (t_0, t_1, t_2, t_3) such that $t_1 > \overline{t}_1^2$ is weakly dominated by the strategy $(t_0, \overline{t}_1^2, t_2, t_3)$. Denote the set of dominated strategies described in this paragraph as $\overline{D}_2^S$. Again, there are other weakly dominated strategies of the Sender (which are not in $\underline{D}_2^S \cup \overline{D}_2^S$) about which we make no claim.

We can generalize this argument to obtain ascending sequences of dominated strategies $\underline{D}_k^R$ and $\overline{D}_k^R$ for odd k and $\underline{D}_k^S$ and $\overline{D}_k^S$ for even k.[14] The argument uses the equilibrium bounds $(\underline{a}_i^k)_{k=0}^\infty$, $(\overline{a}_i^k)_{k=0}^\infty$, $i = 1, 2, 3$, and $(\underline{t}_i^k)_{k=0}^\infty$, $(\overline{t}_i^k)_{k=0}^\infty$, $i = 1, 2$ constructed in Section 4.4.

At stage k of the process, (i) every strategy (a_1, a_2, a_3) such that $a_i < \underline{a}_i^k$ for at least one i is weakly dominated by the strategy $(\max\{a_1, \underline{a}_1^k\}, \max\{a_2, \underline{a}_2^k\}, \max\{a_3, \underline{a}_3^k\})$, provided that the Sender is restricted to playing strategies such that $t_i \geq \underline{t}_i^{k-1}$ for all i.[15] Similarly, (ii) every strategy (a_1, a_2, a_3) such that $a_i > \overline{a}_i^k$ for at least one i is weakly dominated by the strategy $(\min\{a_1, \overline{a}_1^k\}, \min\{a_2, \overline{a}_2^k\}, \min\{a_3, \overline{a}_3^k\})$,[16] provided that the Sender is restricted to playing strategies such that $t_i \leq \overline{t}_i^{k-1}$ for all i. The inductive argument is analogous to that for $k = 1$. Denote the set of strategies described in (i) and (ii) as $\underline{D}_k^R$ and $\overline{D}_k^R$, respectively. Note that $\underline{D}_{k-2}^R \subset \underline{D}_k^R$ and $\overline{D}_{k-2}^R \subset \overline{D}_k^R$, because sequences $(\underline{a}_i^k)_{k=1}^\infty$, $i = 1, 2, 3$, are increasing, and sequences $(\overline{a}_i^k)_{k=1}^\infty$, $i = 1, 2, 3$, are decreasing.

The strategy of the Sender given by $\underline{t}_1^k$ and $\underline{t}_2^k$ is a best response of the Sender to $(\underline{a}_1^{k-1}, \underline{a}_2^{k-1}, \underline{a}_3^{k-1})$, and if the Receiver plays only strategies (a_1, a_2, a_3) from the complement of $\underline{D}_{k-1}^R$, every strategy (t_0, t_1, t_2, t_3) such that $t_i < \underline{t}_i^k$ for at least one i is weakly dominated by the strategy given by $(t_0, \max\{t_1, \underline{t}_1^k\}, \max\{t_2, \underline{t}_2^k\}, t_3)$. Similarly, the strategy of the Sender given by $\overline{t}_1^k$ and $\overline{t}_2^k$ is a best response of the Sender to $(\overline{a}_1^{k-1}, \overline{a}_2^{k-1}, \overline{a}_3^{k-1})$, and if the Receiver plays only strategies (a_1, a_2, a_3) from the complement of $\overline{D}_{k-1}^R$, every strategy (t_0, t_1, t_2, t_3) such that $t_i > \overline{t}_i^k$ for at least one i is weakly dominated by the strategy $(t_0, \min\{t_1, \overline{t}_1^k\}, \min\{t_2, \overline{t}_2^k\}, t_3)$. The inductive argument is analogous to that for $k = 2$. Denote the set of dominated strategies described in this paragraph as $\underline{D}_k^S$ and

[14] Recall that $\overline{t}_i^{k+1} = \overline{t}_i^k$ and $\underline{t}_i^{k+1} = \underline{t}_i^k$ when k is even and $\overline{a}_i^{k+1} = \overline{a}_i^k$ and $\underline{a}_i^{k+1} = \underline{a}_i^k$ when k is odd. Hence we do not delete R's strategies when k is even or S's strategies when k is odd.

[15] Note that $(\max\{a_1, \underline{a}_1^k\}, \max\{a_2, \underline{a}_2^k\}, \max\{a_3, \underline{a}_3^k\}) \notin \underline{D}_{k-2}^R \cup \overline{D}_{k-2}^R$.

[16] Note that $(\min\{a_1, \overline{a}_1^k\}, \min\{a_2, \overline{a}_2^k\}, \min\{a_3, \overline{a}_3^k\}) \notin \underline{D}_{k-2}^R \cup \overline{D}_{k-2}^R$.

$\overline{D}_k^S$. Again, $\underline{D}_{k-2}^S \subset \underline{D}_k^S$ and $\overline{D}_{k-2}^S \subset \overline{D}_k^S$ by the monotonicity of sequences $(\underline{t}_i^k)_{k=1}^\infty$ and $(\overline{t}_i^k)_{k=1}^\infty$.

Because $(\underline{t}_i^k)_{k=0}^\infty$ and $(\overline{t}_i^k)_{k=0}^\infty$ converge to t_i^*, and $(\underline{a}_i^k)_{k=0}^\infty$ and $(\overline{a}_i^k)_{k=0}^\infty$ converge to a_i^* for $i = 1, 2, 3$, only the largest equilibrium belongs to the complement of the sets $\bigcup_{k \text{ odd}} (\underline{D}_k^R \cup \overline{D}_k^R)$ and $\bigcup_{k \text{ even}} (\underline{D}_k^S \cup \overline{D}_k^S)$. The largest equilibrium cannot be deleted under this or any other procedure of iterated deletion of weakly dominated strategies, because each equilibrium action a_i^*, $i = 1, 2, 3$, is the Receiver's unique best response to message m_i given the Sender's equilibrium strategy, and the Sender's equilibrium strategy is the unique best response to the Receiver's equilibrium strategy.

The discussion thus far leaves several issues unresolved. What forces lead the one- and two-interval equilibria to be deleted when there are three messages? What happens when there are more than three messages? What happens when there are only two messages? (How) do the arguments depend on the order of deletion of weakly dominated strategies?

Continue to think in terms of the deletion process that we have outlined. Assume that there are more than three messages. Arguments analogous to those for three messages imply that the only type-action mapping left under our IDWDS procedure would have $a_i^* = 0$ and $t_i^* = 0$ except the three highest i's. No type of the Sender would choose a message that induces action 0, because the three-interval equilibrium satisfies the NITS condition. In contrast, the one- and two-interval equilibria do not satisfy NITS. In particular, if a two-interval equilibrium were to survive with three (or more) available messages, then the Receiver's equilibrium responses would have to be 0, 2/10 and 7/10. Because NITS fails for the two-interval equilibrium, type 0 prefers action 0 to action 0.2. This property is key to our argument. We can show that if there is an unused message, then all lower messages are unused. But then any surviving equilibrium must satisfy NITS. Therefore, provided that there are at least N^* messages, only equilibria that satisfy NITS can survive.

Naturally, the process could not converge to the largest equilibrium if $N < N^*$. In this case, we can show that when the (RC) holds, the only type-action mapping left under our IDWDS procedure would be the (unique) equilibrium type-action mapping with N distinct actions.

The question of how the arguments depend on the order of deletion seems more involved. We will address this question in a companion paper.

7 Discussion

The literature contains different theoretical arguments that suggest why, under upward bias, the equilibrium with N^* actions is salient. Under their regularity condition, CS demonstrate that there is an essentially unique equilibrium type-action mapping with N^* actions and that, under some conditions, this equilibrium is ex ante preferred to all other equilibria by both the Sender and the Receiver.

Mensch [26] notes that monotonicity restrictions in cheap-talk games can lead to the kind of selection that we describe. Rather than impose monotonicity of strategies, Mensch imposes a monotonicity condition on off-path beliefs. This condition directly implies that the Receiver must respond to unsent messages with actions strictly lower than those on equilibrium path.

Milgrom and Roberts [27] and Vives [34] study the class of supermodular games introduced by Topkis [32]. In a supermodular game, each player's strategy set is partially ordered and there are strategic complementarities that cause a player's best response to be increasing in opponents' strategies. Milgrom and Roberts [27] demonstrate that supermodular games have a largest and smallest equilibrium and that these extreme equilibria can be obtained by iterating the best-response correspondence. Our argument uses similar techniques. There are two differences. Our game is not a supermodular game. In particular, it does not satisfy the increasing difference condition of Milgrom and Roberts. In addition, Milgrom and Roberts study the implications of deletion of strictly dominated strategies. Our analysis uses weak dominance. Sobel [31] shows how Milgrom and Roberts's general arguments extend to a broader class of games and a more restrictive solution concept. He points out that the monotonic cheap-talk game satisfies a weak form of supermodularity that makes it possible to bound the set of strategies that survive deletion of weakly dominated strategies using arguments similar to ours. Sobel does not provide conditions under which the process leads to a unique prediction.

Words have commonly accepted meanings. When there are no conflicts of interest, it is natural to assume that agents will use words in conventional ways. In strategic situations, however, sophisticated agents will not take words at face value. Standard models of cheap talk abstract from the conventional meaning of words in order to focus on strategic problems. A limitation of this approach is that meaning is determined completely endogenously. An equilibrium type-action mapping determines the minimum number of distinct messages that the Sender must use, but does not specify which message is associated with which action. If there is to be a connection between the equilibrium use of messages and exogenous meaning, then we must impose additional assumptions. The literature has approached this issue in several ways.

Farrell [13] introduced the first attempt to refine the equilibrium set in cheap-talk games. Farrell's notion of neologism-proof equilibrium models the idea that messages have commonly accepted meanings and that players are able to use these statements provided that they were consistent with strategy constraints. This general idea does refine the set of equilibria in cheap-talk games, but lacks general existence properties.[17]

Chen [7] and Kartik [19] assume that the message space is equal to the type space, which suggests a natural correspondence between types and messages. They make this connection operational by modifying the game. Chen assumes that with positive probability the Sender sends a message equal to her type (and with positive probability the Receiver interprets the message literally). Kartik assumes that the Sender has a cost of "lying." These perturbations create an exogenous meaning for messages. In these models, the limits of equilibria in monotonic strategies as the perturbations vanish converge to an equilibrium that satisfies NITS.[18] Furthermore, the limit equilibrium involves the use of "inflated" messages. Hence these arguments are alternative ways to make the same selection that we make. Our result imposes the monotonicity condition directly on the

[17]In particular, typically no equilibrium is neologism-proof in the uniform-quadratic special case of the CS model.

[18]Chen, Kartik, and Sobel [8] introduce the NITS criterion, which we described in Section 2.

game and makes a selection without perturbations.

Dilmé [11] also provides an argument that selects equilibrium outcomes with communication. Dilmé studies cheap-talk games in which payoffs are perturbed. He then looks for equilibria of the underlying game that are robust, where a robust equilibrium is close to some equilibrium in every nearby game. He shows that in games with a upward bias satisfying the standard regularity condition, only the equilibrium with the maximal number of actions induced is robust. He extends this result to more general cheap-talk games. Dilmé's selection generally coincides with the outcomes we select.[19] His approach has a superficial similarity to Chen, Kartik, and Sobel [8, Section 4.4], in that both operate by perturbing payoffs. Chen, Kartik, and Sobel study a particular kind of signaling cost introduced in Kartik [19] and impose an equilibrium refinement (restriction to monotonic strategies), while Dilmé uses the freedom to specify signaling costs to attain a selection result. Dilmé approach is also related to solution concepts like strategic stability (Kohlberg and Mertens [21]) or truly perfect equilibria (Van Damme [33]) that require robustness with respect to a large family of perturbations. In addition to reaching similar conclusions, the source of Dilmé's results is similar to ours. Both approaches exploit the fact that there are a limited number of specifications of off-path behavior that are consistent with equilibrium. For example, in a cheap-talk model in which the Sender is upward biased, equilibrium requires that off-path actions either agree with on-path actions or are strictly lower than the lowest on-path action. Furthermore, when a regularity condition holds, only the equilibrium with the maximal number of actions induced can be supported using low off-path responses. Dilmé's argument, like ours, operates by showing that some messages must lead to low off-path actions.

Forges and Sémirat [14] study a finite cheap talk game with upward bias. They examine general versions of the following procedure: Start with the finest partition. Let Receiver best reply. Find the highest type, if any, that prefers the action of a higher partition element to the action of its own partition element. Move this type to the next highest partition. Continue. They show that the limit is an undefeated equilibrium in the sense of Matthews, Okuno-Fujiwara, and Postlewaite. Similar to us, the authors present an adjustment process that converges to the "largest" equilibrium for some initial condition. The studies differ because they make different assumptions about the cardinality of the type space. Restricting to a finite type space allows Forges and Sémirat to start with an initial condition that is fully revealing. Forges and Sémirat focus on adjustment processes in which some types change to better replies (but adjustments are typically not best replies to the opponent's previous strategy). The processes in Forges and Sémirat start with a particular initial condition and always converge to the largest equilibrium. In our model, the largest equilibrium always survives iterated deletion of weakly dominated strategies and is the limit of best response dynamic from some initial conditions.

In the context of CS games where the sender has an upward bias, Gordon [16] studies a selection from the composed best response dynamics, defined directly on the set of interval partitions outcomes with any finite number of intervals, without keeping track of messages. Under regularity conditions, the only equilibrium that is stable for this

[19] Dilmé does not resolve message indeterminacy.

dynamics is the one with the maximal number of intervals. Theorem 1 and Corollary 1 generalize this result by showing that convergence is in fact global. Gordon [16] also obtains results on local stability of equilibria for other biases.

Antić and Persico [1] study a game in which the players make a costly investment that can alter ideal points prior to playing a cheap-talk game. They study equilibria of the two-stage game that satisfy a forward-induction refinement. A fixed cheap-talk game can be viewed as a two-stage game in which players face infinite costs associated with changing their biases. Antić and Persico identify conditions on the underlying cheap-talk game and the investment-cost function that imply that only an outcome that satisfies NITS is the limit of refined equilibria of the two-stage game as the investment costs grow to infinity. This argument selects the same type-action mapping as Chen [7] and Kartik [19] by examining limits of equilibria, but the logic of the arguments appears to be different. Chen and Kartik perturb payoffs, while Antić and Persico perturb strategy spaces. Furthermore, Chen and Kartik's selection, like ours, resolves the message-indeterminacy problem while Antić and Persico do not.

Clark and Fudenberg [9] introduce an equilibrium refinement (justified communication equilibrium) for signaling games with both cheap and costly signals. Justified communication equilibria are stable outcomes of learning processes. Assuming that Receivers trust cheap-talk messages initially, they show that these messages must satisfy off-path credibility conditions in stable outcomes. They provide conditions under which cheap-talk messages influence equilibrium outcomes in interesting classes of signaling games. Justified communication equilibria coincide with perfect Bayesian equilibria in cheap-talk games.

Blume [3] and [4] propose refinements for finite cheap-talk games based on Kalai and Samet's [18] concept of persistent equilibrium. In particular, Blume [4] demonstrates that these perturbations to the Sender's messages determine the relationship between types and messages in the equilibria selected by his refinement. These perturbations, like the initial conditions in our dynamic arguments, solve the message-indeterminacy problem and select informative equilibria in games with partial common interest.

Blume [5] introduces a concept of "language equilibrium" in cheap-talk games. He takes as given a distinguished Receiver strategy, called a pre-existing language. Given a cheap-talk game and a pre-existing language, he constructs games that use a subset of the strategy set of the original game by iterating best replies to the pre-existing language and then judiciously adding certain strategies to the resulting limit. He calls equilibria of these games language equilibria and establishes their existence. Similar to our approach, language equilibria resolve message indeterminacy. Blume also shows that in finite versions of the CS setting, language equilibria feature language inflation.

Olszewski [29] investigates the stability of equilibria in cheap-talk game with respect to the introduction of new messages and shows through examples that this idea destabilizes "implausible" equilibria. The initial conditions of our adaptive processes act like new messages do in Olszewski's paper. Hence the approaches share the feature of investigating conditions under which the introduction of novel interpretations of messages (either through the addition of new message that the Receiver interprets randomly as in Olszewski or a rich initial condition that the Receiver responds to optimally as in our paper) and adaptive dynamics can select equilibria.

Lo [24] imposes restrictions on the set of strategies available to agents in a discrete cheap-talk game and then studies the outcomes that survive deletion of weakly dominated strategies.[20] Like Lo, we impose restrictions on strategies and study the implications of IDWDS. Our results differ from hers because we impose only the restrictions that messages are linearly ordered, that higher sender types send weakly higher signals and that the receiver takes weakly higher actions for higher signals. These restrictions do not eliminate any equilibrium outcomes of the original game. Lo makes further restrictions on the strategy space and shows that these can actually lead to outcomes that are not equilibria of the original game.

There are several criticisms of IDWDS. It is well known that, unlike deletion of strongly dominated strategies, the order of deletion may matter. In some games with large strategy spaces, equilibrium may fail to exist in weakly undominated strategies.[21] The process of eliminating weakly dominated strategies may introduce new equilibria. It also leads to strong predictions that are not behaviorally accurate in common settings (like the centipede game).

Some of the technical problems with IDWDS may hold in our setting, but we conjecture a selection result independent of the order of deleting strategies that are interim weakly dominated.

Appendix A (Proofs)

Corollary 1. *Assume $N \geq N^*$. If there exists a unique equilibrium type-action mapping that satisfies NITS, then any robust best-response sequence converges to an equilibrium with this type-action mapping.*

Proof. We first state a property of the limit equilibria $(\underline{a}^*, \underline{t}^*)$ and $(\overline{a}^*, \overline{t}^*)$ that were defined in Section 4.4.

Claim 1. *No two messages induce the same action $a > 0$ in equilibrium $(\underline{a}^*, \underline{t}^*)$, and no two messages induce the same action $a > 0$ in equilibrium $(\overline{a}^*, \overline{t}^*)$.*[22]

Proof. We prove the result for $(\bar{a}^*, \overline{t}^*)$. The same argument applies to $(\underline{a}^*, \underline{t}^*)$. Suppose that action $\bar{a}_i > 0$. Because $\bar{a}_i^k \geq \bar{a}_i$ for all k and $\bar{a}_i^k = a^R(\overline{t}_i^{k-1}, \overline{t}_i^k)$, there exists a $\delta > 0$ such that $\overline{t}_i^k > \delta$ for all k. By construction, $u^S(\bar{a}_i^k, \overline{t}_i^{k+1}) = u^S(\bar{a}_{i+1}^k, \overline{t}_i^{k+1})$. It follows that $a^S(\overline{t}_i^{k+1}) \in (\bar{a}_i^k, \bar{a}_{i+1}^k)$. Furthermore,

$$\bar{a}_i^k = a^R(\overline{t}_{i-1}^k, \overline{t}_i^k) \leq a^R(\overline{t}_i^k) \ \leq a^R(\overline{t}_i^k, \overline{t}_{i+1}^k) \leq \bar{a}_{i+1}^k.$$

Because $a^S(t) > a^R(t)$ for all t, there exists an $\varepsilon > 0$ such that $u^S(t) - u^R(t) > \varepsilon$ for all $t \in [0, 1]$. It follows that if $a^R(\overline{t}_i^k) \in [\bar{a}_i^k, \bar{a}_{i+1}^k]$ and $a^S(\overline{t}_i^{k+1}) \in (\bar{a}_i^k, \bar{a}_{i+1}^k)$, for all k and the

[20] Lo [25] applies similar arguments to study cheap-talk extensions of games with complete information.

[21] A simple example is a first-price auction in which two surplus-maximizing agents bid for an item with known, common value. The only equilibrium of the game involves both players bidding the common value, but this strategy is weakly dominated by bidding less.

[22] If $N > N^*$, then there exist multiple messages that induce action $a = 0$. No message can induce the action $a = 1$ because we assume an upward bias.

sequence $(\bar{a}^k, \bar{t}^k)$ converges, then there exists K such that $\bar{a}^k_{i+1} - \bar{a}^k_i > \varepsilon/2$ for $k \geq K$. Because ε is independent of k, $\bar{a}^*_{i+1} > \bar{a}^*_i$. □

When $N \geq N^*$, Claim 1 implies that the equilibria $\underline{t}^* = (\underline{t}^*_0, \ldots, \underline{t}^*_N)$, $\underline{a}^* = (\underline{a}^*_1, \ldots, \underline{a}^*_N)$, and $\overline{t}^* = (\overline{t}^*_0, \ldots, \overline{t}^*_N)$, $\overline{a}^* = (\overline{a}^*_1, \ldots, \overline{a}^*_N)$ must satisfy NITS. This is so because either some messages induce action zero, or all messages induce positive actions. In the former case, NITS is satisfied because the Sender has the option of inducing action zero, but no type of the Sender chooses this option. In the latter case, $N = N^*$ and NITS is satisfied by Proposition 1 in Chen, Kartik, and Sobel [8].

Thus, if there is a unique equilibrium that satisfies NITS, then $\underline{t}^* = \overline{t}^*$ and $\underline{a}^* = \overline{a}^*$. This yields Corollary 1 by Theorem 1, because the sequences of best responses $(t^k_i)^\infty_{k=0}$ and $(a^k_i)^\infty_{k=0}$ starting from any initial conditions $(t^0_0, \ldots, t^0_N)$ and $(a^0_1, \ldots, a^0_N)$ are sandwiched between $(\underline{t}^k_i)^\infty_{k=0}$ and $(\overline{t}^k_i)^\infty_{k=0}$, and $(\underline{a}^k_i)^\infty_{k=0}$ and $(\overline{a}^k_i)^\infty_{k=0}$, respectively. □

We note a consequence of Claim 1 that we use in the proof of Corollary 2.

Claim 2. *If $i \leq N - N^*$, then $\overline{t}^*_i = \underline{t}^*_i = 0$.*

Proof. If R plays strategy a in an equilibrium that satisfies NITS, upward bias implies that either a_i is equal to an action that is induced with positive probability or is strictly less than all actions induced with positive probability. We know that $\overline{a}^*$ and $\underline{a}^*$ are actions from equilibria that satisfy NITS. Furthermore, no two messages induce the same positive action in equilibrium. Also at most N^* actions are induced. The remaining messages must therefore induce actions lower than the action induced with positive probability. By monotonicity, the lowest $N - N^*$ actions are induced with probability zero. □

Theorem 2. *There exists a procedure of iterated deletion of interim weakly dominated strategies such that the sets $\widetilde{\mathcal{S}}^*$ and $\widetilde{\mathcal{A}}^*$ consist of the strategies $(t_0, \ldots, t_N)$ and $(a_1, \ldots, a_N)$ such that*

$$\underline{t}^*_i \leq t_i \leq \overline{t}^*_i \text{ for } i = 0, \ldots, N \text{ and } \underline{a}^*_i \leq a_i \leq \overline{a}^*_i \text{ for } i = 1, \ldots, N. \tag{5}$$

Proof. We need to describe a procedure that in every round eliminates interim weakly dominated strategies and retains only the strategies between the equilibrium bounds defined in Section 4.4. Our procedure will eliminate all strategies other than the strategies such that:

$$t_i \leq \overline{t}^*_i \text{ and } a_i \leq \overline{a}^*_i,\ i = 1, \ldots, N. \tag{8}$$

An analogous procedure eliminates all strategies other than the strategies such that:

$$t_i \geq \underline{t}^*_i \text{ and } a_i \geq \underline{a}^*_i,\ i = 1, \ldots, N. \tag{9}$$

By applying the two procedures simultaneously, we obtain a procedure that retains only the strategies between the equilibrium bounds.

We will first describe a procedure that eliminates some strategies that may not weakly dominated. For these strategies there will exist strategies which yield a weakly higher payoff against any strategy of the opponent, but which not necessarily yield a strictly higher payoff against some strategy of the opponent. We will point out the place in which strategies that are not weakly dominated are eliminated, and then we will describe a more involved procedure that in every round eliminates interim weakly dominated strategies, and retains only the strategies between the equilibrium bounds defined in Section 4.4.

Lemma 1. *No strategy of the Sender that satisfies the first parts of (8) and (9), and no strategy of the Receiver that satisfies the second parts of (8) and (9) can be interim dominated when the opponent can use all strategies satisfying these two conditions.*

This will show that the strategies satisfying (8) and (9) are the only elements of sets $\widetilde{\mathcal{S}}^*$ and $\widetilde{\mathcal{A}}^*$ for our IDIWDS procedure.

Proof. We will prove Lemma 1 for the Sender's strategies; the argument is analogous for the Receiver's strategies. Take any $(t_0, \ldots, t_N)$ that satisfies the first parts of (8) and (9). If $\overline{t}_i^* > 0$ for some i, then t_i is indifferent between actions $a_i < a_{i+1}$ for some $(a_1, \ldots, a_N)$ that satisfies the second parts of (8) and (9). Indeed, $\underline{t}_i^* \leq t_i$ weakly prefers $\underline{a}_{i+1}^*$ to $\underline{a}_i^*$,[23] and $\overline{t}_i^*$ is indifferent between $\overline{a}_i^*$ and $\overline{a}_{i+1}^*$. Because $u^S(\cdot)$ has strictly positive mixed partial, this implies that t_i weakly prefers $\underline{a}_{i+1}^*$ to $\underline{a}_i^*$ and $\overline{a}_i^*$ to $\overline{a}_{i+1}^*$. Thus, by the intermediate value theorem, there exist a convex combination of $\left(\underline{a}_i^*, \underline{a}_{i+1}^*\right)$ and $\left(\overline{a}_i^*, \overline{a}_{i+1}^*\right)$, denoted by (a_i, a_{i+1}) such that t_i is indifferent between a_i and a_{i+1}. Thus, if $(t_0', \ldots, t_N')$ weakly interim dominates $(t_0, \ldots, t_N)$, then $t_i' = t_i$ for any i such that $\overline{t}_i^* > 0$. If $\overline{t}_i^* = 0$, then $t_i' = t_i = 0$ as well for any $(t_0', \ldots, t_N')$ that satisfies the first parts of (8) and (9). Thus $t_i' = t_i$ for all i, and $(t_0', \ldots, t_N')$ satisfying these two conditions cannot weakly interim dominate $(t_0, \ldots, t_N)$. □

The procedure that achieves (8) deletes in rounds $2k-1$ and $2k$ the strategies of the Receiver such that $a_i > \overline{a}_i^k$ for some $i = 1, \ldots, N$, and deletes the strategies of the Sender such that $t_i > \overline{t}_i^k$ for some $i = 1, \ldots, N$. This procedure retains only the strategies satisfying (8), because $\overline{t}_i^k \to_k \overline{t}_i^*$ and $\overline{a}_i^k \to_k \overline{a}_i^*$ for all i. It remains to show that in each round we eliminate interim weakly dominated strategies, provided that players can use in that round only the strategies not eliminated in the previous rounds. We will show this by induction.

For $k = 1$, it can be that $a_i > \overline{a}_i^k$ only when $i = 1$, because $\overline{a}_i^k = 1$ for $i > 1$. Any strategy $(a_1, a_2, \ldots, a_N)$ with $a_1 > \overline{a}_1^1$ is interim weakly dominated by the strategy $(\overline{a}_1^1, a_2, \ldots, a_N)$. Indeed, the comparison of the two strategies reduces to the payoff generated by the lowest action against the lowest interval $[0, t_1]$ of the Sender's strategy $(t_0, t_1, \ldots, t_N)$. This payoff is strictly greater for the latter strategy than for the former strategy when $t_1 > 0$, because the Receiver's best response to $[0, t_1]$ is always no higher than $\overline{a}_1^1$, which is the Receiver's best response to $[0, 1]$. When $t_1 = 0$, the payoffs are equal.

No strategy of the Sender is deleted for $k = 1$. This completes the first inductive step. However, to present our arguments for the Sender's strategies in the simplest non-trivial case, consider $k = 2$. It can be that $t_i > \overline{t}_i^2$ only when $i = 1$, because $\overline{t}_i^2 = 1$ for $i > 1$. The comparison of strategies $(t_0, t_1, \ldots, t_N)$ and $(t_0, \overline{t}_1^2, \ldots, t_N)$ reduces to comparing which cutoff t_1 or $\overline{t}_1^2$ is better against the Receiver's strategies $(a_1, a_2, \ldots, a_N)$ not eliminated in the first round. Because $a_1 \leq \overline{a}_1^1$ and type $\overline{t}_1^2$ is by definition indifferent between $\overline{a}_1^1$ and 1, cutoff $t_1 > \overline{t}_1^2$ is always worse than cutoff $\overline{t}_1^2$.

The inductive steps are similar. Consider a $k \geq 1$, and a strategy $(a_1, \ldots, a_N)$ such that $a_i \leq \overline{a}_i^k$ for all i and $a_i > \overline{a}_i^{k+1}$ for some i. Because $\overline{a}_i^{k+1} = 1$ for $i > k+1$,

[23] $\underline{t}_i^* > 0$ is indifferent between $\underline{a}_i^*$ to $\underline{a}_{i+1}^*$.

$a_i > \overline{a}_i^{k+1}$ implies $i \leq k+1$. Let i^* be the lowest index for which $a_{i^*} > \overline{a}_{i^*}^{k+1}$. Define a strategy $(b_1, \ldots, b_N)$ by letting $b_{i^*} = \overline{a}_{i^*}^{k+1}$, and $b_j = a_j$ for $j \neq i^*$. Then, $b_j \leq \overline{a}_j^k$ for all j, because $a_j \leq \overline{a}_j^k$ for all j and $\overline{a}_{i^*}^{k+1} \leq \overline{a}_{i^*}^k$. The strategy $(b_1, \ldots, b_N)$ is monotonic. Indeed, $b_{i^*-1} \leq b_{i^*}$ (when $i^* > 1$), because $b_{i^*-1} = a_{i^*-1} \leq \overline{a}_{i^*-1}^{k+1}$ and $b_{i^*} = \overline{a}_{i^*}^{k+1}$; $b_{i^*} \leq b_{i^*+1}$, because $b_{i^*} = \overline{a}_{i^*}^{k+1} < a_{i^*} \leq a_{i^*+1} = b_{i^*+1}$; and $b_j \leq b_{j+1}$ for all other j, because $a_j \leq a_{j+1}$. The comparison of the Receiver's payoffs from playing strategies $(a_1, \ldots, a_N)$ and $(b_1, \ldots, b_N)$ reduces to the comparison of the payoffs of actions a_{i^*} and $\overline{a}_{i^*}^{k+1}$ against intervals $[t_{i^*-1}, t_{i^*}]$ of the Sender's strategies $(t_0, t_1, \ldots, t_N)$ such that $t_{i^*-1} \leq \overline{t}_{i^*-1}^k$ and $t_{i^*} \leq \overline{t}_{i^*}^k$. Because $(\overline{a}_1^{k+1}, \ldots, \overline{a}_N^{k+1})$ is the Receiver's best response to $(\overline{t}_0^k, \overline{t}_1^k, \ldots, \overline{t}_N^k)$, $\overline{a}_{i^*}^{k+1}$ is weakly preferred to a_{i^*} against any such $(t_0, t_1, \ldots, t_N)$. Furthermore, it is strictly preferred against $(t_0, t_1, \ldots, t_N)$ when $t_{i^*-1}^k < t_{i^*}^k$. It can happen that $t_{i^*-1}^k$ is equal to $t_{i^*}^k$ for all strategies $(t_0, t_1, \ldots, t_N)$ such that $t_i \leq \overline{t}_i^k$ only when $\overline{t}_{i^*-1}^k = \overline{t}_{i^*}^k = 0$. In this case, $\overline{a}_{i^*}^{k+1}$ is strictly interim preferred to a_{i^*}, but $\overline{a}_{i^*}^{k+1}$ is not strictly preferred to a_{i^*}. This problem requires modifying our procedure. At the end of the proof, we will modify the procedure so that eliminated strategies are both weakly dominated and interim weakly dominated.

Consider now a strategy $(t_0, \ldots, t_N)$ such that $t_i \leq \overline{t}_i^k$ for all i and $t_i > \overline{t}_i^{k+1}$ for some i. Let i^* be the lowest index for which $t_{i^*} > \overline{t}_{i^*}^{k+1}$. Define a strategy $(s_0, \ldots, s_N)$ by letting $s_{i^*} = \overline{t}_{i^*}^{k+1}$, and $s_j = t_j$ for $j \neq i^*$. Then, $s_j \leq \overline{t}_j^k$ for all j, because $t_j \leq \overline{t}_j^k$ for all j and $\overline{t}_{i^*}^{k+1} \leq \overline{t}_{i^*}^k$. To show the strategy $(s_0, \ldots, s_N)$ is monotonic, observe that $0 = s_0 \leq s_1$ and if $i^* > 1$, then $s_{i^*-1} \leq s_{i^*}$ because $s_{i^*-1} = t_{i^*-1} \leq \overline{t}_{i^*-1}^{k+1}$ and $s_{i^*} = \overline{t}_{i^*}^{k+1}$. Furthermore, $s_{i^*} \leq s_{i^*+1}$, because $s_{i^*} = \overline{t}_{i^*}^{k+1} < t_{i^*} \leq t_{i^*+1} = s_{i^*+1}$. Finally $s_j \leq s_{j+1}$ for all other j, because $t_j \leq t_{j+1}$. The comparison of the Sender's payoffs from playing strategies $(t_0, \ldots, t_N)$ and $(s_0, \ldots, s_N)$ reduces to the comparison of the payoffs from locating the i^*-th cutoff at t_{i^*} and s_{i^*} against the Receiver's strategies $(a_1, \ldots, a_N)$ such that $a_{i^*} \leq \overline{a}_{i^*}^k$. Because $a_{i^*} \leq \overline{a}_{i^*}^k$, $s_{i^*} = \overline{t}_{i^*}^{k+1}$ yields a payoff higher than $t_{i^*} > \overline{t}_{i^*}^{k+1}$ against $(a_1, \ldots, a_N)$. This payoff is strictly higher for $(\overline{a}_1^k, \ldots, \overline{a}_N^k)$, because (i) $(\overline{t}_1^{k+1}, \ldots, \overline{t}_N^{k+1})$ is a best response to $(\overline{a}_1^k, \ldots, \overline{a}_N^k)$ and (ii) $\overline{a}_{i^*}^k > 0$, which implies that $\overline{a}_{i^*}^k < \overline{a}_{i^*+1}^k$. To see $\overline{a}_{i^*}^k > 0$, observe that $0 \leq \overline{t}_{i^*}^{k+1} < t_{i^*} \leq \overline{t}_{i^*}^k$, which implies that $0 < \overline{t}_{i^*}^k = \overline{t}_{i^*}^{k-1}$, which in turn implies that $\overline{a}_{i^*}^k > 0$.

We will now describe a modified procedure. Given $\varepsilon > 0$, define $(\overline{t}_0^k(\varepsilon), \overline{t}_1^k(\varepsilon), \ldots, \overline{t}_N^k(\varepsilon))$ and $(\overline{a}_1^k(\varepsilon), \ldots, \overline{a}_N^k(\varepsilon))$, for $k = 0, 1, \ldots$, as in Section 4.4, except that $\overline{t}_i^k(\varepsilon) \neq 0$ for $i > 0$; instead it will be placed within ε of 0. This is possible if ε is sufficiently small. We construct the sequences inductively.

Let

$$0 = \overline{t}_0^0(\varepsilon) < \overline{t}_1^0(\varepsilon) = \cdots = \overline{t}_N^0(\varepsilon) = 1 \text{ and } \overline{a}_1^0(\varepsilon) = \cdots = \overline{a}_N^0(\varepsilon) = 1.$$

Let $(\overline{a}_1^{k+1}(\varepsilon), \ldots, \overline{a}_N^{k+1}(\varepsilon))$ be the interim best response of the Receiver to the strategy $(\overline{t}_0^k(\varepsilon), \overline{t}_1^k(\varepsilon), \ldots, \overline{t}_N^k(\varepsilon))$ of the Sender.

Define $\underline{j}^k$ to be N if $u^S(\overline{a}_j^{k+1}(\varepsilon), 0) \geq u^S(\overline{a}_j^k(\varepsilon), 0)$ for all j and

$$\min\{j : u^S(\overline{a}_j^{k+1}(\varepsilon), 0) < u^S(\overline{a}_j^k(\varepsilon), 0)\}$$

otherwise. Define $\overline{t}_{\underline{j}^k}^{k+1}(\varepsilon) = 1$ if $\underline{j}^k = N$ and otherwise to be the type indifferent between

$\overline{a}_j^{k+1}(\varepsilon)$ and $\overline{a}_j^k(\varepsilon)$. Hence $\overline{t}_{\underline{j}^k}^{k+1}(\varepsilon) > 0$. For $i = \underline{j}^k - 1, \underline{j}^k - 2, \ldots, 1$, let

$$\overline{t}_i^{k+1}(\varepsilon) \in (0, \min\{\varepsilon, \overline{t}_i^k(\varepsilon)/2, \overline{t}_{i+1}^{k+1}(\varepsilon)\}).$$

This is possible because $\varepsilon, \overline{t}_i^k(\varepsilon)/2$, and $\overline{t}_{i+1}^{k+1}(\varepsilon) > 0$. If type 0 prefers $\overline{a}_{i+1}^k(\varepsilon)$ to $\overline{a}_i^k(\varepsilon)$ for some k, then this is so for all greater k when ε is small enough. Indeed, either $\overline{a}_i^{k+1}(\varepsilon) = \overline{a}_i^k(\varepsilon)$ (for odd k) or $\overline{a}_i^{k+1}(\varepsilon) < \overline{a}_i^k(\varepsilon)$ for all i. Hence $\underline{j}^k$ is non-decreasing. Consequently, $\overline{t}_i^k(\varepsilon)$ are strictly positive, strictly increasing in i, decreasing in k, and $\lim_{k\to\infty} \overline{t}_i^k(\varepsilon) = 0$ if $i \leq \lim_{k\to\infty} \underline{j}^k$.

By construction, this modified procedure satisfies condition (8) and avoids the problem described in the first paragraph of the inductive steps for the original procedure. The only argument (for the original procedure) that requires a change concerns why $s_{i^*} = \overline{t}_{i^*}^{k+1}(\varepsilon)$ yields a payoff higher than $t_{i^*} > \overline{t}_{i^*}^{k+1}(\varepsilon)$ against all $(a_1, \ldots, a_N)$ such that $a_{i^*} \leq \overline{a}_{i^*}^k(\varepsilon)$, and this payoff is strictly higher for some such strategies. If $u^S(\overline{a}_i^k(\varepsilon), 0) > u^S(\overline{a}_{i+1}^k(\varepsilon), 0)$, the argument requires no change; and if $u^S(\overline{a}_i^k(\varepsilon), 0) \leq u^S(\overline{a}_{i+1}^k(\varepsilon), 0)$, then $u^S(\overline{a}_i^k(\varepsilon), \overline{t}_{i^*}^{k+1}(\varepsilon)) < u^S(\overline{a}_{i+1}^k(\varepsilon), \overline{t}_{i^*}^{k+1}(\varepsilon))$, and so s_{i^*} yields higher payoffs than any $t_{i^*} > \overline{t}_{i^*}^{k+1}(\varepsilon)$. □

Corollary 2. *Assume $N \geq N^*$. If there exists a unique equilibrium type-action mapping that satisfies NITS, then there is a procedure of iterated deletion of weakly dominated and interim weakly dominated strategies that retains only this type-action mapping. Furthermore, the surviving strategy uses only the highest N^* messages with positive probability.*

Proof. The first part of Corollary 2 follows from Theorem 2. The second part follows from Claim 2. □

Appendix B

We will first prove that any equilibrium using robust best responses must satisfy NITS if there are at least N^* messages. If exactly N^* actions are induced with positive probability, then the equilibrium must satisfy NITS. So, suppose that fewer actions are induced with positive probability. In an equilibrium of a monotonic game, two messages used with positive probability cannot induce the same action. So, there is a message used with probability zero. Let m_j be the smallest message used with positive probability. If a message m_i where $i < j$ is used with probability zero, and induces an action smaller than that induced by message m_j, then NITS must be satisfied. If NITS were violated and the action a_i played in response to m_i where $i < j$ were equal to the action induced by m_j, then the Sender's best response to an action profile obtained from the equilibrium action profile by slightly decreasing a_i would have an interval of types of length bounded away from zero sending m_i. This would violate robustness. If $i > j$, then the Sender's best response to an action profile obtained from the equilibrium action profile by slightly increasing a_i would have an interval of types of length bounded away from zero sending m_i. This would again violate robustness.

We present now an example of the basic cheap-talk game which has two equilibria that satisfy NITS. Whenever there exist multiple equilibria satisfying NITS, type-action mappings in the best-response dynamics can converge to the type-action mapping induced by either of them. Indeed, the type-action mapping induced by an NITS equilibrium (with off-path messages inducing the best response to the lowest type) is a constant best-response sequence. An additional feature that our example illustrates is that the number of partition intervals in different NITS equilibria can be different.

Suppose the distribution of the Sender's types is uniform, and the utilities of the Sender and the Receiver are $u^S(a,t) = -(a - t - c(t))^2$ and $u^R(a,t) = -(a-t)^2$, where $c(t) > 0$ for all t. So, the only departure from the uniform-quadratic example is that the bias $c(t)$ depends on t. Assume that the bias satisfies the following properties:

$$c(0) = 7/96;\ c(8/96) = 2/96;\ c(16/96) = 8/96;\ c(20/96) = 14/96;\ c(24/96) = 14/96.$$

When $c(\cdot)$ satisfies these conditions, $a^S(t) = t + c(t)$ restricted to $t \in \{0,\ 8/96,\ 16/96,\ 20/96,\ 24/96\}$ is strictly increasing. Hence it is possible to extend the definition of $c(t)$ so that $a^S(t) = t + c(t)$ is strictly increasing for all $t \in [0,1]$.

This game has an equilibrium in which the types from $[0, 20/96)$ induce action $\underline{a}_2^* = 10/96$, and the types from $(20/96, 1]$ induce action $\underline{a}_3^* = 58/96$. If there are three messages $m_1 < m_2 < m_3$, then m_1 is an off-path message which induces action $\underline{a}_1^* = 0$, and m_i induces action $\underline{a}_i^*$ for $i = 2, 3$.

The game also has an equilibrium, in which the types from $[0, 8/96)$ induce action $\overline{a}_1^* = 4/96$, the types from $(8/96, 24/96)$ induce action $\overline{a}_2^* = 16/96$, and the types from $(24/96, 1]$ induce action $\overline{a}_3^* = 60/96$. Both equilibria satisfy NITS because both $\underline{a}_2^*$ and $\overline{a}_1^*$ are closer to $a^S(0) = c(0)$ than 0.

References

[1] Nemanja Antić and Nicola Persico. Equilibrium selection through forward induction in cheap talk games. Technical report, Northwestern University, August 2020.

[2] Dieter Balkenborg, Josef Hofbauer, and Christoph Kuzmics. The refined best-response correspondence in normal form games. *International Journal of Game Theory*, 44(1):165–193, 2015.

[3] Andreas Blume. Equilibrium refinements in sender-receiver games. *Journal of Economic Theory*, 64(1):66–77, October 1994.

[4] Andreas Blume. Neighborhood stability in sender–receiver games. *Games and Economic Behavior*, 13(1):2–25, 1996.

[5] Andreas Blume. Meaning in communication. Technical report, University of Arizona, May 2022.

[6] Yi-Chun Chen, Ngo Van Long, and Xiao Luo. Iterated strict dominance in general games. *Games and Economic Behavior*, 61(2):299–315, 2007.

[7] Ying Chen. Perturbed communication games with honest senders and naive receivers. *Joumal of Economic Theory*, 146(2):401–424, March 2011.

[8] Ying Chen, Navin Kartik, and Joel Sobel. Selecting cheap-talk equilibria. *Econometrica*, 76(1):117–136, January 2008.

[9] Daniel Clark and Drew Fudenberg. Justified communication equilibrium. *American Economic Review*, 111(9):3004–34, 2021.

[10] Vincent P. Crawford and Joel Sobel. Strategic information transmission. *Econometrica*, 50(6):1431–1451, November 1982.

[11] Francesc Dilmé. Robust information transmission. *American Economic Review: Insights*, 2022.

[12] Martin Dufwenberg and Mark Stegeman. Existence and uniqueness of maximal reductions under iterated strict dominance. *Econometrica*, 70(5):2007–2023, 2002.

[13] Joseph Farrell. Meaning and credibility in cheap-talk games. *Games and Economic Behavior*, 5(4):514–531, October 1993.

[14] Françoise Forges and Stéphan Sémirat. Signaling future actions and the potential for sacrifice. Technical report, Université Paris-Dauphine, August 2023.

[15] Sidartha Gordon. On infinite cheap talk equilibria. Technical report, Université de Montréal, August 2010.

[16] Sidartha Gordon. Iteratively stable cheap talk. Technical report, Université de Montréal, May 2011.

[17] Jerry R. Green and Nancy L. Stokey. A two-person game of information transmission. *Journal of Economic Theory*, 135(1):90–104, 2007.

[18] Ehud Kalai and Dov Samet. Persistent equilibria in strategic games. *International Journal of Game Theory*, 13(3):129–144, 1984.

[19] Navin Kartik. Strategic communication with lying costs. *Review of Economic Studies*, 76(4):1359–1395, 2009.

[20] Navin Kartik and Joel Sobel. Effective communication in cheap-talk games. Technical report, UCSD, in preparation 2015.

[21] Elon Kohlberg and Jean-François Mertens. On the strategic stability of equilibria. *Econometrica*, 54(5):1003–1037, September 1986.

[22] Barton L Lipman. A note on the implications of common knowledge of rationality. *Games and Economic Behavior*, 6(1):114–129, 1994.

[23] Melody Pei-yu Lo and Wojciech Olszewski. Learning in cheap-talk games. Technical report, Northwestern University, 2018.

[24] Pei-yu Lo. Common knowledge of language and iterative admissibility in a sender-receiver game. Technical report, Brown University, June 2007.

[25] Pei-yu Lo. Language and coordination games. Technical report, University of Hong Kong, October 2009.

[26] Jeffrey Mensch. On the existence of monotone pure-strategy perfect Bayesian equilibrium in games with complementarities. *Journal of Economic Theory*, 187:105–126, 2020.

[27] Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6):1255–1277, November 1990.

[28] Roger B. Myerson and Jörgen W. Weibull. Tenable strategy blocks and settled equilibria. *Econometrica*, 83(3):943–976, May 2015.

[29] Wojciech Olszewski. Rich language and refinements of cheap-talk equilibria. *Journal of Economic Theory*, 128(1):164–186, 2006.

[30] Wojciech Olszewski. A result on convergence of sequences of iterations with applications to best-response dynamics. *Mathematics of Operations Research*, 47(3):2333–2343, 2022.

[31] Joel Sobel. Iterated weak dominance and interval-dominance supermodular games. *Theoretical Economics*, 14(1):71–102, January 2019.

[32] Donald M. Topkis. Equilibrium points in nonzero-sum n-person submodular games. *Siam Journal of Control and Optimization*, 17:773–787, 1979.

[33] Eric Van Damme. *Stability and Perfection of Nash Equilibria.* Springer Verlag, Berlin, 1987.

[34] Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321, 1990.