

Equilibrium Selection Through Forward Induction in Cheap Talk Games

Nemanja Antić and Nicola Persico

August 26, 2020

Abstract

This paper provides a refinement that uniquely selects the ex-ante Pareto dominant equilibrium in a cheap talk game. The refinement works by embedding any cheap talk game into a class of two-stage games where: in stage 1 sender and receiver choose their biases at a cost, and in stage 2 the cheap talk game is played. For such games, we show that a forward induction logic can be invoked to select the ex-ante Pareto-dominant equilibrium in the second stage. Games with fixed biases (the conventional cheap talk games) are then treated as limiting cases of this larger class of games.

Nemanja Antić: Managerial Economics and Decision Sciences Department, Kellogg School of Management, Northwestern University, nemanja.antic@kellogg.northwestern.edu;

Nicola Persico: Managerial Economics and Decision Sciences Department, Kellogg School of Management, Northwestern University, and NBER, n-persico@kellogg.northwestern.edu

We would like to thank Joel Sobel for suggesting that we study this problem.

1 Introduction

Cheap talk games are ubiquitous in applied theory. However, cheap talk games have multiple equilibria and this presents a problem for analyzing comparative statics. Applied papers usually restrict attention to the most informative equilibrium, and justify this restriction by invoking the fact that this equilibrium is ex-ante Pareto dominant, i.e., it is the best equilibrium for both sender and receiver. This paper provides a refinement that leverages this intuition and uniquely selects the ex-ante Pareto dominant equilibrium.¹

Our refinement works by embedding any cheap talk game into a larger space: the class of games where the relative biases of sender and receiver are chosen endogenously before the cheap talk game is played.² For such games, we show that a forward induction logic can be invoked to select the ex-ante Pareto-dominant equilibrium in the second stage (Section 3.2). Games with fixed biases (the conventional cheap talk games) are then treated as limiting cases of this larger class of games.

Our model is as follows. Before the cheap talk stage, every agent i obtains a certain quantity of q_i at a cost $c_i(q_i)$. After paying this cost, each agent is assigned the following payoff function in the cheap talk game:

$$U(a, q_i, \omega),$$

where a is the action taken by the receiver, the quantity q_i encodes the heterogeneity across agents, and ω represents the unknown (to the receiver) state of the world. Then the q_i 's become publicly observable. Then the sender learns ω and the cheap talk stage takes place.

Given any constellation of (q_R, q_S) chosen in the first stage, the second-stage cheap talk game has multiple equilibria. But, in the first stage, agents can use their choice of q_i to “compellingly signal” their expectation that the ex-ante Pareto-dominant equilibrium will be played.

The limiting case with exogenously fixed biases \bar{q}_i can be represented as the limit of sequences of games with cost functions $\text{cost } c_i(\cdot)$ that increasingly penalize any choice $q_i \neq \bar{q}_i$. Our refinement applies to every game in the sequence, and thus selects the Pareto-dominant equilibrium in the limit.

2 Graphical intuition for the result

The result is proved as follows. Fix any first-stage choice (q_i^*, q_{-i}^*) and assume by contradiction that a Pareto-dominated equilibrium Ω_1 is played in the cheap talk stage. We construct a specific deviation \tilde{q}_i such that, among all cheap-talk equilibria that are possible following (\tilde{q}_i, q_{-i}^*) , exactly one of them makes agent i strictly better off, and any

¹Chen, Kartik, and Sobel (2008) is the only other available refinement, to our knowledge. The pros and cons of that refinement are discussed in Section 6.

²Games where the conflict of interest between sender and receiver is determined endogenously prior to the cheap talk phase are of applied interest in their own right. See, e.g.: Antic and Persico (2020), forthcoming; Argenziano et al., (2016); Austen-Smith, (1994); Deimen and Szalay (2019a, 2019b); Rantakari (2017). These papers, however, do not discuss equilibrium selection.

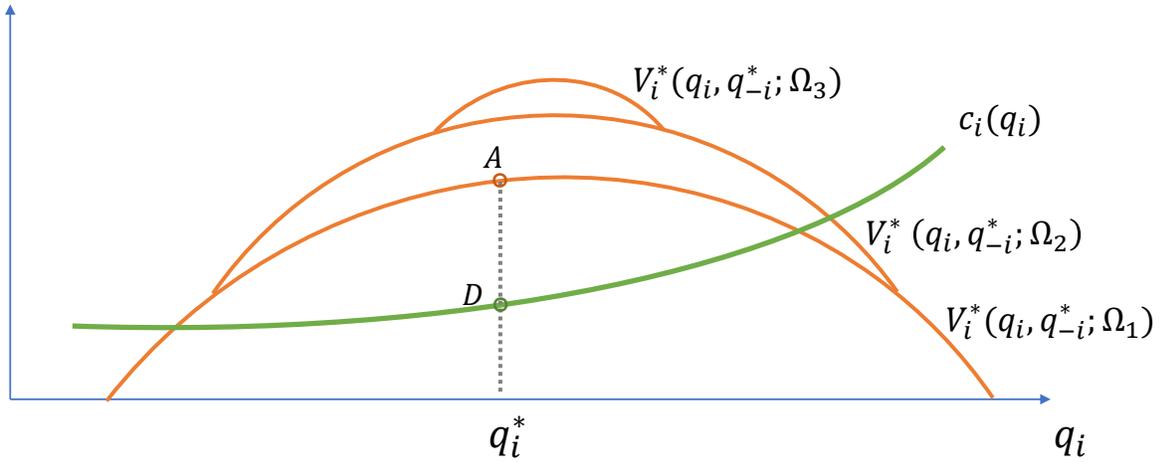


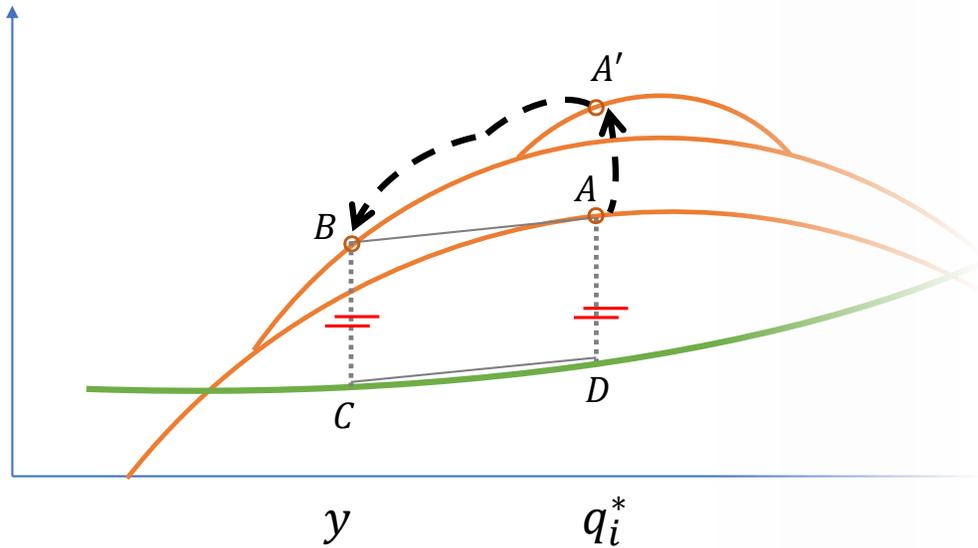
Figure 1: Agent i 's best-response problem. The convex green function represents i 's cost of choosing a certain level of q_i . Every concave orange dome represents i 's ex-ante payoff in the cheap talk game following q_i , if equilibrium Ω is played. Higher domes correspond to equilibria that deliver a higher ex-ante payoff to agent i .

other equilibrium makes agent i strictly worse off, relative to i 's payoff at $(q_i^*, q_{-i}^*, \Omega_1)$. The deviation \tilde{q}_i , if it exists, is “compelling” to the other player because it unequivocally coordinates her on the only one among many possible second-stage equilibria that player i “could have wished to achieve” by deviating. If such a deviation exists, the candidate triple (q_R^*, q_S^*, Ω_1) is deemed inconsistent with forward induction.

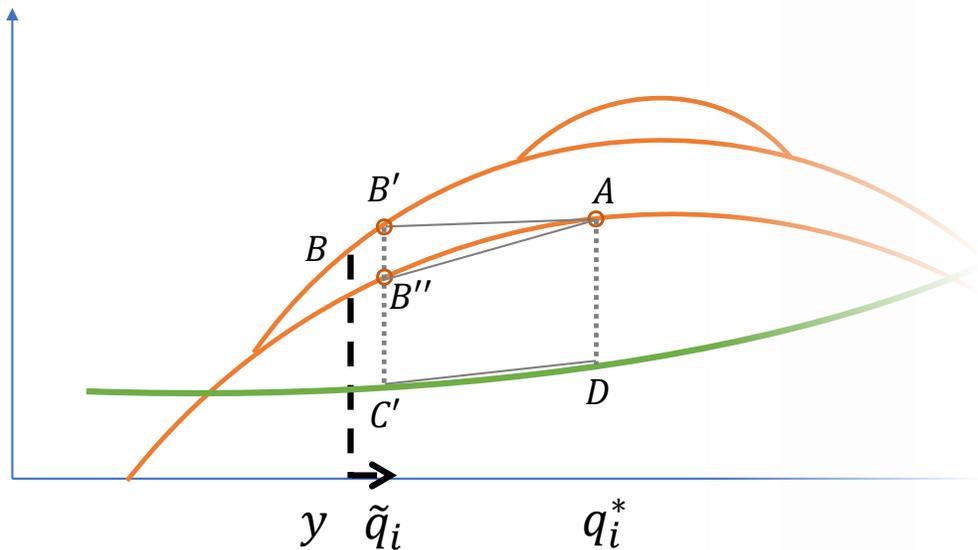
Next we provide a graphical intuition for how the deviation \tilde{q}_i is constructed. The figures that follow are visually patterned after the quadratic cheap talk game in Crawford and Sobel (1982), but the qualitative properties used in the proof hold much more generally.

For a given q_{-i}^* , Figure 1 illustrates the elements of agent i 's best-response problem. The convex green function represents i 's cost of choosing a certain level of q_i . Every concave dome represents i 's ex-ante payoff $V_i(q_i, q_{-i}^*; \Omega)$ in the cheap talk game following q_i , if equilibrium Ω is played. Higher domes correspond to equilibria that deliver a higher ex-ante payoff to agent i . Given q_i^* , for example, Figure 1 indicates that three (non-babbling) equilibria exist and that equilibrium Ω_1 is not Pareto-dominant. The segment \overline{AD} represents i 's *total* payoff, i.e., the payoff after subtracting the cost c_i , if equilibrium Ω_1 is played following (q_i^*, q_{-i}^*) . We now show that the point A is inconsistent with forward induction.

To construct the deviation \tilde{q}_i , first move up vertically on the graph from point A to point A' (refer to Figure 2 step 1). The segment $\overline{A'D}$ represents i 's *total* payoff if the Pareto-dominant equilibrium Ω_3 is played following (q_i^*, q_{-i}^*) . Next, move left along the upper envelope of the orange domes (this corresponds to picking out the best equilibrium following any $q_i < q_i^*$). Initially, that is, for $q_i = q_i^* - \varepsilon$, agent i 's total payoff exceeds the total payoff at q_i^* (the latter is represented by the segment \overline{AD}). As we continue moving



Step 1: go straight up, then move to the left along the upper envelope until the point y is reached



Step 2: move back slightly towards q_i^* , and you've found \tilde{q}_i .

Figure 2: This figure illustrates the two-step procedure used to find a “compelling” deviation \tilde{q}_i that agent i can use to coordinate the other agent away from the Pareto-dominated outcome $(q_i^*, q_{-i}^*, \Omega_1)$. The point \tilde{q}_i is such that, among all cheap-talk equilibria that are possible given (\tilde{q}_i, q_{-i}^*) , exactly one of them (Ω_2 , corresponding to payoff B') makes agent i strictly better off (this is because $\overline{B'C'} > \overline{AD}$); and all other cheap talk equilibria make agent i strictly worse off (this is because $\overline{B''C'} < \overline{AD}$).

down and to the left along the upper envelope, we will eventually encounter a point $q_i = y$ at which i 's total payoff \overline{BC} equals the total payoff \overline{AD} . Set $\tilde{q}_i = y + \varepsilon$: on the graph (Figure 2 step 2) this corresponds to moving back slightly toward q_i^* . The point \tilde{q}_i is our “compelling” deviation: indeed, among all cheap-talk equilibria that are possible given (\tilde{q}_i, q_{-i}^*) , exactly one of them (point B') makes agent i strictly better off relative to i 's payoff at $(q_i^*, q_{-i}^*, \Omega_1)$: this is because $\overline{B'C'} > \overline{AD}$; and all other equilibria make agent i strictly worse off (this is because $\overline{B''C'} < \overline{AD}$).

3 Model

The state of the world is $\omega \in [0, 1]$, where $\omega \sim f$ and $f > 0$ is absolutely continuous. Agents are indexed by $i \in \{S, R\}$ with S being the sender and R the receiver.

The timing is as follows:

1. Agents $i = R, S$ simultaneously and independently select q_i at a cost $c_i(q_i)$. The θ_i 's are publicly observed.
2. The sender privately learns ω and engages in cheap talk with the receiver.
3. The receiver chooses action a .
4. Payoffs accrue.

After paying the cost $c_i(q)$, an agent i who obtains q experiences the following utility:

$$U(a, q, \omega). \tag{1}$$

Note that, in general, the game between S and R is not symmetric because $c_S(\cdot) \neq c_R(\cdot)$. This asymmetry will, in general, cause the agents' choices $q_S^* \neq q_R^*$ to not align, resulting in a conflict of interest between sender and receiver in the cheap talk game. The functions $U(\cdot, \cdot, \cdot)$, and $c_i(\cdot)$ are twice continuously differentiable.³

3.1 Notation for the cheap talk game

The cheap talk game takes q_R and q_S as parameters. A vector $\Omega = [\omega_0, \omega_1, \dots, \omega_N]$ with $\omega_0 = 0, \omega_N = 1$, and $\omega_k < \omega_{k+1}$ is called a partition of the state space.⁴ The elements ω_k are called partition cutoffs. Given a partition Ω (not necessarily an equilibrium one), define the function:

$$V_i(q_i, q_{-i}; \Omega) = \sum_{k=1}^N \int_{\omega_{k-1}}^{\omega_k} U(a_k^*, q_i, \omega) dF(\omega), \tag{2}$$

³This setup is similar to that in our forthcoming paper (Antic and Persico 2020).

⁴Formally, Ω is a set of disjoint intervals of the form: $\{[\omega_0, \omega_1), [\omega_1, \omega_2), \dots, [\omega_{N-1}, \omega_N]\}$; for short, we notate Ω as the vector of interval cutoffs.

where

$$a_k^* \in \arg \max_a \int_{\omega_{k-1}}^{\omega_k} U(a, q_R, \omega) dF(\omega) , \quad (3)$$

for $k = 1, \dots, N$. A partition $\Omega_N = [0, \omega_1^*, \dots, \omega_{N-1}^*, 1]$ that solves:

$$U(a_k^*, q_S, \omega_k^*) = U(a_{k+1}^*, q_S, \omega_k^*) \quad (4)$$

for $k = 1, \dots, N-1$ is said to be an N -partition equilibrium in the cheap talk game. When evaluated at an equilibrium, the function (2) is written as:

$$V_i^*(q_i, q_{-i}; \Omega_N) , \quad (5)$$

and it represents agent i 's ex-ante payoff in this equilibrium.

3.2 Forward induction criterion

We provide a definition of forward induction, and assumptions sufficient to select a unique cheap-talk equilibrium in the second stage. All proofs are provided in Appendix A.

Denote by $\widehat{\Omega}(q_R, q_S)$ any correspondence that, to any pair (q_R, q_S) , associates an N -dimensional vector which is N -partition cheap talk equilibrium given (q_R, q_S) . Note that no continuity assumptions are placed on the correspondence $\widehat{\Omega}$ (for instance, N need not be the same for nearby points in the domain). The next definition works by restricting “admissible” correspondences $\widehat{\Omega}$ through a criterion in the spirit of forward induction (van Damme 1989).⁵ The definition is interpreted as an equilibrium selection criterion for second-stage equilibria.

We say that $(q_R^*, q_S^*, \widehat{\Omega})$ is a pure-strategy equilibrium in the sequential game if, given that the equilibrium $\widehat{\Omega}(q_R, q_S)$ is played in the second stage, (q_R^*, q_S^*) is a pure-strategy Nash equilibrium in the first stage.

Definition 1 (equilibrium selection through forward induction) *A pure-strategy equilibrium in the sequential game $(q_R^*, q_S^*, \widehat{\Omega})$ is **consistent with forward induction** if no agent i and deviation \tilde{q}_i exist, such that exactly one cheap talk equilibrium among those that exist at (\tilde{q}_i, q_{-i}^*) makes agent i strictly better off and all other cheap talk equilibria at (\tilde{q}_i, q_{-i}^*) make her strictly worse off.*

This definition starts from an equilibrium in the sequential game, and checks for specific deviations \tilde{q}_i such that, among all second-stage equilibria that are possible following the deviation (not restricted to $\widehat{\Omega}(\tilde{q}_i, q_{-i}^*)$), exactly one of them makes agent i better off. The deviation \tilde{q}_i , if it exists, is “compelling” to the other player because it unequivocally coordinates her on the only one among many possible second-stage equilibria that player i “could have wished to achieve” by deviating. If such a deviation exists, the candidate triple $(q_R^*, q_S^*, \widehat{\Omega})$ is deemed inconsistent with forward induction.

⁵ The particular definition we use is inspired by Fudenberg and Tirole’s (1991) definition.

This criterion for selecting equilibria is conservative in that equilibria are only eliminated by “compelling” deviations which *cannot be misinterpreted by a rational player who believes her opponent to be rational*. Despite being conservative, under the following assumptions this criterion will rule out all but one second-stage equilibrium.

3.3 Assumptions

The following assumptions are sufficient to prove our main result. We view these as mild assumptions, i.e., not very restrictive.

Assumption 1 (*one-to-oneness*) Fix any (q_R, q_S) . If $\Omega \neq \Omega'$ are two distinct cheap-talk equilibria at (q_R, q_S) , then $V_i^*(q_i, q_{-i}; \Omega) \neq V_i^*(q_i, q_{-i}; \Omega')$ for all i .

This assumption says that, for any given (q_R, q_S) , second-stage equilibria are one-to-one with payoff levels. In Crawford and Sobel’s (1982) model this property holds when their condition (M) is satisfied. In Figure 1 this means that every dome is associated with a distinct cheap talk equilibrium. This is a mild condition.

Let $V_i^{\text{sup}}(q_R, q_S)$ denote agent i ’s highest payoff across all cheap talk equilibria at (q_R, q_S) .

Assumption 2 (*continuous upper envelope*) The function $V_i^{\text{sup}}(q_R, q_S)$ is continuous for all i .

This is a mild assumption. It says that the upper envelope in Figure 1 is continuous.⁶

Assumption 3 (*worse option is available*) For every i there exists a \underline{q}_i such that

$$V_i^{\text{sup}}(\underline{q}_i, q_{-i}^*) - c_i(\underline{q}_i) \leq V_i(q_R^*, q_S^*, \widehat{\Omega}) - c_i(q_i^*).$$

This assumption says that, for any first-stage equilibrium actions (q_R^*, q_S^*) , player i has a “worse option” \underline{q}_i available, at which even the best cheap talk equilibrium at $(\underline{q}_i, q_{-i}^*)$ (not restricted to $\widehat{\Omega}(\underline{q}_i, q_{-i}^*)$) gives a (weakly) worse payoff than the equilibrium one. In Figure 1 this assumption checks out: choosing q_i small enough results in negative payoffs. In general, a \underline{q}_i with the required property can always be found if the cost $c_i(\cdot)$ grows fast enough, simply by picking a very large \underline{q}_i . Alternatively, any \underline{q}_i is suitable such that the only equilibrium at $(\underline{q}_i, q_{-i}^*)$ is babbling.⁷

Assumption 4 (*finite number of second-stage equilibria are pervasive*) For any given (q_i, q_{-i}) , there is a point z arbitrarily close to q_i such that the set of cheap-talk equilibria at (z, q_{-i}) is finite.

⁶In Appendix A we show that this property holds in Crawford and Sobel’s (1982) quadratic example. See that appendix for a formal definition of the function V_i^{sup} .

⁷This is shown in Appendix A. In Crawford and Sobel’s (1982) quadratic example, the latter property holds if, for any given q_{-i} , agent i ’s action set includes a q_i such that $|b| = |q_R - q_S| > 1/4$.

In Figure 1 this assumption checks out: for any given choice of q_i the number of equilibria is finite (3, at most). This property holds under the assumptions of Crawford and Sobel’s (1982) Theorem 1. It is easy to see that this property holds in their quadratic example because, unless $b = 0$, the cheap talk game has a finite number of equilibria.

4 Main result

If the above assumptions hold, then the following result selects the Pareto-dominant cheap-talk equilibrium, if one exists. We say a cheap-talk equilibrium Ω is Pareto-dominant if all players weakly prefer Ω to any other cheap talk equilibrium Ω' .

Proposition 1 (selection through forward induction) *Let $(q_R^*, q_S^*, \widehat{\Omega})$ be an equilibrium in the sequential game.*

1. *If for any (q_R, q_S) the equilibrium selected by $\widehat{\Omega}(\cdot, \cdot)$ Pareto-dominates all other cheap-talk equilibria that exist at (q_R, q_S) , then $(q_R^*, q_S^*, \widehat{\Omega})$ is consistent with forward induction.*
2. *If $\widehat{\Omega}(q_R^*, q_S^*)$ does not Pareto-dominate all other cheap-talk equilibria at (q_R^*, q_S^*) , then there is a cost function \tilde{c}_i arbitrarily close to c_i in the uniform norm such that $(q_R^*, q_S^*, \widehat{\Omega})$ remains an equilibrium in the sequential game but it is not consistent with forward induction.*

Proof. This is a restatement of Corollary 1 in Appendix A. ■

Proposition 1 says the following. Take any equilibrium of the first-stage game (q_R^*, q_S^*) that is computed under the stipulation that second-stage equilibrium selection (that is, the cheap talk equilibrium $\widehat{\Omega}(\cdot, \cdot)$ evaluated at any point (q_R, q_S)) is Pareto-dominant “on and off path;” then, this equilibrium is consistent with forward induction (part 1).⁸ Furthermore, forward induction requires the second-stage (cheap talk) equilibrium to be Pareto-dominant “on path” (part 2, approximately).

This result is applicable only if a Pareto-dominant cheap-talk equilibrium exists. Crawford and Sobel (1982) provide sufficient conditions for all cheap-talk equilibria to be Pareto-ranked. Therefore, our result applies to that class of cheap-talk games.

5 Extension to games with exogenous conflict of interest

The argument derived in this paper can be used to select the best equilibrium in the game with exogenous conflict of interest (i.e., q_i^* is exogenously set equal to \bar{q}_i), as follows. The shape of the cost functions $c_i(\cdot)$ have been restricted only mildly by our assumptions. In

⁸An example in Appendix A suggests that this condition is not only sufficient, but also necessary.

particular, we need not assume $c_i(\cdot)$ is monotone. Consequently, Proposition 1 also applies to cost functions, where the cost is prohibitively high except for in a small neighborhood of some exogenous value \bar{q}_i . Such cost functions result in equilibrium choices $q_i^* \approx \bar{q}_i$, corresponding to a scenario where the endogeneity in the choice of q_i is almost absent. The limiting case of $q_i^* \equiv \bar{q}_i$ (i.e., fully exogenous preferences in the cheap talk game) can be approximated by a sequence of games where the continuous cost functions $c_i(\cdot)$ increasingly penalize any choice $q_i \neq \bar{q}_i$. Intuitively, in games along the sequence the choice of q_i becomes progressively “less endogenous.” Since Proposition 1 applies to every game in the sequence, then in the limit game fully exogenous preferences, the most informative equilibrium is the limit of the sequence of forward-induction proof equilibria.

6 Contribution to the literature and conclusions

Refinement criteria for signaling games such as Kohlberg and Mertens’ (1986) stability and Cho and Kreps’ (1987) intuitive criterion, do not have power in cheap talk games because messages are free, unlike costly actions in signaling games. Some refinements, such as Farrell’s (1993) neologism-proofness, have the power to refine some cheap talk equilibria, but they run into an existence problem: in many popular cheap talk examples, such as Crawford and Sobel’s (1982) quadratic example, no equilibria are neologism-proof.⁹

The most attractive refinement to date is Chen, Kartik, and Sobel’s (2008) “no-incentive-to-separate” (NITS) criterion. An equilibrium satisfies NITS if the type-0 sender could not benefit from credibly identifying himself, if he could. This criterion can be microfounded by viewing the cheap talk game as the limit of games with small lying costs. Monotone equilibria in these games converge to the NITS equilibrium in the cheap talk game, as the lying costs converge to zero uniformly. The NITS refinement is attractive because at least one equilibrium always exists that satisfies NITS and, under Crawford-Sobel’s “condition M,” the only equilibrium that satisfies NITS is the most-informative one. Unlike NITS, our refinement operates at the ex ante stage, i.e., before the sender learns the state of the world. In contrast, in NITS the sender contemplates “causing the equilibrium to switch” after having observed the signal. The two refinements are complementary, in our view: in some applications it may be more natural to assume that there is a cost of lying in the cheap talk game; in other applications, it may be more natural (and indeed, organically desirable) to contemplate the possibility that the agents’ biases are determined at an earlier stage of play.

Our analysis builds on the notion of forward induction. The literature on forward induction has progressed greatly since van Damme’s (1989) seminal paper, with particular advances made on its epistemic foundations. This paper does not seek to push the foundational frontier. Rather, our contribution is applied: we show that an elementary implementation of the forward induction idea can be used to achieve unique selection.

Finally: while Proposition 1 is stated in the context of cheap talk games, the selection

⁹Other papers build on the ideas of neologism proofness. Rabin (1990) runs into multiplicity problems: the selection criterion is not stringent enough to select a unique equilibrium. Matthews, Okuno-Fujiwara, and Postlewaite (1991), like Farrell (1993), runs into existence problems.

logic applies to more general settings. In the appendix, the result is proved without necessarily assuming that the second-stage game is a cheap talk game, provided a Pareto-dominant equilibrium exists in the second-stage game.

References

- [1] Antić, Nemanja, and Nicola Persico (2020). “Cheap Talk with Endogenous Conflict of Interest.” Forthcoming, *Econometrica*.
- [2] Argenziano, R., Severinov, S., & Squintani, F. (2016). “Strategic information acquisition and transmission.” *American Economic Journal: Microeconomics*, 8(3), 119-55.
- [3] Austen-Smith, D. (1994). “Strategic Transmission of Costly Information.” *Econometrica*, 62, p. 955-63.
- [4] Chen, Y. , Kartik, N. and Sobel, J. (2008) “Selecting Cheap-Talk Equilibria.” *Econometrica*, 76: 117-136.
- [5] Cho, I. K., and Kreps, D. M. (1987). “Signaling games and stable equilibria.” *The Quarterly Journal of Economics*, 102(2), 179-221.
- [6] Crawford, V. and Sobel, J. (1982) “Strategic Information Transmission.” *Econometrica*, 50, issue 6, p. 1431-51.
- [7] Deimen, Inga, and Dezső Szalay (2019a). “Delegated expertise, authority, and communication.” *American Economic Review* 109.4 (2019): 1349-74.
- [8] Deimen, Inga, and Dezső Szalay (2019b). “Information and communication in organizations.” *AEA Papers and Proceedings*. Vol. 109. 2019.
- [9] Fudenberg, D. and J. Tirole (1991) *Game Theory* MIT Press.
- [10] Farrell, J. (1993). “Meaning and credibility in cheap-talk games.” *Games and Economic Behavior*, 5(4), 514-531.
- [11] Kartik, N. (2007). “Information Transmission with Almost-Cheap Talk,” Manuscript, UC San Diego.
- [12] Kohlberg, E. and Mertens, J-F. (1986). “On the Strategic Stability of Equilibria.” *Econometrica*, 54(5), 1003-1037.
- [13] Rantakari, H. (2017). “Managerial Influence and Organizational Performance.” Manuscript, University of Rochester.
- [14] van Damme, Eric (1989) “Stable Equilibria and Forward Induction,” *Journal of Economic Theory*, 48(2), pp. 476-496.

A Appendix: Equilibrium Selection Through Forward Induction

This section studies a two-stage sequential game where, in the first stage, two agents indexed by $i = R, S$ simultaneously and independently select q_i at a cost $c_i(q_i)$. The second stage is black-boxed through a payoff function V and an index Ω that denotes a second-stage equilibrium given $\mathbf{q} = (q_R, q_S)$. Agent i 's payoff in the entire game is:

$$W_i(\mathbf{q}, \Omega) = V_i(q_i, q_{-i}; \Omega) - c_i(q_i),$$

where Ω belongs to $\mathcal{S}(\mathbf{q})$, a (non-empty) set of second-stage equilibria. The functions $\{V_i\}_{i=R,S}$ and the set $\mathcal{S}(\mathbf{q})$ are independent of the shape of $c_i(\cdot)$ because, by the second stage, the cost c_i is sunk. The functions $\{V_i, c_i\}_{i=R,S}$ and the set $\mathcal{S}(\mathbf{q})$ are the primitives for this analysis and are taken to be fixed throughout.

In this setting, we provide a definition of forward induction, and assumptions sufficient to select a unique equilibrium Ω in the second stage. The game described in Section 3 is a special case where $\mathcal{S}(\mathbf{q})$ represents the set of *all* equilibria in the cheap talk game with biases given by \mathbf{q} , and:

$$W_i(\mathbf{q}, \Omega) = V_i^*(q_i, q_{-i}; \Omega) - c_i(q_i),$$

where we have made the change of variables $q_i = q_i(\theta)$ and $c_i(q_i) = c_i(\theta)$.

When $\mathcal{S}(\mathbf{q})$ has more than one element, checking whether a first-stage choice of q is an equilibrium requires one to specify which second-stage equilibria should be used to evaluate a different choice of q_i . With this issue in mind, we define a suitable notion of equilibrium in the sequential game. Note that throughout this section we restrict attention to equilibria q^* in pure strategies.

Definition 2 (equilibrium in the sequential game) *For every q , let $\mathfrak{s}(\mathbf{q})$ denote an element of $\mathcal{S}(\mathbf{q})$. We say that $(\mathbf{q}^*, \mathfrak{s})$ is an equilibrium in the sequential game if, for all i , and all \tilde{q}_i :*

$$W_i(\mathbf{q}^*, \Omega^*) \geq W_i((\tilde{q}_i, q_{-i}^*), \mathfrak{s}(\tilde{q}_i, q_{-i}^*)), \text{ where } \Omega^* = \mathfrak{s}(\mathbf{q}^*).$$

This definition parameterizes a (pure-strategy) equilibrium in the sequential game by a selection $\mathfrak{s}(\mathbf{q})$ of second-stage equilibria that are used to evaluate first-stage choices. In the game described in Section 3, an example of $\mathfrak{s}(\mathbf{q})$ is: “the cheap talk equilibrium with 2 partition elements if that exists given \mathbf{q} , else the babbling equilibrium;” then, first-stage choices would be evaluated by restricting attention to those cheap talk equilibria. In particular, agent i would evaluate her deviations according to the selection \mathfrak{s} .

The next definition may be interpreted as an equilibrium selection criterion for second-stage equilibria. It works by restricting “admissible” selections $\mathfrak{s}(\mathbf{q})$ through a criterion in the spirit of forward induction (van Damme 1989).¹⁰

¹⁰We acknowledge that Van Damme invokes genericity and finiteness assumptions in his setting, whereas our analysis assumes that \mathbf{q} is selected from a continuum.

Definition 3 (equilibrium selection through forward induction) *An equilibrium $(\mathbf{q}^*, \mathfrak{s})$ in the sequential game is consistent with forward induction if no agent i and deviation \tilde{q}_i exist, such that exactly one element $\Omega \in \mathcal{S}(\tilde{q}_i, q_{-i}^*)$ makes agent i strictly better off and all other $\Omega' \in \mathcal{S}(\tilde{q}_i, q_{-i}^*)$ make her strictly worse off.*

This definition starts from a (pure-strategy) equilibrium in the sequential game, and checks for specific deviations \tilde{q}_i such that, among all second-stage equilibria that are possible following the deviation (not restricted to $\mathfrak{s}(\tilde{q}_i, q_{-i}^*)$), exactly one of them makes agent i better off. Such a deviation \tilde{q}_i , if it exists, is “compelling” to the other player because it unequivocally coordinates her on the only one among many possible equilibria in $\mathcal{S}(\tilde{q}_i, q_{-i}^*)$ that player i “could have wished to achieve” by deviating.

This criterion for selecting equilibria is conservative in that equilibria are only eliminated by “compelling” deviations which *cannot be misinterpreted by a rational player who believes her opponent to be rational*. Despite being conservative, under some assumptions this criterion will rule out all but one second-stage equilibrium in $\mathcal{S}(\mathbf{q})$.

Definition 4 (upper envelope of the payoff correspondence) *The upper envelope of W_i is the function:*

$$W_i^{\text{sup}}(\mathbf{q}) = \sup_{\Omega \in \mathcal{S}(\mathbf{q})} W_i(\mathbf{q}, \Omega).$$

The upper envelope function selects the upper limit among all the equilibrium payoffs that are possible for agent i given a first-stage choice \mathbf{q} . If the set $\mathcal{S}(\mathbf{q})$ of second-stage equilibria is finite, as is the case when $b \neq 0$ in Crawford and Sobel’s (1982) quadratic example, the sup operator may be replaced by max.¹¹ In that example, W_i^{sup} is attained by the equilibrium with the largest number of partition elements given b .¹²

A.1 Assumptions

Assumption 5 (one-to-oneness) *For every \mathbf{q} , $\Omega \neq \Omega'$ implies $W_i(\mathbf{q}, \Omega) \neq W_i(\mathbf{q}, \Omega')$.*

This assumption says that, for any given \mathbf{q} , second-stage equilibria are one-to-one with payoff levels. In Crawford and Sobel’s (1982) quadratic example this property holds because, for given specification of the bias parameter b , equilibrium payoffs are one-to-one with the number of cutoffs in the equilibrium partition.¹³

Assumption 6 (continuous upper envelope) *The function $W_i^{\text{sup}}(\mathbf{q})$ is continuous for all i .*

¹¹Crawford and Sobel (1982) theorem 1 implies that this is true very generally as long as the sender and receiver prefer different actions at all states of the world.

¹²This is true more generally in their setting, provided their condition (M) holds.

¹³This is also true more generally when their condition (M) is satisfied.

Loosely speaking this assumption says that i 's best-equilibrium payoff is continuous in \mathbf{q} . This assumption holds in Crawford and Sobel's (1982) quadratic example. To see this, fix the cardinality N of an equilibrium partition and then the equilibrium payoff in the N -partition equilibrium is continuous in b . Since in the region $b \neq 0$ there is a finite number of cheap talk equilibria indexed by N , the function $W_i^{\text{sup}}(\mathbf{q})$ is the upper envelope of a finite number of continuous functions, and is thus itself continuous. The function $W_i^{\text{sup}}(\mathbf{q})$ is also continuous at $b = 0$ because as $b \rightarrow 0$ the payoff of the best equilibrium approaches the full-information payoff.

Assumption 7 (*worse option is available*) Suppose $(\mathbf{q}^*, \mathfrak{s})$ is an equilibrium in the sequential game. Then for every i there exists a \underline{q}_i such that

$$W_i^{\text{sup}}(\underline{q}_i, q_{-i}^*) \leq W_i(\mathbf{q}^*, \Omega^*), \text{ where } \Omega^* = \mathfrak{s}(\mathbf{q}^*).$$

This assumption says that, for any equilibrium \mathbf{q}^* , player i has a “worse option” \underline{q}_i available, at which even the best equilibrium in $\mathcal{S}(\underline{q}_i, q_{-i}^*)$ gives a (weakly) worse payoff than the equilibrium one.

A \underline{q}_i with the required property can always be found if the cost $c_i(\cdot)$ grows fast enough, simply by picking a very large \underline{q}_i . Alternatively, in the game described in Section 3, any \underline{q}_i is suitable such that $\mathcal{S}(\underline{q}_i, q_{-i}^*)$ only contains the babbling equilibrium.¹⁴ In Crawford and Sobel's (1982) quadratic example, the assumption holds if, for any given q_{-i} , agent i 's action set includes a q_i such that $|b| = |q_R - q_S| > 1/4$.

Assumption 8 (*finite number of second-stage equilibria are pervasive*) For any given q_{-i} , the set of points z such that the set $\mathcal{S}(z, q_{-i})$ has finite cardinality, is dense in \mathbb{R} .

This assumption means that for any \mathbf{q} there is an arbitrarily close $\tilde{\mathbf{q}}$ with a finite number of second-stage equilibria. This property holds in Crawford and Sobel's (1982) quadratic example because $b = 0$ is the only value at which the cheap talk game has an infinite number of equilibria.¹⁵

¹⁴To see this, write:

$$\begin{aligned} W_i(\mathbf{q}^*, \Omega^*) &\geq W_i(q_i, q_{-i}^*, \mathfrak{s}(q_i, q_{-i}^*)) \\ &= \sup_{\Omega \in \mathcal{S}(q_i, q_{-i}^*)} W_i(q_i, q_{-i}^*, \Omega) = W_i^{\text{sup}}(q_i, q_{-i}^*), \end{aligned}$$

where the inequality holds because $(\mathbf{q}^*, \mathfrak{s})$ is an equilibrium, and the first equality holds because the only element in the set $\mathfrak{s}(q_i, q_{-i}^*)$ is the babbling equilibrium.

¹⁵This property also holds under the more general assumptions of Crawford and Sobel's (1982) Theorem 1.

A.2 Results

We say a second-stage equilibrium Ω is Pareto-dominant if all players weakly prefer Ω to any other equilibrium Ω' .

Lemma A.1 (*Pareto-dominant second-stage equilibrium is consistent with forward induction*) Consider any equilibrium in the sequential game $(\mathbf{q}^*, \mathfrak{s})$. If $\mathfrak{s}(\mathbf{q})$ selects a Pareto-dominant second-stage equilibrium in $\mathcal{S}(\mathbf{q})$ for all \mathbf{q} , then $(\mathbf{q}^*, \mathfrak{s})$ is consistent with forward induction.

Proof. Because $(\mathbf{q}^*, \mathfrak{s})$ is an equilibrium, for all i we have:

$$W_i(\mathbf{q}^*, \Omega^*) \geq W_i((\tilde{q}_i, q_{-i}^*), \mathfrak{s}(\tilde{q}_i, q_{-i}^*)) \geq W_i((\tilde{q}_i, q_{-i}^*), \Omega),$$

where $\Omega^* = \mathfrak{s}(\mathbf{q}^*)$ and, by Pareto-dominance of \mathfrak{s} , Ω is any element of $\mathcal{S}(\tilde{q}_i, q_{-i}^*)$. This means that there is no \tilde{q}_i and element of $\mathcal{S}(\tilde{q}_i, q_{-i}^*)$ that make agent i strictly better off than $(\mathbf{q}^*, \mathfrak{s}(\mathbf{q}^*))$. By definition, this means that $(\mathbf{q}^*, \mathfrak{s})$ is consistent with forward induction. ■

In Crawford and Sobel (1982), the equilibrium with the largest number of partition elements is Pareto-dominant if assumption (M) holds.¹⁶ If \mathfrak{s} selects this equilibrium for every \mathbf{q} then, by the above lemma, any equilibrium $(\mathbf{q}^*, \mathfrak{s})$ in the sequential game is consistent with forward induction.

To see why the above lemma requires $\mathfrak{s}(\mathbf{q})$ to be Pareto-dominant for all \mathbf{q} , consider the following simple example in the quadratic setting of Crawford and Sobel (1982). Set both cost functions $\{c_i\} \equiv 0$, and $q_R^* = 2/10, q_S^* = 1/10$. Because $b = q_R^* - q_S^* = 1/10$, the best cheap-talk equilibrium has two partition elements. Denote this equilibrium by $\Omega_2 = \mathfrak{s}(q_R^*, q_S^*)$. Suppose $\mathfrak{s}(\mathbf{q})$ picks out the babbling equilibrium for all values of \mathbf{q} such that $q_R - q_S \neq 1/10$ (note that this violates Pareto-dominance). Given this choice of \mathfrak{s} , $(\mathbf{q}^*, \mathfrak{s})$ is an equilibrium. However, $(\mathbf{q}^*, \mathfrak{s})$ is not consistent with forward induction. To see this, observe that by increasing q_S slightly above q_S^* , the sender's best-equilibrium payoff exceeds that at $(\mathbf{q}^*, \mathfrak{s})$, but the second-highest equilibrium payoff (babbling is the only other one that exists) does not. Thus, $(\mathbf{q}^*, \mathfrak{s})$ is not consistent with forward induction, but is an equilibrium.

Losely speaking, this discussion shows that the forward induction requirement tends to “select/require” high-payoff outcomes in the second stage, both on and off the equilibrium path, that is, over the entire domain of the function $\mathfrak{s}(\cdot)$. The next lemma says that, generically, the only second-stage equilibria that are consistent with forward induction are Pareto efficient.

Lemma A.2 (*generically, only Pareto-dominant second-stage equilibria are consistent with forward induction*) If there is an equilibrium in the sequential game $(\mathbf{q}^*, \mathfrak{s})$

¹⁶See their theorems 3 and 5.

such that $W_i(\mathbf{q}^*, \mathfrak{s}(\mathbf{q}^*)) < W_i^{\text{sup}}(\mathbf{q}^*)$ for some i , then there is a cost function \tilde{c}_i arbitrarily close to c_i in the uniform norm, for which:

1. $(\mathbf{q}^*, \mathfrak{s})$ remains an equilibrium in the sequential game
2. $(\mathbf{q}^*, \mathfrak{s})$ is not consistent with forward induction.

Proof. Fix any $\{V_i, c_i\}_{i=R,S}$, $\mathcal{S}(\mathbf{q})$, $\mathfrak{s}(\mathbf{q})$. Take an equilibrium $(\mathbf{q}^*, \mathfrak{s})$ such that for some i :

$$W_i(\mathbf{q}^*, \Omega^*) < W_i^{\text{sup}}(\mathbf{q}^*), \quad (6)$$

where $\Omega^* = \mathfrak{s}(\mathbf{q}^*)$. We now look for a deviation for player i that demonstrates a violation forward induction either for the cost function c_i or for a “nearby” one \tilde{c}_i .

Step 1: identifying the “test neighborhood” \tilde{y}

By Assumption 7, there exists a \underline{q}_i such that

$$W_i^{\text{sup}}(\underline{q}_i, q_{-i}^*) \leq W_i(\mathbf{q}^*, \Omega^*),$$

which, in conjunction with (6), implies $\underline{q}_i \neq q_i^*$. Assume without loss of generality that $\underline{q}_i > q_i^*$ (the other case is treated symmetrically). By the mean value theorem, which applies because W_i^{sup} is continuous by Assumption 6, there is at least one $y \geq q_i^*$ such that:

$$W_i^{\text{sup}}(y, q_{-i}^*) = W_i(\mathbf{q}^*, \Omega^*).$$

Let \tilde{y} denote the infimum in the set of such y 's. This infimum belongs to the set because the set is closed, and thus:

$$W_i^{\text{sup}}(\tilde{y}, q_{-i}^*) = W_i(\mathbf{q}^*, \Omega^*), \quad (7)$$

which, in conjunction with (6), implies $\tilde{y} \neq q_i^*$.

Step 2: identifying a “test deviation” \tilde{q}_η arbitrarily close to \tilde{y}

By Assumption 8, the interval (q_i^*, \tilde{y}) contains an increasing sequence $\{\tilde{q}_\eta\}_{\eta=1}^\infty$ converging to \tilde{y} such that, for all η , $S(\tilde{q}_\eta, q_{-i}^*)$ is a finite set. Fix any η . Because the set of second-stage equilibria $S(\tilde{q}_\eta, q_{-i}^*)$ is finite, the set $W_i((\tilde{q}_\eta, q_{-i}^*), S(\tilde{q}_\eta, q_{-i}^*))$ of associated payoffs has some finite cardinality N_η . By assumption 5 the payoffs in this set may be strictly ordered as follows:

$$w_1(\eta) < w_2(\eta) < \dots < w_{N_\eta}(\eta) = W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*). \quad (8)$$

Step 3: constructing the “arbitrarily close” cost function $\tilde{c}_i(q; \eta)$

Fix any η . Denote

$$\Delta(\eta) = W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) - W_i(\mathbf{q}^*, \Omega^*) > 0, \quad (9)$$

where the inequality follows from the fact that for every $z \in (q_i^*, \tilde{y})$ we have $W_i^{\text{sup}}(z, q_{-i}^*) > W_i^{\text{sup}}(\tilde{y}, q_{-i}^*) = W_i(\mathbf{q}^*, \Omega^*)$ (both the inequality and the equality follow from the definition of \tilde{y}).

We use the quantity $\Delta(\eta)$ to define an ancillary “upper bound” function that is close to c_i :

$$\bar{c}_i(q; \eta) = c_i(q) + \Delta(\eta) - \frac{1}{2} \min[\Delta(\eta), w_{N\eta}(\eta) - w_{N\eta-1}(\eta)]. \quad (10)$$

The function $\bar{c}_i(\cdot; \eta)$ is greater than $c_i(\cdot)$, but only by a “small amount” less than $\Delta(\eta)$. Now we construct the main object of interest. Let $\tilde{c}_i(q; \eta)$ be any continuous function such that:

$$\begin{aligned} \tilde{c}_i(q; \eta) &\in [c_i(q), \bar{c}_i(q; \eta)] \\ \tilde{c}_i(\tilde{q}_\eta; \eta) &= \bar{c}_i(\tilde{q}_\eta; \eta) \\ \tilde{c}_i(q_i^*; \eta) &= c_i(q_i^*). \end{aligned}$$

Step 4: showing that the “test deviation” \tilde{q}_η in conjunction with the cost function $\tilde{c}_i(q; \eta)$ triggers a violation of forward induction

Fix any η . The payoff function of an agent who is endowed with the cost function $\tilde{c}_i(q; \eta)$ instead of $c_i(q)$ is denoted by:

$$\widetilde{W}_i(\mathbf{q}, \Omega) = V_i(q_i, q_{-i}; \Omega) - \tilde{c}_i(q_i; \eta).$$

Denote the ordered elements of the set of payoffs $\widetilde{W}_i((\tilde{q}_\eta, q_{-i}^*), S(\tilde{q}_\eta, q_{-i}^*))$ by:

$$\tilde{w}_1(\eta) < \tilde{w}_2(\eta) < \dots < \tilde{w}_{N\eta}(\eta),$$

with generic element:

$$\tilde{w}_n(\eta) = w_n(\eta) + c_i(\tilde{q}_\eta) - \tilde{c}_i(\tilde{q}_\eta; \eta). \quad (11)$$

Now note that:

$$\begin{aligned} \tilde{w}_{N\eta}(\eta) &= W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) + c_i(\tilde{q}_\eta) - \tilde{c}_i(\tilde{q}_\eta; \eta) \\ &= W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) + c_i(\tilde{q}_\eta) - \bar{c}_i(\tilde{q}_\eta; \eta) \\ &= W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) - \Delta(\eta) + \frac{1}{2} \min[\Delta(\eta), w_{N\eta}(\eta) - w_{N\eta-1}(\eta)] \\ &= W_i(\mathbf{q}^*, \Omega^*) + \frac{1}{2} \min[\Delta(\eta), w_{N\eta}(\eta) - w_{N\eta-1}(\eta)] \\ &> W_i(\mathbf{q}^*, \Omega^*) \\ &= \widetilde{W}_i(\mathbf{q}^*, \Omega^*). \end{aligned}$$

In the above formulas, the first equality comes from the definition of $\tilde{w}_n(\eta)$ (eq. 11), after using (8) to substitute for $w_{N\eta}(\eta)$. The next equality holds because $\tilde{c}_i(\tilde{q}_\eta; \eta) = \bar{c}_i(\tilde{q}_\eta; \eta)$ by construction. The equality in line 3 results from substituting for \bar{c}_i using (10), and the next equality follows from substituting for $\Delta(\eta)$ using (9). The strict inequality holds because both arguments of the min operator are strictly positive. The final equality holds because $\tilde{c}_i(q_i^*; \eta) = c_i(q_i^*)$ by construction.

Proceeding as in the previous paragraph:

$$\begin{aligned}
\tilde{w}_{N\eta-1}(\eta) &= w_{N\eta-1}(\eta) + c_i(\tilde{q}_\eta) - \tilde{c}_i(\tilde{q}_\eta; \eta) \\
&= w_{N\eta-1}(\eta) - \Delta(\eta) + \frac{1}{2} \min[\Delta(\eta), w_{N\eta}(\eta) - w_{N\eta-1}(\eta)] \\
&\leq w_{N\eta-1}(\eta) - \Delta(\eta) + \frac{1}{2} [w_{N\eta}(\eta) - w_{N\eta-1}(\eta)] \\
&< w_{N\eta-1}(\eta) - \Delta(\eta) + [w_{N\eta}(\eta) - w_{N\eta-1}(\eta)] \\
&= W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) - \Delta(\eta) \\
&= W_i(\mathbf{q}^*, \Omega^*) \\
&= \tilde{W}_i(\mathbf{q}^*, \Omega^*).
\end{aligned}$$

In the above formulas, the equality in line 5 uses (8) to substitute for $w_{N\eta}(\eta)$.

In sum, we have shown that

$$\tilde{w}_{N\eta-1}(\eta) < \tilde{W}_i(\mathbf{q}^*, \Omega^*) < \tilde{w}_{N\eta}(\eta). \quad (12)$$

Step 5: wrap up

We have shown that, for every η :

Under $\tilde{c}_i(q; \eta)$, the pair $(\mathbf{q}^, \mathfrak{s})$ remains an equilibrium in the sequential game.*

This is true because q_i^* remains a best response to q_{-i}^* (this follows because, by construction $\tilde{c}_i(q_i^*; \eta) = c_i(q_i^*)$ and $\tilde{c}_i(q; \eta) \geq c_i(q)$).

Under $\tilde{c}_i(q; \eta)$, the pair $(\mathbf{q}^, \mathfrak{s})$ is not consistent with forward induction.*

This is true because equation (12). Note that while this equation is a statement about *payoffs*, also has implications for *second-stage equilibria* in light of Assumption 5. Indeed, equation (12) shows that the deviation \tilde{q}_η is such that exactly one element of $\mathcal{S}(\tilde{q}_\eta, q_{-i}^*)$ makes agent i strictly better off and all others make her strictly worse off. Thus, the deviation \tilde{q}_η is used to show that the equilibrium $(\mathbf{q}^*, \mathfrak{s})$ is not consistent with forward induction.

It remains to show that $\tilde{c}_i(q; \eta)$ and $c_i(q)$ can be made arbitrarily close by an opportune choice of η . To verify this, observe that, for all q ,

$$0 \leq \tilde{c}_i(q; \eta) - c_i(q) < \Delta(\eta).$$

Substitute (7) into (9) to get:

$$\Delta(\eta) = W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) - W_i^{\text{sup}}(\tilde{y}, q_{-i}^*).$$

By construction, the sequence $\{\tilde{q}_\eta\}_{\eta=1}^\infty$ converges to \tilde{y} , and by continuity of W_i^{sup} (Assumption 6) we get:

$$\lim_{\eta \rightarrow \infty} \Delta(\eta) = \lim_{\eta \rightarrow \infty} W_i^{\text{sup}}(\tilde{q}_\eta, q_{-i}^*) - W_i^{\text{sup}}(\tilde{y}, q_{-i}^*) = 0.$$

This shows that a suitable choice of η makes $\Delta(\eta)$ the uniform distance between $\tilde{c}_i(q; \eta)$ and $c_i(q)$ arbitrarily small. ■

These two lemmas apply directly to the cheap talk game with endogenous conflicts of interest described in Section 3.

Corollary 1 (*application to cheap talk games with endogenous conflict of interest*) Suppose the set $\{V_i, c_i\}_{i=R,S}$, $\mathcal{S}(\mathbf{q})$ is generated by the game in Section 3. Consider any equilibrium $(\mathbf{q}^*, \mathfrak{s})$.

1. If $\mathfrak{s}(\mathbf{q})$ Pareto-dominates all other cheap-talk equilibria for all \mathbf{q} , then $(\mathbf{q}^*, \mathfrak{s})$ is consistent with forward induction.
2. If $\Omega^* = \mathfrak{s}(\mathbf{q}^*)$ does not Pareto-dominate all other cheap-talk equilibria at \mathbf{q}^* , then there is an i and a cost function \tilde{c}_i arbitrarily close to c_i in the uniform norm such that $(\mathbf{q}^*, \mathfrak{s})$ remains an equilibrium in the sequential game but it is not consistent with forward induction.