Robustness of Full Revelation in Multisender Cheap Talk*

Margaret Meyer, Inés Moreno de Barreda and Julia Nafziger

First draft: February 2013 This draft: March 23, 2016

Abstract

This paper studies information transmission in a two-sender, multidimensional cheap talk setting where there are exogenous restrictions on the feasible set of policies for the receiver. Such restrictions, which are present in most applications, can, by limiting the punishments available to the receiver, prevent the existence of fully revealing equilibria (FRE). We focus on FRE that are i) robust to small mistakes by the senders, in that small differences between the senders' messages result in only small punishments by the receiver, and ii) independent of the magnitudes of the senders' bias vectors. For convex policy spaces in an arbitrary number of dimensions, we prove that if there exists a FRE satisfying property ii), then there exists one satisfying both i) and ii). Thus the requirement of robustness is, under these assumptions, not restrictive. For convex policy spaces in two dimensions, we provide a simple geometric condition, the Local Deterrence Condition, on the directions of the senders' biases relative to the frontier of the policy space, that is necessary and sufficient for the existence of a FRE satisfying i) and ii). We also provide a specific policy rule, the Min Rule, for the receiver that supports a FRE satisfying i) and ii) whenever one exists. The Min Rule is the anonymous rule that punishes incompatible reports in the least severe way, subject to maintaining the senders' incentives for truthtelling, no matter how large their biases. We characterize necessary and sufficient conditions for collusion-proofness of a FRE supported by the receiver using the Min Rule and show that if such a FRE is not collusion-proof, then no other FRE satisfying ii) can be collusion-proof. We extend our existence results to convex policy spaces in more than two dimensions and to non-convex two-dimensional spaces. Finally, our necessary and sufficient condition, as well as our specific policy rule, can be easily adapted if the receiver is uncertain about the directions of the biases and/or if the biases vary with the state of the world.

*Meyer: Nuffield College and Dept. of Economics, Oxford University and CEPR, margaret.meyer@nuffield.ox.ac.uk. Moreno de Barreda: Dept. of Economics, Oxford University, ines.morenodebarreda@economics.ox.ac.uk. Nafziger: Dept. of Economics and Business, Aarhus University, jnafziger@econ.au.dk. We thank Elizabeth Baldwin, Tom Cunningham, Péter Esö, Navin Kartik, Kohei Kawamura, Paul Klemperer, Patrick Legros, Gilat Levy, John Morgan, Ronny Razin, Joel Sobel, and Peter Norman Sørensen for helpful comments, as well as seminar audiences at Copenhagen, ESSET Gerzensee, LSE, Oxford, the Spanish Economic Association, the Transatlantic Theory Workshop, UAB, and ESEM Toulouse.

1 Introduction

In sender-receiver games with cheap talk, the decision-maker (receiver) has imperfect information about the consequences of a policy and elicits reports from better-informed experts (the senders), whose preferences are not perfectly aligned with those of the decision-maker (i.e. the experts are "biased"). The advice transmitted by the senders is costless but unverifiable (hence, "cheap talk"), and the receiver cannot commit himself in advance to how he will respond to the senders' advice.[1] Cheap talk games with two biased experts have been used, for example, in organizational economics to analyze the interaction between a CEO and division managers, and in political science to study the transmission of information from legislative committees to the legislature as a whole.[2] In both of these contexts, as well as in most other settings to which cheap-talk models have been applied, the receiver typically faces constraints on the feasible set of policies—these may stem from limited budgets, from physical restrictions on what is possible (within a given time frame), or from legal constraints.

In this paper, we analyze a two-sender cheap-talk model in which the senders observe common information that is unavailable to the receiver and the receiver faces exogenous restrictions on the feasible set of policies. When the receiver can consult two equally informed senders, the receiver has the potential to extract all information from the senders, by comparing the senders' messages and punishing any discrepancy between them. However, the exogenous restrictions on the feasible set of policies can, by limiting the set of possible responses by the receiver to incompatible reports by the senders, prevent the existence of fully revealing equilibria. Our objective is to provide simple geometric conditions, on the shape of the feasible set of policies relative to the directions of the senders' bias vectors, that are necessary and sufficient for the existence of equilibria that are not only fully revealing but have additional desirable properties.

The first such desirable property of a fully revealing equilibrium (FRE) is robustness to small mistakes by the senders. Even if the commonly informed senders attempt to report truthfully to the receiver, noise in the communication process might result in the receiver receiving incompatible messages. Yet if the discrepancy between the incompatible messages is small, it is reasonable to think that the receiver would wish to respond by choosing a policy that is close to each of the messages. We formulate a natural and analytically tractable definition of robustness of an equilibrium to small mistakes that explicitly requires that small differences between the senders' messages should result in only small adjustments in the receiver's response, or, phrased differently, in only small punishments by the receiver.

In our model, the receiver and the senders all have quadratic utility functions, and sender i's ideal point differs from the receiver's by a vector, b_i, sender i's bias vector.[3] The second desirable property of the FRE's that we seek is independence of the magnitudes of the

[1]For the seminal paper in this literature see Crawford and Sobel (1982).

[2]For the former application, see Alonso and Matouschek (2008) and for the latter, Gilligan and Krehbiel (1989) and Krishna and Morgan (2001a;b)

[3]In Section 5 we sketch how our results can be extended to general quasi-concave preferences.

senders' biases. In many settings the receiver will know the direction of the divergence in interests between himself and each of the senders but will be less certain of the intensity of each of these divergences. We seek fully revealing equilibria such that the *same set of strategies* remain a FRE no matter how large the magnitudes of the senders' biases. Besides being invariant to the intensity of senders' preferences, such equilibria are also appealing because of the relative tractability of their characterization.

We begin by focusing on convex policy spaces in an arbitrary number of dimensions. We prove in Proposition 3 that whenever there exists a fully revealing equilibrium that is independent of the magnitudes of the biases, there also exists a robust FRE that is independent of these magnitudes. In other words, when biases can be arbitrarily large, if small deviations cannot be deterred with small punishments, then they cannot be deterred with any feasible punishments. Moreover, we show that for convex policy spaces that are two-dimensional or multidimensional and compact, it is sufficient for existence of a FRE (robust or not) that small deviations can be deterred with small punishments. These results are extremely useful, because they show that a) robustness is, perhaps surprisingly, not a restrictive requirement on a FRE that is independent of the bias magnitudes; and b) in the two-dimensional or compact multidimensional cases, we need only ensure that local deviations can be punished.

Section 3.1 then focuses on the case where the policy space is a convex subset of $\mathbb{R}^2$. Proposition 4 identifies a simple geometric condition, the *Local Deterrence Condition*, on the directions of the senders' bias vectors relative to the frontier of the policy space, that is necessary and sufficient for all small deviations to be deterrable with small punishments, and hence, given Proposition 3, for the existence of a FRE (robust or not) that is independent of the magnitudes of the biases. The proposition also provides a specific policy rule for the receiver, the *Min Rule*, that supports a robust FRE whenever one exists. To describe this rule, observe that as the senders' biases become arbitrarily large, their indifference curves approach hyperplanes. Using the coordinate system defined by these limiting preferences of the senders, the Min Rule specifies that, given any two reported states, the receiver chooses the component-wise minimum of these reports. The Min Rule is anonymous in that it selects the same policy in response to a pair of incompatible reports no matter which sender sent which report. This rule is the anonymous rule that punishes incompatible reports in the least severe way, subject to deterring the senders from misreporting, no matter how large their biases. If the Min Rule is feasible, it deters deviations from truthful reporting in a manner robust to small mistakes. Proposition 4 shows that the Min Rule is feasible, for all pairs of reports, if and only if the Local Deterrence Condition is satisfied, or in other words, if and only if a (robust) FRE exists.

For convex policy spaces in two dimensions, we also prove that the Local Deterrence Condition remains necessary and sufficient for existence of a FRE that is robust even when the biases have known finite sizes. This is true because, when the receiver is constrained to use small punishments, it is only the orientations, not the magnitudes, of the senders' bias vectors that determine whether or not they have incentives to deviate from truthtelling.

Since the senders have common information and could attempt to mislead the receiver by both making the same false report, another desirable property of a FRE is collusion-proofness. We will say that a FRE is collusion-proof if, whenever there is a feasible policy that both senders prefer to the one the receiver would choose if they were truthful, collusion on this preferred policy would not be sefl-enforcing, because one of the senders would find it profitable to unilaterally deviate. In Section 3.2, we show that the Min Rule is the *best* punishment rule the receiver could use in order to prevent collusion. Specifically, we show that if a FRE supported by the receiver using the Min Rule is not collusion-proof, then no other FRE that is independent of the magnitudes of the biases can be collusion-proof. The intuition for this result is that the Min Rule prescribes the least severe punishment subject to maintaining the senders' incentives, and hence it makes deviations from the collusive report most attractive. We then characterize necessary and sufficient conditions for collusion-proofness of a FRE supported by the Min Rule.

To illustrate these results, suppose that the receiver is an executive who has to allocate funds from a budget to two different projects, each one overseen by a manager. The maximum amount of funds that can be allocated is exogenously fixed, and each project must receive a non-negative level of funding. The managers (senders) have common information about the returns to the two projects, and hence about which feasible allocation of funds to each project the receiver would prefer, if he had access to the senders' information. Whatever the best allocation of funds for the receiver (whether or not it exhausts the whole budget), each sender would prefer that a strictly higher level of funding be allocated to the project he oversees. Figure 1 illustrates the receiver's feasible set of policies (allocations of funds) and the directions of the senders' bias vectors.

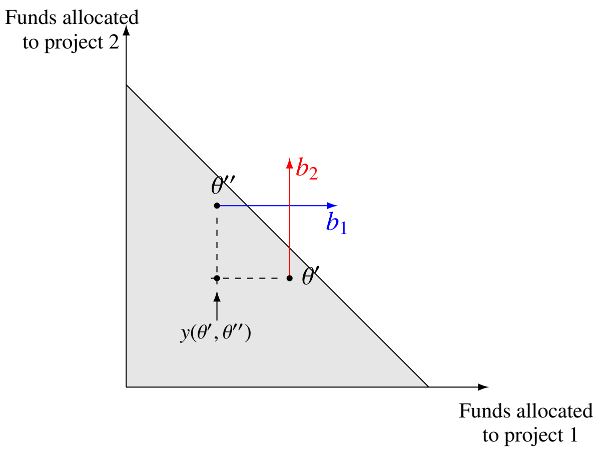


Figure 1: Allocation of funds given a budget constraint. The shaded area represents the feasible allocations. Given a pair of reports (θ', θ''), the Min Rule implements $y(\theta', \theta'') = (\min\{\theta'_1, \theta''_1\}, \min\{\theta'_2, \theta''_2\})$.

Since the policy space is convex, Proposition 3 implies that a robust FRE independent of the magnitudes of the biases exists whenever a FRE independent of these magnitudes does. Moreover, since the policy space is also a subset of $\mathbb{R}^2$, to prove existence it is enough to check whether all small deviations can be deterred with small punishments. This is precisely what the Local Deterrence Condition establishes. Since the frontier of the policy space here has only three distinct orientations, it is easy to confirm that this condition is satisfied in this setting.[4] Proposition 4 provides a way to construct a robust FRE that is independent of the biases' magnitudes: the receiver uses the Min Rule to respond to any pair of incompatible reports. For the orthogonal bias directions shown, the Min Rule specifies that the receiver chooses the component-wise minimum of the senders' reports with respect to the Euclidean coordinates.[5] It is clear from Figure 1 that the allocation specified by the Min Rule is feasible in this setting for all pairs of reported allocations, and this rule implements a robust FRE that is independent of the magnitudes of the biases.

Furthermore, in this setting, Proposition 6 implies that the robust FRE supported by the Min Rule is, whatever the magnitudes of the biases, collusion-proof: Given any plan by the senders to make a common false report overstating the amount of funding the receiver should allocate to both projects, the divergence in bias directions between the two senders means that each of them could gain by unilaterally deviating from the collusive plan.

In Section 3.3, we extend our results to convex policy spaces of any dimension larger than two. The key observation here is that the only directions of conflict between the senders and the receiver are the ones in the (two-dimensional) plane spanned by the senders' bias vectors. Proposition 7 shows that, for existence of a FRE (robust or not) that is independent of the magnitudes of the biases, it is necessary and sufficient to look at the projection of the policy space onto the subspace of conflict of interest and see whether a FRE can be constructed there. The reason is that, when the magnitudes of the biases can be arbitrarily large, no given shift of the receiver's action in a direction orthogonal to the plane of the biases can be certain to serve as a punishment for a deviating sender. Therefore, in response to incompatible reports, the receiver must choose as a punishment an action whose projection onto the plane of the biases is worse for both senders. Such an action exists if and only if the projection of the policy space onto the plane of the biases satisfies the Local Deterrence Condition identified in Proposition 4 for the two-dimensional case.

Since increasing returns or indivisibilities may cause the set of feasible policies for the receiver to be non-convex, it is important to examine when a robust FRE exists for non-convex policy spaces. This we do in Section 4. We identify an additional geometric condition, the *Global Deterrence Condition*, on the directions of the senders' biases relative to the frontier of the *convex hull* of the policy space, that together with the Local Deterrence Condition identified in Proposition 4, is necessary and sufficient for existence of a robust FRE that is independent of the magnitudes of the biases. The Local Deterrence Condition is necessary and sufficient for small deviations to be deterrable with small punishments, but for non-convex policy spaces, this is not sufficient for existence of a robust FRE: large deviations might not be deterrable even if small ones are. The Global Deterrence Condition is necessary and sufficient for existence of a (not necessarily robust) FRE. When the policy space is convex, the Local and the Global Deterrence Conditions coincide.

In Section 5, we relax the assumptions that the directions of the senders' biases are

[4]Sections 3.1 and 3.2, in particular Figures 9a and 10b, provide more details for this example.

[5]That is, if Sender 1 reports that the receiver's optimal allocation of funds to project 1 is θ_1' and to project 2 is θ_2', while Sender 2's respective reports are θ_1'' and θ_2'', then the receiver allocates $min\{\theta_1', \theta_1''\}$ to project 1 and $min\{\theta_2', \theta_2''\}$ to project 2.

(i) common knowledge and (ii) independent of the realization of the state. We prove that when the receiver does not know the actual biases but knows only the minimal closed cone in which they are certain to lie, and this minimal cone is the same for all states, then the necessary and sufficient condition for existence of a robust FRE that is independent of the bias magnitudes is the same Local Deterrence Condition identified in Proposition 4 in Section 3.1, except that the known biases b_1 and b_2 there are replaced by the least aligned possible realizations of the biases. Furthermore, whenever a robust FRE exists, it is supported by the receiver using the same punishment rule, the Min Rule, defined previously, except that now the punishment is computed using the least aligned possible bias realizations.

1.1 Related Literature

The closest papers to ours are Battaglini (2002) and Ambrus and Takahashi (2008), both of which analyze the existence of fully revealing equilibria when the receiver can consult two equally informed senders.

Battaglini (2002) assumes that the policy space is the whole of $\mathbb{R}^q$. He observes that each sender's preferences are aligned with those of the receiver in the subspace orthogonal to the sender's bias vector, and therefore the receiver can extract truthful information when the sender's influence is restricted to those dimensions. As long as the two senders' biases are linearly independent, the receiver can combine the two truthful reports to extract all the information. This construction supports a FRE that is independent of the magnitudes of the biases and also robust to small mistakes. However, Battaglini's construction breaks down when there are restrictions on the receiver's feasible set of policies, since there are pairs of reports for which the receiver's response rule in this construction is infeasible.

Our construction of a FRE supported by the Min Rule relies on the same coordinate system defined by the normal vectors to the senders' biases. Yet there are crucial differences between our construction and Battaglini's. First, the Min Rule is an anonymous rule, whereas the receiver's strategy in Battaglini's construction critically depends on which sender made which report; consequently, even when $Y = \mathbb{R}^2$, the two strategies do not always coincide. Second, as we show in Section 5, our construction of a robust FRE using the Min Rule can be extended to a setting where the receiver does not know the actual directions of the biases but knows only the minimal closed cone in which they are certain to lie. Such an extension is not possible with Battaglini's construction, since the receiver's uncertainty about the directions of the biases prevents him from identifying the subspaces in which each sender's preferences are aligned with his. Most importantly, while Battaglini's construction is infeasible for restricted policy spaces even when a FRE exists, we show that whatever the form of the policy space, the feasibility of the Min Rule is both necessary as well as sufficient for the existence of a robust FRE.

Ambrus and Takahashi (2008) consider the case of compact and convex policy spaces. They show that there exists a FRE for any magnitudes of the biases if and only if, as the biases become large, the senders have a common least-preferred policy.[6] While this char-

[6] In this paper, we seek existence of a fully revealing equilibrium that is independent of the magnitudes of

acterization result is elegant, their equilibrium construction involves the receiver punishing any discrepancies between the senders' reports by choosing their common least-preferred policy.[7] As they themselves acknowledge, the use of extreme punishments after even small deviations is unappealing, since such deviations could in practice arise from small mistakes by the senders. In response to this criticism they introduce a robustness concept called *continuity on the diagonal* that can be shown to be equivalent to our robustness definition.[8] However, they present only negative findings concerning the existence of robust fully revealing equilibria.

Their negative findings contrast strikingly with our main result, Proposition 3, which states that once we have secured the existence of a FRE that is independent of the magnitudes of the biases, then robustness comes for free. The reason for this striking contrast is that their examples of nonexistence of a robust FRE are all ones in which a FRE that is independent of the magnitudes of the biases does not exist. Our Proposition 5 shows that if there does not exist a FRE that is independent of the magnitudes of the biases, then no robust FRE exists, even for small magnitudes of the biases.

Finally,[9] Ambrus and Lu (2014) and Rubanov (2015) construct equilibria in a unidimensional policy space that are arbitrarily close to full revelation.[10] Both papers show that their equilibria survive the introduction of a small probability of the senders observing a random state which is independent of the true state. The equilibria in the two papers involve constructing complex partitions of the policy space (different for different senders) and having each sender report the element of his partition in which his observation lies. Asymptotically, incentives for truthful reporting follow because any change in a sender's message (holding fixed the messages of the others) results in a sufficiently large change in the receiver's action. Importantly, the equilibria in these two papers do not satisfy our robustness concept: The senders might observe states that are arbitrarily close to each other, yet the receiver's action in response to the equilibrium messages might be far away from the senders' observations. Our notion of robustness differs from theirs in that we do not explicitly model the occurrence of mistakes, or in other words, the players in our model are unaware of the possibility of mistakes. Yet crucially, our equilibrium concept requires

---

the biases. In general, this is a stronger requirement than Ambrus and Takahashi (2008)'s requirement that a FRE exist for any magnitudes of the biases, since we, unlike them, require the *same* equilibrium strategies to remain an equilibrium no matter how large the bias magnitudes. In fact, if the policy space is not compact, then it is always possible to construct a FRE that varies with the magnitudes of the biases, no matter how large they become, whereas there might not exist a FRE that is independent of these magnitudes. On the other hand, when the policy space is compact and convex, the two requirements are equally strong, and the geometric condition in Proposition 8 of Ambrus and Takahashi (2008) can be shown to be equivalent to our Local Deterrence Condition. Nevertheless, in the general non-compact (convex) case, these two geometric conditions are not equivalent, but our Local Deterrence Condition remains necessary and sufficient for existence of a (robust) FRE that is independent of the magnitudes of the biases.

[7]For small biases, even in states where this policy is not the least preferred policy for both senders, it is nevertheless always worse for both of them than the policy the receiver would choose if he had full information.

[8]See Lemma 3 in the Appendix.

[9]There is a small recent experimental literature on multi-sender cheap talk. Lai et al. (2015), in particular, discuss robustness, but because their state and policy space is discrete, our concept of robustness cannot be applied in their setting.

[10]Ambrus and Lu (2014) require the policy space to be larger, the closer their equilibrium is to be to full revelation, whereas Rubanov (2015) requires the number of senders to grow for his asymptotic result.

that the response of the receiver be close to the senders' reports whenever these reports are themselves close: this ensures that small mistakes by the senders do not lead to large responses/punishments by the receiver. Our motivation for imposing continuity on the receiver's out-of-equilibrium actions is shared by Friedman and Samuelson (1990; 1994), who impose a similar restriction on strategies in repeated games. They argue (1994, p.56), "In many circumstances strategies associating severe penalties with arbitrarily small deviations are implausible."

2 The Model

We analyze a game of cheap talk between two senders, S_1, S_2, and a receiver, R. Both senders perfectly observe $\theta \in \Theta$, the realization of a real random variable $\hat{\theta}$. We will refer to the realization θ as the *state*, and to $\Theta \subseteq \mathbb{R}^p$ as the *state space*, which has dimension $p \geq 1$. The prior distribution of $\hat{\theta}$ is given by F and is commonly known. After observing θ, each sender S_i sends a costless and unverifiable message $m_i \in \mathcal{M}_i$ to the receiver[11], who then chooses a policy y from a set of feasible policies Y, a closed subset of Θ.[12] Y is referred to as the *policy space* and has dimension $q \leq p$. We will refer to the pair (Θ, Y) as the *environment* of the game.[13]

Given the state θ and the chosen policy y, the receiver's utility is $u^R(y, \theta) = -(y - \theta)^2$ and each sender i's utility is $u^{S_i}(y, \theta) = -(y - \theta - b_i)^2$. The vector $b_i \in \mathbb{R}^p$ is referred to as the bias vector of sender S_i. Unless otherwise specified, the direction of b_i is assumed to be independent of the state θ and common knowledge among the players.[14] Given these utilities, the ideal policy for the receiver is to match the state, however since Y might be a strict subset of Θ, such policy might not be feasible. We denote by $y^*(\theta) \in \arg\min_{y \in Y}(y-\theta)^2$ the optimal feasible policy for the receiver when the state is θ. Clearly, in the particular case in which $Y \equiv \Theta$, $y^*(\theta) = \theta$.[15]

A pure strategy for sender S_i will be denoted by $s_i : \Theta \longrightarrow \mathcal{M}_i$, and a pure strategy for the receiver will be denoted by $y : \mathcal{M}_1 \times \mathcal{M}_2 \longrightarrow Y$. Given messages m_1, m_2, $\mu(m_1, m_2)$ denotes the receiver's belief about $y^*(\hat{\theta})$ after receiving m_1, m_2. We denote by $y^\mu(m_1, m_2) \in Y$ an optimal policy for the receiver given belief $\mu(m_1, m_2)$, and $\mu(\cdot)$ will be referred to as the belief function of the receiver. Since the senders' payoffs depend on the receiver's choice of policy, it is more convenient to work directly with the receiver's beliefs over the optimal feasible policy $y^*(\hat{\theta})$ than with his beliefs over $\hat{\theta}$. The equilibrium concept we use is *Perfect Bayesian Equilibrium*.

[11]The set of possible messages $\mathcal{M}_i$ is sufficiently large so that sender S_i could if he wished reveal the state.

[12]The interpretation is that Θ is the set of ideal policies, but only those in $Y \subseteq \Theta$ are feasible. Clearly in the case $\Theta \subset Y$, we can without loss of generality ignore all those policies in Y that are not in Θ.

[13]The shape of the policy space Y plays a key role in the analysis. As we will see in Proposition 2, for a fixed Y, it makes no difference to our results whether $Y = \Theta$ or $Y \subset \Theta$.

[14]These assumptions are relaxed in Section 5.

[15]When the policy space Y is a strict subspace of Θ and Y is non-convex, the set $\arg\min_{y \in Y}(y - \theta)^2$ might not be a singleton. In such a case, we will focus on one particular optimal feasible policy and label this $y^*(\theta)$; which optimal policy is singled out in this way is irrelevant.

Definition 1. The strategies (s_1, s_2, y) constitute a *Perfect Bayesian Equilibrium* if there exists a belief function $\mu(\cdot)$ such that:

(i) s_i is optimal given s_{-i} and y for $i \in \{1, 2\}$.

(ii) $y(m_1, m_2) = y^\mu(m_1, m_2)$ for each $(m_1, m_2) \in \mathcal{M}_1 \times \mathcal{M}_2$

(iii) If $s_1^{-1}(m_1) \cap s_2^{-1}(m_2) \neq \emptyset$ then $\mu(m_1, m_2)$ is derived from Bayes' rule.

We will focus on a particular type of equilibrium in which the receiver perfectly learns the optimal feasible policy from the messages of the senders. We will say that the strategies (s_1, s_2) are *fully revealing* if for all $\theta \in \Theta$, the conditional probability of the random variable $y^*(\hat{\theta})$ given messages $s_1(\theta)$ and $s_2(\theta)$ has mass one on $y^*(\theta)$. In particular, $y^\mu(s_1(\theta), s_2(\theta)) = y^*(\theta)$. An equilibrium with fully revealing strategies is called a *fully revealing equilibrium* (FRE).

Our goal in this paper is to characterize the conditions under which there exist equilibria that are not only fully revealing but that also satisfy two additional desirable properties: i) robustness to small mistakes by the senders and ii) independence of the magnitudes of the senders' biases. The next two subsections formally define each of these properties of a FRE.[16]

2.1 Robustness

In a Perfect Bayesian Equilibrium, no restriction is imposed on beliefs in response to out-of-equilibrium messages, i.e. messages such that $s_1^{-1}(m_1) \cap s_2^{-1}(m_2) = \emptyset$. This implies that after receiving incompatible messages, the receiver could choose any feasible policy to punish the deviation. However, when the discrepancy between two incompatible messages is small, it might be reasonable to think that the receiver's chosen policy should be close to each of the messages, since small discrepancies might be due not to deliberate misrepresentation by the senders but rather to small mistakes.

We now formulate a definition of robustness of an equilibrium to small mistakes that explicitly captures this reasoning.

Definition 2. Given some fully revealing strategies (s_1, s_2), a belief function $\mu(\cdot)$ *deters local deviations with local punishments* if for any $\theta \in \Theta$ and any $\epsilon > 0$, there exist $\delta > 0$ and $\delta' > 0$ such that if: *(i)* $\theta', \theta'' \in B(\theta, \delta) \cap \Theta$ and *(ii)* $y^*(\theta'), y^*(\theta'') \in B(y^*(\theta), \delta') \cap Y$, then

$$
\begin{gathered}
y^\mu(s_1(\theta'), s_2(\theta'')) \in B(y^*(\theta), \epsilon) \cap Y, \\
|y^\mu(s_1(\theta'), s_2(\theta'')) - (\theta'' + b_1)| \geq |y^*(\theta'') - (\theta'' + b_1)|, \text{ and} \\
|y^\mu(s_1(\theta'), s_2(\theta'')) - (\theta' + b_2)| \geq |y^*(\theta') - (\theta' + b_2)|.
\end{gathered}
$$

[16]The restriction to pure strategies when the policy space is convex is without loss of generality: if there exists a robust FRE that is independent of the magnitudes of the biases in mixed strategies, there exists a robust FRE that is independent of the magnitudes of the biases in pure strategies. When the policy space is non-convex, this restriction is no longer without loss of generality. In Section 4, when analyzing the case of non-convex policy spaces, we allow the receiver to use mixed strategies.

A fully revealing equilibrium (s_1, s_2, y) supported by a belief function that deters local deviations with local punishments is called a *robust* fully revealing equilibrium.[17][18]

Definition 2 requires that after any pair of incompatible reports (θ', θ'') such that the two messages are close and also map to optimal feasible policies that are close, the receiver's belief, and hence his optimal response, satisfy three conditions: it must be close to the optimal feasible policies corresponding to the reports, and it must ensure that sender 1 (resp., 2) in state θ'' (resp., θ') could not gain by deviating to a report of θ' (resp., θ''). Figure 2 illustrates these requirements.

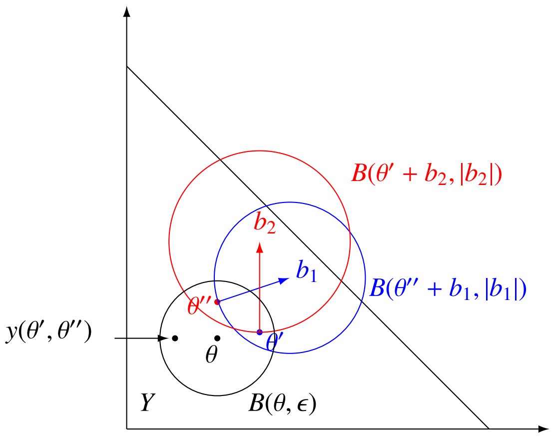


Figure 2: Robust punishment: the policy $y(\theta', \theta'')$ deters the deviation (θ', θ''), since $y(\theta', \theta'')$ lies outside the preferred sets of both senders (blue circle for sender 1 and red circle for sender 2). This policy is also robust since it is close to both θ' and θ''.

If Y is convex and/or $\Theta = Y$, then condition (*ii*) in Definition 2 is superfluous: in either case, whenever θ' and θ'' are close, $y^*(\theta')$ and $y^*(\theta'')$ are also close. Condition (*ii*) is relevant when Y is non-convex and $Y \subset \Theta$, because in this case, small changes in θ for $\theta \notin Y$ could result in large changes in the receiver's optimal feasible policy $y^*(\theta)$.

There are two interpretations of the type of small mistakes to which Definition 2 requires the receiver to respond with only small punishments. That is, there are two interpretations for why, even if senders do not intend to mislead the receiver, the receiver might nevertheless receive incompatible reports. First, there might be some noise in the communication process, with the result that the receiver might not interpret the messages exactly as the senders intended. Second, even if the communication process were noiseless, the senders might not perceive the state perfectly accurately, and their errors might not be perfectly correlated. Under either interpretation of mistakes, our analysis would apply when the senders and the receiver were unaware that these mistakes might happen. Our robust-

[17] Strategies (s_1, s_2) can together be fully revealing even if each sender's report by itself does not fully reveal the optimal feasible policy. Battaglini's (2002) construction of a fully revealing equilibrium for an unrestricted multidimensional state space is an example of this possibility. We have stated Definition 2 in a way that allows for this possibility.

[18] This concept of robustness can be shown to coincide with the concept of diagonal continuity introduced by Ambrus and Takahashi (2008). However this representation of the concept is easier to work with.

ness requirement ensures that as the size of the mistakes goes to zero, the outcome in the presence of mistakes approaches the outcome when mistakes never occur.

In a fully revealing equilibrium (FRE), the receiver perfectly learns the optimal feasible policy from the pair of messages, and neither sender has an incentive to try to mislead the receiver by sending a different message. Using a similar argument to the Revelation Principle we can, without loss of generality, concentrate on fully revealing equilibria in which each sender truthfully reports the optimal feasible policy given his observation. The strategies (s_1, s_2) are *truthful* if $\mathcal{M}_1 = \mathcal{M}_2 = Y$ and $s_i(\theta) = y^*(\theta)$. An equilibrium with truthful strategies is called a *truthful equilibrium*.

Lemma 1 is an extension of Lemma 1 in Battaglini (2002) that incorporates our notion of robustness. It considerably simplifies our subsequent analysis.

Lemma 1. *For any (robust) fully revealing equilibrium there exists a (robust) truthful equilibrium that is outcome-equivalent to it.*

Proof: All proofs are in the Appendix.

2.2 Independence of the Magnitudes of the Biases

Given a bias vector $b \in \mathbb{R}^p$, we define the magnitude of the bias as the Euclidean norm of the vector, $|b|$. We focus on fully revealing equilibria in which strategies are independent of the magnitudes of the biases:

Definition 3. Given biases $b_1, b_2 \in \mathbb{R}^p$ and a (robust) FRE (s_1, s_2, y^μ), the equilibrium is *independent of the magnitudes of the biases* if it remains a (robust) FRE for all biases $t_1 b_1, t_2 b_2$, with $t_1, t_2 \in (0, +\infty)$.

Note that if there exists a FRE for bias vectors (b_1, b_2), then the same strategies constitute a FRE for any biases $(t_1 b_1, t_2 b_2)$, where $0 < t_1, t_2 \leq 1$. The reason is that, for $0 < t_i \leq 1$, the set of policies that, in state θ, a sender with bias b_i prefers to θ contains the set of policies preferred to θ by a sender with bias $t_i b_i$: formally, $B(\theta + tb, |tb|) \subseteq B(\theta + b, |b|)$. Therefore, the same punishment that will deter deviations by a sender with bias b_i will deter deviations by one with smaller bias $t_i b_i$. (See Figure 3a.) Requiring the FRE to be independent of the magnitudes of the biases is thus equivalent to requiring the *same strategies* to remain a FRE even as the magnitudes of the biases become arbitrarily large. Note that such a FRE would thus remain a FRE even if the receiver were *uncertain* about the magnitudes of the biases.[19]

Besides being invariant to the intensity of senders' preferences, such equilibria are also appealing because of the relative tractability of their characterization. As observed by Levy and Razin (2007), the indifference curves of a sender with a very large bias are very close to hyperplanes orthogonal to the bias vector, and such a sender's ranking of policies is approximately independent of the true state of the world. This observation provides a very simple characterization of the set of policy choices for the receiver that would deter the sender from misreporting the state. (See Figure 3b.)

[19]Section 5 shows how our characterization results can be extended when the receiver is also, to some degree, uncertain about the directions of the senders' biases.

(a) The ball $B(\theta + b, |b|)$ represents the set of policies preferred to θ by a sender with bias b. Scaling down the bias to tb with $0 < t \leq 1$ implies that $B(\theta + tb, |tb|) \subseteq B(\theta + b, |b|)$. Hence any policy that deters deviations from a sender with bias b also deters deviations from a sender with smaller bias tb.

(b) As the magnitude of the bias increases, the indifference curve passing through θ converge to the hyperplane orthogonal to b and passing through θ, denoted $h(b, b\theta)$. The shaded half space represents the set of policies that are never preferred to θ independently of the magnitude of the bias.

Figure 3: Independence of the magnitudes of the biases.

Requiring the equilibrium to be independent of the magnitudes of the biases might be seen as a strong requirement. However, we will show in Proposition 5 that when the policy space is two-dimensional and robustness is required, the existence of a FRE becomes no more likely even if we drop the requirement of independence of the magnitudes of the biases.

Before proceeding, we introduce two additional pieces of notation that will be used throughout the paper. Given a (bias) vector $b \in \mathbb{R}^p$ and a scalar $k \in \mathbb{R}$, we define $H(b, k) \equiv \{x \in \mathbb{R}^p \mid bx > k\}$ and $h(b, k) \equiv \{x \in \mathbb{R}^p \mid bx = k\}$. In words, $H(b, k)$ is the half-space composed of all the points in $\mathbb{R}^p$ whose inner product with b is greater than k, and $h(b, k)$ is the boundary of $H(b, k)$. For any point $y \in H(b, b\theta)$, we have $by > b\theta$, and there exists a scalar $t > 0$ large enough that $y \in B(\theta + tb, t|b|)$, i.e. in state θ, the policy y is preferred to θ by a sender with bias tb. (See Figure 3b.) In order for a given policy choice of the receiver to deter misreporting independently of the magnitude of the sender's bias, the receiver's choice must lie in $\mathbb{R}^2 \setminus H(b, b\theta)$, which corresponds to the shaded half-space in Figure 3b.

2.3 Preliminary Results

Propositions 1 and 2 below allow us to abstract from specifying particular belief functions when proving the existence or nonexistence of robust fully revealing equilibria that are independent of the magnitudes of the senders' biases. Proposition 1 deals with the case in which the policy space coincides with the state space, and Proposition 2 extends the result to the more general case in which the policy space is a subset of the state space.

The first part of Proposition 1 provides a necessary and sufficient condition for the existence of a fully revealing equilibrium that is independent of the magnitudes of the biases.[20]

[20]This condition coincides with the condition that Ambrus and Takahashi (2008) show in their Proposition 7 to

The second part of Proposition 1 establishes a necessary and sufficient condition for the existence of a belief that deters local deviations with local punishments and is independent of the magnitudes of the biases. Finally, we show that the two conditions together are not only necessary but also sufficient for the existence of a robust FRE that is independent of the magnitudes of the biases.

Proposition 1. *Suppose $Y \equiv \Theta \subseteq \mathbb{R}^p$. Given $b_1, b_2 \in \mathbb{R}^p$,*

(i) There exists a fully revealing equilibrium that is independent of the magnitudes of the biases if and only if

$$\text{for any } (\theta', \theta'') \in Y, \quad Y \nsubseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta') \tag{1}$$

(ii) There exists some fully revealing strategies and a belief function that deters local deviations with local punishments and is independent of the magnitudes of the biases if and only if

$$\text{for any } \theta \in Y \text{ and any } \epsilon > 0, \text{ there exists } \delta > 0 \text{ such that for any } (\theta', \theta'') \in B(\theta, \delta) \cap Y, \\ B(\theta, \epsilon) \cap Y \nsubseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta') \tag{2}$$

(iii) Conditions (1) and (2) are necessary and sufficient for the existence of a robust fully revealing equilibrium that is independent of the magnitudes of the biases.

When condition (1) holds, the receiver's policy rule $y^\mu(\theta', \theta'')$ in a truthful equilibrium will satisfy $y^\mu(\theta', \theta'') = \theta'$ if $\theta' = \theta''$ and $y^\mu(\theta', \theta'') \in Y \setminus H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ if $\theta' \neq \theta''$. Such a rule is feasible and ensures that sender 1 (resp., 2) has no incentive to deviate to a report of θ' (resp., θ'') when the true state is θ'' (resp., θ'), even for arbitrarily large magnitudes of the biases. Condition (2) implies that a feasible punishment can be found arbitrarily close to the reports when these converge to each other. Henceforth we will refer to $\mathbb{R}^p \setminus H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ as the "punishment region" for the deviation (θ', θ'') and to $Y \setminus H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ as the "feasible punishment region". See Figure 4.

Proposition 2 states that there exists a (robust) fully revealing equilibrium for $Y \subseteq \Theta$, if and only if there exists a (robust) fully revealing equilibrium when the space state is reduced to coincide with the policy space. In other words, when determining whether or not a (robust) FRE exists, we can ignore those states that cannot be implemented as a policy. Given a state space Θ with dimension p and a policy space Y with dimension $q \leq p$, denote by S^Y the q-dimensional subspace of $\mathbb{R}^p$ such that $Y \subseteq S^Y$. Given a vector $b \in \mathbb{R}^p$, we denote by b^Y the projection of b onto S^Y.

---

be equivalent, for the case of compact Y, to the existence of a FRE for arbitrarily large biases. Note, however, that our requirement that there exist a FRE that is independent of the magnitudes of the biases is in general stronger than Ambrus and Takahashi (2008)'s requirement that there exists a FRE for arbitrarily large biases: we require the *same* equilibrium strategies to remain an equilibrium no matter how large the magnitudes of the biases, whereas they require that for any magnitude of the biases there exists a FRE (which might depend on the magnitudes). For instance, when the policy space is not compact, there always exists a FRE for arbitrarily large biases, so Ambrus and Takahashi (2008)'s Proposition 7 does not hold in this case; in contrast, when Y is not compact, there might not exist a FRE that is independent of the bias magnitudes. Our Proposition 1 is valid whether or not Y is compact.

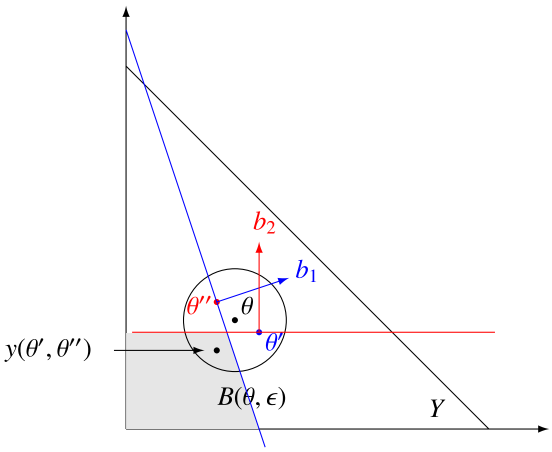


Figure 4: The shaded area represents the feasible punishment region given the incompatible pair of reports (θ', θ''). The intersection of the shaded area with the ball $B(\theta, \epsilon)$ represents the *local* punishment region.

Proposition 2. *Given $\Theta \subseteq \mathbb{R}^p$, $Y \subseteq \mathbb{R}^q$ with $Y \subseteq \Theta$ and given $b_1, b_2 \in \mathbb{R}^p$ and their projections $b_1^Y, b_2^Y \in S^Y$, the following two statements are equivalent:*

(i) For the environment (Θ, Y) and biases (b_1, b_2), there exists a (robust) fully revealing equilibrium that is independent of the magnitudes of the biases.

(ii) For the environment (Y, Y) and biases (b_1^Y, b_2^Y), there exists a (robust) fully revealing equilibrium that is independent of the magnitudes of the biases.

Given Proposition 2, the shape of the state space Θ is irrelevant (as long as $Y \subseteq \Theta$), and all that matters for the existence of a (robust) FRE is the shape of the policy space, relative to the projections of the senders' bias vectors onto the minimal subspace, S^Y, containing the policy space. Therefore, when proving existence results for (robust) FRE's, we can without loss of generality focus on the case in which $\Theta \equiv Y$, and hence $y^*(\theta) = \theta$. Proposition 1, which is stated for the case $\Theta \equiv Y$, will be our primary tool. For the sake of simplicity, and with some abuse of language and notation, for the rest of the paper we will refer to the projections of the p-dimensional biases onto the relevant q-dimensional subspace S^Y such that $Y \subseteq S^Y$ as *the biases* and denote them directly by b_1, b_2.

Finally, we discuss two special cases where, for any number of dimensions and any shape of Y, it is straightforward to draw conclusions about the existence of a robust fully revealing equilibrium that is independent of the magnitudes of the biases. First, if the senders' bias vectors are in exactly the same direction (i.e. $b_1 = tb_2$ for some strictly positive scalar t), then there always exists a robust FRE that is independent of the magnitudes of the biases. In it, the receiver responds to any discrepancy between the messages by choosing whichever of the two reported states leads to a smaller inner product with (each of) the bias vectors. In other words, the receiver's chosen policy coincides with whichever of the reported states would be less preferred by both senders, if both biases were sufficiently large. Such a strategy for the receiver ensures that neither sender can strictly gain by deviating from truthful reporting, no matter how large his bias. Furthermore, since the receiver's chosen policy always coincides with one of the senders' messages, this FRE satisfies our definition of robustness. (See Figure 5a.)

Second, if the biases are exactly opposite (i.e. $b_1 = tb_2$ for some strictly negative scalar t), then it follows from part (i) of Proposition 1 and Proposition 2 that there exists a FRE that is independent of the magnitude of the biases "if and only if Y is included in a lower dimensional hyperspace that is orthogonal to the direction of the biases" (Ambrus and Takahashi (2008, p.13)). In addition, it follows from part (ii) Proposition 1 that when a FRE exists in this case, a robust FRE exists as well: a (truthful) robust FRE is supported by a response function for the receiver such that $y(\theta', \theta'') = \lambda\theta' + (1 - \lambda)\theta''$, for $\lambda \in [0, 1]$. (See Figure 5b.)

(a) Aligned biases: $y(\theta', \theta'') = \theta''$ provides a feasible robust punishment. The shaded area is the punishment region given (θ', θ'').

(b) Opposite biases: the punishment region given (θ', θ'') is empty.

Figure 5: Special Cases: aligned biases and opposite biases.

For the remainder of the paper, we will exclude these two special cases and assume that b_1 and b_2 are linearly independent.

3 Convex Policy Space

We begin by focusing on convex policy spaces in an arbitrary number of dimensions. Proposition 3 below shows that, when focusing on strategies that are independent of the magnitudes of the biases, whenever there exists a fully revealing equilibrium (FRE), there also exists a robust FRE. In fact, if small deviations cannot be deterred with small punishments, then they cannot be deterred with any feasible punishments. Moreover, we show that for convex state spaces that are two-dimensional or multidimensional and compact, it is sufficient for existence of a FRE that is independent of the magnitude of the biases (robust or not) that small deviations can be deterred (with small punishments). These preliminary results are extremely useful, because they show that a) robustness is, perhaps surprisingly, not a restrictive requirement on a FRE when it is independent of the magnitudes of the biases and the policy space is convex; and b) in the two-dimensional or compact multidimensional cases, we need only ensure that local deviations can be punished.

Proposition 3. *Given $Y \subseteq \mathbb{R}^q$ convex and $b_1, b_2 \in \mathbb{R}^q$ linearly independent, the following statements are equivalent:*

(i) There exists a fully revealing equilibrium that is independent of the magnitudes of the biases.

(ii) There exists a robust fully revealing equilibrium that is independent of the magnitudes of the biases.

When we further assume that a) $Y \subseteq \mathbb{R}^2$ or that b) $Y \subseteq \mathbb{R}^q$ and Y is compact, then the following statement is also equivalent to the previous two:

(iii) Local deviations can be deterred with local punishments that are independent of the magnitudes of the biases.

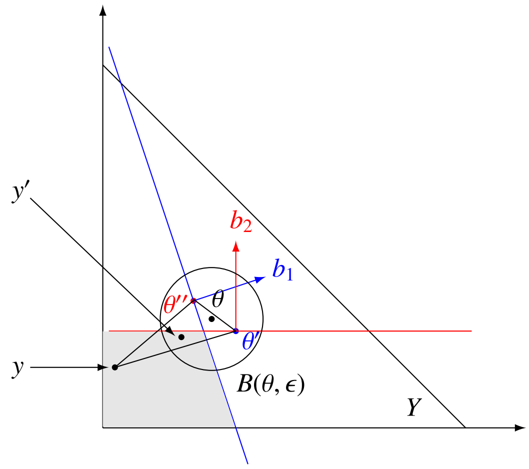


Figure 6: Y convex: A local punishment exists whenever a punishment exists. The grey area represents the feasible punishment region given (θ', θ'').

The intuition behind the first equivalence in Proposition 3 is illustrated in Figure 6. Given an incompatible pair of reports (θ', θ''), consider the feasible punishment region $Y \setminus (H(b_1, b_1\theta'') \cup H(b_2, b_2\theta'))$ (shaded area). If there exists a FRE, then the feasible punishment region has to be non empty. Denote by y a feasible punishment. Given that Y is convex, and θ', θ'' and y are all feasible policies, the triangle of convex combinations of these three policies is also feasible. As θ' and θ'' converge to θ, there are policies, such as y', that belong to that triangle and that lie in the intersection of the feasible punishment region and the ball $B(\theta, \epsilon)$, so are feasible local punishments.

Figure 7 illustrates the intuition for why, when $Y \subseteq \mathbb{R}^2$, local deterrence of local deviations is sufficient for existence of a FRE. Suppose there were a (large) deviation, such as (θ', θ''), that could not be deterred, i.e., the punishment region (dark grey area) does not intersect the policy space. Then it would be possible to construct a local deviation $(\theta', \tilde{\theta}'')$, by choosing $\tilde{\theta}''$ along the segment $[\theta', \theta'']$ and close to θ', and by the convexity of Y, the whole of the punishment region for that deviation would still be infeasible.

The result of Proposition 3 might be surprising in the light of Ambrus and Takahashi's (2008; Section 4.2)'s negative findings regarding the existence of robust fully revealing equilibria. They provided examples in which a robust FRE did not exist even though for sufficiently small magnitudes of the biases it was possible to construct a FRE. However, in all of these examples, a FRE did not exist for sufficiently large biases.

In Proposition 5 in Section 3.1, we show that for two-dimensional spaces, once the equilibrium is required to be robust, the magnitudes of the biases do not play any role.

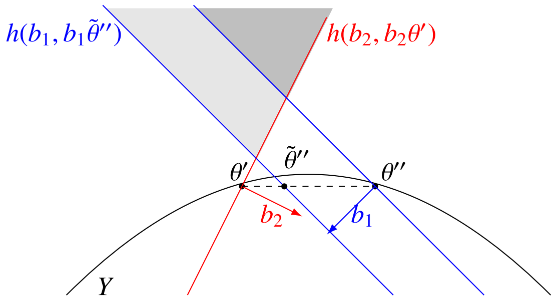


Figure 7: $Y \subseteq \mathbb{R}^2$ and convex: When a large deviation such as (θ', θ'') cannot be deterred, there is a local deviation, $(\theta', \tilde{\theta}'')$, that cannot be deterred with local punishments. The shaded areas correspond to the punishment regions given (θ', θ'') (dark grey) and given $(\theta', \tilde{\theta}'')$ (light and dark grey).

Intuitively, even a small bias can be regarded as extremely large when we consider local punishments after very small disagreements among senders. And hence, if there does not exist a FRE that is independent of the magnitudes of the biases, then no robust FRE exists, even for small magnitudes of the biases.

In the next subsection, we provide a local necessary and sufficient condition for the existence of a robust FRE independent of the magnitudes of the biases that is constructive in the sense that it provides a policy rule that implements a robust FRE whenever one exists.

3.1 Policy Space a Subset of $\mathbb{R}^2$

In this section, we focus on the case where the policy space is two-dimensional. We start by defining a particular policy rule for the receiver that, if feasible, implements a robust FRE.

Definition 4. Given $b_1, b_2 \in \mathbb{R}^2$ linearly independent, the *Min Rule* is the function from $\mathbb{R}^2 \times \mathbb{R}^2$ to $\mathbb{R}^2$ that, for every pair of reports $(\theta', \theta'') \in \mathbb{R}^2$, implements the point $\theta' \wedge_{\{b_1,b_2\}} \theta''$ in $\mathbb{R}^2$ defined by:

$$b_1(\theta' \wedge_{\{b_1,b_2\}} \theta'') = \min\{b_1\theta', b_1\theta''\}$$
$$b_2(\theta' \wedge_{\{b_1,b_2\}} \theta'') = \min\{b_2\theta', b_2\theta''\}.$$

In words, the Min Rule selects the policy which is the coordinate-wise minimum of the senders' reports, using the coordinate system formed by the normal vectors to the biases. If these reports coincide, the Min Rule selects the common report. It follows from the definition that the inner product of the policy selected by the Min Rule with each bias vector is weakly smaller than the inner product of each of the reports with that bias. Hence, the chosen policy $\theta' \wedge_{\{b_1,b_2\}} \theta''$ is, regardless of the magnitudes of the biases, weakly worse for both senders than both of the reports provided, so can act as a punishment for the deviation $\theta' \neq \theta''$.[21] Figure 8 illustrates the Min Rule for two different scenarios. Note that the

[21] To see this, denote by $y = \theta' \wedge_{\{b_1,b_2\}} \theta''$ so $b_i y \leq b_i\theta$ for $i = 1, 2$. Then $(\theta + b_i - y)^2 = b_i^2 + (\theta - y)^2 + 2b_i(\theta - y) > b_i^2 = (\theta + b_i - \theta)^2$. In other words, for both senders, y is farther from $\theta + b_i$ than θ is, and hence neither sender has an incentive to deviate from truthfully reporting θ.

policy selected by the Min Rule can coincide with one of the reports, as in Figure 8a, or can implement a completely distinct point, as in Figure 8b.

(a) Min Rule when the Senders have the same ranking over reports. In this case both senders prefer θ' to θ''.

(b) Min Rule when the Senders have different rankings over reports. In this case Sender 1 prefers θ' to θ'', whereas Sender 2 prefers θ'' to θ'.

Figure 8: Min Rule for the receiver

The Min Rule is anonymous in that it selects the same policy in response to (θ', θ'') and (θ'', θ'). Among all anonymous rules, the Min Rule prescribes the punishment that is least severe for each of the senders, subject to deterring both of them from deliberately misreporting, no matter how large their biases. In particular, as the two reports converge to each other, the policy selected by the Min Rule also converges to the same point, and hence if the punishment is feasible, it constitutes a robust punishment that is independent of the magnitudes of the biases.

It remains to determine when it is feasible for the receiver to respond to a deviation according to the Min Rule. Proposition 4 below provides a simple geometric condition, which we term the *Local Deterrence Condition*, that is necessary and sufficient for the existence of a robust FRE. The proposition also shows that the Min Rule is a feasible policy if and only if the Local Deterrence Condition is satisfied. In other words, the feasibility of punishing deviations according to the Min Rule is necessary as well as sufficient for the existence of a robust FRE that is independent of the magnitudes of the biases.

Before stating the characterization, we need to introduce some additional notation. Given $S \subset \mathbb{R}^2$ closed and convex, we denote the frontier of S by $Fr(S)$. We say that a point $s \in Fr(S)$ is *smooth* if there exists a unique tangent hyperplane to $Fr(S)$ at s. Any point in $Fr(S)$ that is not smooth will be called a *kink*. The set of smooth points in the frontier is denoted by $\widetilde{Fr}(S)$. For any $s \in \widetilde{Fr}(S)$, we denote by $n_S^{In}(s)$ the unit normal vector to $Fr(S)$ at s in the *inward* direction to S: $n_S^{In}(s)$ is the unique vector that satisfies $n_S^{In}(s)(s' - s) \geq 0$ for all $s' \in S$.

Given $b_1, b_2 \in \mathbb{R}^2$, define $C(b_1, b_2) \equiv \{b \in \mathbb{R}^2 \mid b = \alpha b_1 + \beta b_2 \mid \alpha, \beta > 0\}$, the *open* convex cone spanned by the vectors b_1, b_2. Similarly, adapting notation used for intervals in $\mathbb{R}$, define $C[b_1, b_2] \equiv \{b \in \mathbb{R}^2 \mid b = \alpha b_1 + \beta b_2 \mid \alpha, \beta \geq 0\}$, the *closed* convex cone spanned by the vectors b_1, b_2 and $C(b_1, b_2] \equiv \{b \in \mathbb{R}^2 \mid b = \alpha b_1 + \beta b_2,$ with $\alpha \geq 0, \beta > 0\}$, the convex cone that includes the extreme direction b_2 but not b_1.

Proposition 4. *Given $Y \subseteq \mathbb{R}^2$ convex and $b_1, b_2 \in \mathbb{R}^2$ linearly independent, the following statements are equivalent:*

(i) There exists a (robust) fully revealing equilibrium that is independent of the magnitudes of the biases.

(ii) For every $\theta \in \widetilde{Fr}(Y)$, $\quad n_Y^{In}(\theta) \notin C(b_1, b_2)$. *(Local Deterrence Condition)*

(iii) For every $\theta', \theta'' \in Y$, $\quad \theta' \wedge_{\{b_1,b_2\}} \theta'' \in Y$. *(Feasibility of Min Rule)*

Condition (*iii*) says that responding to any deviation according to the Min Rule is a feasible strategy for the receiver, and hence by using the Min Rule, the receiver can deter deviations in a robust way. Whenever the reports (θ', θ'') of the senders do not agree, the receiver's action is rationalized by a belief that allocates mass one to $\theta' \wedge_{\{b_1,b_2\}} \theta'' \in Y$.

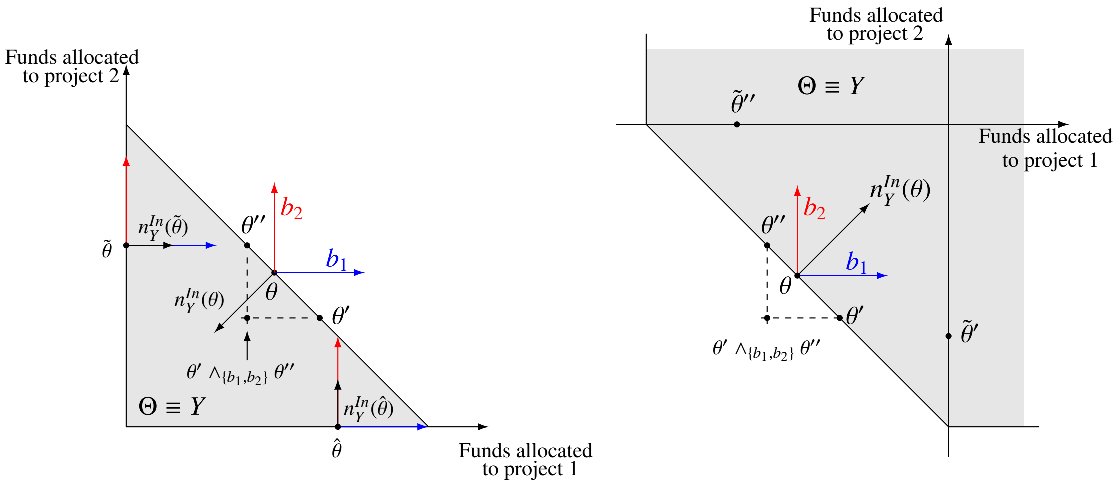
The Local Deterrence Condition (condition (ii)) is a condition on the directions of the senders' bias vectors relative to the frontier of the policy space. For a given smooth point on the frontier, $\theta \in \widetilde{Fr}(Y)$, the condition $n_Y^{In}(\theta) \notin C(b_1, b_2)$ is satisfied if and only if, in state θ, there exists, close to θ, a feasible policy for the receiver that, no matter how large the magnitudes of the biases, is worse for both senders than the policy $y = \theta$.

To understand the intuition behind the proposition, note that by Proposition 3, to prove the existence of a robust FRE it is enough to check whether local deviations can be punished locally. Whether this is the case is exactly what the Local Deterrence Condition establishes. Clearly, if a pair of incompatible reports converge to a point in the *interior* of the policy space, punishing locally is never problematic, since the punishment region always intersects the policy space in a neighbourhood of the point. However, if a pair of incompatible reports converge to a point θ on the *frontier* of the policy space, then punishing locally requires that, close to θ, there exist a feasible policy that, for any bias magnitudes, would make both senders worse off than the policy $y = \theta$.

In what follows we illustrate the results of Proposition 4 through a couple of examples. Consider the funding allocation game introduced in the Introduction and depicted in Figure 9a. The receiver has to allocate funds from a budget to two different projects, each one overseen by one of the senders. The maximum amount of funds that can be allocated is exogenously fixed, and each project must receive a non-negative level of funding. The optimal allocation of funding depends on some information known to the senders but not to the receiver. Relative to the receiver, each sender is biased towards the project he oversees. Since the inward normal vector to smooth points on the frontier of the policy space has only three distinct orientations in this example, it is particularly easy to confirm that the Local Deterrence Condition is satisfied in this setting, as illustrated in Figure 9a. It follows that all local deviations can be punished locally.

The Min Rule for the receiver explicitly identifies a feasible local punishment in response to any local deviation. For the orthogonal biases shown, the Min Rule chooses the component-wise minimum of the senders' reports with respect to the Euclidean coordinates: $\theta' \wedge_{\{b_1,b_2\}} \theta'' = (\min\{\theta'_1, \theta''_1\}, \min\{\theta'_2, \theta''_2\})$. It is also easy to confirm directly from Figure 9a that in this setting the punishment specified by the Min Rule is always feasible.

Note that big deviations are also deterred by the receiver using the Min Rule. For example, if each sender claims the whole budget for his own project, the Min Rule prescribes that the receiver allocate zero funds to each of them. But by Proposition 3, we do not need to explicitly check whether big deviations can be deterred: as long as local deviations are deterrable, big deviations will also be deterrable.



(a) Allocation of funds given a budget constraint. The shaded area represents the feasible allocations.

(b) Allocation of funds across departments given a maximum total budget cut. The shaded area corresponds to the feasible allocations.

Figure 9: Illustration of robust FRE in the case of allocation of funds/cuts.

Figure 9b illustrates why the Local Deterrence Condition is necessary for existence of a robust FRE. In Figure 9b, the receiver, a local government, needs to determine the allocation of funding for two departments; depending on the state of the world, the optimal allocation might involve some budget cuts up to a maximum total level of cuts. As can be seen in the figure, the Local Deterrence Condition is violated at θ, a smooth point on the segment of the frontier representing the maximum total level of cuts. The implication of this violation is that incompatible reports along the frontier, such as (θ', θ''), cannot be deterred, since doing so in a manner independent of the magnitudes of the biases would require imposing as a punishment deeper total cuts than the maximum level allowed.

One might argue that violations of the senders' incentives for truthtelling along the frontier are a minor problem if, for instance, the probability of those states arising is close to zero.[22] However, the fact that local deviations along the frontier are not deterrable implies that more general and bigger deviations are not deterrable either. Consider, in Figure 9b, the pair of reports $(\tilde{\theta}', \tilde{\theta}'')$, which correspond to each sender arguing that some cuts are needed but not in his own department. There is no feasible punishment that would deter such a deviation independently of the magnitudes of the biases.

To introduce the final result of this subsection recall that, as Proposition 3 showed, requiring of a FRE that is independent of the magnitudes of the biases that it also be robust

[22] As argued above, local deviations from a point in the interior of the policy space are always deterrable regardless of the shape of the policy space.

does not restrict the circumstances under which it exists. We now show, in Proposition 5 below, that existence of a robust FRE becomes no more likely even if we drop the requirement that it be independent of the magnitudes of the biases. The reason is that when the receiver is constrained to use small punishments, then whether the senders have incentives to deviate from truthtelling depends only on the orientations, not the magnitudes, of their bias vectors.

Proposition 5. *Given $Y \subseteq \mathbb{R}^2$ convex and $b_1, b_2 \in \mathbb{R}^2$ linearly independent, the following statements are equivalent:*

(i) There exists a robust fully revealing equilibrium.

(ii) There exists a (robust) fully revealing equilibrium that is independent of the magnitudes of the biases.

Proposition 5, together with Proposition 4, implies that the Local Deterrence Condition is a necessary and sufficient condition for the existence of a robust FRE for given biases b_1, b_2.

3.2 Collusion Proofness

Since the senders have common information and could potentially attempt to mislead the receiver by both making the same false report, another desirable property of a fully revealing equilibrium is collusion-proofness. We will say that a FRE is collusion-proof if, whenever there is a feasible policy $\hat{\theta}$ that is at least weakly preferred by both senders to the true state θ, colluding by both reporting $\hat{\theta}$ would not be self-enforcing for the senders, because at least one of them would have a profitable unilateral deviation. To state this definition formally, define $\overline{B}(x, r)$ to be the *closed* ball with centre x and radius r.

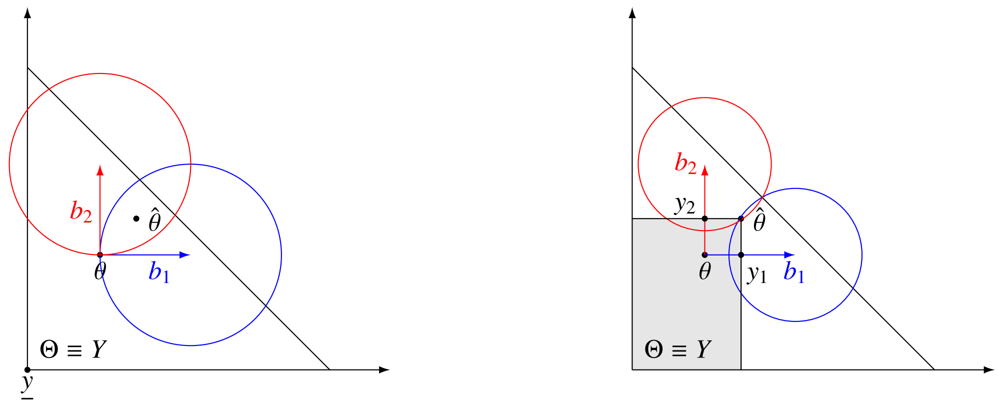
Definition 5. Given bias vectors b_1, b_2 and given a FRE (s_1, s_2, y^μ), we say that the equilibrium is *collusion-proof*, if for all $\theta \in Y$ and all $\hat{\theta} \in Y$ such that

$$\hat{\theta} \in \overline{B}(\theta + b_1, |b_1|) \cap \overline{B}(\theta + b_2, |b_2|),$$

there exists $\tilde{\theta} \in Y$ such that either:

$$|\theta + b_1 - y^\mu(\tilde{\theta}, \hat{\theta})| < |\theta + b_1 - \hat{\theta}| \quad \text{or} \quad |\theta + b_2 - y^\mu(\hat{\theta}, \tilde{\theta})| < |\theta + b_2 - \hat{\theta}|.$$

Given a FRE, the specific beliefs held by the receiver following incompatible reports will affect whether or not the FRE is collusion-proof, since these beliefs will determine the receiver's response if one of the senders were to deviate from the collusive plan. Intuitively, the more severely the receiver punishes incompatible reports, the less incentive either sender will have to deviate from the collusive plan, and hence the less likely a FRE is to be collusion-proof. To illustrate, consider the same funding allocation game depicted in Figure 9a, which is reproduced in Figure 10a. For this game, there exists a non-robust FRE that is independent of the magnitudes of the biases, in which the receiver responds to any incompatible reports by choosing policy $\underline{y}$: policy $\underline{y}$ is such that, for any true state $\theta \neq \underline{y}$, it is strictly worse for both senders than θ. It follows that whenever the true state θ is such that there exists a policy $\hat{\theta}$ that is preferred by both senders to θ, colluding by both reporting $\hat{\theta}$ would be self-enforcing. So a FRE supported by such a harsh punishment strategy by the receiver is not collusion-proof.



(a) The FRE supported by the extreme punishment $\underline{y}$ is not collusion-proof: in state θ, collusion on $\hat{\theta}$ is self-enforcing.

(b) The FRE supported by the Min Rule is collusion-proof.

Figure 10: The severity with which the receiver punishes incompatible reports affects whether or not a FRE is collusion-proof.

Now, in the same funding allocation game, consider the robust FRE supported by the receiver using the Min Rule. Figure 10b illustrates the argument. Suppose that the true state is θ, and consider whether collusion on the report $\hat{\theta}$, which is preferred by both senders, would be self-enforcing. For either sender, if he expected the other sender to report $\hat{\theta}$ and the receiver to respond using the Min Rule, then the set of policies which he could induce the receiver to choose is indicated by the shaded region in the figure. (Recall that the Min Rule is an anonymous rule.) Sender 1 could increase his payoff by deviating from the collusive report of $\hat{\theta}$ to the report y_1, thereby inducing the policy y_1. Similarly, Sender 2 could gain by deviating from $\hat{\theta}$ to y_2, thereby inducing the receiver to choose y_2. In fact, in the environment depicted in Figure 10b, whenever there is a feasible policy $\hat{\theta}$ that is weakly preferred by both senders (and at least strictly preferred by one of them) to the true state θ, a plan to collude by both reporting $\hat{\theta}$ would give both senders an incentive to unilaterally deviate. So the FRE supported by the receiver using the Min Rule is collusion-proof.

As we noted when we defined the Min Rule in Definition 4, this is the anonymous rule that, given incompatible reports, prescribes the *least* severe punishment for each of the senders, subject to deterring deliberate misreporting, no matter how large their biases. Intuitively, then, we would expect that the Min Rule should be the strategy for the receiver that would make deviations from the collusive report *most* attractive. The next lemma formalizes the benefit of the Min Rule in deterring collusion by the senders.

Lemma 2. *If a FRE supported by the receiver using the Min Rule is not collusion-proof, then no other FRE that is independent of the magnitude of the baises can be collusion-proof.*

Two things follow from Lemma 2: It is sufficient to focus on the FRE supported by the Min Rule in order to see whether a collusion-proof FRE exists, and whenever a collusion-proof FRE exists, there exists a *robust* FRE that is collusion-proof.

The dimension and shape of Y are both important in determining whether or not a FRE supported by the Min Rule is collusion-proof. When $Y \subseteq \mathbb{R}$, then it follows from the discussion at the end of Section 2.3 that a robust FRE exists if and only if the senders' bias vectors point in the same direction. Yet even when a robust FRE exists with Y one-dimensional, no such FRE, even one supported by the Min Rule, is collusion-proof.[23]

On the other hand, suppose Y is the whole of $\mathbb{R}^2$. Since the Min Rule is clearly always feasible, by Proposition 4 there exists a robust FRE supported by it. Here, as long as the senders' bias vectors are linearly independent, the robust FRE supported by the Min Rule is collusion-proof. This follows from the same logic illustrated in Figure 10b. Suppose the senders planned to collude by both reporting some $\hat{\theta}$ that they both preferred to the true state θ. Any sender i for whom $\hat{\theta}$ did not lie on the line through θ in the direction of b_i would, in state θ, strictly prefer the receiver to choose the projection, y_i, of $\hat{\theta}$ onto this line—since the policy space is the whole of $\mathbb{R}^2$ and the receiver is expected to use the Min Rule, Sender i could induce the receiver to choose y_i by deviating from reporting $\hat{\theta}$ to reporting y_i. And the fact that the senders' bias vectors are linearly independent guarantees that the projections y_1 and y_2 are distinct policies, so at least one sender will have a strict incentive to deviate.[24]

To characterize exactly how the shape of the set of feasible policies determines whether or not a FRE supported by the Min Rule is collusion-proof, we now introduce some further notation that will be used only in this section. Given a smooth point on the frontier $\theta \in \widetilde{Fr}(Y)$, we denote by $n_Y^{Out}(\hat{\theta})$ the *outward* normal vector to Y at $\hat{\theta}$, and by $t_Y(\hat{\theta})$ either of the two unit tangent vectors to Y at $\hat{\theta}$. For a generic $\hat{\theta}$ on the frontier of Y, we define the *Polar Cone* of Y at $\hat{\theta}$ as the convex cone:

$$PC_Y(\hat{\theta}) \equiv \{n \in \mathbb{R}^2 \mid n(\theta - \hat{\theta}) \leq 0, \text{ for all } \theta \in Y\}$$

Note that if $\hat{\theta}$ is a smooth point, then $PC_Y(\hat{\theta}) \equiv n_Y^{Out}(\hat{\theta})$, but if $\hat{\theta}$ is a kink point, then $PC_Y(\hat{\theta})$ has a non-empty interior.

The following proposition characterizes when the robust FRE in which the receiver uses the Min Rule is collusion-proof.

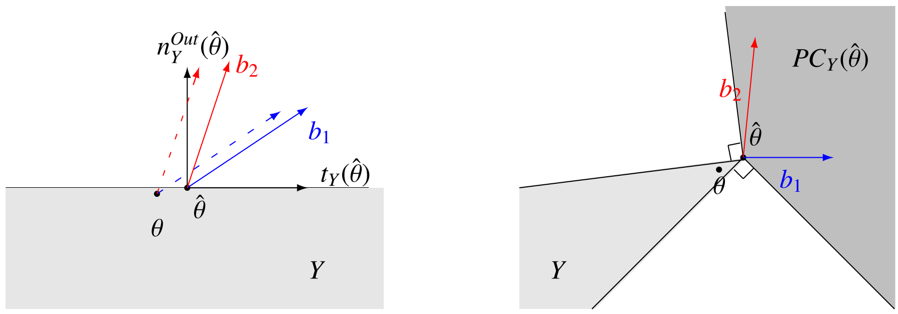
Proposition 6. *Let $Y \subseteq \mathbb{R}^2$ be convex and have a non-empty interior, and let the biases $b_1, b_2 \in \mathbb{R}^2$ be linearly independent. Whatever the magnitudes of the biases, the robust FRE supported by the Min Rule is collusion-proof if and only if the following two conditions are satisfied:*

[23]Given any finite biases pointing in the same direction, find a state θ and a candidate collusive report $\hat{\theta}$ sufficiently close to θ such that $\hat{\theta}$ lies, for each sender, strictly between θ and his ideal point $\theta + b_i$. Then neither sender would have an incentive to deviate from the collusive report $\hat{\theta}$, since any robust response rule for the receiver that supported a FRE would lead the receiver, in response to such a deviation, to choose a policy farther away than $\hat{\theta}$ from both senders' ideal points.

[24]If $b_1 = tb_2$ for some strictly positive scalar t, then even for $Y = \mathbb{R}^2$, there is no collusion-proof FRE. The argument parallels that for $Y \subseteq \mathbb{R}$.

(*i*) *There does not exist a* $\hat{\theta} \in \widetilde{Fr}(Y)$ *such that* $b_1, b_2 \in C(n_Y^{Out}(\hat{\theta}), t_Y(\hat{\theta})]$.

(*ii*) *There does not exist a* $\hat{\theta} \in Fr(Y)$ *such that* $b_1, b_2 \in int(PC_Y(\hat{\theta}))$.



(a) Failure of condition (i): In state θ, collusion on $\hat{\theta}$ is self-enforcing. The shaded region represents the policy space Y.

(b) Failure of condition (ii): In state θ, collusion on $\hat{\theta}$ is self-enforcing. The light grey region represents the policy space Y, and the dark grey region represents the polar cone $PC_Y(\hat{\theta})$.

Figure 11: The necessary and sufficient conditions for collusion-proofness

Proposition 6 shows that, given linearly independent bias vectors, there are only two situations in which a robust FRE supported by the receiver using the Min Rule fails to be collusion-proof.[25] These two situations are illustrated in the two panels of Figure 11. When condition (i) is violated at some $\hat{\theta}$, as in Figure 11a, then whatever the magnitudes of the biases, there must exist a state θ sufficiently close to $\hat{\theta}$, such that collusion on $\hat{\theta}$ would be self-enforcing in state θ: the frontier of Y renders infeasible the type of profitable unilateral deviations from $\hat{\theta}$ that were possible in Figure 10b. Violation of condition (ii), illustrated in Figure 11b, is equivalent to existence of a policy $\hat{\theta}$ which, for sufficiently large biases, is both senders' preferred policy in Y, regardless of the true state—it follows that in this case, for any finite biases, there must exist a θ sufficiently close to $\hat{\theta}$ that collusion on $\hat{\theta}$ would be self-enforcing in state θ, since in state θ, $\hat{\theta}$ is the best feasible policy for both senders.

Note that given Lemma 2, conditions (i) and (ii) together with the Local Deterrence Condition are necessary and sufficient for the existence of a robust FRE that is collusion-proof.

3.3 Multidimensional Spaces

The results in Proposition 4 extend easily to higher dimensions. For b_1, b_2 linearly independent, the only directions of conflict between the senders and the receiver are the ones in

[25]Both conditions (i) and (ii) depend only on the directions and not on the magnitudes of the biases. If they are satisfied, the FRE supported by the Min Rule is collusion-proof whatever the magnitudes of the biases. If, however, one of the conditions fails, then our proof shows that whatever the magnitudes of the biases, there is some true state θ and some report $\hat{\theta}$ such that collusion on $\hat{\theta}$ would be self-enforcing. The smaller the magnitudes of the biases, the closer to the true state the collusive report would have to be, in order to be strictly preferred by both senders to the true state. Nevertheless, whenever senders with given biases would have an incentive, in state θ, to collude on the report $\hat{\theta}$, so would senders whose biases had larger magnitudes.

the plane spanned by these two vectors. Thus, senders will not have incentives to deviate by misreporting dimensions of the state orthogonal to this plane. On the other hand, the receiver could potentially utilize these dimensions of no conflict to punish inconsistent messages. However, this strategy cannot be guaranteed to work for the receiver if the senders' biases can be arbitrarily large. Proposition 7 shows that it is necessary and sufficient to project the policy space onto the plane of the bias vectors and to check whether condition (*ii*) in Proposition 4 is satisfied by this two-dimensional projection.

Given $b_1, b_2 \in \mathbb{R}^q$ linearly independent, denote by $\Pi_b \subset \mathbb{R}^q$ the plane spanned by these two vectors. Denote by $Proj_b : \mathbb{R}^q \longrightarrow \Pi_b$ the orthogonal projection onto Π_b. We will denote by x_b a generic element of Π_b and by B_b and H_b the two-dimensional balls and half-spaces in the plane Π_b. Finally θ_b will denote a generic element of $Y_b \equiv Proj_b(Y)$.

Proposition 7. *Given $Y \subseteq \mathbb{R}^q$ convex and compact*[26]*, and $b_1, b_2 \in \mathbb{R}^q$ linearly independent, the following statements are equivalent:*

(i) There exists a (robust) fully revealing equilibrium that is independent of the magnitudes of the biases.

(ii) For every $\theta_b \in \widetilde{Fr}(Y_b)$, $n_{Y_b}(\theta_b) \notin C(b_1, b_2)$.

(iii) For every $\theta', \theta'' \in Y$, $\quad \theta'_b \wedge_{\{b_1,b_2\}} \theta''_b \in Y_b$.

Proposition 7 implies that for the existence in high-dimensional spaces of a FRE (robust or not) that is independent of the magnitudes of the biases, it is necessary and sufficient to look at the projection of the policy space onto the subspace of conflict of interest and see whether a FRE can be constructed there. The reason is that when the equilibrium is required to exist regardless of the magnitudes of the biases, then no given shift of the receiver's action in a direction orthogonal to the plane of the biases can be certain to serve as a punishment for a deviating sender. Therefore, to be sure that a deviation is punished, the receiver needs to choose an action whose projection on the plane of the biases is worse for both senders. Such an action exists if and only if the projection of the policy space onto the plane of the biases satisfies the Local Deterrence Condition in Proposition 4. Although condition (*iii*) does not directly pin down a single policy as a punishment to a deviation, it says that there exists a feasible policy whose projection is $\theta'_b \wedge_{\{b_1,b_2\}} \theta''_b$ and which would therefore deter such a deviation. Finally, as the two original reports converge to each other, the convexity of Y implies that the set of policies in Y whose projection onto Π_b is $\theta'_b \wedge_{\{b_1,b_2\}} \theta''_b$ contains points close to θ' and θ''.

Note that Proposition 5 does not extend to higher dimensions. If the magnitudes of the biases have known finite upper bounds, the version of the Local Deterrence Condition that appears in Proposition 7 (condition (*ii*)) is sufficient for existence of a robust FRE but not necessary. In this case, the receiver might be able to exploit the dimensions orthogonal

[26]For $Y \subseteq \mathbb{R}^q$ closed but not necessarily compact, there exists a (robust) FRE that is independent of the magnitudes of the biases if and only if there exists a (robust) FRE that is independent of the magnitudes of the biases in the projection Y_b. However, because the projection of a closed set Y does not need to be closed, we cannot simply use the characterization provided in Proposition 4. By requiring Y to be compact, we ensure that the proj ection Y_b is closed and hence that Porposition 4 can be applied to Y_b.

to the biases for punishments. For example, if the state space were unrestricted in one dimension orthogonal to the plane of the biases and bias magnitudes had known finite upper bounds, then a robust FRE would always exist.

4 Non-Convex Policy Spaces

This section considers the case where the policy space is non-convex, for example because of increasing returns to some resources or indivisibilities. We identify an additional geometric condition, the *Global Deterrence Condition*, on the directions of the senders' biases relative to the frontier of the *convex hull* of the policy space, that together with the Local Deterrence Condition identified in Proposition 4, is necessary and sufficient for existence of a robust FRE that is independent of the magnitudes of the biases. The form of the Global Deterrence Condition differs from that of the Local Deterrence Condition *only* in that the frontier of the convex hull of the policy space Y replaces the frontier of Y itself.

When the policy space is non-convex, it is possible for the belief of the receiver to be such that two or more policies are optimal. To see a simple example of this, consider a policy space in which everything is feasible except policies that are within ϵ of zero, i.e. $Y = \mathbb{R}^2 \setminus B(0, \epsilon)$. If the belief of the receiver is such that the expected state is 0, then any policy such that $|y| = \epsilon$ is optimal. In such a case, the restriction to pure strategies from the receiver is not without loss of generality, and hence for this section we allow the receiver to use mixed strategies.

Before presenting our result, we need to generalize our definition of a *smooth* point on the frontier as well as that of an *inward* normal vector to the frontier at a smooth point.[27] Given an arbitrary set S, a point $s \in Fr(S)$ is a *smooth* point if *locally* there is a unique tangent hyperplane to $Fr(S)$ at s, or more precisely, if there exists an $\epsilon > 0$ such that, for any $0 < \delta < \epsilon$, $B(s, \delta) \cap Fr(S)$ has a unique tangent hyperplane at s. As before, we will denote the set of smooth points on the frontier of S by $\widetilde{Fr}(S)$. The *inward* normal vector to the frontier of S at a smooth point s, is then the normal vector to $Fr(S)$, $n_S(s)$, that satisfies the condition that there exists an $\epsilon > 0$ such that for any $0 < \delta < \epsilon$, $s + \delta n_S(s) \in S$.

Proposition 8. *Suppose $Y \subseteq \mathbb{R}^2$ is compact and Fr(Y) has finitely many kinks. Given $b_1, b_2 \in \mathbb{R}^2$ linearly independent, the following statements are generically equivalent:*

(i) There exists a robust fully revealing equilibrium that is independent of the magnitudes of the biases.

(ii)
1. *For every $\theta \in \widetilde{Fr}(co(Y))$, $n^{In}_{co(Y)}(\theta) \notin C(b_1, b_2)$ (Global Deterrence Condition).* *and*
2. *For every $\theta \in \widetilde{Fr}(Y)$, $n^{In}_Y(\theta) \notin C(b_1, b_2)$ (Local Deterrence Condition).*

In the proof of the proposition we show that the Global Deterrence Condition (GDC) and the Local Deterrence Condition (LDC) are together necessary and sufficient for the

[27] Recall that for a convex set S we defined an *inward* normal vector to a smooth point $s \in \widetilde{Fr}(S)$, as the only vector $n^{In}_S(s)$ such that $n^{In}_S(s)(s' - s) \geq 0$ for all $s' \in S$. This definition does not apply to non-convex sets.

existence of a robust FRE in *pure* strategies as long as there is not a very special sort of kink point in the frontier of the policy space. At such a special kink point, the frontier is locally non-convex and locally linear on the two sides of the kink, and furthermore the inward normal vectors to these linear pieces coincide with the bias vectors. If such a kink point exists, then the GDC and LDC are not sufficient for the existence of a robust FRE in *pure* strategies, but a robust FRE in *mixed* strategies does exist. Finally, we argue that if a robust FRE in pure strategies does not exist, then allowing the receiver to use mixed strategies will not *generically* help in supporting an equilibrium.

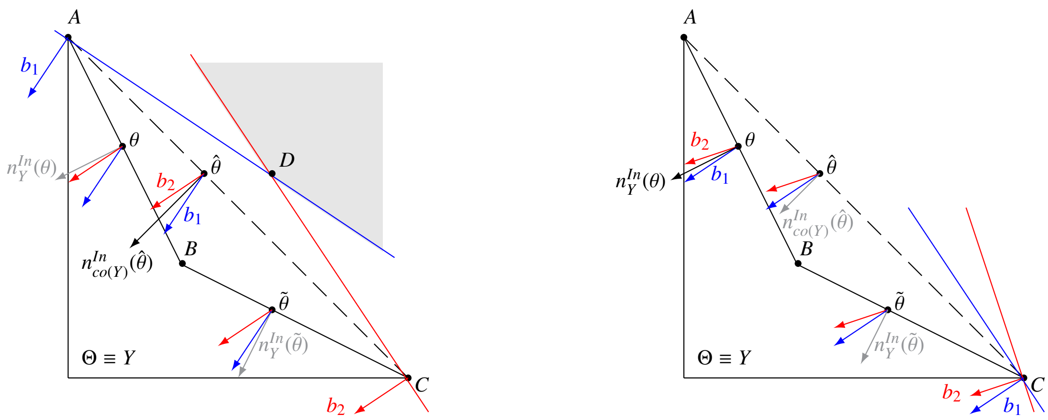
The LDC is necessary and sufficient for small deviations to be deterrable with small punishments, whether the biases have known finite magnitudes or whether they can be arbitrarily large.[28] As we have mentioned before, when only local punishments are considered, the senders' incentives to deviate from truthtelling depend only on the orientations, not the magnitudes, of their biases. When Y is convex, this condition is necessary and sufficient for existence of a robust FRE, as Proposition 4 shows. When Y is non-convex, however, deterrence of small deviations with small punishments is no longer sufficient for existence of a robust FRE. The GDC is necessary and sufficient for all deviations, including large ones, to be deterrable with feasible punishments, when the biases can be arbitrarily large. Since for non-convex Y, deterrability of large deviations depends in general on the magnitudes as well as the orientations of the biases, existence of a robust FRE that is independent of the magnitudes of the biases implies, but is not in general implied by, existence of a robust FRE for biases of known finite magnitudes. We illustrate these points with two examples, displayed in Figures 12a and 12b.

Figure 12a presents an example in which all local deviations can be deterred with local punishments, because the LDC is satisfied. However, at points along the dashed line connecting A to C, which is part of the frontier of $co(Y)$, the GDC is violated. To see the consequence of this violation, observe that if Sender 1 reports C and Sender 2 reports A, and the magnitudes of the biases are very large, then there is no feasible response for the receiver that would suffice to deter Sender 1, in state A, from deviating to a report of C, and that would also deter Sender 2, in state C, from deviating to a report of A—any response that would deter both of these deviations would have to lie in the punishment region represented by the shaded area. Hence, a FRE that is independent of the magnitudes of the biases does not exist.

Figure 12a shows that when non-convexities are present, deterrence of local deviations with local punishments no longer implies that all deviations are deterrable. In particular, in Proposition 3, condition (*iii*) no longer implies condition (*i*) if the assumption of convexity of Y is dropped.

Figure 12b displays an example in which there exists a FRE that is independent of the magnitudes of the biases, because the GDC is satisfied. For very large magnitudes of the biases, point C is the least-preferred point in $Y \equiv \Theta$ for both senders, so it can be used by

[28] Note that the set of smooth points $\widetilde{Fr}(S)$ might be empty. This is the case if the set S is finite for example. If $\widetilde{Fr}(S)$ is empty, robustness does not impose any restriction.



(a) Small deviations can be deterred with small punishments, but the large deviation (A,C) cannot be deterred.

(b) There exists a FRE that is independent of the magnitudes of the biases, but small deviations from θ cannot be deterred locally.

Figure 12: Conditions for Robustness of FRE for non-convex Y.

the receiver to punish any discrepancies in the senders' reports.[29] However, no robust FRE exists. To see why, note that along segment AB on the frontier of Y, the LDC is violated; as a consequence, it is not possible to deter local deviations along segment AB with local punishments. This example shows that in Proposition 3, condition (i) no longer implies condition (iii) if the assumption of convexity is dropped.

5 Uncertain and State-Dependent Biases

We can generalize the approach and results of the previous Section 3 to accommodate some uncertainty about the directions of the senders' biases. In particular, we can relax the assumptions that the directions of the senders' biases are (i) common knowledge and (ii) independent of the realization of the state.

Suppose that the players have a common prior joint distribution G over (θ, b_1, b_2). Each sender observes θ and his own bias vector, while the receiver does not observe any of these realizations. The definition of a fully revealing equilibrium remains unchanged.

Proposition 9. *Given $Y \subseteq \mathbb{R}^2$ convex, suppose that there exists a closed convex cone $C[\underline{b}, \overline{b}] = \{\alpha \underline{b} + \beta \overline{b} \mid \alpha, \beta \geq 0\}$, such that for all $\theta \in \Theta$, the supports of the conditional distributions of the bias directions b_1 and b_2 given θ are both contained in $C[\underline{b}, \overline{b}]$. Then conditions ($i$) and ($ii$) are equivalent and imply (iii):*

(i) For all $\theta \in \widetilde{Fr}(Y)$, $\quad n_Y^{In}(\theta) \notin C(\underline{b}, \overline{b})$.

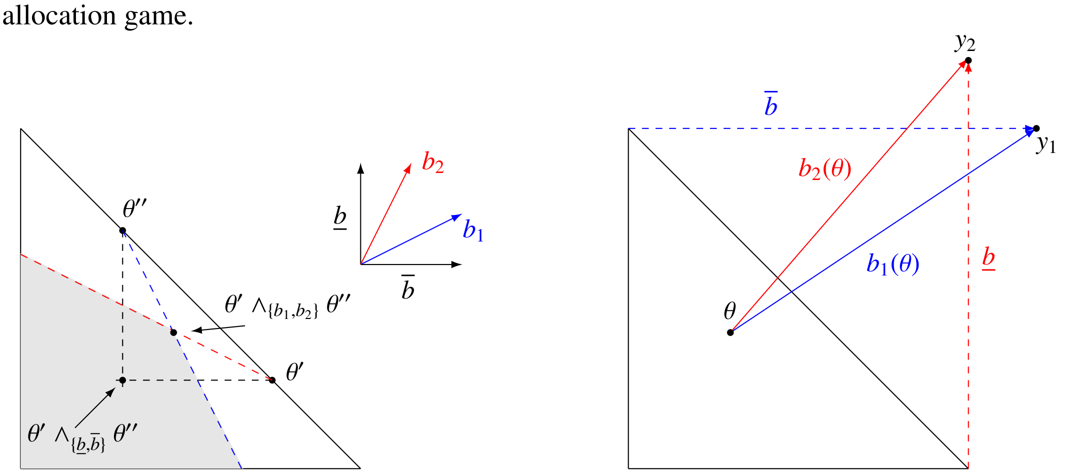
(ii) For all $\theta', \theta'' \in Y$, $\quad \theta' \wedge_{\{\underline{b}, \overline{b}\}} \theta'' \in Y$.

[29] To be more precise, $b_i C \leq b_i \theta$ for $i = 1, 2$ for any $\theta \in Y$. To see that C serves as a punishment for all deviations, even for senders with small biases, note that $b_i C \leq b_i \theta$ implies that $|\theta + b_i - C|^2 = |b_i|^2 + |\theta - C|^2 + 2b_i(\theta - C) > |b_i|^2 = |\theta + b_i - \theta|^2$. In other words, for both senders, C is farther from $\theta + b_i$ than θ is, and hence neither sender has an incentive to deviate from truthfully reporting θ.

(iii) There exists a (robust) fully revealing equilibrium that is independent of the magnitudes of the biases.

Moreover, if the conditional distribution of the bias directions (b_1, b_2) *given* θ *assigns positive density to* $(\underline{b}, \overline{b})$ *for all* $\theta \in Y$*, then* (*iii*) *implies* (*i*) *and* (*ii*).

The sufficiency part of Proposition 9 says that when the receiver does not know the actual biases but knows only that they are certain to lie in a given closed convex cone $C[\underline{b}, \overline{b}]$, then two equivalent sufficient conditions for the existence of a robust FRE that is independent of the magnitudes of the biases are the Local Deterrence Condition and the feasibility of the Min Rule, just as in Proposition 4, except that here the known biases b_1 and b_2 are replaced by the least aligned possible realizations, $\underline{b}$ and $\overline{b}$. Each of these conditions ensures that for all true states on the frontier of Y, the receiver can find local punishments that would deter local deviations, whether the realized values of (b_1, b_2) were $(\underline{b}, \overline{b})$ or $(\overline{b}, \underline{b})$. (Recall that the Min Rule is an anonymous rule.) This in turn implies that for any more closely aligned realizations of the biases, local deviations would continue to be deterred by these same local punishments, as illustrated in Figure 13a for the funding allocation game.



(a) Funding allocation game when the senders' uncertain biases have support within $C[\underline{b}, \overline{b}]$.

(b) Funding allocation game when the senders have state-independent preferences: y_1, y_2 represent their ideal points.

Figure 13: Funding allocation game when the biases are uncertain and/or state-dependent.

The necessity part of Proposition 9 shows that if there exists a *minimal* closed convex cone containing the biases and this minimal cone is the same for all states θ, then these sufficient conditions are also necessary. The necessity part of the proposition holds, for example, if the distribution of the biases, on the support $C[\underline{b}, \overline{b}]$, is independent of the realization of the state.

The sufficiency result of Proposition 9 can also be applied to cases in which the biases are determinitic but depend on the realization of the state. Consider, for example, the same funding allocation setting as in Figure 13a, and now suppose senders S_1 and S_2 have *fixed* ideal points, y_1 and y_2, independent of the realization of the state. Such state-independent quadratic preferences can expressed as follows:

$$u^{S_i}(y,\theta) = -(y-y_i)^2 = -(y-\theta-\underbrace{(y_i-\theta)}_{b_1(\theta)})^2$$

Therefore these preferences correspond to each sender S_i having a deterministic state-dependent bias $b_i(\theta) = (y_i - \theta)$. (See Figure 13b.) If these biases lie within a closed convex cone $C[\underline{b}, \overline{b}]$, and no inward normal vector to a smooth point on the frontier lies within the cone, then it follows from Proposition 9 that there exists a robust FRE. Furthermore, it can be implemented by the receiver using the Min Rule defined with respect to the biases $\{\underline{b}, \overline{b}\}$.

Finally, to conclude our analysis, note that the sufficiency result of Proposition 9 can be easily extended to general convex preferences, using a similar argument to the case of state-dependent biases. Suppose for simplicity that $\Theta \equiv Y$, with $Y \subseteq \mathbb{R}^2$ and convex. Denote by $u^R(y,\theta)$ the quasi-concave utility function representing the receiver's preferences, such that $\theta = \arg\max_{y\in Y} u^R(y,\theta)$, that is, the ideal policy of the receiver in state θ is $y = \theta$. Denote by $u^{S_i}(y,\theta)$ the quasi-concave utility function of sender S_i. For every realization of the state $\theta \in Y$, define the bias vector of S_i as the normal vector to S_i's indifference curve through the policy $y = \theta$, in state θ.[30] Now suppose that there exists a closed convex cone $C[\underline{b}, \overline{b}]$ such that for both senders, $b_i(\theta) \in C[\underline{b}, \overline{b}]$ for all $\theta \in Y$. Then if the Local Deterrence Condition is satisfied for the pair of biases $\{\underline{b}, \overline{b}\}$, there exists a robust FRE that can be implemented using the Min Rule defined with respect to $\{\underline{b}, \overline{b}\}$. Figure 14 illustrates the intuition.

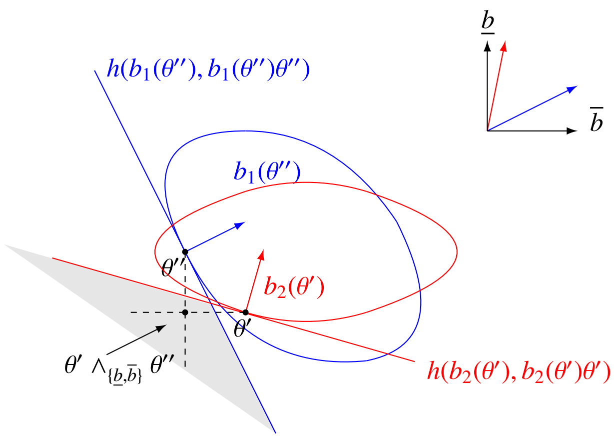


Figure 14: Quasi-concave utilities and $Y \subseteq \mathbb{R}^2$, convex. In state θ'', S_1 would prefer a policy in the direction of $b_1(\theta'')$, and hence would be punished for deviating by the receiver choosing a policy to the south-west of $h(b_1(\theta''), b_1(\theta'')\theta'')$. Analogously, in state θ', S_2 would be punished for deviating by the receiver choosing a policy to the south of $h(b_2(\theta'), b_2(\theta')\theta')$. The grey area represents the punishment region for the deviation (θ', θ''). The Min Rule, defined for $\{\underline{b}, \overline{b}\}$, delivers a policy that lies in the punishment region.

[30] More formally, S_i's bias in state θ, $b_i(\theta)$, is defined as the gradient vector, with respect to y, to the indifference curve of S_i at the policy $y = \theta$.

A Appendix

Proof of Lemma 1:

Consider a robust fully revealing equilibrium (s_1, s_2, y^{μ}) supported by the belief function $\mu(\cdot)$ and consider the following strategies: $\tilde{s}_i : \Theta \longrightarrow Y$, such that $\tilde{s}_i(\theta) = y^*(\theta)$; $\tilde{y} : Y \times Y \longrightarrow Y$, such that $\tilde{y}(y', y'') = y^{\mu}(s_1(y'), s_2(y''))$ and the belief function $\tilde{\mu}(y', y'') = \mu(s_1(y'), s_2(y''))$. We show that the restriction of $(\tilde{s}_1, \tilde{s}_2, \tilde{y})$ to the environment (Y, Y) is a robust truthful equilibrium. An argument identical to the one in Proposition 2 shows that in fact $(\tilde{s}_1, \tilde{s}_2, \tilde{y})$ is also a robust truthful equilibrium in the environment (Θ, Y). Consider the report (θ', θ'') with $\theta' \neq \theta''$, $\theta', \theta'' \in Y$. Then:

$$|\tilde{y}(\theta', \theta'') - (\theta'' + b_1)| = |y^{\mu}(s_1(\theta'), s_2(\theta'')) - (\theta'' + b_1)| \leq |b_1|$$
$$|\tilde{y}(\theta', \theta'') - (\theta' + b_2)| = |y^{\mu}(s_1(\theta'), s_2(\theta'')) - (\theta' + b_2)| \leq |b_2|$$

where the inequality follows from the fact that (s_1, s_2, y^{μ}) is an equilibrium. Finally, if $\theta \in Y$ and $\epsilon > 0$, by the robustness of (s_1, s_2, y^{μ}) there exists a $\delta > 0$ such that if $\theta', \theta'' \in B(\theta, \delta)$,

$$\tilde{y}(\theta', \theta'') = y^{\mu}(s_1(\theta'), s_2(\theta'')) \in B(\theta, \epsilon).$$

hence $(\tilde{s}_1, \tilde{s}_2, \tilde{y})$ is a robust truthful equilibrium.

Proof of Proposition 1:

(i): $\Rightarrow$) Suppose there exist $\theta', \theta'' \in Y$ such that $Y \subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. Then $y(s_1(\theta'), s_2(\theta'')) \in H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. In particular, denoting $y \equiv y(s_1(\theta'), s_2(\theta''))$, either $b_1(y - \theta'') > 0$ or $b_2(y - \theta') > 0$. Suppose that $b_1(y - \theta'') > 0$ and consider $t_1 > \frac{|y-\theta''|^2}{2b_1 \cdot (y-\theta'')}$. Then $y(s_1(\theta'), s_2(\theta'')) \in B(\theta'' + t_1 b_1, t_1|b_1|)$ which implies that for the sender 1with bias $t_1 b_1$ has an incentive to deviate to $s_1(\theta')$ given θ''. The symmetric argument could be made if $b_2(y - \theta') > 0$ with $t_2 > \frac{|y-\theta'|^2}{2b_2 \cdot (y-\theta')}$.

$\Leftarrow$) Consider truthful strategies and the following belief function $\mu(\cdot)$ such that $\mu(\theta, \theta)$ allocates mass one on θ and $\mu(\theta', \theta'')$ with $\theta' \neq \theta'' \in Y$, puts mass one in an element of $Y \setminus H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. Given a report (θ', θ''), $\mu(\theta', \theta'') \notin H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ so in particular $\mu(\theta', \theta'') \notin B(\theta'' + t_1 b_1, t_1|b_1|)$ and $\mu(\theta', \theta'') \notin B(\theta' + t_2 b_2, t_2|b_2|)$. So none of the two senders has an incentive to deviate.

(ii): $\Rightarrow$) Suppose there exist some fully revealing strategies (s_1, s_2) and a belief function $\mu(\cdot)$ that deters local deviation with local punishments, then for any $\theta \in Y$ and any $\epsilon > 0$ there exists $\delta > 0$ such that for every $\theta', \theta'' \in B(\theta, \delta)$, $\bar{\mu}(s_1(\theta'), s_2(\theta'')) \in B(\theta, \epsilon) \cap Y \setminus (B(\theta' + t_1 b_1, t_1|b_1|) \cup B(\theta + t_2 b_2, t_2|b_2|))$ for any $t_1, t_2 \geq 0$. Hence $B(\theta, \epsilon) \cap Y \nsubseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$.

$\Leftarrow$) By the argument used in the proof of Lemma 1 we can focus on truthful strategies. For any $\theta \in \Theta$ define $\mu(\theta, \theta)$ a belief that allocates mass one to θ. If $\theta \neq \theta' \in \Theta$ define $\mu(\theta, \theta')$ a belief that allocates mass one to an element of $\arg\min_{s \in Y \setminus (H(b_1, b_1\theta') \cup H(b_2, b_2\theta))} |s - \theta|$, if

$Y \not\subseteq H(b_1, b_1\theta') \cup H(b_2, b_2\theta)$, and any arbitrary belief if $Y \subseteq H(b_1, b_1\theta') \cup H(b_2, b_2\theta)$. To see that this belief function deters local deviation with local punishments consider any $\theta \in \Theta$ and any $\epsilon > 0$, by hypothesis, for $\tilde{\epsilon} = \epsilon/3$ there exists $0 < \delta < \tilde{\epsilon}$ such that for all $\theta', \theta'' \in B(\theta, \delta) \cap Y$, $B(\theta, \tilde{\epsilon}) \cap Y \not\subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. Consider any $\hat{\theta} \in B(\theta, \tilde{\epsilon}) \cap Y \not\subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. And $|\overline{\mu}(\theta', \theta'') - \theta| \leq |\mu(\theta', \theta'') - \theta'| + |\theta' - \theta| \leq |\hat{\theta} - \theta'| + |\theta' - \theta| \leq |\hat{\theta} - \theta| + 2|\theta' - \theta| < 3\tilde{\epsilon} = \epsilon$, hence $\overline{\mu}(\theta', \theta'') \in B(\theta, \epsilon) \setminus (H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')) \subset B(\theta, \epsilon) \setminus (B(\theta'' + t_1 b_1, t_1|b_1|) \cup B(\theta' + t_2 b_2, t_2|b_2|)$.

(iii): The necessity is given by parts *(i)* and *(ii)*. To see the sufficiency, consider truthful strategies and the belief specified in the previous paragraph. Note that given condition (2), $Y \setminus (H(b_1, b_1\theta') \cup H(b_2, b_2\theta)) \neq \emptyset$ for any $\theta \neq \theta' \in Y$. □

Proof of Proposition 2:

$(i) \Rightarrow (ii)$: Note that $Y \not\subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ if and only if $Y \not\subseteq H(b_1^Y, b_1^Y\theta'') \cup H(b_2^Y, b_2^Y\theta')$. Hence, (i) is equivalent to the existence, for the environment (Θ, Y), of a (robust) FRE for all biases $(t_1 b_1^Y, t_2 b_2^Y)$, $t_1, t_2 \geq 0$. Statement (ii) follows trivially from this.

$(ii) \Rightarrow (i)$: Suppose there exists a (robust) FRE in (Y,Y) for all biases $(t_1 b_1^Y, t_2 b_2^Y)$, $t_1, t_2 \geq 0$. By Lemma 1, there exists a truthful (robust) equilibrium outcome-equivalent to it. Denote the truthful equilibrium by (s_1, s_2, y^μ) where for all $\theta \in Y$, $s_i(\theta) = y^*(\theta) = \theta$. For $\theta \in \Theta$ we define the following strategies: $\tilde{s}_i(\theta) = y^*(\theta)$. We claim that $(\tilde{s}_1, \tilde{s}_2, y^\mu)$ is a (robust) FRE in (Θ, Y).

Consider the out-of-equilibrium messages (y', y'') where $y' \neq y''$ and denote by $x = y^\mu(y', y'')$ the receiver's policy after the report (y', y''). By Proposition 1,

$$b_1^Y(y'' - x) \geq 0, \qquad b_2^Y(y' - x) \geq 0. \tag{3}$$

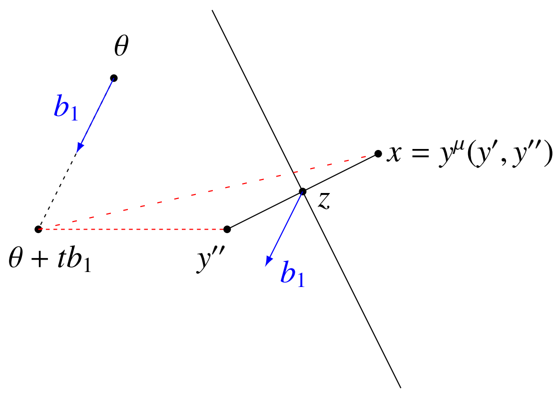


Figure 15:

For sender S_1, we need to show that for any $\theta \in \Theta$ such that $y^*(\theta) = y''$, $|\theta + tb_1 - y''| \leq |\theta + tb_1 - x|$ for all $t > 0$. For any $\theta \in \Theta$ with $y^*(\theta) = y''$, y'' is the closest point in Y to θ and therefore $|\theta - y''| \leq |\theta - x|$. In particular, if z is the midpoint of the segment $[x, y'']$, $\theta \in H(y'' - x, z(y'' - x))$ or analogously, $\theta(y'' - x) \geq z(y'' - x)$. Note that since

$y'', x \in Y$, $y'' - x \in S^Y$ and hence $b_1(y'' - x) = b_1^Y(y'' - x) \geq 0$ where the last inequality follows by (3). Therefore $(\theta + tb_1)(y'' - x) \geq z(y'' - x)$ for all $t > 0$, or in other words $|\theta + tb_1 - y''| \leq |\theta + tb_1 - x|$ for all $t > 0$. A similar argument for S_2 shows that for any $\theta \in \Theta$ such that $y^*(\theta) = y'$, $|\theta + tb_2 - y'| \leq |\theta + tb_2 - x|$ for all $t > 0$. Therefore $(\tilde{s}_1, \tilde{s}_2, y^\mu)$ is a FRE in (Θ, Y). (See Figure 15.)

Finally if we further assume that the initial equilibrium (s_1, s_2, y^μ) is robust, then by Proposition 1, for any $y \in Y$ and $\epsilon > 0$, there exists $\delta > 0$ such that for all $y', y'' \in B(y, \delta) \cap Y$, $y^\mu(y', y'') \in B(y, \epsilon)$. In particular, for any $\theta \in \Theta$ and $\theta', \theta'' \in \Theta$ such that $y^*(\theta'), y^*(\theta'') \in B(y^*(\theta), \delta)$, we have that $y^\mu(y^*(\theta'), y^*(\theta'')) \in B(y^*(\theta), \epsilon)$. □

Proof of Proposition 3:

By Proposition 2, we can restrict attention to the case in which $\Theta \equiv Y$.

$(ii) \Rightarrow (i)$ is trivial.

$(i) \Rightarrow (ii)$ We argue in two steps. First, we prove that if a local deviation from $\theta \in Y$ cannot be deterred with a local punishment, then there exists $\epsilon > 0$ such that

$$B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$$

where $\overline{S}$ denotes the closure of S. Note that this statement is independent of whether Y is convex or not. Second, we use the first result and the convexity of Y to show that if a local deviation cannot be deterred with a local punishment, it cannot be deterred with any punishment and hence a fully revealing equilibrium does not exist.

STEP 1: If local deviations from $\theta \in Y$ cannot be deterred with local actions then by Proposition 1 there exists $\epsilon > 0$ such that for every $\delta > 0$ there exist $\theta'_\delta, \theta''_\delta \in B(\theta, \delta) \cap Y$ such that $B(\theta, \epsilon) \cap Y \subseteq H(b_1, b_1\theta''_\delta) \cup H(b_2, b_2\theta'_\delta)$. We show that for that same ϵ, $B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$. Suppose that $B(\theta, \epsilon) \cap Y \not\subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$. Then there exists $\tilde{\theta} \in B(\theta, \epsilon) \cap Y$, such that $b_1\tilde{\theta} < b_1\theta$ and $b_2\tilde{\theta} < b_2\theta$. Define $\tilde{\delta} = min\{\frac{|b_1(\theta - \tilde{\theta})|}{|b_1|}, \frac{|b_2(\theta - \tilde{\theta})|}{|b_2|}\}$ and denote $\tilde{\theta}', \tilde{\theta}'' \in B(\theta, \tilde{\delta})$ the corresponding $\theta'_{\tilde{\delta}}$ and $\theta''_{\tilde{\delta}}$ such that $B(\theta, \epsilon) \cap Y \subseteq H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}')$. But by the definition of $\tilde{\delta}$, $b_1\tilde{\theta} < b_1\tilde{\theta}''$ and $b_2\tilde{\theta} < b_2\tilde{\theta}'$ and hence $\tilde{\theta} \in B(\theta, \epsilon) \cap Y \setminus \left(H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}')\right)$ which is a contradiction. See Figure 16.

STEP 2: Suppose that local deviations from $\theta \in Y$ cannot be deterred with a local punishment. By Step 1 there exists $\epsilon > 0$ such that $B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$. Define $\theta' \in \arg\min\{b_2\tilde{\theta} \mid \tilde{\theta} \in \overline{B}(\theta, \epsilon) \cap Y\}$ and $\theta'' \in \arg\min\{b_1\tilde{\theta} \mid \tilde{\theta} \in \overline{B}(\theta, \epsilon) \cap Y\}$. Clearly $B(\theta, \epsilon) \cap Y \subset H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ and hence either $b_1\theta > b_1\theta''$ or $b_2\theta > b_2\theta'$. Without loss of generality assume that $b_1\theta > b_1\theta''$. We show now that the deviation $\{\theta', \theta''\}$ cannot be deterred in Y. (See Figure 17.)

Suppose there exists $\hat{\theta} \in Y$ such that $b_1\hat{\theta} \leq b_1\theta'' < b_1\theta$ and $b_2\hat{\theta} \leq b_2\theta' \leq b_2\theta$. Below we prove that $b_2\hat{\theta} < b_2\theta$. As a consequence, by the convexity of Y there exists a $\lambda \in (0, 1)$ such that $\lambda\hat{\theta} + (1 - \lambda)\theta \in B(\theta, \epsilon) \cap Y$ and $b_1(\lambda\hat{\theta} + (1 - \lambda)\theta) < b_1\theta$, $b_2(\lambda\hat{\theta} + (1 - \lambda)\theta) < b_2\theta$ which

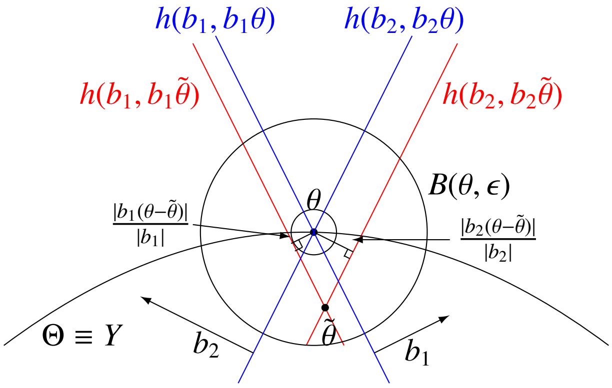


Figure 16:

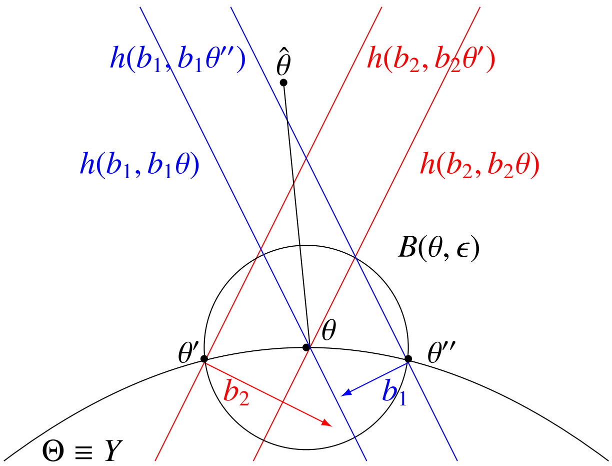


Figure 17: Every fully revealing equilibrium is robust: if a local deviation cannot be deterred with a local punishment, it cannot be deterred with any punishment.

contradicts that $B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$.

To see that $b_2\hat{\theta} < b_2\theta$, suppose $b_2\hat{\theta} = b_2\theta = b_2\theta' = \min\{b_2\tilde{\theta} \mid \tilde{\theta} \in B(\theta, \epsilon) \cap Y\}$ (See Figure 18). Since local deviations from θ cannot be deterred, by Proposition 1 there exists an $\tilde{\epsilon} > 0$ such that for any $\delta > 0$ there exists $\theta'_\delta, \theta''_\delta \in B(\theta, \delta) \cap Y$ such that $B(\theta, \tilde{\epsilon}) \cap Y \subset H(b_1, b_1\theta''_\delta) \cup H(b_2, b_2\theta'_\delta)$. Consider $\delta < \min\{\epsilon, \tilde{\epsilon}\frac{|b_1 n_2|}{|b_1|}\}$, where n_2 is the unit normal vector to b_2 such that $b_1 n_2 < 0$. In particular, since $\delta < \tilde{\epsilon}\frac{|b_1 n_2|}{|b_1|}$, $b_1\theta''_\delta > b_1(\theta + \tilde{\epsilon} n_2)$ and hence there exists $\mu \in (0, 1)$ such that $\mu\hat{\theta} + (1-\mu)\theta \in B(\theta, \tilde{\epsilon}) \cap Y$ and $b_1\theta''_\delta > b_1(\mu\hat{\theta} + (1-\mu)\theta)$. Moreover, since $\delta < \epsilon$ and $b_2\hat{\theta} = b_2\theta = b_2\theta' = \min\{b_2\tilde{\theta} \mid \tilde{\theta} \in B(\theta, \epsilon) \cap Y\}$, $b_2(\mu\hat{\theta} + (1-\mu)\theta) \leq b_2\theta'_\delta$. But this contradicts that $B(\theta, \tilde{\epsilon}) \cap Y \subset H(b_1, b_1\theta''_\delta) \cup H(b_2, b_2\theta'_\delta)$.

$(ii) \Rightarrow (iii)$ is trivial.

$(iii) \Rightarrow (i)$ Suppose there exists $\theta', \theta'' \in Y$ such that $Y \subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. Then $h(b_1, b_1\theta'') \cap h(b_2, b_2\theta') \cap Y = \emptyset$ and for Y compact or $Y \subseteq \mathbb{R}^2$, there exist[31]

[31] If Y is compact then the minimum is reached within the set. This is also the case if $Y \subseteq \mathbb{R}^2$ because $h(b_1, b_1\theta'') \cap h(b_2, b_2\theta') \cap Y = \emptyset$ implies that the sets $Y \cap h(b_2, b_2\theta')$ and $Y \cap h(b_1, b_1\theta'')$ are closed, bounded (from below) half-lines and hence they have a minimum. For general $Y \subseteq \mathbb{R}^q$, even if $Y \cap h(b_i, b_i\theta)$ is closed and bounded from below, it might be the case that the minimum is never reached.

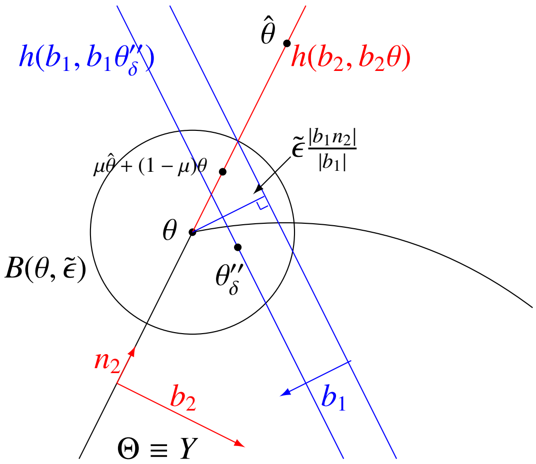


Figure 18:

$$\tilde{\theta}' \in \arg\min_y \{b_1 y \mid y \in Y, b_2 y = b_2\theta'\} \text{ and}$$
$$\tilde{\theta}'' \in \arg\min_y \{b_2 y \mid y \in Y, b_1 y = b_1\theta''\}.$$

Note that $H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}') = H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ and hence $Y \subseteq H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}')$. We show that for any $\lambda \in (0,1)$, $Y \subseteq H(b_1, b_1(\lambda\tilde{\theta}' + (1-\lambda)\tilde{\theta}'')) \cup H(b_2, b_2\tilde{\theta}')$ and therefore given $\tilde{\theta}'$, for every $\delta > 0$ there exists a $\lambda \in (0,1)$ such that $\lambda\tilde{\theta}' + (1-\lambda)\tilde{\theta}'' \in B(\tilde{\theta}', \delta) \cap Y$ and

$$Y \subseteq H(b_1, b_1(\lambda\tilde{\theta}' + (1-\lambda)\tilde{\theta}'')) \cup H(b_2, b_2\tilde{\theta}')$$

and hence a local deviation from $\tilde{\theta}'$ cannot be deterred[32]. See Figure 19.

Suppose that there exists $\lambda \in (0,1)$ and $\hat{\theta} \in Y$ such that $\hat{\theta} \notin H(b_1, b_1(\lambda\tilde{\theta}' + (1-\lambda)\tilde{\theta}'')) \cup H(b_2, b_2\tilde{\theta}')$. Since $\hat{\theta} \in H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}')$,

$$\begin{aligned} &\hat{\theta} \in H(b_1, b_1\tilde{\theta}'') \\ &\hat{\theta} \notin H(b_2, b_2\tilde{\theta}') \\ &\hat{\theta} \notin H(b_1, b_1(\lambda\tilde{\theta}' + (1-\lambda)\tilde{\theta}'')) \end{aligned}$$

Moreover since $\tilde{\theta}', \tilde{\theta}'' \in Y \subseteq H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}')$, $b_2\tilde{\theta}' < b_2\tilde{\theta}''$ and $b_1\tilde{\theta}'' < b_1\tilde{\theta}'$. Putting all the inequalities together we find that

$$b_2\hat{\theta} \le b_2\tilde{\theta}' < b_2\tilde{\theta}'' \tag{4}$$
$$b_1\tilde{\theta}'' < b_1\hat{\theta} < b_1\tilde{\theta}' \tag{5}$$

By (4) there exists $\mu \in (0,1]$ such that $b_2(\mu\hat{\theta} + (1-\mu)\tilde{\theta}'') = b_2\tilde{\theta}'$ and by convexity $\mu\hat{\theta} + (1-\mu)\tilde{\theta}'' \in Y$. But by (5), $b_1(\mu\hat{\theta} + (1-\mu)\tilde{\theta}'') < b_1\hat{\theta} < b_1\tilde{\theta}'$ which contradicts the definition of $\tilde{\theta}'$. $\square$

[32]Note that for any $\epsilon > 0$, $B(\tilde{\theta}', \epsilon) \cap Y \subseteq Y \subseteq H(b_1, b_1(\lambda\tilde{\theta}' + (1-\lambda)\tilde{\theta}'')) \cup H(b_2, b_2\tilde{\theta}')$, and hence local deviations from $\tilde{\theta}'$ cannot be deterred with local punishments.

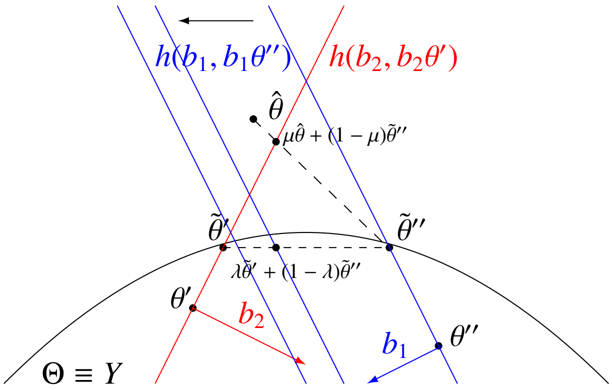


Figure 19: In $\mathbb{R}^2$, if a deviation cannot be deterred, there is a local deviation that cannot be deterred with local actions.

Proof of Proposition 4:

By Proposition 2, we can restrict attention to the case $\Theta \equiv Y$.

$(i) \Rightarrow (ii)$: Suppose there exists $\theta \in \widetilde{Fr}(\theta)$ such that $n_Y^{In}(\theta) \in C(b_1, b_2)$. Then $Y \subset \bar{H}(b_1, b_1\theta) \cup \bar{H}(b_2, b_2\theta)$. Moreover, since $n_Y^{In}(\theta) \neq b_1$ and $n_Y^{In}(\theta) \neq b_2$, for any $\delta > 0$ there exists $\theta' \in Y \cap B(\theta, \delta)$ and $\theta'' \in Y \cap B(\theta, \delta)$ such that $b_2\theta' < b_2\theta$ and $b_1\theta'' < b_1\theta$. But then $Y \subset \bar{H}(b_1, b_1\theta) \cup \bar{H}(b_2, b_2\theta) \subset H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ and local deviations from θ cannot be deterred, which contradicts (i).
$(ii) \Rightarrow (iii)$: Suppose that there exist $\theta', \theta'' \in Y$ such that $x \equiv \theta' \wedge_{\{b_1,b_2\}} \theta'' \notin Y$. In particular $\theta' \wedge_{\{b_1,b_2\}} \theta'' \notin \{\theta', \theta''\}$. Without loss of generality assume that $b_1 x = b_1\theta''$ and $b_2 x = b_2\theta'$. Consider any $\tilde{\theta} \in \widetilde{Fr}(Y)$ that lies in the interior of the triangle formed by θ', θ'' and x.[33] See Figure 20. In particular, since Y is convex, $h(n_Y(\tilde{\theta}), n_Y(\tilde{\theta})\tilde{\theta})$ is a hyperplane separating Y from x, and

$$n_Y(\tilde{\theta})(\theta' - \tilde{\theta}) \geq 0 \tag{6}$$
$$n_Y(\tilde{\theta})(\theta'' - \tilde{\theta}) \geq 0 \tag{7}$$
$$n_Y(\tilde{\theta})(x - \tilde{\theta}) < 0. \tag{8}$$

Moreover, since b_1, b_2 span $\mathbb{R}^2$ there exists $\alpha, \beta \in \mathbb{R}$ such that $n_Y(\tilde{\theta}) = \alpha b_1 + \beta b_2$. Substituting this into equations (6), (7), (8), and then substracting (8) from (6) and (7), we obtain

$$0 < \alpha b_1(\theta' - x) - \beta b_2(\theta' - x) = \alpha b_1(\theta' - \theta'') \tag{9}$$
$$0 < \alpha b_1(\theta'' - x) - \beta b_2(\theta'' - x) = \beta b_2(\theta'' - \theta'), \tag{10}$$

where the equalities follow by the definition of x. And given that $b_1\theta' > b_1\theta''$ and $b_2\theta' < b_2\theta''$, (9) and (10) imply $\alpha > 0$ and $\beta > 0$, respectively. Hence $n_Y(\tilde{\theta}) \in C(b_1, b_2)$, which

[33]Note that $Fr(Y)$ has at most a countable number of kinks. Since Y is convex, $Fr(Y)$ is locally the graph of a concave (convex) function and hence the derivative of this function is monotonic, and it has at most a countable number of jumps.

contradicts (ii).

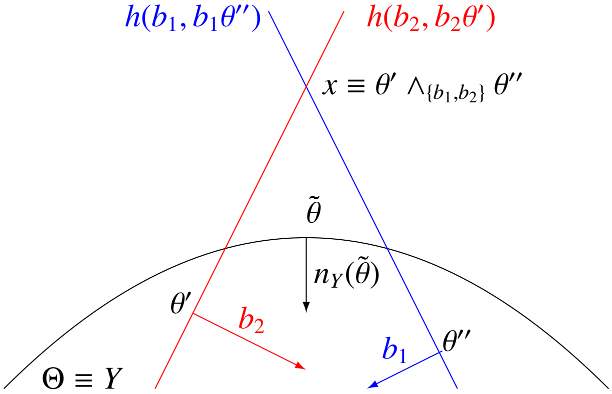


Figure 20: Relationship between the Local Deterrence Condition and the Min Rule

$(iii) \Rightarrow (i)$: Suppose that for all $\theta', \theta'' \in Y$, $y \equiv \theta' \wedge_{\{b_1,b_2\}} \theta'' \in Y$. By the definition of $\theta' \wedge_{\{b_1,b_2\}} \theta''$, $b_1 y \leq b_1\theta''$ and $b_2 y \leq b_2\theta'$. Therefore $y \notin H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ and it is a punishment for the deviation (θ', θ'') for arbitrarily large biases. Moreover as θ', θ'' converge to a point θ, $y = \theta' \wedge_{\{b_1,b_2\}} \theta''$ also converges to θ and the equilibrium is robust. □

Proof of Proposition 5:

We show that (i) implies the local deterrence condition (ii) of Proposition 4. Suppose there exists $\theta \in \widetilde{Fr}(Y)$ such that $n_Y(\theta) \in C(b_1, b_2)$. Since Y is convex $Y \subseteq \overline{H}(n_Y(\theta), n_Y(\theta)\theta)$. We can find $\epsilon > 0$ such that

$$B(\theta, \epsilon) \cap \overline{H}(n_Y(\theta), n_Y(\theta)\theta) \subset \overline{B}(\theta + b_1/2, |b_1|/2) \cup \overline{B}(\theta + b_2/2, |b_2|/2) \tag{11}$$

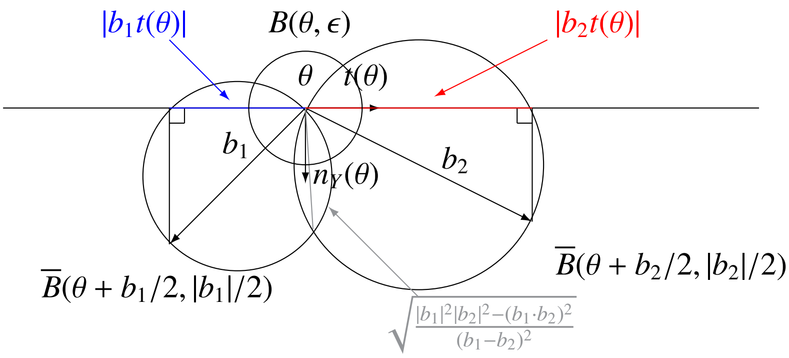


Figure 21: Robustness implies the Local Deterrence Condition.

More precisely, if we denote by $t(\theta)$ a unit normal vector to $n_Y(\theta)$, any $0 < \epsilon \leq$

$\min\{|b_1 t(\theta)|, |b_2 t(\theta)|, \sqrt{\frac{|b_1|^2|b_2|^2-(b_1\cdot b_2)^2}{(b_1-b_2)^2}}\}$ will satisfy (11).[34] See Figure 21.
Moreover, for any $\delta > 0$,

$$B(\theta - b_1/2, |b_1|/2) \cap Y \cap B(\theta, \delta) \neq \emptyset$$
$$B(\theta - b_2/2, |b_2|/2) \cap Y \cap B(\theta, \delta) \neq \emptyset$$

Consider $\tilde{\epsilon} = \min\{\epsilon, |b_1|/2, |b_2|/2\}$. Then for any $\delta > 0$ consider θ' an arbitrary element of $B(\theta - b_2/2, |b_2|/2) \cap Y \cap B(\theta, \delta)$ and θ'' an arbitrary element of $B(\theta - b_1/2, |b_1|/2) \cap Y \cap B(\theta, \delta)$. In what follows, we show that $B(\theta, \tilde{\epsilon}) \cap Y \subset B(\theta'' + b_1, |b_1|) \cup B(\theta' + b_2, |b_2|)$ and hence local deviations from θ cannot be deterred locally.
Consider $\tilde{\theta} \in B(\theta, \tilde{\epsilon}) \cap Y$, then since $\tilde{\epsilon} \leq \epsilon$, $\tilde{\theta} \in \overline{B}(\theta + b_1/2, |b_1|/2) \cup \overline{B}(\theta + b_2/2, |b_2|/2)$. Suppose $\tilde{\theta} \in \overline{B}(\theta + b_1/2, |b_1|/2)$, then

$$|\tilde{\theta} - (\theta'' + b_1)| \leq |\tilde{\theta} - (\theta + \frac{b_1}{2})| + |\theta - \frac{b_1}{2} - \theta''| < \frac{|b_1|}{2} + \frac{|b_1|}{2} = |b_1|$$

which implies that $\tilde{\theta} \in B(\theta'' + b_1, |b_1|)$. The case $\tilde{\theta} \in \overline{B}(\theta + b_2/2, |b_2|/2)$ is analogous. $\square$

Proof of Lemma 2:

Since the FRE supported by the Min Rule is not collusion-proof, there exists a state θ and a collusion agreement $\hat{\theta}$ (weakly preferred by both senders to θ), such that neither sender unilaterally gain from deviating. In other words, for any possible deviation $\tilde{\theta}$ from the agreement, both senders weakly prefer $\hat{\theta}$ to $\hat{\theta} \wedge \tilde{\theta}$ in state θ.
Consider a FRE that is independent of the magnitudes of the biases. We show that for the same state θ, the same collusion agreement $\hat{\theta}$ is self-enforcing. In order to see this suppose, without loss of generality, that sender S_1 could deviate from the agreement and induce a policy y that he preferred to $\hat{\theta}$, that is, a policy $y \in \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|) \cap Y$. Since the FRE is independent of the magnitudes of the biases, and $\hat{\theta}$ is weakly preferred to θ by S_1,

$$b_1\theta \leq b_1 y \leq b_1\hat{\theta}. \tag{12}$$

Consider the policy $y' = y \wedge_{\{b_1,b_2\}} \hat{\theta}$. Note that y' is a feasible policy because $y, \hat{\theta} \in Y$, hence the existence of a FRE that is independent of the magnitudes of the biases implies by Proposition 4 that $y \wedge_{\{b_1,b_2\}} \hat{\theta} \in Y$. We consider the following two cases:
Case 1: If $y' \in \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$ then S_1 could have deviated in the FRE supported by the Min Rule to policy y (or y'), and gained by doing so, which contradicts the hypothesis.
Case 2: If $y' \notin \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$ the rest of the proof will show case S_2 can gain by deviating to y (or y') which he can induce through the Min Rule and hence contradicting the hypothesis again.
In order to see this, if $y' \notin \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$ it must be that $b_2\hat{\theta} < b_2 y$ (otherwise $y \wedge_{\{b_1,b_2\}} \hat{\theta} = y \in \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$). In particular, $y' \in h(b_2, b_2\hat{\theta})$. Moreover, given that

[34] The last number in this minimum corresponds to the length of the common chord of the two balls. It is derived using standard trigonometry.

$\hat{\theta}$ is also weakly preferred to θ by S_2, it has to be that

$$b_2\theta < b_2\hat{\theta}. \tag{13}$$

Denote by x the projection of θ into $h(b_2, b_2\hat{\theta})$. Our goal is to prove that y' lies in between $\hat{\theta}$ and x in $h(b_2, b_2\hat{\theta})$, and since x is the point that S_2 out of all the point in $h(b_2, b_2\hat{\theta})$, we would have that S_2 prefers y' to $\hat{\theta}$.

Denoting by $\underline{\theta}$ the point in $\bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$ with lowest inner product with b_1, i.e. $\underline{\theta} = \theta + (1 - |\theta + b_1 - \hat{\theta}|)b_1$. We show that $b_2\hat{\theta} < b_2\underline{\theta}$: If that wasn't the case, i.e., if $b_2\underline{\theta} < b_2\hat{\theta}$, the point $\breve{\theta} \equiv \arg\min\{b_1\tilde{y} \mid \tilde{y} \in \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|), b_2\tilde{y} \geq b_2\hat{\theta}\}$ satisfies that $b_2\breve{\theta} = b_2\hat{\theta}$. In other words, considering the policy space $\tilde{Y} \equiv \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|) \cap \bar{H}(b_2, b_2\hat{\theta})$, there exists a FRE that is independent of the magnitudes of the biases in which any incompatible reports are punished by $\breve{\theta}$. By Proposition 4, the Min Rule in such a policy space is always feasible. In particular, $y' = y \wedge_{\{b_1, b_2\}} \hat{\theta} \in \tilde{Y} \subseteq \bar{B}(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$. But this contradicts our hypothesis in Case 2. Summing up, we have the following inequalities:

$$b_2\theta < b_2\hat{\theta} < b_2\underline{\theta}. \tag{14}$$

Note that by the definition of $\underline{\theta}$ and (12) we had

$$b_1\theta < b_1\underline{\theta} < b_1\hat{\theta}. \tag{15}$$

In particular, $h(b_2, b_2\hat{\theta})$ has to intersect the segment $[\theta, \underline{\theta}]$ at some interior point z and x (the projection of θ into $h(b_2, b_2\hat{\theta})$) has to satisfy that $b_1x < b_1z < b_1\underline{\theta} \leq b_1y = b_1y'$ and therefore $y' \in [x, \hat{\theta}]$ (see figure 22). But this concludes our argument since x is the preferred point in $h(b_2, b_2\hat{\theta})$ by S_2, hence the quasi concavity of preferences imply that y' is preferred to $\hat{\theta}$ by S_2. But then S_2 could have gained by deviating to y (or y') in the FRE supported by the Min Rule.

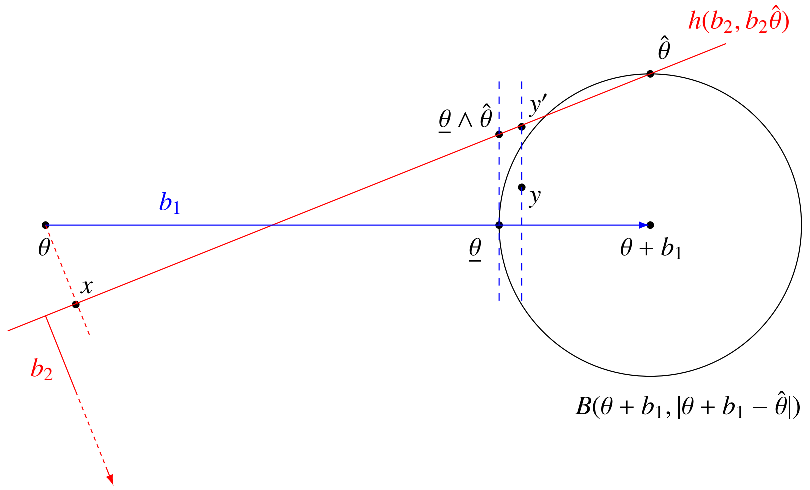


Figure 22: If $y' \equiv y \wedge \hat{\theta} \notin B(\theta + b_1, |\theta + b_1 - \hat{\theta}|)$, then $y' \in B(\theta + b_2, |\theta + b_2 - \hat{\theta}|)$

Proof of Proposition 6:

We introduce some notation that will be used in the proof. Given a close convex set $S \subset \mathbb{R}^2$ and a point $\hat{\theta} \in Fr(S)$, we denote by $t_S^+(\hat{\theta})$, $t_S^-(\hat{\theta})$ the two unit tangent vectors to S at $\hat{\theta}$ such that $C[t_S^+(\hat{\theta}), t_S^-(\hat{\theta})]$ is the smallest cone with vertex $\hat{\theta}$ that contains S, i.e.,

(i) $S \subseteq \{y = \hat{\theta} + v \mid v \in C[t_S^+(\hat{\theta}), t_S^-(\hat{\theta})]\}$

(ii) for any cone C such that $S \subseteq \{y = \hat{\theta} + v \mid v \in C\}$ we have that $C[t_S^+(\hat{\theta}), t_S^-(\hat{\theta})] \subseteq C$

Note that if $\hat{\theta}$ is a smooth point of S then $t_S^+(\hat{\theta}) = -t_S^-(\hat{\theta})$, but this is not the case if $\hat{\theta}$ is a kink point. Given $t_S^+(\hat{\theta}), t_S^-(\hat{\theta})$ an alternative definition of the polar cone to S at $\hat{\theta}$ is:

$$PC_S(\hat{\theta}) \equiv \{n \in \mathbb{R}^2 \mid nt_S^+(\hat{\theta}) \leq 0, nt_S^-(\hat{\theta}) \leq 0\}. \tag{16}$$

Before presenting the proof of the Proposition, we prove two claims about $PC_S(\hat{\theta})$ that will be used in the proof. Claim 1 states that given a convex set S and a point in the frontier $\hat{\theta}$, the closest point in S to any point in the set $X \equiv \{x = \hat{\theta} + n \mid n \in PC_S(\hat{\theta})\}$ is $\hat{\theta}$ itself.

Claim 1: Given a convex set S and a point $\hat{\theta} \in Fr(S)$, then for all $x = \hat{\theta} + n$ where $n \in PC_S(\hat{\theta})$, $n \neq 0$, we have that,

$$\hat{\theta} = \arg\min_{\theta \in S} |\theta - x|$$

Proof of Claim 1: Consider $x = \hat{\theta} + n$ with $n \in PC_S(\hat{\theta})$ and any point $\theta \in S$. Define by y the projection of $\theta - \hat{\theta}$ on to the line with direction n passing through $\hat{\theta}$. Since $n \in PC_S(\hat{\theta})$, $n(\theta - \hat{\theta}) \leq 0$ and hence y and x lie in the same line on opposite sides of $\hat{\theta}$. Therefore by Pythagoras theorem:

$$|\theta - x|^2 = |\theta - y|^2 + |y - x|^2 > |\theta - y|^2 + |\hat{\theta} - x|^2 \geq |\hat{\theta} - x|^2$$

□

Claim 2: Given a convex set $S \subset \mathbb{R}^2$ and a point $\hat{\theta} \in Fr(S)$, if $b \in int(PC_S(\hat{\theta}))$, then there exists $\epsilon > 0$ such that for all $\theta \in B(\hat{\theta}, \epsilon)$, $\theta + b - \hat{\theta} \in int(PC_S(\hat{\theta}))$.

Proof of Claim 2: For $b \in int(PC_S(\hat{\theta}))$, $bt_S^+(\hat{\theta}) < 0$ and $bt_S^-(\hat{\theta}) < 0$ by (16). Define $\epsilon = \min\{-bt_S^+(\hat{\theta}), -bt_S^-(\hat{\theta})\} > 0$. Consider $\theta \in B(\hat{\theta}, \epsilon)$, then $(\theta + b - \hat{\theta})t_S^+(\hat{\theta}) = (\theta - \hat{\theta})t_S^+(\hat{\theta}) + bt_S^+(\hat{\theta}) < \epsilon - \epsilon = 0$. Analogously, $(\theta + b - \hat{\theta})t_S^-(\hat{\theta}) < 0$ and hence by (16)

$$\theta + b - \hat{\theta} \in int(PC_S(\hat{\theta})).$$

□

We start now with the proof of Proposition 6. We first show that conditions (i) and (ii) are necessary for collusion-proofness:
Suppose there exists a $\hat{\theta} \in \widetilde{Fr}(Y)$ such that $b_1, b_2 \in C(n_Y^{Out}(\hat{\theta}), t_Y(\hat{\theta})]$ with $t_Y(\hat{\theta})$ one of the two unit normal vectors to Y at $\hat{\theta}$. Suppose that $b_2 \in C(b_1, n_Y^{Out}(\hat{\theta}))$. The opposite

case in which $b_1 \in C(b_2, n_Y^{Out}(\hat{\theta}))$ is symmetric. Since Y has non-empty interior and $b_1 \in C(n_Y^{Out}(\hat{\theta}), t_Y(\hat{\theta})]$, there exists $\epsilon_1 > 0$ such that for all $0 < \epsilon < \epsilon_1$, $\hat{\theta} - \epsilon b_1 \in Y$. Moreover, since $b_2 \in C(b_1, n_Y^{Out}(\hat{\theta}))$, $b_2 = \alpha b_1 + \beta n_Y^{Out}(\hat{\theta})$ with $\alpha > 0, \beta > 0$. Therefore there exists $\epsilon_2 > 0$ such that for all $0 < \epsilon < \epsilon_2$, $b_2 - \epsilon b_1 = (\alpha - \epsilon)b_1 + \beta n_Y(\hat{\theta}) \in C(b_1, n_Y^{Out}(\hat{\theta}))$. Define $\tilde{\epsilon} = \min\{\epsilon_1, \epsilon_2, 1\}$ and $\theta = \hat{\theta} - \tilde{\epsilon} b_1$. Since $\tilde{\epsilon} \leq \epsilon_1$, $\theta \in Y$. We now show that at θ the senders can collude at $\hat{\theta}$.

Denote by S the punishment region if both senders report $\hat{\theta}$, i.e., S is the set of policies that a sender deviating from the collusive report $\hat{\theta}$ could induce, given that the receiver uses the Min Rule. Since $b_2 \in C(b_1, n_Y^{Out}(\hat{\theta}))$, for all $y \in Y$ such that $b_1 y \leq b_1 \hat{\theta}$,

$$b_2 y = (\alpha b_1 + \beta n_Y^{Out}(\hat{\theta}))y = \alpha b_1 y + \beta n_Y^{Out}(\hat{\theta})y \leq \alpha b_1 \hat{\theta} + \beta n_Y^{Out}(\hat{\theta})\hat{\theta} = b_2 \hat{\theta}$$

where the last inequality uses the facts that $b_1 y \leq b_1 \hat{\theta}$ and that by the definition of $n_Y^{Out}(\hat{\theta})$, $n_Y^{Out}(\hat{\theta})(y - \hat{\theta}) \leq 0$. Therefore the punishment region can be written as $S = Y \setminus H(b_1, b_1\hat{\theta})$ and $PC_S(\hat{\theta}) = C[b_1, n_Y^{Out}(\hat{\theta})]$. Note that in particular $\theta \in S$ and hence either sender could induce θ if he preferred it to $\hat{\theta}$. Now,

$$\theta + b_1 - \hat{\theta} = (1 - \tilde{\epsilon})b_1 \in PC_S(\hat{\theta})$$
$$\theta + b_2 - \hat{\theta} = b_2 - \tilde{\epsilon} b_1 \in PC_S(\hat{\theta})$$

hence, by Claim 1, $\hat{\theta}$ is the closest point to both $\theta + b_1$ and $\theta + b_2$ in S and neither of the senders would benefit by deviating from $\hat{\theta}$. See Figure 23.

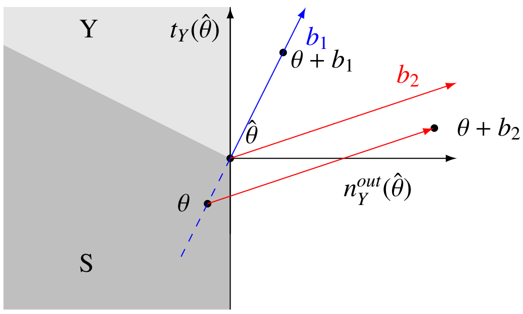


Figure 23: When $b_1, b_2 \in C(n_Y^{Out}(\hat{\theta}), t_Y(\hat{\theta})]$ we can find a $\theta \in Y$ such that both senders would like to collude at $\hat{\theta}$ and neither of the senders have an incentive to deviate from the collusive report.

Suppose now that there exists $\hat{\theta} \in Fr(Y)$ such that $b_1, b_2 \in int(PC_Y(\hat{\theta}))$. By Claim 2, there exists ϵ_i>0, $i \in \{1, 2\}$, such that for all $\theta \in B(\hat{\theta}, \epsilon_i)$, $\theta + b_i - \hat{\theta} \in int(PC_Y(\hat{\theta}))$. Define $\epsilon = \min\{\epsilon_1, \epsilon_2\}$, since Y has non-empty interior, there exists $\theta \in B(\hat{\theta}, \epsilon) \cap Y$ with $\theta \neq \hat{\theta}$. Then $\theta + b_1 - \hat{\theta}, \theta + b_2 - \hat{\theta} \in int(PC_Y(\hat{\theta}))$ and hence by Claim 1, $\hat{\theta}$ is the closest point in Y to both $\theta + b_1$ and $\theta + b_2$. Therefore at θ both senders want to collude to $\hat{\theta}$ and neither sender has an incentive to deviate from the collusive report.

We prove now the other implication, i.e. that conditions (i) and (ii) are sufficient for collusion-proofness:

Suppose that the FRE supported by the Min Rule is not collusion-proof for some magni-

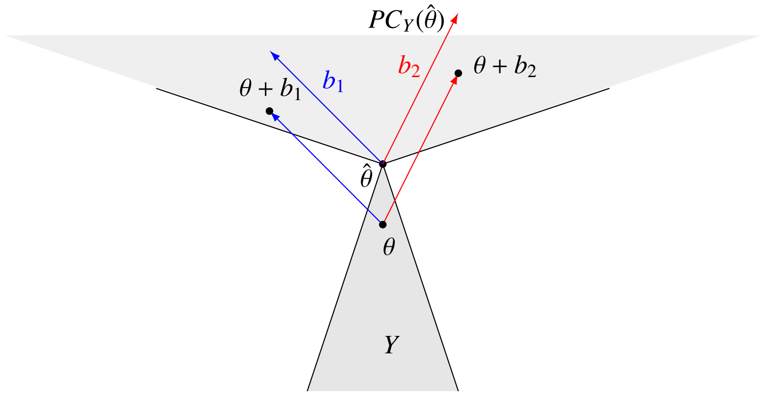


Figure 24: If $b_1, b_2 \in PC_Y(\hat\theta)$, we can find θ such that both senders would like to collude at $\hat\theta$ and neither sender has an incentive to deviate from the collusive report.

tudes of the biases $t_1 = |b_1|$, $t_2 = |b_2|$. Then there exists $\theta, \hat\theta \in Y$ such that:

$$(a) \qquad \hat\theta \in \bar{B}(\theta + b_1, |b_1|) \cap \bar{B}(\theta + b_2, |b_2|)$$

$$(b) \quad \text{For all } \tilde\theta \in Y \text{ with } b_1\tilde\theta \leq b_1\hat\theta, \text{ and } b_2\tilde\theta \leq b_2\hat\theta, \text{ then } \tilde\theta \notin \bar{B}(\theta + b_1, |\theta + b_1 - \hat\theta|) \cup \bar{B}(\theta + b_2, |\theta + b_2 - \hat\theta|)$$

Condition (a) states that in state θ, both senders prefer the policy $\hat\theta$ to θ. Condition (b) states that in state θ, neither sender can gain by deviating from the collusive report $\hat\theta$, given that the receiver will respond to a deviation using the Min Rule; the receiver's use of the Min Rule restricts a deviating sender to inducing only policies $\tilde\theta$ such that $b_1\tilde\theta \leq b_1\hat\theta$, and $b_2\tilde\theta \leq b_2\hat\theta$.

Condition (a) can be rewritten as

$$\theta \in Y \cap \bar{B}(\hat\theta - b_1, |b_1|) \cap \bar{B}(\hat\theta - b_2, |b_2|)$$

Note that $Y \cap \bar{B}(\hat\theta - b_1, |b_1|) \cap \bar{B}(\hat\theta - b_2, |b_2|) \subset Y \setminus (H(b_1, b_1\hat\theta) \cup H(b_2, b_2\hat\theta))$. Hence, denoting by n_1 and n_2 the unit normal vector to b_1 and b_2 respectively such that $n_1 b_2 < 0$, $n_2 b_1 < 0$, condition (a) implies that $\theta - \hat\theta \in C[n_1, n_2]$. In particular

$$\text{either} \quad n_1(\theta - \hat\theta) > 0 \quad \text{or} \quad n_2(\theta - \hat\theta) > 0. \tag{17}$$

Consider $i \in \{1, 2\}$ such that $n_i(\theta - \hat\theta) > 0$. Then condition (b) implies that for any $0 < \epsilon \leq n_i(\theta - \hat\theta)$, $\hat\theta + \epsilon n_i$ is not feasible; if there were an ϵ with $0 < \epsilon \leq n_i(\theta - \hat\theta)$ such that $\tilde\theta = \hat\theta + \epsilon n_i \in Y$, we would have that $b_i\tilde\theta = b_i\hat\theta$, $b_j\tilde\theta < b_j\hat\theta$ (since $n_i b_j < 0$) and $|\theta + b_i - \tilde\theta| < |\theta + b_i - \hat\theta|$. To see this last inequality, note that $\theta + b_i$ lies on the orthogonal line to $\tilde\theta - \hat\theta$ passing through θ. Denote by y the intersection between the line passing by $\hat\theta$ and $\tilde\theta$ and the orthogonal line passing by θ and $\theta + b_i$. The distance between $\tilde\theta$ and y is $n_i(\theta - \hat\theta) - \epsilon$, whereas the distance between $\hat\theta$ and y is $n_i(\theta - \hat\theta)$. By Pythagoras,

$$|\theta + b_i - \tilde\theta|^2 = |\theta + b_i - y|^2 + (n_i(\theta - \hat\theta) - \epsilon)^2 < |\theta + b_i - y|^2 + n_i(\theta - \hat\theta)^2 = |\theta + b_i - \hat\theta|^2.$$

Therefore sender S_i would be able to induce $\tilde\theta$ given the receiver use of the Min Rule, and he would gain by deviating from $\hat{\theta}$ to $\tilde{\theta}$. See Figure (25).

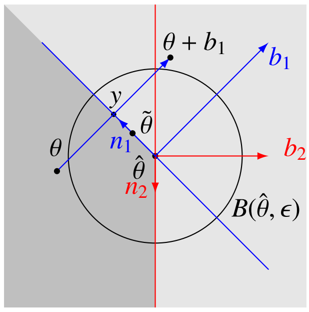


Figure 25: If the collusive report $\hat{\theta}$ is in the interior of Y, at least one sender (in this case S_1) has an incentive to deviate from it.

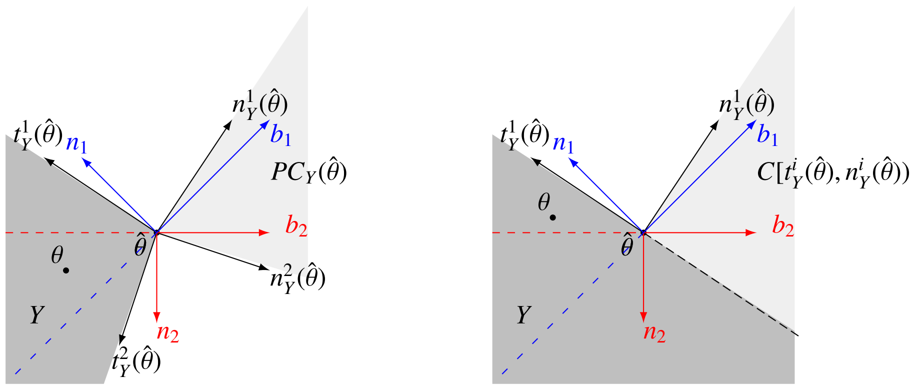
Hence a necessary condition for $\hat{\theta}$ to be a collusive report given θ is that for any $i \in \{1, 2\}$:

$$\text{if} \quad n_i(\theta - \hat{\theta}) > 0 \quad \text{then} \quad \hat{\theta} + \epsilon n_i \notin Y, \text{ for any } 0 < \epsilon \leq n_i(\theta - \hat{\theta}) \tag{18}$$

In other words, the ray departing from $\hat{\theta}$ in the direction of n_i cannot be feasible. In particular, $\hat{\theta}$ has to be in the frontier of Y.

Consider $n_i(\theta - \hat{\theta}) > 0$, (recall that by (17) we know that there exists such i), then (18) implies that one and only one of the tangent vectors $t_Y^+(\hat{\theta})$ and $t_Y^-(\hat{\theta})$ belongs to $C[\theta - \hat{\theta}, n_i)$. Denote such tangent vector by $t_Y^i(\hat{\theta})$, and the *outward*[35] normal vector to $t_Y^i(\hat{\theta})$ by $n_Y^i(\hat{\theta})$ (See Figure 26). Then,

$$n_i \in C(t_Y^i(\hat{\theta}), n_Y^i(\hat{\theta})). \tag{19}$$



(a) If $n_1(\theta - \hat{\theta}) > 0$ and $n_2(\theta - \hat{\theta}) > 0$, then $\hat{\theta}$ is a kink point and $b_1, b_2 \in int(PC_Y(\hat{\theta}))$.

(b) If $n_1(\theta - \hat{\theta}) > 0$ and $n_2(\theta - \hat{\theta}) \leq 0$, then $b_1, b_2 \in C[t_Y^i(\hat{\theta}), n_Y^i(\hat{\theta}))$.

Figure 26

Suppose that $n_j(\theta - \hat{\theta}) > 0$ for $j \neq i$. Then, following the same argument as above, the ray departing from $\hat{\theta}$ in the direction n_j cannot be feasible and $\hat{\theta}$ has to be a kink point (See Figure 26a). Denoting by $t^j(\hat{\theta}) \in \{t_Y^+(\hat{\theta}), t_Y^-(\hat{\theta})\} \setminus t_Y^i(\hat{\theta})$, the remaining tangent vector, it has

[35] By *outward* we refer to the normal vector to $t_Y^i(\hat{\theta})$ that exits from Y.

to be that $t^j_Y(\hat{\theta}) \in C[\theta - \hat{\theta}, n_j)$, and denoting by $n^j(\hat{\theta})$ its normal outward vector, we have that $n_j \in C(t^j_Y(\hat{\theta}), n^j_Y(\hat{\theta}))$. This together with (19) implies that $b_1, b_2 \in C(n^i(\hat{\theta}), n^j(\hat{\theta}))$. But $C(n^i(\hat{\theta}), n^j(\hat{\theta})) \equiv int(PC_Y(\hat{\theta}))$, hence $b_1, b_2 \in int(PC_Y(\hat{\theta}))$ which contradicts condition (ii) of the proposition.

Lastly, suppose that $n_j(\theta - \hat{\theta}) \leq 0$ while $n_i(\theta - \hat{\theta}) > 0$. This together with (19) implies that

$$b_1, b_2 \in C(n^i_Y(\hat{\theta}), -t^i_Y(\hat{\theta})]. \tag{20}$$

If $\hat{\theta}$ is a smooth point then $n^i_Y(\hat{\theta}) = n^{Out}_Y(\hat{\theta})$ and $-t^i_Y(\hat{\theta}) = t_Y(\hat{\theta})$. Hence, condition (i) of the proposition is violated (See Figure 26b). If $\hat{\theta}$ is a kink point, suppose first that b_2 and $-t^i_Y(\hat{\theta})$ are linearly independent, then by moving along the frontier in the direction of $t^i_Y(\hat{\theta})$, there exists a smooth point $\bar{\theta}$ sufficiently close to $\hat{\theta}$ such that $t_Y(\bar{\theta}) \in C(b_2, -t^i_Y(\hat{\theta}))$ and hence $b_1, b_2 \in C(n^{Out}_Y(\bar{\theta}), t_Y(\bar{\theta})]$ and condition (i) of the proposition is violated. If b_2 has the same direction as $-t^i_Y(\hat{\theta})$, then it has to be that $n_j(\theta - \hat{\theta}) = 0$. Consider $\bar{\theta} = \hat{\theta} + \epsilon n_j$ for $0 < \epsilon < |\hat{\theta} - \theta|$. $\bar{\theta}$ is a smooth point in the frontier and $n^{Out}_Y(\bar{\theta}) = n^i_Y(\hat{\theta})$, $t_Y(\bar{\theta}) = -t^i_Y(\hat{\theta})$. Hence $b_1, b_2 \in C(n^{Out}_Y(\bar{\theta}), t_Y(\bar{\theta})]$ which contradicts condition (i) of the proposition. □

Proof of Proposition 7:

By Proposition 3, it is enough to show the equivalence for fully revealing equilibria. Given Proposition 2 we can focus on the case $\Theta \equiv Y$. We show that for $Y \subseteq R^q$, a fully revealing equilibrium that is independent of the magnitudes of the biases exists if and only if, for the two-dimensional state space Y_b, a fully revealing equilibrium that is independent of the magnitudes of the biases exists, where now the biases are regarded as two-dimensional vectors in Π_b. Since Y is compact, Y_b is closed and the equivalences claimed in the proposition then follow from Proposition 4.

Given $\tilde{\theta} \in \mathbb{R}^q$, define $\tilde{\theta}_b \equiv Proj_b(\tilde{\theta})$. Then

$$\tilde{\theta} \in H(b, b\theta) \iff b\tilde{\theta} > b\theta \iff b\tilde{\theta}_b > b\theta_b \iff \tilde{\theta}_b \in H_b(b, b\theta_b). \tag{21}$$

Suppose there does not exist a fully revealing equilibrium for all biases $(t_1 b_1, t_2 b_2)$ with $t_1, t_2 \geq 0$. By Proposition 1, $Y \subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$ for some $\theta', \theta'' \in Y$. Define $\theta'_b \equiv Proj_b(\theta')$ and $\theta''_b \equiv Proj_b(\theta'')$. Then it follows from (21) that $Y_b \subseteq H_b(b_1, b_1\theta''_b) \cup H_b(b_2, b_2\theta'_b)$, so by Proposition 1, for the state space Y_b, there does not exist a fully revealing equilibrium for arbitrarily large magnitudes of the biases. The reverse implication is proved analogously, again using (21). □

Proof of Proposition 8:

We begin the proof by showing that the GDC and the LDC in Proposition 8 correspond *essentially* to the necessary and sufficient conditions for the existence of a robust fully revealing equilibrium in pure strategies established in Proposition 1.[36] In the second part of

[36] As will be apparent in the proof, if the frontier has a very particular kind of kink, even if the conditions of Proposition 8 are satisfied, it might not be possible to locally punish with pure strategies some local deviations

the proof we show that mixing by the receiver after out of equilibrium reports will *almost never* support an FRE when an FRE cannot be supported in pure strategies.

To establish the equivalence for pure strategies, note that by Proposition 2, we can restrict attention to the case in which $\Theta \equiv Y$. We will show that the GDC is equivalent to condition (*i*) of Proposition 1 and that the LDC is equivalent to condition (*ii*) of Proposition 1 whenever the policy space does not have a very particular kind of kink point. We deal with this particular kink separately.

Prop.1-(i) $\Rightarrow$ *GDC*: Suppose there exists $\theta \in \widetilde{Fr}(co(Y))$ such that $n_{co(Y)}(\theta) \in C(b_1, b_2)$. Then there exists $\alpha > 0$, $\beta > 0$ such that $n_{co(Y)}(\theta) = \alpha b_1 + \beta b_2$. Moreover, since $\theta \in \widetilde{Fr}(co(Y))$, $h(n_{co(Y)}(\theta), n_{co(Y)}(\theta)\theta)$ is the unique separating hyperplane to $co(Y)$ at θ. Hence neither $h(b_1, b_1\theta)$ nor $h(b_2, b_2\theta)$ are separating hyperplanes of $co(Y)$. In particular, there exist $\theta', \theta'' \in Y$ such that

$$b_1\theta'' < b_1\theta < b_1\theta' \qquad b_2\theta' < b_2\theta < b_2\theta''.$$

We now show that $Y \subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$, which contradicts Prop.1-(*i*). Suppose there exists $\tilde{\theta} \in Y$ such that $\tilde{\theta} \notin H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$, then $b_1\tilde{\theta} \leq b_1\theta'' < b_1\theta$ and $b_2\tilde{\theta} \leq b_2\theta' < b_2\theta$. And hence $n_{co(Y)}(\theta)\tilde{\theta} = \alpha b_1\tilde{\theta} + \beta b_2\tilde{\theta} < \alpha b_1\theta + \beta b_2\theta = n_{co(Y)}(\theta)\theta$ which implies that $n_{co(Y)}(\theta)(\tilde{\theta} - \theta) < 0$ and hence contradicts the definition of $n_{co(Y)}(\theta)$.

GDC $\Rightarrow$ *Prop.1-(i)*: Suppose there exist $\theta', \theta'' \in Y$ such that $Y \subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. Since Y is compact, consider $\tilde{\theta}' \in \arg\min\{b_2 y \mid y \in Y\}$ and $\tilde{\theta}'' \in \arg\min\{b_1 y \mid y \in Y\}$. Since $b_1\tilde{\theta}'' \leq b_1\theta''$ and $b_2\tilde{\theta}' \leq b_2\theta'$, $Y \subseteq H(b_1, b_1\tilde{\theta}'') \cup H(b_2, b_2\tilde{\theta}')$. In particular $b_1\tilde{\theta}' > b_1\tilde{\theta}''$, $b_2\tilde{\theta}'' > b_2\tilde{\theta}'$ and $x = h(b_1, b_1\tilde{\theta}'') \cap h(b_2, b_2\tilde{\theta}') \notin Y$. Moreover, by the definition of $\tilde{\theta}', \tilde{\theta}''$, $Y \subset \overline{H}(b_1, b_1\tilde{\theta}'') \cap \overline{H}(b_2, b_2\tilde{\theta}')$ and x cannot be written as a convex combination of points in Y ($x \notin co(Y)$). Now choose any point $\tilde{\theta} \in \widetilde{Fr}(co(Y))$ such that $\tilde{\theta}$ belongs to the triangle formed by $x, \tilde{\theta}'$ and $\tilde{\theta}''$. Then denoting $n = n_{co(Y)}(\tilde{\theta})$, we have that $n(\tilde{\theta}' - \tilde{\theta}) \geq 0$, $n(\tilde{\theta}'' - \tilde{\theta}) \geq 0$, $n(x - \tilde{\theta}) < 0$ which implies that $n(\tilde{\theta}' - x) > 0$ and $n(\tilde{\theta}'' - x) > 0$. Using $\{b_1, b_2\}$ as a base for $\mathbb{R}^2$ we can write $n = \alpha b_1 + \beta b_2$ and hence $\alpha b_1(\tilde{\theta}' - \tilde{\theta}'') > 0$ and $\beta b_2(\tilde{\theta}'' - \tilde{\theta}') > 0$ which implies $\alpha, \beta > 0$ and therefore $n \in C(b_1, b_2)$ which contradicts the GDC.

Prop.1-(ii) $\Rightarrow$ *LDC*: Suppose there exists $\theta \in \widetilde{Fr}(Y)$ such that $n_Y(\theta) \in C(b_1, b_2)$. Then for any $\delta > 0$, both $B(\theta, \delta) \cap Y \cap H(b_i, b_i\theta) \neq \emptyset$ and $B(\theta, \delta) \cap Y \cap \{y \in \mathbb{R}^2 \mid b_i y < b_i\theta\} \neq \emptyset$ for $i = 1, 2$. Moreover, there exists $\epsilon > 0$ such that

$$B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta). \tag{22}$$

For any $\delta > 0$ consider $\theta' \in B(\theta, \delta) \cap Y \cap \{x \in \mathbb{R}^2 \mid b_2 x < b_2\theta\}$ and $\theta'' \in B(\theta, \delta) \cap Y \cap \{x \in \mathbb{R}^2 \mid b_1 x < b_1\theta\}$. Then $B(\theta, \epsilon) \cap Y \subseteq H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$. To see this, consider $\tilde{\theta} \in B(\theta, \epsilon) \cap Y$. By (22), $\tilde{\theta} \in \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$. Suppose $\tilde{\theta} \in \overline{H}(b_1, b_1\theta)$, then $b_1\tilde{\theta} \geq b_1\theta > b_1\theta''$ so $\tilde{\theta} \in H(b_2, b_2\theta')$. Similarly, if $\tilde{\theta} \in \overline{H}(b_2, b_2\theta)$, then $\tilde{\theta} \in H(b_2, b_2\theta')$. Hence $\tilde{\theta} \in H(b_1, b_1\theta'') \cup H(b_2, b_2\theta')$.

from the kink. In that case, though, we show that local mixing by the receiver deters such local deviations.

LDC $\Rightarrow$ *Prop.1-(ii)*: For this implication we assume that the $Fr(Y)$ does not have a very particular kink point; we assume that there does not exists $\tau > 0$ and $\theta \in Fr(Y)$ such that $Y \cap B(\theta, \tau) = (\overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)) \cap B(\theta, \tau)$. In other words, there is not a locally non-convex kink point[37] such that in a neighbourhood of the kink point the frontier is linear on both sides of the point with inward normal vectors equal to the two biases.

Suppose that local deviations from $\theta \in Y$ cannot be deterred with pure strategies (and hence $\theta \in Fr(Y)$). By Step 1 of Proposition 3, there exists an $\epsilon > 0$ such that $B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$. Moreover for all $\delta > 0$, $B(\theta, \delta) \cap Y \not\subseteq \overline{H}(b_1, b_1\theta) \cap \overline{H}(b_2, b_2\theta)$ because if not θ would be locally the worst point for both senders and a local deviation could be deterred by choosing θ. Since $Fr(Y)$ has a finite number of kinks and the frontier to the sides of θ is not linear with normal vector in $\{b_1, b_2\}$, there exits either:

$\theta' \in Fr(Y) \cap B(\theta, \epsilon) \cap H(b_1, b_1\theta) \setminus \overline{H}(b_2, b_2\theta)$ such that $Fr(Y)$ is differentiable in (θ, θ') and $b_1(\theta' - \theta) > 0$ and $b_2(\theta' - \theta) < 0$, or

$\theta'' \in Fr(Y) \cap B(\theta, \epsilon) \cap H(b_2, b_2\theta) \setminus \overline{H}(b_1, b_1\theta)$ such that $Fr(Y)$ is differentiable in (θ, θ'') and $b_1(\theta'' - \theta) > 0$ and $b_2(\theta'' - \theta) > 0$.

Suppose that we are in the first case (the second case is analogous), then by the mean value theorem there exists $\tilde{\theta} \in Fr(Y)$ between θ and θ' such that $n_Y(\tilde{\theta})(\theta' - \theta) = 0$. Using b_1, b_2 as a base of $\mathbb{R}^2$, we have that $n_Y(\tilde{\theta}) = \alpha b_1 + \beta b_2$ and hence, $0 = n_Y(\tilde{\theta})(\theta' - \theta) = \alpha b_1(\theta' - \theta) + \beta b_2(\theta' - \theta)$. And since $b_1(\theta' - \theta) > 0$ and $b_2(\theta' - \theta) < 0$, we have that both α and β have the same sign. Moreover since $n_Y(\tilde{\theta})$ is the inward normal vector and $B(\theta, \epsilon) \cap Y \subseteq \overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)$, it has to be that both $\alpha, \beta > 0$, and hence $n_Y(\tilde{\theta}) \in C(b_1, b_2)$.

Suppose that a locally non-convex kink point θ exists such that for $\tau > 0$ the frontier $Fr(Y) \cap B(\theta, \tau) = (\overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)) \cap B(\theta, \tau)$ (see Figure 27). Local deviations from θ cannot be deterred locally in pure strategies because $B(\theta, \tau) \cap (Y \setminus (\overline{H}(b_1, b_1\theta) \cup \overline{H}(b_2, b_2\theta)) = \emptyset$. However, we show below that local deviations could be deterred locally if the receiver mixes between two actions.

Formally, we need to define the receiver's belief given any report (θ', θ'') and the best response (that could involve mixing) given those beliefs that would deter deviations from the senders.

In order to define the receiver's belief, denote by β half the outside angle of $Fr(Y)$ at θ.[38] Define $\bar{\delta} = \tau \cos^3(\beta) \sin(\beta)$. For any $\theta', \theta'' \in B(\theta, \bar{\delta})$ define $\hat{y}(\theta', \theta'') = \theta + \hat{b}\frac{1}{\cos^2(\beta)\sin(\beta)} \max\{|\theta' - \theta|, |\theta'' - \theta|\}$ where $\hat{b}$ is the unit bisector vector to $\{-b_1, -b_2\}$. Figure 27 illustrates $\bar{\delta}$ and for given (θ', θ'') $\hat{y}(\theta', \theta'')$. Note that the definition of $\hat{y}(\theta', \theta'')$ only depends on the maximum distance from θ', θ'' to θ. Hence for example, given $\tilde{\theta}''$ in Figure 27, $\hat{y}(\theta', \tilde{\theta}'')$ coincides with $\hat{y}(\theta', \theta'')$. Moreover $\hat{y}(\theta', \theta'')$ always lies on the bisector line that is equidistant to the two sides of the frontier.

[37] We say that a kink point $\hat{\theta}$ is locally *non-convex* if for all $\epsilon > 0$, there exist $\theta', \theta'' \in B(\hat{\theta}, \epsilon) \cap Fr(Y)$ such that there exists $\lambda \in (0, 1)$, $\lambda\theta' + (1 - \lambda)\theta'' \notin Y$.

[38] The formula for such an angle is given by $\beta = 90° - \frac{1}{2}\cos^{-1}\left(\frac{b_1 b_2}{|b_1||b_2|}\right)$. Note that $0 < \beta < 90°$ and hence $\cos(\beta) > 0$.

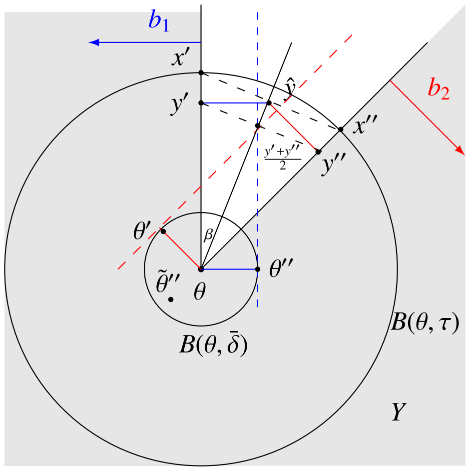


Figure 27: The point θ is a non-convex kink such that the frontier is linear on the two sides of θ, and the inward normal vectors to the frontier coincide with b_1, b_2. A local deviation from θ cannot be deterred by a pure strategy from the receiver. However, the receiver can deter deviations by locally mixing between y' and y'' with equal probability. This can be rationalised by some receiver's beliefs succedenth that the expected state is $\hat{y}$.

We define the receiver's expected beliefs as follows:

$$\hat{\mu}(\theta',\theta'') = \begin{cases} \theta' \wedge_{b_1,b_2} \theta'' & if \quad \theta' \wedge_{b_1,b_2} \theta'' \in Y \\ \hat{y}(\theta',\theta'') & if \quad \theta' \wedge_{b_1,b_2} \theta'' \notin Y, \quad \theta',\theta'' \in B(\theta,\bar{\delta}) \\ \arg\min_{y\in Y} b_1 y & otherwise \end{cases}$$

Note that by the definition of $\hat{y}(\theta',\theta'')$, if $\theta',\theta'' \in B(\theta,\bar{\delta})$, $|\hat{y}(\theta',\theta'')-\theta| \leq \frac{\bar{\delta}}{\cos^2(\beta)\sin(\beta)} = \tau\cos(\beta)$ and hence $\hat{\mu}(\theta',\theta'') \in co(Y)$ and the expected beliefs are well defined.[39]

In order to prove that local deviations can be deterred locally, consider any $\epsilon > 0$ with $\epsilon < \tau$, and define $\delta = \epsilon\cos^3(\beta)sin(\beta) < \bar{\delta}$. For any $\theta',\theta'' \in B(\theta,\delta)$, either $\theta' \wedge_{b_1,b_2} \theta'' \in Y$ (and note that in this case $\theta' \wedge_{b_1,b_2} \theta'' \in B(\theta,\epsilon)$) or, $\theta' \wedge_{b_1,b_2} \theta'' \notin Y$, and the expected belief of the receiver is given by $\hat{y}(\theta',\theta'')$. In this last case, an optimal response for the receiver is to mixed with equal probability between $y'(\theta',\theta'')$ and $y''(\theta',\theta'')$, the projections of $\hat{y}(\theta',\theta'')$ into the frontier (see Figure 27).

We show now, that such response from the receiver deters deviations from both senders. For simplicity we drop the arguments of $\hat{y}, y'$ and y'' below. Sender S_1 has not an incentive to deviate if:

$$\begin{array}{rrcl} & \frac{1}{2}(y'-(\theta''+b_1))^2 + \frac{1}{2}(y''-(\theta''+b_1))^2 & \geq & b_1^2 \\ \Leftrightarrow & \frac{1}{2}(y'-\theta'')^2 + \frac{1}{2}(y''-\theta'')^2 & \geq & \frac{1}{2}b_1(\frac{y'+y''}{2}-\theta'') \end{array}$$

But $b_1(\frac{y'+y''}{2}-\theta'') \leq 0$ by construction, and hence this inequality is always satisfied. The argument for Sender S_2 is symmetric. Moreover it is straight forward to see that the punishments are robust in the sense that as θ' and θ'' converge to each other, either they converge to θ and y' and y'' also converge to θ or, $\theta' \wedge_{b_1,b_2} \theta''$ becomes feasible and it also converges

[39] If $\theta' \wedge_{b_1,b_2} \theta'' \in Y$ the receiver allocates mass one to $\theta' \wedge_{b_1,b_2} \theta''$. When $\theta' \wedge_{b_1,b_2} \theta'' \notin Y$ but $\theta',\theta'' \in B(\theta,\bar{\delta})$, the receiver allocates equal probability to the two point in the frontier that are at a distance $\frac{|\hat{y}(\theta',\theta'')-\theta|}{\cos(\beta)}$ to θ (denoted by x', x'' in Figure 27). Finally, in the remaining cases, the receiver allocates mass one to $\arg\min_{y\in Y} b_1 y$.

to the limit point.

We have therefore seen that the GDC and LDC are sufficient for the existence of a robust fully revealing equilibrium. For the necessity implication, we would need to argue that mixing from the receiver generically does not help in supporting the existence of a robust fully revealing equilibrium, if the conditions are not satisfied. In order to see that we show that if the GDC is not satisfied, then it is not possible to construct a FRE in mixed strategies. Finally, if the GDC is satisfied but the LDC fails, it is not possible to support a robust FRE unless the section of the non-convex part of the frontier where the LDC is violated forms part of a circumference.

Suppose that condition (*ii*-1) in Proposition 8 is violated. Since $co(Y)$ is convex, by Proposition 4 there is not a fully revealing equilibrium in the environment $(\Theta, co(Y))$. Given that $co(Y)$ is compact, define $\underline{Y}_1 = \arg\min_{\theta \in co(Y)} b_1\theta$ and $\underline{Y}_2 \in \arg\min_{\theta \in co(Y)} b_2\theta$. Since there is not a FRE in $co(Y)$, it has to be the case that $\underline{Y}_1 \cap \underline{Y}_2 = \emptyset$. Moreover, by the definition of $co(Y)$ it has to be the case that $\underline{Y}_i \cap Y \neq \emptyset$. Consider $\theta' \in \underline{Y}_1 \cap Y$ and $\theta'' \in \underline{Y}_2 \cap Y$. Clearly $\theta' \neq \theta''$ and any mixing in Y by the receiver is strictly preferred by at least one of the senders to either θ' or θ''.

Finally suppose that condition (*ii*-1) is satisfied but condition (*ii*-2) is not. Consider the point $\theta \in \widetilde{Fr}(Y)$ such that $n_Y^{In}(\theta) \in C(b_1, b_2)$. Using a similar argument to the proof of Proposition 4, it is easy to see that it is not possible to punish small deviations robustly using pure strategies. If $Fr(Y)$ is locally convex at θ, and the receiver is constrained to punish locally, it will never be optimal for her to mix.

Suppose that $Fr(Y)$ is locally non convex at θ. In order to robustly deter deviations from such a point with mixed strategies there needs to exists three sequences $\{\theta'_n\}_{n\in\mathbb{N}}, \{\theta''_n\}_{n\in\mathbb{N}} \subset Fr(Y)$ and $\{\hat{y}_n\}_{n\in\mathbb{N}} \subset co(Y)$ such that:

1. For each $n \in \mathbb{N}$, $\theta'_n \neq \theta''_n$ and $\theta'_n, \theta''_n \in \arg\min_{y\in Y}\{|\hat{y}_n - y|\}$.
2. Both $\{\theta'_n\}_{n\in\mathbb{N}}$ and $\{\theta''_n\}_{n\in\mathbb{N}}$ converge to θ.

The sequence $\{\hat{y}_n\}_{n\in\mathbb{N}}$ represents the sequence of expected beliefs for the receiver, and for each $\hat{y}_n$, there must be at least two distinct points, θ'_n and θ''_n, in the receiver's best response set.

Such requirements are very strong since it requires $Fr(Y)$ to be tangent to a circle (centred at $\hat{y}_n$) at both θ'_n and θ''_n for all $n \in \mathbb{N}$, i.e., as they converge to θ. Moreover, this has to be satisfied for all $\theta \in \widetilde{Fr}(Y)$ for which the condition (*ii*-2) is violated, in particular for a continuum of points. We believe that this condition can only be satisfied if the non convex part of the frontier at which the condition is violated is an arc of a circumference itself, and hence it is non-generic. □

Proof of Proposition 9:

(*i*) $\Leftrightarrow$ (*ii*): This follows from Proposition 4.

(*ii*) $\Rightarrow$ (*iii*): Consider first the case $\Theta = Y$. We show that if for any pair of reports (θ', θ'') in Y such that $\theta' \neq \theta''$, the receiver responds by choosing $y^\mu(\theta', \theta'') = \theta' \wedge_{\{\underline{b},\bar{b}\}} \theta''$, this

response deters both senders from deviating, whatever the realizations of $b_1, b_2 \in C[\underline{b}, \overline{b}]$, and therefore the truthful strategies (s_1, s_2) together with y^μ constitute a robust FRE.
Since $b_1, b_2 \in C[\underline{b}, \overline{b}]$, there exist $\alpha_1, \alpha_2, \beta_1, \beta_2 \geq 0$ such that, $b_i = \alpha_i \underline{b} + \beta_i \overline{b}$, for $i = 1, 2$. Then we have

$$\begin{aligned} b_1(\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta'') &= (\alpha_1 \underline{b} + \beta_1 \overline{b})(\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta'') = \alpha_1 \underline{b}(\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta'') + \beta_1 \overline{b}(\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta'') \\ &= \alpha_1 \min\{\underline{b}\theta', \underline{b}\theta''\} + \beta_1 \min\{\overline{b}\theta', \overline{b}\theta''\} \\ &= \min\{b_1\theta', b_1\theta'', \alpha_1 \underline{b}\theta' + \beta_1 \overline{b}\theta'', \alpha_1 \underline{b}\theta'' + \beta_1 \overline{b}\theta'\} \\ &\leq \min\{b_1\theta', b_1\theta''\} \end{aligned} \tag{23}$$

Analogously, $b_2(\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta'') \leq \min\{b_2\theta', b_2\theta''\}$. Therefore, when the realized biases are (b_1, b_2), the strategy $y^\mu(\theta', \theta'') = \theta' \wedge_{\{\underline{b},\overline{b}\}} \theta''$ deters S_1 in state θ'' from reporting θ' and S_2 in state θ' from reporting θ''. (Since the rule $y^\mu(\theta', \theta'') = \theta' \wedge_{\{\underline{b},\overline{b}\}} \theta''$ is anonymous with respect to the senders, it also deters them, when the biases are (b_1, b_2), from generating the incompatible pair (θ'', θ').) Note that whenever θ', θ'' converge to θ, $\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta''$ also converges to θ, and hence $\theta' \wedge_{\{\underline{b},\overline{b}\}} \theta''$ deters local deviations with local actions. Furthermore, observe that the inequality (23) and the analogous inequality for b_2 hold for any $b_1, b_2 \in C[\underline{b}, \overline{b}]$ independently of whether those values of the biases belong to the support of the conditional distribution of the biases given the realization of the state.
Consider now $Y \subsetneq \Theta$, and for any $\theta \in \Theta$ define $\tilde{s}_i(\theta) = y^*(\theta)$. We show that for any realization of the biases (b_1, b_2), $(\tilde{s}_1, \tilde{s}_2, y^\mu)$ is a robust FRE in (Θ, Y) for arbitrarily large biases.
Given $y', y'' \in Y$ denote by $x = y^\mu(y', y'') = y' \wedge_{\{\underline{b},\overline{b}\}} y''$. For sender S_1 we need to show that for any $\theta \in \Theta$ such that $y^*(\theta) = y''$, $|\theta + tb_1 - y''| \leq |\theta + tb_1 - x|$ for all $t > 0$ and for all $b_1 \in C[\underline{b}, \overline{b}]$. Consider any such $\theta \in \Theta$ with $y^*(\theta) = y''$, that is, y'' is the closest point in Y to θ. In particular $|\theta - y''| \leq |\theta - x|$. Define z as the midpoint of the segment $[x, y'']$. Then $\theta(y'' - x) \geq z(y'' - x)$, and by (23), $(\theta + tb_1)(y'' - x) \geq z(y'' - x)$ for all $t > 0$ and all $b_1 \in C[\underline{b}, \overline{b}]$, or in other words $|\theta + tb_1 - y''| \leq |\theta + tb_1 - x|$ for all $t > 0$ and all $b_1 \in C[\underline{b}, \overline{b}]$. A similar argument for S_2 shows that for any $\theta \in \Theta$ such that $y^*(\theta) = y'$, $|\theta + tb_2 - y'| \leq |\theta + tb_2 - x|$ for all $t > 0$ and all $b_2 \in C[\underline{b}, \overline{b}]$. Therefore $(\tilde{s}_1, \tilde{s}_2, y^\mu)$ is a FRE in (Θ, Y).
$(iii) \Rightarrow (i)$: Given Y, if for $\Theta \supsetneq Y$ there exists a robust FRE then for $\Theta = Y$ there exists a robust FRE. Given that for all $\theta \in Y$, the realization of biases $(\underline{b}, \overline{b})$ has positive probability, then Proposition 4 implies that condition (i) must hold. □

Equivalence of Robustness and Continuity on the Diagonal

As mentioned in the introduction Ambrus and Takahashi (2008) define the following notion of continuity in the responses of the receiver:

Definition 6 (Ambrus and Takahashi (2008))**.** A fully revealing equilibrium (s_1, s_2, y) is *continuous on the diagonal* if

$$\lim_{n\to\infty} y(s_1(\theta_1^n), s_2(\theta_2^n)) = y^*(\theta)$$

for any sequence $\{(\theta_1^n, \theta_2^n)\}_{n\in\mathbb{N}}$ of pairs of states such that $\lim_{n\to\infty} y^*(\theta_1^n) = \lim_{n\to\infty} y^*(\theta_2^n) = y^*(\theta)$.

We show in the lemma below that this notion of continuity is equivalent to our definition of robustness.

Lemma 3. *A fully revealing equilibrium (s_1, s_2, y) is robust if and only if it is continuous on the diagonal.*

Proof. $\Rightarrow$) Consider any pair of sequences $\{(\theta_1^n, \theta_2^n)\}_{n\in\mathbb{N}} \subset \Theta$ such that $\lim_{n\to\infty} y^*(\theta_1^n) = \lim_{n\to\infty} y^*(\theta_2^n) = y^*(\theta)$. Since μ deters local deviations with local actions, for every $\epsilon > 0$ there exists a $\delta > 0$ such that for all $y^*(\theta'), y^*(\theta'') \in B(y^*(\theta), \delta) \cap Y$, $y(s_1(\theta'), s_2(\theta'')) \in B(y^*(\theta), \epsilon)$. Now, $\lim_{n\to\infty} y^*(\theta_1^n) = \lim_{n\to\infty} y^*(\theta_2^n) = y^*(\theta)$ implies that for that $\delta > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $y^*(\theta_1^n), y^*(\theta_2^n) \in B(y^*(\theta), \delta) \cap Y$, which implies that $y(s_1(\theta_1^n), s_2(\theta_2^n)) \in B(y^*(\theta), \epsilon)$ and hence the equilibrium is continuous on the diagonal.

$\Leftarrow$) We argue by contradiction. Suppose that μ does not deter local deviations with local actions. Then there exists $\theta \in \Theta$ and $\epsilon > 0$ such that for all $n \in \mathbb{N}$ there exists θ_1^n, θ_2^n such that $y^*(\theta_1^n), y^*(\theta_2^n) \in B(y^*(\theta), \frac{1}{n}) \cap Y$ with

$$y(s_1(\theta_1^n), s_2(\theta_2^n)) \notin B(y^*(\theta), \epsilon) \setminus \left(B(\theta_1^n + b_2, |b_2|) \cup B(\theta_2^n + b_1, |b_1|)\right).$$

Note that for any n such that $\frac{1}{n} < \epsilon$, $\theta_1^n \neq \theta_2^n$, because if $\theta_1^n = \theta_2^n$, $y(s_1(\theta_1^n), s_2(\theta_2^n)) = y^*(\theta_1^n) \in B(y^*(\theta), \epsilon) \setminus \left(B(\theta_1^n + b_2, |b_2|) \cup B(\theta_2^n + b_1, |b_1|)\right)$. Since (s_1, s_2, y) is an equilibrium, $y(s_1(\theta_1^n), s_2(\theta_2^n)) \notin B(\theta_1^n + b_2, |b_2|) \cup B(\theta_2^n + b_1, |b_1|)$, otherwise either sender 1 would have an incentive to deviate to $s_1(\theta_1^n)$ when θ_2^n is realized, or sender 2 would have an incentive to deviate to $s_2(\theta_2^n)$ when θ_1^n is realized. Hence $y(s_1(\theta_1^n), s_2(\theta_2^n)) \notin B(\theta, \epsilon)$, which contradicts the diagonal continuity of the equilibrium. □

References

Alonso, R., W. D. and Matouschek, N. (2008). When Does Coordination Require Centralization?, *American Economic Review* **98**: 145–179.

Ambrus, A. and Lu, S. E. (2014). Almost Fully Revealing Cheap Talk with Imperfectly Informed Senders, *Games and Economic Behavior* **88**: 174–189.

Ambrus, A. and Takahashi, S. (2008). Multi-Sender Cheap Talk with Restricted State Spaces, *Theoretical Economics* **3**: 1–27.

Battaglini, M. (2002). Multiple Referrals and Multidimensional Cheap Talk, *Econometrica* **70**: 1379–1401.

Crawford, V. P. and Sobel, J. (1982). Strategic Information Transmission, *Econometrica* **50**: 1431–1451.

Friedman, J. W. and Samuelson, L. (1990). Subgame Perfect Equilibrium with Continuous Reaction Functions, *Games and Economic Behavior* **2**: 304–324.

Friedman, J. W. and Samuelson, L. (1994). Continuous Reaction Functions in Duopolies, *Games and Economic Behavior* **6**: 55–82.

Gilligan, T. W. and Krehbiel, K. (1989). Asymmetric Information and Legislative Rules with a Heterogeneous Committee, *American Journal of Political Science* **33**: 459–490.

Krishna, V. and Morgan, J. (2001a). A Model of Expertise, *Quarterly Journal of Economics* **116**: 747–775.

Krishna, V. and Morgan, J. (2001b). Asymmetric Information and Legislative Rules: Some Amendments, *American Political Science Review* **95**: 435–452.

Lai, E. K., Lim, W. and Wang, J. T. (2015). An Experimental Analysis of Multidimensional Cheap Talk, *Games and Economic Behavior* **91**: 114–144.

Levy, G. and Razin, R. (2007). On The Limits of Communication in Multidimensional Cheap Talk: A Comment, *Econometrica* **75**: 885–893.

Rubanov, O. (2015). Asymptotic Full Revelation in Cheap Talk with a Large Number of Senders. Mimeo.