Communicating about Endogenous Issues*

Elliot Lipnowski† Doron Ravid‡
Yale University *University of Michigan*

August 2026

Abstract

Limited attention forces organizations to decide not only how much to discuss, but also which issues merit discussion. We study strategic communication about a multidimensional decision when a receiver can respond only along a few endogenously chosen issues. Players agree on the ideal action but prioritize different errors. In equilibrium, communicated and omitted issues must be statistically unrelated and separable according to the sender's preferences. Thus, the sender's priorities determine the agenda; the receiver's do not. Synchronized priorities raise the receiver's best equilibrium payoff but lower his worst, so he may prefer a less synchronized sender.

1 Introduction

Organizations cannot discuss every aspect of every decision. A CEO cannot learn every local condition known to a division manager; a school board cannot deliberate over every detail known to a union representative; and a policymaker cannot absorb every implication of an expert's analysis. Limited attention therefore forces people to communicate using summaries. But a summary does more than leave information out. It also groups the underlying facts into a small number of issues.

This observation has deep roots in organization theory. March and Simon (1958) use the term *uncertainty absorption* for communication in which an inference drawn

*We would like to thank Ricardo Alonso, Jeff Ely, Ben Golub, Daniel Gottlieb, Faruk Gul, Yoram Halevy, Alex Jakobsen, Emir Kamenica, Navin Kartik, Marina Halac, Michael Lipnowski, Alessandro Pavan, Wolfgang Pesendorfer, Phil Reny, Joel Sobel, and various seminar and conference audiences for extremely helpful feedback.

†website: https://elliotlipnowski.com, email: elliot.lipnowski@yale.edu
‡website: https://doronravid.com, email: doronr@umich.edu

from evidence is transmitted in place of the evidence itself. An information-processing view emphasizes the need to match communication capacity to organizational uncertainty and interdependence (Tushman and Nadler, 1978). Research on strategic issue diagnosis studies how data and stimuli are translated into focused issues that organize attention (Dutton et al., 1983), while the interpretation-systems and attention-based views emphasize how organizations give information meaning and direct decision makers toward particular "issues and answers" (Daft and Weick, 1984; Ocasio, 1997).

Together, these theories suggest that organizational summaries solve two problems at once: they economize on attention by limiting how much is discussed, and they organize attention by determining what the discussion is about. This distinction is especially pertinent for multidimensional decision problems, which rarely come with a unique list of issues. The same facts can be summarized using different aggregates, comparisons, or categories, each of which makes some aspects of the decision transparent and others obscure. This paper studies which divisions of the decision—that is, which issues—will be transmitted in an organization when communication is strategic.

For concreteness, consider a firm owner deciding how much money to allocate to the development of each of two products. The firm's manager knows this ideal pair, but the owner has the time and energy to understand only one figure. That figure could for instance describe the ideal budget for the first product, the ideal total budget for developing the two products, or the ideal difference between the two product budgets. Each report uses the same communication capacity. Nevertheless, each frames the owner's decision in a different way.

What figure should we expect the manager to communicate to the owner? Our analysis shows the answer depends on the manager's perspective. For a heuristic illustration, suppose the manager cares only about assigning the division the right total budget, while the owner cares mainly about the first product. Intuitively, if the owner were to use the manager's report to set the budget of the first product alone, the manager would inflate that report when the second product needs more money. In effect, she uses the first product's budget to partially correct the total. By contrast, if the owner uses the report only to set the division's total budget, the manager should have no reason to slant it: errors in either direction move the total away from its ideal level. Thus, it is the manager's perspective on the decision that determines which issues can be credibly communicated.

Our paper formalizes this logic within a model of strategic communication with bounded rationality. A receiver consults an informed sender about a multidimensional action. Both players would like the action to match an unknown state perfectly. Although they perfectly agree on the ideal action in each state, they have different priorities, and hence evaluate imperfectly tailored actions differently. For example, both the owner and the manager agree on the ideal budget for each product, but one cares more about getting a particular product's budget right while the other cares more about setting the correct the total. Finally, the receiver is boundedly rational: he must choose a low-dimensional interpretation of the sender's message.

Formally, the state and action are vectors $\boldsymbol{\omega}, \mathbf{a} \in \mathbb{R}^n$. We normalize the state's mean to zero and assume it has finite variance. The sender and receiver have quadratic losses,

$$(\mathbf{a}-\boldsymbol{\omega})^\top S(\mathbf{a}-\boldsymbol{\omega}) \qquad \text{and} \qquad (\mathbf{a}-\boldsymbol{\omega})^\top R(\mathbf{a}-\boldsymbol{\omega}),$$

where the positive-definite matrices S and R describe their respective priorities. The sender observes $\boldsymbol{\omega}$ and sends a message $\mathbf{m} \in \mathbb{R}^k$, where $k < n$. The receiver interprets messages through a full-rank linear decision rule D, taking action $\mathbf{a} = D\mathbf{m}$. The k columns of D are the issues the receiver understands the sender to be discussing. So k measures the capacity of the communication channel, while the range of D describes its substantive orientation. Thus, the receiver can choose how to interpret the sender, but he must settle on a k-dimensional way of doing so.

We show equilibrium communication is shaped by the sender's *perspective*. A perspective is a basis that decomposes the decision into statistically unrelated issues. A sender perspective is one for which she finds issues also unrelated in a payoff sense: in these coordinates, her loss is a weighted sum of squared errors, with one weight for each issue and no cross-issue spillovers. Every selection of k issues from a sender perspective can be communicated in equilibrium. In such an equilibrium, the sender reports the realized values of those issues, whereas the receiver takes her report on them at face value and takes the prior-optimal action on the remaining issues. Conversely, every equilibrium outcome can be represented in this way. Hence, equilibrium communication is shaped by the sender's priorities and the underlying state distribution, but not by the receiver's priorities.

This result follows from the players' optimality conditions. If the receiver's action space mixes dimensions to which the sender assigns different weights, the sender wants to tilt her report toward the dimensions she considers more important. Truthful communication is sustainable only when the communicated and uncommunicated components are separable from the sender's point of view. Meanwhile, the receiver faces a signal-extraction problem with the sender's equilibrium report as the signal. The receiver's incentives therefore require the communicated and residual components to be statistically unrelated, while the sender's incentives determine which decompositions are credible.

Let us return to the owner and manager, and consider the implications of our result when the two products' ideal budgets are independent and identically distributed. Suppose the manager evaluates errors in the total budget more heavily than errors in the difference between the product budgets, and cares about these two issues separably. Her perspective then consists of exactly these two issues: the total budget and the difference. If the owner can contemplate only one issue, some equilibrium exists in which the manager communicates the ideal total, and another exists in which she communicates the ideal difference. In contrast, no equilibrium has the manager simply reporting the first product's ideal budget, even though the owner cares more about the first product than the second.

Would the owner be better off with a manager whose perspective is more similar to his own? We show the answer depends on whether the manager expects his favorite or least favorite equilibrium. For a heuristic demonstration, suppose the owner places weights 98 and 2 on matching the ideal budget of the first and second product, respectively. Because either issue in the manager's perspective mixes the two products symmetrically, the owner obtains a payoff of 50 in either equilibrium. But what if the owner instead consults a manager who shares his perspective? Such a manager could communicate either product's ideal budget, giving the owner a payoff of 98 in one equilibrium and 2 in the other.

We show the above paragraph demonstrates a more general phenomena: a sender who is better for the receiver than another under best-equilibrium selection tends to be worse under adversarial selection. The broad intuition is that improving the receiver's best-case payoff requires the sender's perspective to better separate the receiver's high- and low-priority concerns. Such a separation stretches the set of equilibrium payoffs: it makes especially useful conversations possible while also making it possible to focus on minutiae.

We formalize this idea by defining a synchronization order between sender perspectives, and establishing for every given receiver perspective a three way equivalence between: (1) one sender being more synchronized with that perspective than another sender; (2) one sender giving a higher best-equilibrium payoff than another for all receivers with that perspective; and (3) one sender giving a lower worst-equilibrium payoff than another to all receivers with that perspective. Thus, while seeing eye-to-eye is better when the players talk about high-priority issues, having different worldviews provides insurance against coordinating on low-priority topics.

Finally, we allow the receiver to choose how much attention to devote to the sender. Under favorable equilibrium selection, the receiver learns about his favorite k sender issues. Attention therefore has decreasing marginal returns, and an interior capacity can be optimal. Under adversarial selection, he expects to hear about his least valuable sender issues. Thus, attention instead has increasing marginal returns, so the receiver either listens fully or not at all. Moreover, the average value of an issue is invariant to the sender's perspective in this case. Consequently, the receiver's attention choice under adversarial selection does not depend on whom he consults.

Related literature. We build on the cheap-talk framework of Crawford and Sobel (1982) and Green and Stokey (2007). In multidimensional cheap talk, equilibrium often determines the directions along which information is conveyed. Multiple experts gain influence over endogenous directions of local agreement in Battaglini (2002), while Levy and Razin (2007) show how cross-issue spillovers can obstruct that logic for a single sender. Chakraborty and Harbaugh (2007) study comparative statements across ordered issues, and Chakraborty and Harbaugh (2010) obtain endogenous directions that balance a sender's exaggeration incentives across dimensions. Most directly related, Bizzotto and Hancart (2026) study common-interest communication in which a two-dimensional state is summarized by a one-dimensional score. With a Gaussian prior, their equilibrium linear scores are projections onto the eigenvectors of the covariance-adjusted common loss matrix—equivalently given their bivariate setting, the ex ante best and worst linear scores. We instead characterize all rank-k equilibrium action rules under differing priorities: the sender's loss matrix shapes the equilibrium set, whereas the receiver's does not.

Our dimensional restriction also differs from other technological constraints on communication. With common interest, Jäger et al. (2011) and Rodriguez (2026) study finite languages in multidimensional environments. With sender commitment and possible misalignment, Le Treust and Tomala (2019) bound information-theoretic capacity, while Aybas and Turkel (2026) fix the number of messages. Work on limited verification instead specifies which primitive facts can be checked (Glazer and Rubinstein, 2004; Carroll and Egorov, 2019; Antić and Chakraborty, 2025; Ball and Gao, 2026). In our model, each message coordinate can carry arbitrarily fine information; the restriction is on the dimension of the receiver's continuous response, and equilibrium determines which linear combinations become issues.

Our low-dimensional assumption on receiver behavior connects as well to rational inattention, decision simplification, and organizational design. Most closely within this strand, Hu (2020) studies communication of correlated multidimensional information to a rationally inattentive agent when the players agree on the appropriate state-contingent actions but differ in the relative importance they assign to different dimensions. There, the principal controls the information provided to the agent; here, communication is cheap talk, and reporting incentives determine which linear combinations of the state can constitute separate issues. More generally, information can be suppressed to redirect attention (Lipnowski et al., 2020), and persuasion changes when receivers may disregard costly information (Bloedel and Segal, 2021). Scarce attention can focus coordination on a few tasks (Dessein et al., 2016), while sparse attention and mental budgeting simplify individual choice (Gabaix, 2014; Kőszegi and Matějka, 2020). Related organizational work studies communication networks, knowledge hierarchies, task representations, technical languages, codes, and modular organization (Bolton and Dewatripont, 1994; Garicano, 2000; Cronin and Weingart, 2007; Cremer et al., 2007; Wernerfelt, 2004; Koçak and Puranam, 2022; Matouschek et al., 2025). Our players already share labels and coordinates; the friction is that only some groupings can be treated as separate issues without introducing misreporting incentives.

Finally, our paper relates to work on how conventional or natural meanings discipline strategic communication (Farrell, 1993; Rabin, 1994; Reny, 2025; Blume, 2026). Whereas this literature begins with some pre-existing language and studies how its meanings constrain its strategic use and interpretation, our message coordinates have no exogenous substantive labels. The receiver's linear response gives each coordinate an action interpretation, whereas equilibrium determines the state summary statistic it conveys; thus, the issues that message coordinates come to represent are themselves equilibrium objects.

2 Endogenous issues

A receiver (R, he) consults with a sender (S, she) about what action $a \in \mathbb{R}^n$ he should take. R's optimal action is summarized by a stochastic state $\boldsymbol{\omega} \in \mathbb{R}^n$. S agrees it is best that R's action matches $\boldsymbol{\omega}$ as closely as possible. However, the two players disagree about how they should trade off the errors in the different dimensions, and so may differ in the way they rank suboptimal actions. Specifically, S and R suffer a loss that is quadratic in the difference between a and $\boldsymbol{\omega}$, with each player's loss having different weights—in other words, the respective payoffs of S and R are

$$\mathbf{u}_S = \boldsymbol{\omega}^\top S\boldsymbol{\omega} - (\boldsymbol{\omega} - \mathbf{a})^\top S(\boldsymbol{\omega} - \mathbf{a}) \quad \text{and} \quad \mathbf{u}_R = \boldsymbol{\omega}^\top R\boldsymbol{\omega} - (\boldsymbol{\omega} - \mathbf{a})^\top R(\boldsymbol{\omega} - \mathbf{a}),$$

where $S, R \in \mathbb{R}^{n\times n}$ are positive definite, and $\mathbf{a}$ is the (endogenous) random variable representing R's action. Note the above payoffs include the strategically irrelevant terms $\boldsymbol{\omega}^\top S\boldsymbol{\omega}$ and $\boldsymbol{\omega}^\top R\boldsymbol{\omega}$ (normalizing the payoff from action zero to zero) to simplify payoff expressions.

The distribution of $\boldsymbol{\omega}$ has finite mean and variance. Essentially without further loss of generality, we take $\boldsymbol{\omega}$ to have a full-rank covariance matrix V (otherwise, one can restrict attention to a lower dimensional subspace), and zero mean (see Section 8).

We view each action a as representing a myriad of activities that can be decomposed and combined in various ways formalized via vector-space operations. Thus, each action can be separate into n different components: given a basis $\{b_1, \ldots, b_n\}$, one can uniquely decompose every action $a \in \mathbb{R}^n$ into a sum $a = \sum_i \beta_i b_i$ for some collection of scalars $\beta_i \in \mathbb{R}$. We interpret a basis as a way of dividing R's action into separate issues, where $a = \sum_i \beta_i b_i$ means R takes action β_i on issue b_i.

The key novelty of our model is the assumption that players can only communicate about a small number $k \in \{0, \ldots, n\}$ of endogenously chosen issues. To formalize this restriction, we use the strategic form of a standard cheap talk game (Crawford and Sobel, 1982; Green and Stokey, 2007) as a baseline. Thus, S sees $\boldsymbol{\omega}$ and chooses a message m to send to R, who then takes an action. However, we modify this strategic form in two ways. First, S's message consists of k real numbers, meaning $m \in \mathbb{R}^k$. And second—and more importantly—we restrict R's strategy space: instead of allowing R to use all maps from messages to actions, we restrict R to using a map that is *linear* and *full-rank*. Consequently, R's strategy is describable via a full-rank matrix, $D \in \mathbb{R}^{n\times k}$. Hence, if S sends message m and R uses strategy D, then R's action equals $a = Dm$.

R's strategy space expresses a type of bounded rationality. In particular, R must interpret S's messages as a recommendation as to what to do on k issues. However, R can choose which k issues he interprets S as actually discussing. Thus, by choosing a full-rank matrix $D = [d_1, \ldots, d_k]$, R is deciding to interpret every S message m as telling him to take action m_i on issue d_i, resulting in the aggregate action $a = Dm = \sum_{i=1}^{k} m_i d_i$.

Hence, a strategy profile is a pair, (C, D), where $C : \mathbb{R}^n \to \mathbb{R}^k$ describes S's chosen message as a function of the state, and $D \in \mathbb{R}^{n\times k}$ has full rank. Let $\mathcal{D}$ be the set of all full-rank $n \times k$ real-matrices. A profile (C, D) is a pure-strategy **Bayes-Nash equilibrium (BNE)** if and only if two conditions hold:[1]

1. S's message is optimal given $\boldsymbol{\omega}$:

$$C(\omega) \in \operatorname{argmin}_{m\in\mathbb{R}^k} \left[(Dm - \omega)^\top S(Dm - \omega)\right] \quad \forall\, \omega \in \mathbb{R}^n, \tag{S-BR}$$

2. R's strategy is optimal:[2]

$$D \in \operatorname{argmin}_{D\in\mathcal{D}} \mathbb{E}\left\{[DC(\boldsymbol{\omega}) - \boldsymbol{\omega}]^\top R\,[DC(\boldsymbol{\omega}) - \boldsymbol{\omega}]\right\}. \tag{R-BR}$$

In what follows, we use the term **equilibrium** as a short-hand for pure-strategy BNE.

Our model shares some limitations with the standard cheap talk game. In particular, R must choose how to interpret S's messages before they are realized. Consequently, S has no way of alerting R that a certain issue is more urgent than the others. In other words, S cannot send R a message that changes the issues he expects S to talk about. Instead, the issues S talks about are determined endogenously through R's expectations and the players' equilibrium behaviors.

One limitation of the usual cheap talk game that we avoid is the possibility of babbling. The reason is that we assume R uses the entire capacity of his communication channel—i.e., that R's linear map must have full rank. Under this assumption, when $k \geq 1$, R's action is always responsive to S's message and some information is

[1] One can show the restriction to pure strategies is without loss.

[2] Whenever S's incentive constraint is satisfied, it follows from our analysis that this R expected value is well defined for any strategy he chooses.

always communicated in equilibrium.[3]

Thus, the only way to get babbling in our model is to have $k = 0$. In this case, the players cannot do anything but babble, and R always chooses 0. At the opposite extreme, when $k = n$, equilibrium communication is perfect: every R strategy has full range, and so S can always induce R to match the state perfectly. We therefore restrict the capacity k to be strictly between 0 and n in what follows, except for Section 7, where we endogenize k.

Before proceeding, we note there are two other game forms whose solutions coincide with the equilibria we analyze in the sequel. The first equivalent model is a standard cheap talk model with the usual strategy space for R, but where we restrict the set of allowable equilibria. Specifically, R is free to choose his action after seeing S's message in $\mathbb{R}^k$. However, we consider only pure-strategy perfect Bayesian equilibria in which R's strategy happens to be a full-rank linear map. For such equilibria to exist and coincide with the equilibria of our model, one must impose the additional assumption that $\boldsymbol{\omega}$ has an elliptical distribution (e.g., Gaussian).[4]

The second equivalent game form is one in which both agents choose linear maps. Here, S and R move simultaneously without any knowledge of the state's realization, with S choosing a full-rank matrix $C \in \mathbb{R}^{k \times n}$ and R choosing a full-rank matrix $D \in \mathbb{R}^{n \times k}$, resulting in action $\mathbf{a} := DC\boldsymbol{\omega}$. The solution concept is pure-strategy Nash equilibrium.

In both of the above formulations, the characterization of equilibrium action rules (and hence payoffs) is exactly the same as in our main model. We therefore welcome a reader to interpret our model in any of these three equivalent ways.

3 An illustrative example

This section illustrates the key tensions in our model using a simple example. To give the example some color, suppose R is a school board making decisions on two dimensions: the school's curriculum and teacher staffing. The head of the teacher's union, S, knows the optimal decision on both of these dimensions. To simplify the analysis, suppose the pair of optimal decision on these two dimensions is Gaussian,

[3] We discuss the full-rank assumption in greater depth in Section 8.

[4] One may be interested in requiring also that S's equilibrium strategy is a full-rank linear map. This property turns out to hold automatically given the structure of R's strategy.

with each dimension having a mean of 0 and a variance of 1. More importantly, the two dimensions are positively correlated, with correlation $\rho \in (0, 1)$. Note that we do not require $\boldsymbol{\omega}$ to be Gaussian in the general model.

We assume R cares more about choosing the correct curriculum, while S cares more about teacher staffing. More concretely,

$$S = \begin{bmatrix} \epsilon & 0 \\ 0 & 100 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 98 & 0 \\ 0 & 2 \end{bmatrix},$$

for some $\epsilon > 0$. That is, S and R's payoffs are respectively given by

$$\mathbf{u}_S = \boldsymbol{\omega}^\top S \boldsymbol{\omega} - \epsilon(\mathbf{a}_1 - \boldsymbol{\omega}_1)^2 - 100(\mathbf{a}_2 - \boldsymbol{\omega}_2)^2 \quad \text{and} \quad \mathbf{u}_R = \boldsymbol{\omega}^\top R \boldsymbol{\omega} - 98(\mathbf{a}_1 - \boldsymbol{\omega}_1)^2 - 2(\mathbf{a}_2 - \boldsymbol{\omega}_2)^2,$$

where dimension 1 represents the choice of curriculum (e.g., the relative emphasis on sciences vs. the arts), and dimension 2 represents changes in teacher staffing. We caricature this situation by considering the case where $\epsilon \approx 0$, and describe the equilibria in the limit as ϵ vanishes.[5]

Due to various attention and time constraints, R can consult with S about only one issue, that is $k = 1$. Our main interest is in understanding which issue S ends up communicating in equilibrium.

If S could have her way, she would use R's attention to ensure he takes the optimal staffing decision. One can create such an equilibrium by having S tell R the value of that dimension, $\mathbf{m} = \boldsymbol{\omega}_2$. For an explanation, note that having seen such a message, R is clearly best choosing the correct teacher staffing, $\mathbf{a}_2 = \mathbf{m} = \boldsymbol{\omega}_2$. Moreover, the correlation between the two dimensions enables R to make inferences about the optimal curriculum. In particular, the example's Gaussian assumption means $\mathbb{E}[\boldsymbol{\omega}_1 | \mathbf{m}] = \rho \boldsymbol{\omega}_2$. Therefore, by setting $D = [\rho, 1]^\top$, R can ensure his optimal action matches the conditional expectation of $\boldsymbol{\omega}$ given $\mathbf{m}$—which is optimal. R's payoff is $100 - 98(1 - \rho^2)$, while S gets her first-best payoff of 100.

Unlike the optimal teacher staffing, S cannot credibly tell R the optimal curriculum—that is, $\mathbf{m} = \boldsymbol{\omega}_1$ cannot occur in equilibrium. To see why, observe that a symmetric argument to the one above implies that R's best reply to $\mathbf{m} = \boldsymbol{\omega}_1$ is $D = [1, \rho]^\top$. Hence, $\mathbf{a}_1 = \mathbf{m}$, and $\mathbf{a}_2 = \rho \mathbf{m}$. Given this R strategy, S's best reply is not to send

[5] More precisely, we will characterize the maps from state to action that arise as the pointwise limit of equilibria as $\epsilon \searrow 0$.

$\mathbf{m} = \boldsymbol{\omega}_1$. Rather, she prefers to send $\mathbf{m} = \boldsymbol{\omega}_2/\rho$: R would then take $\mathbf{a}_2 = \boldsymbol{\omega}_2$, giving S her first-best payoff. This deviation illustrates our environment's key communication friction: S cannot credibly communicate about issues that mix topics she prioritizes unevenly.

To have S communicate about the optimal curriculum, S must clean out the information it provides about teacher staffing. More concretely, suppose S sent the message $\mathbf{m} = \boldsymbol{\omega}_1 - \rho\boldsymbol{\omega}_2$. In this case, R's conditional expectations are given by $\mathbb{E}[\boldsymbol{\omega}_1|\mathbf{m}] = \mathbf{m}$, and $\mathbb{E}[\boldsymbol{\omega}_2|\mathbf{m}] = 0$, and so setting $D = [1, 0]^\top$ is optimal. Moreover, S cannot benefit from deviating: her utility is zero regardless of her message. R's payoff in this equilibrium is $98(1 - \rho^2)$.

These two equilibria demonstrate a more general phenomenon: S can communicate an issue to R in equilibrium if and only if that issue combines only dimensions to which S assigns the same value. Whenever an issue mixes topics that S values unevenly, S benefits from slanting communication in favor of the topics that matter more to her. Conversely, S never gains from miscommunicating an issue that influences R's actions on dimensions that concern S equally: such miscommunication can only lead R to take a wrong action on said issues, thereby reducing S's payoff. Moreover, if R expects S to honestly communicate an issue, then R should always take S's report at face value, making accurately communicating said issue self-enforcing.

The above intuition also suggests that R's priorities play no role in shaping equilibrium communication. Broadly speaking, the reason is that, once R knows S's communication is honest, R faces a simple signal extraction problem, the solution to which is the best linear estimator of the state given S's report. Since this estimator does not depend on R's preferences, neither does the equilibrium set. One can see this independence in the above example: the equilibria we derived remain even if we changed R's loss matrix, since that matrix has no impact on his best replies.[6]

In the sequel, we show that equilibrium communication is shaped by S's preferences and the state's variance-covariance matrix. In particular, equilibria are shaped by an *S*-perspective, which is a decomposition of R's actions into issues that S views as unrelated. The next section defines this object formally and uses it to characterize the model's equilibrium set.

[6]While the argument we presented in the example relied on the example's Gaussian assumption, we show this assumption is not necessary in general.

4 What can you talk about?

This section presents our equilibrium characterization. Our characterization shows the equilibrium is shaped by S's perspective—an object which we will define shortly. Intuitively speaking, an S-perspective provides a way of framing R's decision as n separate choices that S views as unrelated. This perspective depends only on the state's variance-covariance matrix V and S's loss matrix S. As a consequence, we show R's priorities are irrelevant in equilibrium—changing R's loss matrix has no impact on the equilibrium set.

We begin by defining a perspective. Given an $n \times n$ matrix $P = [p_1, \ldots, p_k]$ and an index i, we let $\boldsymbol{\omega}_i^P := p_i \cdot \boldsymbol{\omega}$ and $\mathbf{a}_i^P := p_i \cdot \mathbf{a}$ respectively be the (non-normalized) projection of $\boldsymbol{\omega}$ and $\mathbf{a}$ onto the vector p_i. We say that P is a **perspective** if $\boldsymbol{\omega}_i^P$ and $\boldsymbol{\omega}_j^P$ are uncorrelated for all $i \neq j$, and if $\boldsymbol{\omega}_i^P$ has unit variance for every i. Intuitively, a perspective decomposes R's decision problem into n statistically unrelated choices. Geometrically, P is a perspective if $(\sqrt{V} p_i)_{i=1}^n$ is an *orthonormal basis* for $\mathbb{R}^n$; that is, the vectors $(\sqrt{V} p_i)_{i=1}^n$ have unit length and are pairwise orthogonal (hence linearly independent).

A perspective P is an S**-perspective** if some weights $\sigma_1, \ldots, \sigma_n \in \mathbb{R}_{++}$ exist such that every $\omega, a \in \mathbb{R}^n$ have

$$(a - \omega)^\top S(a - \omega) = \sum_{i=1}^n \sigma_i (a_i^P - \omega_i^P)^2.$$

Thus, an S-perspective separates R's action into n decisions that S views as unrelated to each other—both statistically and in the sense of payoff separability. The weights $\sigma_1, \ldots \sigma_n$ designate how much S cares about each issue. The order and directionality of the issues is irrelevant: we can turn one S-perspective into a different one by permuting columns and multiplying some of them by -1. We refer to two S-perspectives that are related in this way as **issue equivalent**. If all S-perspectives are issue equivalent, we say the S-perspective is **unique**. Geometrically, an S-perspective is an eigendecomposition of $\sqrt{V} S \sqrt{V}$. As such, an S-perspective always exists, and is unique whenever all n weights are distinct—a generic property.[7] Moreover, these weights are unique up to the same permutation as the corresponding columns.

We now demonstrate these definitions in the context of another example.

[7]See Definition 3 for a formal definition of genericity, and Lemma 12 for a proof of genericity.

Example 1. Suppose R is a CEO of a company, and S is a division manager. S's division has two products. R's decision is how much money to dedicate to each product (net of some status quo), where $\mathbf{a}_i$ is the amount R allocates to product i. S knows the optimal allocation $(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$, which is stochastic and has a unit variance-covariance matrix $V = I$. Both S and R would like to give the right amount of resources to each product, but differ in their priorities, as expressed by the different loss matrices:

$$S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, \qquad R = \begin{bmatrix} 98 & 0 \\ 0 & 2 \end{bmatrix}.$$

The matrix representation makes R's priorities quite clear: R cares mostly about product 1 (e.g., because it is R's pet project). Indeed, writing R's utility more explicitly gives

$$\mathbf{u}_R = \boldsymbol{\omega}^\top R \boldsymbol{\omega} - 98(\mathbf{a}_1 - \boldsymbol{\omega}_1)^2 - 2(\mathbf{a}_2 - \boldsymbol{\omega}_2)^2.$$

By contrast, S's priorities are harder to discern from the matrix representation. Expanding the matrix term is not very helpful either:

$$\mathbf{u}_S = \boldsymbol{\omega}^\top S \boldsymbol{\omega} - 3(\mathbf{a}_1 - \boldsymbol{\omega}_1)^2 - 3(\mathbf{a}_2 - \boldsymbol{\omega}_2)^2 - 2(\mathbf{a}_1 - \boldsymbol{\omega}_1)(\mathbf{a}_2 - \boldsymbol{\omega}_2).$$

One can get a clearer picture of S's priorities, however, by finding her perspective. Using the above definition reveals that S has a unique perspective,

$$P = \tfrac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right], \quad \text{and} \quad (\sigma_1, \sigma_2) = (4, 2).$$

Accordingly, one can write S's loss as

$$\mathbf{u}_S = \boldsymbol{\omega}' S \boldsymbol{\omega} - 2\left[(\mathbf{a}_1 + \mathbf{a}_2) - (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2)\right]^2 - \left[(\mathbf{a}_1 - \mathbf{a}_2) - (\boldsymbol{\omega}_1 - \boldsymbol{\omega}_2)\right]^2.$$

These expressions reveal S's true priorities: S's cares more about the total budget for her division than she cares about the way that budget is allocated. □

To get a sense of why S's perspective is central for equilibrium, consider S's problem (S-BR) for a fixed state realization $\omega \in \mathbb{R}^n$. This problem has a strictly concave and differentiable objective, and so its solution is characterized by a first order condition. This condition describes a tangency between two objects. The first object is the range of R's strategy D, which is a k-dimensional linear subspace of $\mathbb{R}^n$. This

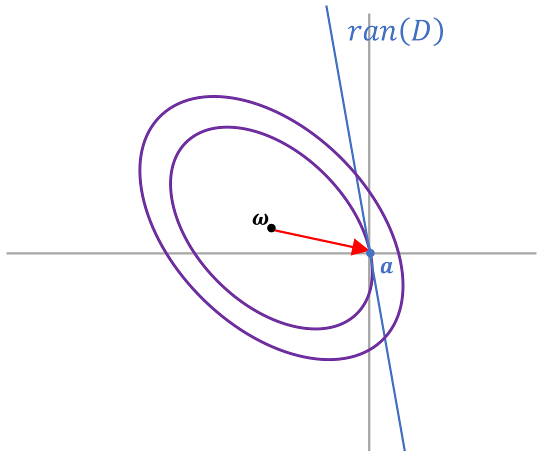


Figure 1: Geometric illustration of the tangency conditions characterizing S-optimality when $n = 2$ and $k = 1$.

range represents S's "budget" set: S can induce any action in this range by choosing an appropriate message. The second object is an S indifference curve. These curves are ellipsoids centered at ω. S's first-order condition says that inducing R to take a is optimal for S only if a is the tangency point between its indifference curve and the range of D (see Figure 1).

S's perspective characterizes an orthogonality between the linear space tangent to S's indifference curve at an action a and the indifference curve's radius vector that goes to a—i.e., the line connecting a to the realized state ω. When V is the identity, one can describe this orthogonality using the standard vector dot product. In particular, the radius vector to a is perpendicular to S's indifference curve's tangent subspace at a if and only if that subspace admits a subset of S's perspective as a basis. Figure 2 demonstrates this tangency in the context of Example 1.

Intuitively, the radius vector to a represents the unique non-compensable direction: no local movement of a would compensate S for a local push of a away from ω in the radius's direction. By contrast, the tangent hyperplane represents all the directions of indifference—that is, all directions to which one can perturb a without meaningfully changing S's losses. The previously mentioned orthogonality means that the indifference directions and the non-compensable direction are separable from each other both in a statistical sense and from the point of view of S's preferences.

As our next result shows, S's perspective shapes equilibrium communication. To present the result formally, say that a strategy profile (C, D) **communicates an**

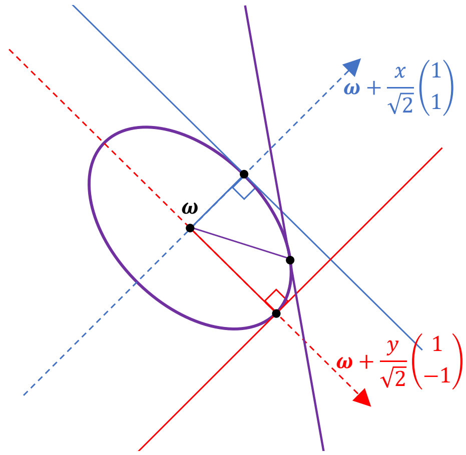


Figure 2: Illustration of the relationship between S's perspective and orthogonality in Example 1. Dashed lines represent shifts in the state in the directions of S's perspective. These directions coincide with the direction of the radius vector to a point on the indifference curve exactly when the vector is perpendicular to the tangent at the point.

S**-perspective** if an S-perspective P exists such that $\mathbf{m} = C(\boldsymbol{\omega})$ and $\mathbf{a} = D\mathbf{m}$ have[8]

$$\mathbf{m}_i = \boldsymbol{\omega}_i^P \quad \text{for } i = 1, \ldots, k,$$

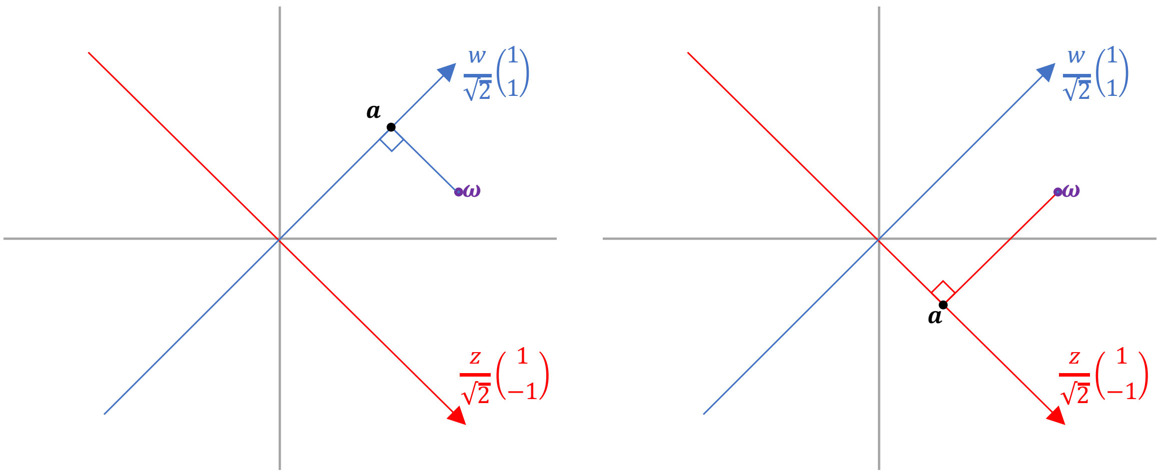
$$\mathbf{a}_i^P = \begin{cases} \mathbf{m}_i & \text{if } i = 1, \ldots, k, \\ 0 & \text{if } i = k+1, \ldots, n. \end{cases}$$

In words, a profile communicates an S-perspective P if S tells R what to do on the first k issues of P; R responds by doing as he is told on those k issues, and taking the (prior-optimal) zero action on the remaining issues. We use the convention that communication is restricted to the first k issues of P. Hence, two profiles that communicate issue-equivalent S-perspectives typically transmit different information.

As a demonstration, consider Example 1. In this example, there are essentially only two strategy profiles that communicate S's perspective. We illustrate these profiles in Figure 3. In the first profile, (C, D), the sender tells the receiver the size of $\boldsymbol{\omega}$'s projection onto $\frac{1}{\sqrt{2}}(1, 1)^\top$. The second profile $(\tilde{C}, \tilde{D})$ has S reporting the size of the projection of $\boldsymbol{\omega}$ onto $\frac{1}{\sqrt{2}}(1, -1)^\top$. These correspond to the respective S-strategies $C(\boldsymbol{\omega}) = \frac{1}{\sqrt{2}}(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2)$ and $\tilde{C}(\boldsymbol{\omega}) = \frac{1}{\sqrt{2}}(\boldsymbol{\omega}_1 - \boldsymbol{\omega}_2)$. In both profiles, R responds

[8]Every perspective admits some strategy profile that communicates it—see Appendix A for an explicit description.

by matching the size of the projection to the appropriate vector, and setting the projection onto the other vector to zero. More precisely, $D_1\mathbf{m} = D_2\mathbf{m} = \frac{1}{\sqrt{2}}\mathbf{m}$, and $\tilde{D}_1\mathbf{m} = -\tilde{D}_2\mathbf{m} = \frac{1}{\sqrt{2}}\mathbf{m}$. Thus, the first profile is equivalent to the division manager communicating the appropriate total budget for the division and the CEO splitting that budget evenly between the two products. The second profile is equivalent to the division manager telling the CEO the correct difference in the ideal budget for the division's two products, and the CEO splitting that difference evenly.



(a) S dictates the total budget, and R splits it equally between the two products

(b) S dictates the relative budget, and R uses the ex-ante correct total budget.

Figure 3: The two equilibrium state-contingent action choices for Example 1.

Proposition 1 says the game's equilibria are characterized by the profiles that communicate S's perspective. To state this result, say that two strategy profiles are **equivalent** if they induce the same joint distribution for $\boldsymbol{\omega}$ and $\mathbf{a}$.

Proposition 1.

(i) Any strategy profile that communicates an S-perspective is an equilibrium.

(ii) Any equilibrium is equivalent to one that communicates an S-perspective.

The proposition highlights how equilibrium communication is shaped by two features of the environment: S's priorities and the state's stochastic structure. Intuitively, S's incentives requires S to view the issues on which R acts as separate from the issues about which he does not act—otherwise, S would distort her messages so as to better tailor R's actions on issues she assigns a higher weight. Meanwhile, R chooses his issue interpretation to make S's communication maximally instrumental; hence, the issues on which R acts are uncorrelated with the issues he leaves idle.

For a more detailed overview of the proposition's logic, consider again S's problem given an arbitrary state realization $\omega \in \mathbb{R}^n$. As explained earlier, the problem's solution is characterized by tangency between S's indifference curve and the range of R's strategy. This tangency is defined by a linear equation, and so implies S's best response is a linear function of the state's realization ω. Geometrically, this linearity is most transparent when $k = n - 1$. In that case, one can decompose any change in the state's realization ω into two linear operations. The first operation is a translation in a direction parallel to the range of D. Such a translation results in an identical translation of the resulting tangency point. The second operation is a shift toward or away from the range of D in the direction of the radius vector of S's indifference curve—a shift that leaves the resulting tangency point unchanged.

Consider now R's best reply. For intuition, suppose the states are uncorrelated and have unit variance, $V = I$. In equilibrium, R expects S's messaging strategy C to be a linear map. One can therefore decompose the state into two parts: its projection onto the map's null space $\boldsymbol{\omega}^N$, and its projection $\boldsymbol{\omega}^R$ onto the null's orthogonal complement, also known as the map's row space. Moreover, S's strategy is invertible when restricted to this row space. It turns out that R's best reply inverts this map, matching $\boldsymbol{\omega}^R$ exactly and estimating $\boldsymbol{\omega}^N$ using its best linear predictor given $\boldsymbol{\omega}^R$. Intuitively, R's action is uncorrelated with $\boldsymbol{\omega}^N$ no matter what strategy he chooses, and hence perfectly matching $\boldsymbol{\omega}^R$ comes at no opportunity cost.[9]

Geometrically, R optimality can be described via an orthogonality condition between the range of R's strategy and the line connecting each state realization to its associated action. When $V = I$, we can express this orthogonality via the usual dot product: the line connecting the state to the action must be perpendicular to the range of R's strategy. In fact, in this case, since the best linear predictor of $\boldsymbol{\omega}^N$ given $\boldsymbol{\omega}^R$ is zero, the range of R's strategy and the row space of S's message strategy must coincide. Figure 4 illustrates this relationship.

The proposition follows from combining the above-mentioned geometric facts. By R incentives, the line connecting a state realization ω to the resulting action a must be orthogonal to R's action space. This space, in turn, needs to be tangent to the

[9] When $V \neq I$, an analogous decomposition applies, using orthogonality with respect to a prior-normalized coordinate system.

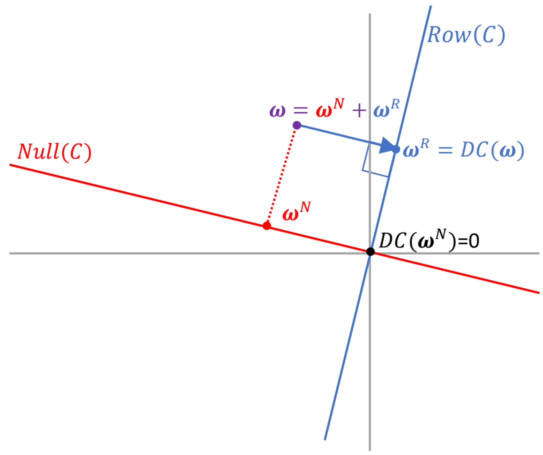


Figure 4: The orthogonality condition characterizing R optimality for $n = 2$ and $k = 1$, specialized to the case of $V = I$. The figure illustrates the decomposition of $\boldsymbol{\omega}$ into an element $\boldsymbol{\omega}^N$ of the null space of C (denoted by $Null(C)$) and an element $\boldsymbol{\omega}^R$ of the row space of C (denoted by $Row(C)$). In this setting, R optimality is equivalent to requiring the line connecting $\boldsymbol{\omega}$ to $\boldsymbol{\omega}^R = DC(\boldsymbol{\omega})$ to be perpendicular to the range of D. Consequently, the row space of C and the range of D coincide.

S-indifference curve going through a. All that remains is to recall the fact mentioned above: a radius vector of an S-indifference curve is orthogonal to a tangent linear subspace if and only if that space admits a subset of an S-perspective as a basis. Replicating equilibrium communication using this basis delivers an equivalent profile that communicates S's perspective.

5 What do you want to talk about?

By Proposition 1, communication must decompose R's decision according to an S-perspective. Subject to this decomposition, however, S can communicate to R about any collection of issues. This section asks: which issues would each player want to talk about?

For a more formal exposition of this section's question, note that an S-perspective P specifies two things. First, the columns of P specify a particular way of framing R's decision; that is, decomposing it into separate issues. Second, the order of the columns specifies which issues are discussed in a P-equilibrium. Our interest in this section is in the second component; that is, which of the issues in P would R like S to communicate, and which would S prefer to communicate to R?

To answer our question, we calculate the players' payoffs from a given equilibrium. To present the calculation's outcome, let Q be an R-perspective with associated weights $\rho_1, \dots, \rho_n$, and P be an S-perspective with associated weights $\sigma_1, \dots, \sigma_n$. For $i, j = 1, \dots n$, define the **coefficient of determination** (also known as the "r^2") between $\boldsymbol{\omega}_i^P$ and $\boldsymbol{\omega}_j^Q$ as

$$\Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q) := \frac{\left(\mathbb{E}[\boldsymbol{\omega}_i^P \boldsymbol{\omega}_j^Q]\right)^2}{\mathbb{V}(\boldsymbol{\omega}_i^P)\mathbb{V}(\boldsymbol{\omega}_j^Q)} = \left(\mathbb{E}[\boldsymbol{\omega}_i^P \boldsymbol{\omega}_j^Q]\right)^2.$$

In words, $\Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q)$ is the percentage of $\boldsymbol{\omega}_j^Q$'s variance one can capture if one were to use the best linear estimate of $\boldsymbol{\omega}_j^Q$ given $\boldsymbol{\omega}_i^P$ (or vice versa). We interpret $\Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q)$ as a measure of informativeness: the higher $\Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q)$ is, the more information issue p_i provides about issue q_j (and vice versa).

Proposition 2. *Fix an S-perspective P with weights $\sigma_1, \dots, \sigma_n$, and an R-perspective Q with weights $\rho_1, \dots, \rho_n$. S and R's payoffs in any equilibrium that communicates P are, respectively,*

$$\mathbb{E}[\mathbf{u}_S] = \sum_{i=1}^{k} \sigma_i \quad \textit{and} \quad \mathbb{E}[\mathbf{u}_R] = \sum_{i=1}^{k} \sum_{j=1}^{n} \rho_j \Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q).$$

The result highlights the way each player evaluates the different issues. S's valuations are straightforward: the benefit to S from communicating issue p_i to R is σ_i. Thus, the weight associated with issue in P gives that issue's value to S.

R's evaluation of different issues is more nuanced. If S communicates to R the issue p_i, R obtains a value of

$$v_i := \sum_{j=1}^{n} \rho_j \Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q). \tag{1}$$

In other words, R prefers S to communicate issues that are provide more information about issues R views as more important. This preference is easiest to see when $P = Q$: in this case, R's benefit from learning issue $p_j = q_j$ is just ρ_j, and so R would like to order P according to his weights.

Next, we illustrate Proposition 2 in the context of Example 1.

Example 1 (continued)**.** Recall this example has $V = I$ and loss matrices

$$S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \text{ and } R = \begin{bmatrix} 98 & 0 \\ 0 & 2 \end{bmatrix}.$$

The corresponding perspectives are

$$P = \tfrac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right], (\sigma_1, \sigma_2) = (4, 2) \quad \text{and} \quad Q = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right], (\rho_1, \rho_2) = (98, 2).$$

By Proposition 1, this example admits at most two equilibria: one which communicates the issue $p_1 = (1, 1)^\top$, and the other that communicates $p_2 = (-1, 1)^\top$. If S communicates the first issue, she can guarantees that R will allocate the correct total budget for her division, resulting in a payoff of 4. This payoff is larger than the payoff of 2 that S would get if she were to communicate the second issue—i.e., the optimal difference in budget allocation between the two products.

It turns out that R does not care which of these two equilibria is played. To see this identity, note the coefficient of determination between the issues in P and in Q are $\Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q) = 0.5$ for any i and j. Therefore, R's payoff is equal to $0.5(2) + 0.5(98) = 50$ in both equilibria.

What if instead R interacted with a sender with loss matrix R? In this case, S and R would have the same perspective, and so every equilibrium is equivalent to S communicating either the issue $q_1 = (1, 0)^\top$ or the issue $q_2 = (0, 1)^\top$. S's and R's payoffs from each of these equilibria are the same: they both get a payoff of $\rho_1 = 98$ from the first equilibrium, and $\rho_2 = 2$ from the second. □

6 Who do you want to talk to?

This section explores the implications of our equilibrium characterization for the question: who would you like to talk to? The answer is clearly relevant when one can actually choose one's conversation partner. But the answer can also be relevant for organizational design; by changing the design of an organization, one can impact the priorities of its members. For an illustration, consider the CEO in Example 1. In addition to replacing the manager, the CEO can change the division manager's priorities (i.e., S) in many ways: he can change what products are under the manager's authority, restructure the manager's contract, change the way the division interacts with other parts of the organization, etc.

The results of this section highlight that while S does not care about R, R's payoff vary with S in a rich way. As previously noted, R evaluates a given S according to the correlation between S's perspective and his own. Equilibrium selection turns out to play a critical role: whether R wants more or less correlation depends on whether equilibrium selection is favorable or adversarial. Consequently, the way R ranks senders depends on which equilibrium R expects to be played. In particular, we show there is a sense in which an S that is good under R-favorable equilibrium selection is often bad under R-adversarial equilibrium selection. A rough intuition is that an S whose perspective is "synchronized" with R's has issues that are highly correlated with R's most important issues and other issues that are highly correlated with R's least important issues.

The asymmetry between S and R is driven by the forces that shape equilibrium communication. In particular, R's preferences are irrelevant in equilibrium; R simply attempts to match the state as best as he can given his information. Therefore:

Corollary 1. *S's equilibrium payoff set does not depend on R.*

Whereas S's equilibrium payoff set is independent of R's loss matrix, R's equilibrium payoff set clearly depends on S—specifically, S's perspective. However, as previously mentioned, R's ranking of different S depends on the equilibrium selection criterion. Indeed, we have already observed an instance of this dependency in Example 1: the CEO would like to replace the division manager with a manager who shares his priorities only if he expects his favorite equilibrium to be played. By contrast, if the CEO expects his worst equilibrium to be played, he prefers the original manager.

Our next result outlines the way in which R's best and worst equilibrium payoffs vary with S. To state this result, given an S, take $\bar{u}_R^k(S,R)$ and $\underline{u}_R^k(S,R)$ respectively to be R's highest and lowest equilibrium payoffs (which our analysis shows exist) when his loss matrix is R and S's matrix is S. Letting $(\rho_i)_{i=1}^n$ be R's weights associated with some R-perspective, take r_L to be the sum of the k lowest weights, r_H be the sum of the k highest weights, and r_M be k times the average weight.

Proposition 3. *For a pair $(\bar{r}, \underline{r}) \in \mathbb{R}^2$, the following are equivalent.*

(i) Some S has $\bar{u}_R^k(S,R) = \bar{r}$ and $\underline{u}_R^k(S,R) = \underline{r}$.

(ii) $r_L \leq \underline{r} \leq r_M \leq \bar{r} \leq r_H$ *and* $r_M = (1-\beta)\underline{r} + \beta\bar{r}$ *for some* $\beta \in \operatorname{co}\left\{\frac{k}{n},\ 1-\frac{k}{n}\right\}$.

The proposition shows that an S with an R-best equilibrium payoff of $\bar{r}$ and a R-worst equilibrium payoff of $\underline{r}$ exist if and only if $\bar{r}$ and $\underline{r}$ can be expressed as a moderate binary splitting of r_M. This moderation is expressed both in terms of its support—$\underline{r}$ and $\bar{r}$ must be between r_L and r_H— and in terms of each point's mass: neither $\bar{r}$ and $\underline{r}$ can have a mass outside $\operatorname{co}\left\{\frac{k}{n},\ 1-\frac{k}{n}\right\}$, which is an interval centered around $\frac{1}{2}$.

Underlying the proposition are two related constraints on the amount of information (as measured by $\Re^2$) the issues in S's perspective P can provide about issues in R's perspective Q, and vice versa. The first constraint is about the amount of information one can learn about Q by observing a single issue from P. In particular, the more information p_i provides about q_j, the less information p_i must provide about other issues in Q. The reason is that knowing $(\boldsymbol{\omega}^Q_m)^n_{m=1}$ is equivalent to knowing the state, and so is sufficient for knowing $\boldsymbol{\omega}^P_i$ as well. Therefore,

$$\sum_j \Re^2(\boldsymbol{\omega}^P_i, \boldsymbol{\omega}^Q_j) = 1. \tag{2}$$

Hence, the higher is $\Re^2(\boldsymbol{\omega}^P_i, \boldsymbol{\omega}^Q_j)$, the lower must be $\sum_{m\neq j} \Re^2(\boldsymbol{\omega}^P_i, \boldsymbol{\omega}^Q_m)$.

The second constraint is about the total amount of information the issues in P provide about a single issue from Q. This constraint is a mirror image of the previous one: because knowing the variables $(\boldsymbol{\omega}^P_m)^n_{m=1}$ is sufficient for knowing the state, knowing these variables is also sufficient for determining $\boldsymbol{\omega}^Q_j$. Therefore,

$$\sum_i \Re^2(\boldsymbol{\omega}^P_i, \boldsymbol{\omega}^Q_j) = 1. \tag{3}$$

An implication of this fact is that the less informative a particular issue p_i is about q_j, the more informative the other issues in P must be about q_j.

A pertinent consequence of the above constraints is that the values R obtains from learning each issue in P—represented by the vector $v = (v_1, \ldots, v_n)^\top$— must be "less spread out" than $\rho = (\rho_1, \ldots, \rho_n)^\top$ in the sense of *majorization.* Formally, given two vectors, $x, y \in \mathbb{R}^n$, we say x **majorizes** y if $y = Dx$ for a bistochastic $n \times n$ matrix D. In particular, this means that every y_i is a weighted average of elements of x.

To connect this definition back to the information constraint, define the **similarity matrix** $\Re^2(P, Q)$ of P and Q to be the matrix whose ij-entry is the coefficient of determination between $\boldsymbol{\omega}_i^P$ and $\boldsymbol{\omega}_j^Q$—that is, $\Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q)$. By the information constraints, (2) and (3), this matrix is bistochastic. Since $v = \Re^2(P, Q)\rho$ (see equation (1)), it follows the vector ρ majorizes v.

There are at least two other equivalent definitions of majorization (see Marshall et al., 2011). The first definition says that x majorizes y if for every $m \in \{1, \ldots, n\}$ the sum of the m largest entries of x is larger than the sum of the m largest entries of y, and the sum of all n entries of both vectors is the same. The second definition is the mirror image of the first: in addition to all n entries of both vectors having an equal sum, the sum of the m smallest entries of x must be smaller than the sum of the smallest m entries of y for every $m \in \{1, \ldots, n\}$.

The fact that ρ majorizes v is the reason Proposition 2's part (i) implies part (ii). For an explanation, note that since ρ majorizes v, the sum of v's least k entries and the sum of v's largest k entries must be less extreme than the corresponding sums for ρ. In addition, the average of v's entries coincides with the average of ρ's, and must also be between the average of v's largest and smallest k entries. This latter observation also delivers that one can think of of $\underline{r}$ and $\bar{r}$ as a splitting of r_M. The fact that the weights associated with this splitting cannot be too extreme follows from noting that r_M is k times the average of n numbers, with the lowest k summing up to $\underline{r}$, and the highest k summing to $\bar{r}$.

To show the converse—namely that Proposition 2's part (ii) implies part (i)—we use condition (ii) to find some v that is majorized by ρ whose highest and lowest k entries sum to $\bar{r}$ and $\underline{r}$, respectively. We then construct a perspective P such that $\Re^2(P, Q)\rho = v$. A key difficulty in this construction is the fact that not every bistochastic matrix can arise as a similarity matrix for some S-perspective. To circumvent this issue, we rely on the Schur-Horn theorem, which uses majorization to characterize the vectors that can arise as the diagonal of a matrix with a fixed set of eigenvalues.

Proposition 3 implies that when $k = n/2$, the information constraints above create a stark tradeoff: the higher the payoff S gives R in R's favorite equilibrium, the lower must be the payoff S gives R in R's worst equilibrium. This tradeoff is immediately apparent from the Proposition's part (ii). When $k = n/2$, this part says that r_M must be the unweighted average of $\underline{r}$ and $\bar{r}$, and so a higher $\bar{r}$ must be accompanied by a lower $\underline{r}$.

The intuition is easiest to see when $n = 2$ and $k = 1$. Suppose S and R have unique perspectives P and Q, respectively, and that S communicates the issue p_1 in R's favorite equilibrium. To increase R's payoff in his favorite equilibrium, one would need to make p_1 more informative about R's favorite issue—say q_1. By the above constraints, this means that p_2 must become less informative about q_1 and more informative about the issue R cares about less, namely q_2. Consequently, R's least-favorite equilibrium—which has S communicating about p_2—becomes worse for R.

This tradeoff above remains when $k \neq n/2$, but is less stark. For example, an S that maximizes R's worst equilibrium payoff must also minimize R's best equilibrium, and vice versa. We document this fact in the corollary below. To state the corollary, let P and Q be an S-perspective and R-perspective, respectively, and let $\rho = (\rho_i)_{i=1}^n$ be the vector of weights associated with R-perspective Q.

Corollary 2. *Given R-perspective Q, the following are equivalent:*

(i) R's worst equilibrium payoff is maximized, $\underline{u}_R^k(S,R) = r_M$.

(ii) R's best equilibrium payoff is minimized, $\bar{u}_R^k(S,R) = r_M$.

(iii) The vector $\Re^2(P,Q)\rho$ *is constant for any S-perspective P.*

Thus, an equivalence exists between S maximizing R's worst equilibrium payoff, S minimizing R's best equilibrium payoff, and all issues in S's perspective P having the same value to R. To get a sense as to why the corollary holds, suppose S and R have a unique perspective. In this case, R's lowest equilibrium payoff is given by the sum of k lowest values of $v = \Re^2(P,Q)\rho$. Since ρ majorizes v by the information constraints, this sum can equal r_M if and only if all entries of v are equal to r_M/k, which, in turn, is equivalent to the sum of the k highest entries of v equaling r_M—in other words, R's highest equilibrium value must be r_M.

Another feature of the $k = n/2$ case that holds generally is that an S that shares a perspective with R both maximizes R's best equilibrium payoff and minimizes R's worst equilibrium payoff across all S loss matrices.

Observation 1. *If S and R share a perspective, then* $\bar{u}_R^k(S,R) = r_H$ *and* $\underline{u}_R^k(S,R) = r_L$.

Intuitively, if S shares a perspective with R, then S can communicate clearly both about the issues R cares about most and the issues R cares about least. The former is helpful for maximizing R's payoff in his favorite equilibrium. The above-mentioned information constraints mean that the latter is helpful for minimizing R's payoff in his least favorite equilibrium: the more information an S-issue provides about the R-issues R cares about less, the less information that issue must provide about the R-issues R cares about more.

It turns out the above intuition holds more generally. In particular, there is a sense in which the better synchronized S's perspective is with R's, the better she is for R in R's favorite equilibrium, and the worse she is in R's worst equilibrium. The converse also holds: if, regardless of R's priorities, one S perspective is always better (worse) than another in R's favorite (worst) equilibrium, then the former perspective must be better synchronized with R's perspective than the latter one.

This relationship is the content of our next result. To state this result, define a **k-average row** of a matrix $M \in \mathbb{R}^{n\times n}$ to be an unweighted average of k distinct rows of M, and a **mixed k-average row** of M to be a convex combination of k-averages.[10]

Proposition 4. *Fix a receiver perspective Q. Consider senders S_1 and S_2 each with a unique perspective P_1 and P_2, respectively. The following are equivalent.*

(i) Any R with perspective Q has $\bar{u}_R^k(S_1, R) \geq \bar{u}_R^k(S_2, R)$.

(ii) Any R with perspective Q has $\underline{u}_R^k(S_1, R) \leq \underline{u}_R^k(S_2, R)$.

(iii) Every k-average row of $\Re^2(P_2, Q)$ is a mixed k-average row of $\Re^2(P_1, Q)$.

Proposition 4 suggests a way of ordering the degree to which two S-perspectives P_1 and P_2 are synced with a given R-perspective Q: say P_1 **more Q-synced than** P_2 (denoted by $P_1 \succsim_Q P_2$) if every k-average row of $\Re^2(P_2, Q)$ is a mixed k-average row of $\Re^2(P_1, Q)$ for all $k \in \{1, \dots, n-1\}$. By the above proposition, whenever P_1 is more Q-synced than P_2, a receiver with perspective Q who gets to choose his equilibrium prefers an S with a unique perspective of P_1 over an S whose unique perspective is P_2 for *any* k. Moreover, this ranking is reversed if R always expects to play his worst equilibrium: in that case, R would prefer the P_2-sender over the P_1- sender, regardless of k.

[10]Given the Birkhoff-von Neumann theorem, a vector is a mixed k-average row of M if and only if it is a convex combination of the rows of M, with each row having weight at most $\frac{1}{k}$.

The synchronization order is related to a known order in majorization theory. A matrix $M \in \mathbb{R}^{n\times n}$ **directionally majorizes** a matrix $\tilde{M} \in \mathbb{R}^{n\times n}$ if Mx majorizes $\tilde{M}x$ for any vector $x \in \mathbb{R}^n$ (Marshall et al., 2011, p. 618,). Applying Theorem 3.12 from Pería et al. (2005) delivers that $P_1 \succsim_Q P_2$ holds if and only if $\Re^2(P_1, Q)$ directionally majorizes $\Re^2(P_2, Q)$.

As one might expect, the perspective Q is more Q-synced than any other perspective, since by Proposition 3 and Observation 1, an S with such a perspective always maximizes R's best equilibrium payoff and minimizes R's worst equilibrium payoff. In other words, the synchonization order always admits a maximum. This order often also admits a minimum: a perspective W satisfying $\Re^2(\boldsymbol{\omega}_i^W, \boldsymbol{\omega}_j^Q) = 1/n$ for all i, j is less Q-synced than any other perspective.[11] Indeed, by Proposition 2, an equilibrium that communicates such a perspective gives R a payoff of $\sum_{i=1}^k \sum_{j=1}^n \rho_j \Re^2(\boldsymbol{\omega}_i^W, \boldsymbol{\omega}_j^Q) = \frac{k}{n}\sum_{j=1}^n \rho_j = r_M$, which by Proposition 3 both minimizes R's best equilibrium utility and maximizes R's worst equilibrium utility.

It is worth noting that the tradeoff between R's payoff under favorable selection and adversarial selection is less stark when R's priorities ρ are known and $k \neq n/2$. We illustrate this fact below.

Example 2. Consider the setting of Example 1, but assume now the division manager (S) has three products under her control. Thus, we assume now that $n = 3$. Moreover, both the division manager and the CEO (R) assign the highest priority to providing the division with the appropriate budget for this third product. Specifically, the loss matrices are given by

$$S = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 1000 \end{bmatrix}, \qquad R = \begin{bmatrix} 98 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1000 \end{bmatrix}.$$

We still assume that the two can only talk about one issue (i.e., $k = 1$), and that states are uncorrelated and have unit variance $V = I$. Under these assumptions, S's

[11] Such a perspective exists if some order-n Hadamard matrix does, which occurs whenever n equals any power of 2 (see Theorem 18.3, Van Lint and Wilson, 2001).

perspective and associated weights are

$$P = \left[\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right], \quad \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 1000 \end{pmatrix},$$

whereas R's perspective and associated weights are

$$Q = \left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right], \quad \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = \begin{pmatrix} 98 \\ 2 \\ 1000 \end{pmatrix}.$$

Calculating the similarity matrix $\mathfrak{R}^2(P, Q)$ and issue-value vector v gives

$$\mathfrak{R}^2(P, Q) = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad v = \begin{pmatrix} 50 \\ 50 \\ 1000 \end{pmatrix}.$$

Therefore, R receives an expected payoff of 50 in both the equilibrium that communicates the issue p_1 and the equilibrium that communicates the issue p_2. The equilibrium that communicates the issue p_3 gives R a payoff of 1000. By Proposition 2, this payoff maximizes R's utility across all equilibria and across all S.

As noted in Observation 1, R could also get an equilibrium payoff of 1000 if he communicated with a sender with a perspective Q. Indeed, in this case the similarity matrix would be the identity $\mathfrak{R}^2(Q, Q) = I$, and so the issue-value vector would be ρ, meaning R's highest equilibrium payoff is 1000 as well.

Note however, that R's worst equilibrium payoff is lower with an S that shares his perspective Q than it is with an S whose perspective is P: whereas the lowest coordinate of ρ equals 2, the lowest coordinate of v equals 50. Therefore, while R would be indifferent between a S with perspective Q and an S with perspective P under favorable equilibrium selection, R would strictly prefer the latter if equilibrium selection were adverserial. □

As demonstrated by the above example, the tradeoff between R's value in his favorite equilibrium and in his worst equilibrium is less stark when $k \neq n/2$ and ρ is fixed. In particular, some senders are worse for R than others under both equilibrium selection criteria. In fact, when R's payoffs are generic, one can find an S that maximizes R's favorite equilibrium payoff without minimizing R's payoff under adversarial selection. We summarize this observation in the following corollary.

Corollary 3. *For generic R, the following are equivalent.*

(i) Some S has $\bar{u}_R^k(S,R) = \bar{u}_R^k(R,R)$ *and* $\underline{u}_R^k(S,R) > \underline{u}_R^k(R,R)$*.*

(ii) Some S_1, S_2 *have* $\bar{u}_R^k(S_1,R) > \bar{u}_R^k(S_2,R)$ *and* $\underline{u}_R^k(S_1,R) > \underline{u}_R^k(S_2,R)$*.*

(iii) We have $k \neq \frac{n}{2}$*.*

Whereas the proof relies on Proposition 3, the corollary's intuition is straightforward. When $k < n/2$, one can find S-perspectives that enable S to clearly communicate about the k issues R cares the most about, but mix together issues R assigns intermediate importance to with the issues R cares least about.[12] Such a mixing serves as a hedge, preventing scenarios in which S communicates only information that R views as minutiae. Consequently, some senders can be replaced by senders who improve R's payoff under both equilibrium selection rules—as long as R is generic. For a non-generic set of R, different issues may have the same value (i.e., $\rho_i = \rho_j$ for some $i \neq j$), limiting the ability to hedge against adversarial equilibrium selection via issue mixing.

7 How much do you want to talk?

In this section, we assess the value of communication resources. Specifically, we consider a richer game in which our main model is preceded by R making an investment decision that determines k. If R describes the decision-making protocols in an organization, then the capacity k could be determined by organizational design choices, representing investments in adaptive decision-making. Taking our model more literally, we can think of a choice of k as a costly attention choice.

The section's results show that endogenizing k means R no longer needs to balance his value in his best and worst equilibrium when choosing between senders. Under adversarial equilibrium selection, R either pays full attention to S (setting $k = n$), or

[12]When $k > n/2$, one can analogously mix together R's k most important issues, leaving the value of communicating all of them unchanged, but improving R's value from his worst equilibrium.

none (i.e., $k = 0$), and R's attention decision is the same regardless of S's perspective. By contrast, under favorable selection R's attention decision can take interior values, and S's identity matters: if R expects to hear the S-issues he cares about most, then R always prefers a more synchronized S. Thus, when k is endogenous, R prefers more synchronized senders under favorable selection, but does not care about S's identity if he expects equilibrium selection to be adversarial.

Greater precision is in order. Suppose the game starts with R making a public choice of $k \in \{0, \cdots, n\}$, and then play proceeds according to our main model. S's payoff is as in our main model, whereas R's payoff is equal to that in our main model net of an investment cost,

$$\gamma k.$$

We assume sequential rationality, in that an equilibrium is played following the investment decision. This equilibrium is determined according to one of two possible selection criteria: either R's best equilibrium is played for every choice of k, or R's worst equilibrium is played.

We begin by analyzing R's decision under favorable selection. Suppose for exposition that S and R are generic, and so have unique perspectives, P and Q, and let ρ be R's associated weights. Suppose further that the issues in P are ordered in terms of their (generically unique) value to the receiver; that is, $v := \mathfrak{R}^2(P, Q)\rho$ is such that $v_1 > \cdots > v_n$. Given this ordering, choosing a capacity of k results in S communicating issues $1, \ldots, k$ in R's favorite equilibrium. Therefore, R's payoff under favorable selection from choosing k is

$$\bar{u}_R^k(S, R) - \gamma k = \sum_{i=1}^{k} v_i - \gamma k = \sum_{i=1}^{k} (v_i - \gamma).$$

Clearly, R strictly benefits (on the margin) from listening to issue k if and only if $v_k > \gamma$. Moreover, capacity exhibits decreasing marginal returns, meaning R benefits from listening to the k^{th} issue only if he also benefits from listening to the i^{th} issue for any $i < k$. Therefore, the optimal capacity under favorable selection is determined by a first-order condition: $v_k \geq \gamma \geq v_{k+1}$.

What if R expects adversarial selection? In this case, listening to k issues gives R a payoff of

$$\underline{u}_R^k(S, R) - \gamma k = \sum_{i=n-k+1}^{n} (v_i - \gamma).$$

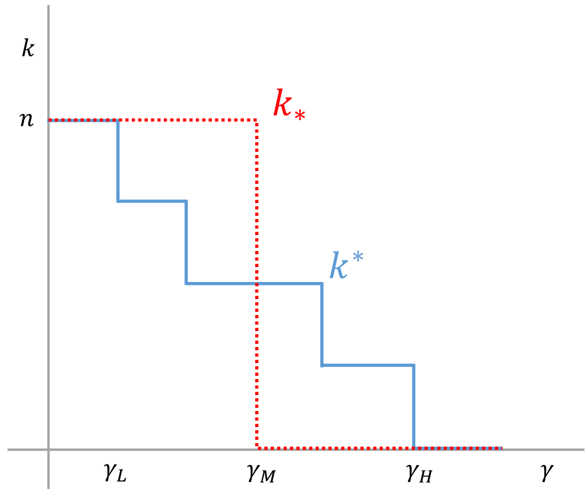


Figure 5: Optimal investment k by R under best-case and worst-case equilibrium, respectively, as a function of the cost parameter γ.

The above objective exhibits *increasing* marginal returns to capacity. Hence, if listening to k issues is better than not listening at all, then listening to $k+1$ issues is even better. So R's optimal capacity under adversarial selection is bang-bang: R either gives S his full attention, setting $k = n$, or does not listen at all. Which of these two options is optimal depends on the costs. Specifically, choosing a capacity of n is optimal if and only if

$$\frac{1}{n}\sum_{i=1}^{n} v_i \geq \gamma.$$

Otherwise, R does not listen to S and so chooses zero capacity.

Thus, whereas R's chosen capacity under favorable selection is determined by the value of the best *marginal* issue, his capacity under adversarial selection depends on the value of the *average* issue. Consequently, R's capacity choice differs between these two equilibrium selection regimes. Figure 5 illustrates these differences, denoting R's optimal investment under best-case selection by k^*, and under worst-case selection by k^*.

When γ is sufficiently low, R sets $k^* = k_* = n$ and pays full attention to S in both regimes. For γ slightly above the value of R's least favorite issue v_n, R reduces his attention under favorable selection $k^* < n$, but keeps paying full attention under adversarial selection ($k_* = n$). As γ increases further, R's attention under favorable selection decreases gradually, going to zero once γ passes the value v_1 of R's favorite S-issue. By contrast, an R expecting the worst equilibrium keeps giving S his full attention until γ passes the value of the average issue $\frac{1}{n}\sum_{i=1}^{n} v_i$, at which stage R's attention drops from $k_* = n$ to $k_* = 0$. Proposition 5 summarizes these differences.

Proposition 5. *Some cutoffs $\gamma_H \geq \gamma_M \geq \gamma_L > 0$ exist (generically, $\gamma_H > \gamma_M > \gamma_L$) such that any R-best-case and R-worst-case optimal investments k^* and k_* satisfy:*

(i) If $\gamma_L < \gamma < \gamma_M$, then $k^* < k_* = n$;

(ii) If $\gamma_M < \gamma < \gamma_H$, then $k^* > k_* = 0$;

(iii) If $\gamma < \gamma_L$ or $\gamma > \gamma_H$, then $k^* = k_* \in \{0, n\}$.

R's ability to change her capacity also changes the way he ranks different senders. If R expects the worst equilibrium, he invests the same amount regardless of S. The reason is that the value of the average issue does not depend on S's perspective:

$$\frac{1}{n}\sum_{i=1}^{n} v_i = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q)\rho_j = \frac{1}{n}\sum_{j=1}^{n}\rho_j \sum_{i=1}^{n} \Re^2(\boldsymbol{\omega}_i^P, \boldsymbol{\omega}_j^Q) = \frac{1}{n}\sum_{j=1}^{n}\rho_j.$$

By contrast, R cares a lot about which S he faces under favorable selection, a fact which we now demonstrate.

Example 1 (continued)**.** Let us return to the setting of Example 1 where $n = 2$, the variance-covariance matrix is the identity $V = I$, and the loss matrices are

$$S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \text{ and } R = \begin{bmatrix} 98 & 0 \\ 0 & 2 \end{bmatrix}.$$

S and R's corresponding perspectives and associated weights are given by

$$P = \tfrac{1}{\sqrt{2}}\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right], (\sigma_1, \sigma_2) = (4, 2) \quad \text{and} \quad Q = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right], (\rho_1, \rho_2) = (98, 2).$$

Suppose further that R publicly chooses the value of k at a cost of $25k$ before the game. Since both p_1 and p_2 gives R a payoff of 50, R chooses to pay full attention to S under both favorable and adversarial equilibrium selection. Thus, we get $k^* = k_* = 2$, and R's total payoff is 50.

What if S's loss is given by $\tilde{S} = R$ instead? Under adversarial selection, R still chooses full attention and gets a payoff of 50. The same is not true under favorable selection. For an explanation, note that if $k = 1$, S can either communicate $\boldsymbol{\omega}_1$ or $\boldsymbol{\omega}_2$ to R, giving R a payoff of 98 under the former, and a payoff of 2 under the latter. Therefore, an R who expects his best equilibrium chooses to hear only the value of $\boldsymbol{\omega}_1$, getting a net payoff of $98 - 25 = 73$. Note this payoff is strictly higher than under S. $\square$

The example compares an S who is minimally synced with R with an S who is as synced with R as possible, and shows R prefers the latter over the former. This ranking, however, is not confined to the extremes of the synchronization order: the more synced S is with R, the better off R is under favorable selection. The reason is that a more synced S is better for R under favorable selection for every k, and so is clearly also better under the optimal k.

Notice though that having a more synced S does not mean that R pays S greater attention. The example is a case in point: R chooses a smaller k^* when faced with the maximally synced S than he chooses when faced with a minimally synced one. Our next result shows this comparative static holds more generally, applying to other pairs of senders who are are strictly ranked via the synchronization order.

Proposition 6. *Suppose R has perspective Q, and let S_1 and S_2 be two sender losses whose unique perspectives are P_1 and P_2, respectively. If $P_1 \succsim_Q P_2$ and $\bar{u}^k_R(S_1, R) \neq \bar{u}^k_R(S_2, R)$ for some k, then some $\gamma_{HH} > \gamma_H \geq \gamma_L > \gamma_{LL} > 0$ exist such that the set of R-best-case optimal investments k^*_1 and k^*_2 for S_1 and S_2, respectively, satisfy:*[13]

(i) If $\gamma_{LL} < \gamma < \gamma_L$, then $k^*_1 < k^*_2$;

(ii) If $\gamma_H < \gamma < \gamma_{HH}$, then $k^*_1 > k^*_2$;

(iii) If $\gamma < \gamma_{LL}$ or $\gamma > \gamma_{HH}$, then $k^*_1 = k^*_2$.

Consequently, the above holds for generic R-weights ρ whenever $P_1 \succ_Q P_2$.

Thus, under favorable selection, a generic R listens more to a more synchronized S when attention is moderately expensive, and listens less when attention is somewhat cheap. Intuitively, the more synchronized an S is with R, the clearer S communicates about the issues R cares least and most about. Consequently, the more synchronized S is with R, the value of the best issue S can communicate increases, while value of the

[13]We order k^*_1 and k^*_2 via the strong set order.

worst issue S can communicate decreases. Therefore, when costs are high, making S more synchronized might result in R switching from not listening to S all, to agreeing to listen to the S-issues R values most. Similarly, when costs are low, making S more synchronized could push R to reduce his attention by dropping the S-issues that are less important to him.

8 What do you want us to discuss?

We conclude by discussing the implications of some of our modeling assumptions.

Partial Rank. The number k parametrizes the capacity of R's attention. In our main analysis, we assume R utilizes all of this capacity, requiring R's strategy to have full rank. This assumption represents a scenario in which R's attention is free at the margin at the time he communicates with S. An alternative model would allow R to choose an arbitrary linear decision rule $D \in \mathbb{R}^{n \times k}$ that need not have full rank; hence capacity up to k is free at the margin, but R may abstain from using it.

Given our analysis, one can characterize the equilibrium outcomes of this alternative game. For each $\tilde{k} \in \{0, \ldots, k\}$ and each S-perspective P, we can consider a strategy profile in which S sends message $\mathbf{m} = (\boldsymbol{\omega}_i^P \mathbf{1}_{i \leq \tilde{k}})_{i=1}^n$ and R chooses action $\mathbf{a}$ such that $\mathbf{a}_i^P = \mathbf{m}_i \mathbf{1}_{i \leq \tilde{k}}$. In words, S fully communicates the first $\tilde{k}$ issues according to her perspective, and R takes these at face value and ignores the remaining dimensions of the message. Our analysis extends to show all of these strategy profiles are equilibria, and every equilibrium has an outcome-equivalent one from this family.[14]

Note that exhausting the capacity on communicating this perspective is a strong Pareto improvement, so every equilibrium with $\tilde{k} < k$ constitutes a coordination failure—the most extreme one having $\tilde{k} = 0$ so that action zero is always chosen.

These coordination failures, however, rely on a fragile kind of self-fulfilling expectations: R ignores some dimensions because he is certain that they are not helpful for listening to S, while S does not use these dimensions because she expects R to ignore them. These expectations are quite fragile, however, as they rely on each player being absolutely certain that the other is wasting some capacity. For an illustration, suppose R has a positive probability of being restricted to full-rank strategies, and

[14] In particular, all equilibria of the original game remain equilibria even though R has new deviation opportunities.

S sends her message in ignorance of this realization. Then one can show that every equilibrium has R using a full-rank decision rule even when he is not forced to (and using the same decision rule in either case). Therefore, one can think of our full-rank restriction to studying equilibria that are robust to the possibility of non-wastefulness.

Uncenteredness We have assumed to this point that the state is centered, that is, it has zero expectation. Suppose now that the state may have a nonzero expectation $\hat{\omega}$. The comparison between this model and our original model depends on what we assume about R's space of available strategies.

First, suppose R's space of strategies is the same as in our original model, meaning R must choose a full-rank linear strategy. In this strategy space the zero action occupies a special role: No matter what strategy R chooses, S always has some message to send (the zero message) that ensures the zero action is played. In our main model this fact—combined with linearity of S's best reply—results in R's action being an unbiased estimate of the state. The same is not true in the uncentered model, since the range of R's strategy might not contain $\hat{\omega}$. In this case, R's decision becomes a biased estimate of the state. This bias obviously has welfare consequences, but it also shapes R's incentives. Whereas in our main model R acts to minimize his exposure to (weighted) variance of the state, in this model he balances this goal against minimizing his exposure to the bias. In fact, a straightforward modification of Appendix A's Lemma 2 proof tells us this model is identical—with the same set of equilibria and resulting expected payoffs—to that with a modified covariance matrix of $V + \hat{\omega}\hat{\omega}^\top$. Hence, imposing a default action different from the default state is equivalent to R being more uncertain about the state in the direction of the imposed bias in decision making.

An alternative model is one in which R chooses a pair $(\hat{a}, D)$, where $\hat{a} \in \mathbb{R}^n$ is a default action and $D \in \mathbb{R}^{n \times k}$ has full rank. This model is essentially equivalent to our general model. Indeed, R incentives require that the expected action be equal to the expected state. As a consequence, equilibrium communication and play is identical to our original model with a modified state $\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} - \hat{\omega}$. Up to outcome equivalence, equilibria take the following form: We decompose the centered state $\tilde{\boldsymbol{\omega}}$ according to some S perspective, R chooses a default action $\hat{\omega}$, and the players communicate this perspective concerning $\tilde{\boldsymbol{\omega}}$. So if R can freely choose his default action, then our centeredness assumption is purely a normalization.

Bias Our model assumes the players have the same ideal action given the state. This assumption allows us to focus on the incentive frictions that stem from our bounded rationality assumption. Nevertheless, one might wonder how such bounded rationality interacts with a misalignment between S's and R's ideal action. Specifically, suppose S's ideal action in state $\boldsymbol{\omega}$ is $\boldsymbol{\omega} + b$ for some nonzero vector $b \in \mathbb{R}^n$.

Suppose that, as discussed in the case of an uncentered state, R may choose a default action in addition to a linear decision rule. Hence, R's strategy is a pair $(\hat{a}, D)$. S's incentives now require the action to be a rank-k linear function of her ideal action $\boldsymbol{\omega}+b$. Meanwhile, R's incentives require that the average action be zero (which is the average state). Hence, any equilibrium admits an outcome-equivalent equilibrium in which the default action is zero. Consequently, the conditions for equilibrium in this modified model are the same as in our main model, but with the added restriction that R's resulting action is simultaneously both a linear function of $\boldsymbol{\omega}$ and a linear function of $\boldsymbol{\omega} + b$. In the language of Proposition 1, all equilibria are (up to outcome equivalence) given by strategy profiles that communicate an S-perspective P such that every $i \in \{1, \dots, k\}$ has $b_i^P = 0$. We can interpret this orthogonality condition as saying S cannot be biased about any communicated issue. Unfortunately, these equilibrium conditions are impossible to satisfy for most S matrices.

Commitment We have assumed the players' play to be strategically simultaneous: R chooses a decision rule without observing how S communicates, and S (knowing the state) decides how to communicate without observing R's decision rule. Other natural communication protocols would allow one of the players to commit to their behavior.

First, suppose S could commit to a strategy, and R would respond with his choice of a full-rank $n \times k$ matrix. Such a setting would amount to a Bayesian persuasion model (Kamenica and Gentzkow, 2011) in which R is subject to our imposed bounded rationality constraint. The next result establishes that the solution to this persuasion problem is in fact an equilibrium. Hence, the sole benefit of commitment power to S is the ability to select an equilibrium.

Proposition 7. *A strategy profile is S-preferred if and only if it is outcome equivalent to an S-preferred equilibrium.*

Note the above result says that neither player's incentive constraints in our game impose costs on S. If S could choose any strategy profile at all in our game, she would effectively be choosing among all state-contingent action choices, subject to the constraint that all chosen actions live in some k-dimensional subspace of $\mathbb{R}^n$. The proposition shows S's preferred strategy profile simply communicates the k most important issues from her perspective, which is an equilibrium of our model. We remark that Corollary 1 from Bizzotto and Hancart (2026) essentially yields a specialization of this result for case in which $S = R$, there are only $n = 2$ dimensions, and R's capacity is $k = 1$.

The situation is quite different if R can commit. If R first publicly commits to a strategy, the interaction is one of *subspace delegation*: R announces a k-dimensional set of actions from which S can freely choose. This delegation model can be seen as an organizational design problem. Maintaining too much flexibility in an organization is infeasible, but an organization can potentially decide which kinds of flexibility to invest in.[15]

The next result shows that R benefits from this commitment power in most specifications of our model. In particular, S commitment and R commitment have very different implications for issue communication.

Proposition 8. *R strictly benefits from commitment for generic* (S, R).

For intuition, let us return to CEO and manager of Example 1. Recall that the CEO (R) chiefly wants to choose the correct budget for division 1, whereas the manager (S) mainly wants to set the correct total budget. One equilibrium communicates the ideal total budget, one communicates the ideal difference between the two divisions' budgets, and both give a payoff of 50 to the CEO. But what if the CEO could commit to only take the manager's advice on division 1, and to stick with the ex-ante optimal budget for the other division—i.e., $D = [1\ 0]$? A direct computation shows that the manager would respond by sending message $\mathbf{m} = \boldsymbol{\omega}_1 + \frac{1}{3}\boldsymbol{\omega}_2$. She inflates the requested budget for division 1 when the ideal budget of division 2 is high, because better targeting the total budget outweighs the added distortion on relative budgets. Meanwhile, this strategy profile gives the CEO an expected payoff of $\frac{784}{9} \approx 87.1$.

The CEO is able to attain the above payoff substantially higher than his equilibrium payoff of 50 only because he is using his commitment power. For an explanation,

[15]This problem is similar to Frankel (2016) in that it studies a variant of Holmstrom (1980) with a multidimensional state. Relative to Frankel (2016), our subspace delegation problem replaces an additive bias with a discrepancy in priorities and puts further structure on admissible delegation sets.

note that the CEO's best linear estimate of $\boldsymbol{\omega}_2$ when hearing a message of $\mathbf{m}$ is a strictly positive multiple of $\mathbf{m}$, and that his best estimate of $\boldsymbol{\omega}_1$ deflates $\mathbf{m}$ toward zero, because $\mathbf{m}$ equals $\boldsymbol{\omega}_1$ plus independent "noise". Hence, the CEO would like to deviate in both coordinates of D. Committing not do so enables him to benefit from superior advice.

References

N. Antić and A. Chakraborty. Selected facts. Working paper, November 2025, 2025.

Y. C. Aybas and E. Turkel. Persuasion with coarse communication. *Economic Theory*, 2026. doi: 10.1007/s00199-026-01734-z. Published online 27 June 2026.

A. Baker. *Matrix groups: An introduction to Lie group theory*. Springer Science & Business Media, 2003.

I. Ball and X. Gao. Checking cheap talk, 2026. Version 3, revised March 2026.

M. Battaglini. Multiple Referrals and Multidimensional Cheap Talk. *Econometrica*, 70(4):1379–1401, 2002. ISSN 0012-9682. doi: 10.1111/1468-0262.00336.

J. Bizzotto and N. Hancart. Simple communication, 2026. Version 4, revised June 2026.

A. W. Bloedel and I. Segal. Persuading a rationally inattentive agent. 2021.

A. Blume. Meaning in communication games. *International Journal of Game Theory*, 55(1):3, 2026. doi: 10.1007/s00182-025-00973-z. Article 3.

P. Bolton and M. Dewatripont. The firm as a communication network. *The Quarterly Journal of Economics*, 109(4):809–839, 1994.

G. Carroll and G. Egorov. Strategic communication with minimal verification. *Econometrica*, 87(6):1867–1892, 2019.

A. Chakraborty and R. Harbaugh. Comparative cheap talk. *Journal of Economic Theory*, 132(1):70–94, 2007.

A. Chakraborty and R. Harbaugh. Persuasion by cheap talk. *American Economic Review*, 100(5):2361–2382, 2010.

V. P. Crawford and J. Sobel. Strategic information transmission. *Econometrica*, 50 (6):1431–1451, 1982.

J. Cremer, L. Garicano, and A. Prat. Language and the theory of the firm. *The Quarterly Journal of Economics*, 122(1):373–407, 2007.

M. A. Cronin and L. R. Weingart. Representational gaps, information processing, and conflict in functionally diverse teams. *Academy of Management Review*, 32(3): 761–773, 2007. doi: 10.5465/amr.2007.25275511.

R. L. Daft and K. E. Weick. Toward a model of organizations as interpretation systems. *Academy of Management Review*, 9(2):284–295, 1984. doi: 10.5465/amr. 1984.4277657.

W. Dessein, A. Galeotti, and T. Santos. Rational inattention and organizational focus. *American Economic Review*, 106(6):1522–1536, 2016.

J. E. Dutton, L. Fahey, and V. K. Narayanan. Toward understanding strategic issue diagnosis. *Strategic Management Journal*, 4(4):307–323, 1983. doi: 10.1002/smj. 4250040403.

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

A. Frankel. Delegating multiple decisions. *American economic journal: Microeconomics*, 8(4):16–53, 2016.

X. Gabaix. A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710, 2014.

L. Garicano. Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5):874–904, 2000. doi: 10.1086/317671.

J. Glazer and A. Rubinstein. On optimal rules of persuasion. *Econometrica*, 72(6): 1715–1736, 2004.

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

B. Holmstrom. On the theory of delegation. Technical report, Discussion Paper, 1980.

P. K. Hu. Multidimensional information and rational inattention. 2020.

G. Jäger, L. P. Metzger, and F. Riedel. Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals. *Games and economic behavior*, 73(2):517–537, 2011.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, October 2011.

Ö. Koçak and P. Puranam. Separated by a common language: How the nature of code differences shapes communication success and code convergence. *Management Science*, 68(7):5287–5310, 2022. doi: 10.1287/mnsc.2021.4157.

B. Kőszegi and F. Matějka. Choice simplification: A theory of mental budgeting and naive diversification. *The Quarterly Journal of Economics*, 135(2):1153–1207, 2020.

S. G. Krantz and H. R. Parks. *A primer of real analytic functions*. Springer Science & Business Media, 2002.

M. Le Treust and T. Tomala. Persuasion with limited communication capacity. *Journal of Economic Theory*, 184:104940, 2019. doi: 10.1016/j.jet.2019.104940.

G. Levy and R. Razin. On the limits of communication in multidimensional cheap talk: a comment. *Econometrica*, 75(3):885–893, 2007.

E. Lipnowski, L. Mathevet, and D. Wei. Attention management. *American Economic Review: Insights*, 2(1):17–32, 2020. doi: 10.1257/aeri.20190165.

J. G. March and H. A. Simon. *Organizations*. John Wiley & Sons, New York, 1958.

A. W. Marshall, I. Olkin, and B. C. Arnold. Inequalities: Theory of majorization and its applications. *Springer Series in Statistics (*, 2011.

N. Matouschek, M. Powell, and B. Reich. Organizing modular production. *Journal of Political Economy*, 133(3):986–1046, 2025.

W. Ocasio. Towards an attention-based view of the firm. *Strategic Management Journal*, 18(S1):187–206, 1997.

F. D. M. Pería, P. G. Massey, and L. E. Silvestre. Weak matrix majorization. *Linear Algebra and its Applications*, 403:343–368, 2005.

M. Rabin. A model of pre-game communication. *Journal of Economic Theory*, 63(2): 370–391, 1994.

P. J. Reny. Natural language equilibrium i: Off-path conventions. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2025-64), 2025.

N. Rodriguez. Language constraints in multi-sender communication games. *International Journal of Game Theory*, 55:24, 2026. doi: 10.1007/s00182-026-00998-y. Article 24.

W. Rudin. *Real and complex analysis*. McGraw-Hill, Inc., 1987.

M. L. Tushman and D. A. Nadler. Information processing as an integrating concept in organizational design. *Academy of Management Review*, 3(3):613–624, 1978. doi: 10.5465/amr.1978.4305791.

L. Van den Dries. *Tame topology and o-minimal structures*, volume 248. Cambridge university press, 1998.

J. H. Van Lint and R. M. Wilson. *A course in combinatorics*. Cambridge university press, 2001.

B. Wernerfelt. Organizational languages. *Journal of Economics & Management Strategy*, 13(3):461–472, 2004. doi: 10.1111/j.1430-9134.2004.00019.x.

A Proofs for Main Results

Definition 1. For any positive definite matrix $M \in \mathbb{R}^{n\times n}$:

1. Let $\langle\cdot,\cdot\rangle_M$ denote the inner product $\mathbb{R}^n$ given by $\langle x, y\rangle := x^\top My$.
2. Say $A \in \mathbb{R}^{n\times n}$ is an M**-projection** if any of these equivalent conditions hold:
   - A is an orthogonal projection on the inner product space $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_M)$.
   - A is idempotent (i.e., $A^2 = A$) with $A^\top M = MA$.
   - $\sqrt{M}A\sqrt{M}^{-1}$ is idempotent and symmetric.
   - $\sqrt{M}A\sqrt{M}^{-1}$ is an orthogonal projection on the $\mathbb{R}^n$ with the dot product.
   - $A^\top MA = MA$.

Notation 1. *Left multiplication by any matrix $C \in \mathbb{R}^{k\times n}$ is an S strategy, which we also denote by C. Conversely, any linear S strategy C can be represented as left multiplication by a matrix, and we also let $C \in \mathbb{R}^{k\times n}$ denote this matrix.*

Lemma 1. *Let (C, D) be a strategy profile. Then C is an S best response to D if and only if C is linear and $A := DC \in \mathbb{R}^{n\times n}$ is an S-projection of rank k. Moreover, a unique S best response exists for any R strategy.*

Proof. Given any $\omega \in \mathbb{R}^n$, we have:

$$
\begin{aligned}
& C(\omega) \in \operatorname{argmax}_{m\in\mathbb{R}^k} \left\{\omega^\top S\omega - (\omega - Dm)^\top S(\omega - Dm)\right\} \\
\iff\quad & DC(\omega) \in \operatorname{argmin}_{a\in \operatorname{ran} D} \left\{(\omega - a)^\top S(\omega - a)\right\} \\
\iff\quad & DC(\omega) \text{ is the } \langle\cdot,\cdot\rangle_S\text{-nearest point to } \omega \text{ in } \operatorname{ran} D \\
\iff\quad & DC(\omega) \text{ is the } \langle\cdot,\cdot\rangle_S\text{-orthogonal projection of } \omega \text{ onto } \operatorname{ran} D.
\end{aligned}
$$

Therefore, C is an S best response to D if and only if C is linear and $DC(\cdot)$ is the $\langle\cdot,\cdot\rangle_S$-orthogonal projection map onto $\operatorname{ran}D$.

Because left multiplication by D is an injective linear map on $\mathbb{R}^k$, it follows that this map has an inverse, which is linear. Therefore, given the above equivalence, S has a unique best response to D, namely, the composition of this inverse with the $\langle\cdot,\cdot\rangle_S$-orthogonal projection of ω onto $\operatorname{ran}D$ (a linear map).

Given $C \in \mathbb{R}^{k\times n}$, let $A := DC \in \mathbb{R}^{n\times n}$. We have argued C is an S best response to D only if A is the S-projection onto $\text{ran}D$—or, equivalently, an S-projection with $\text{ran}A = \text{ran}D$. But because $\text{ran}A$ is a linear subspace of $\text{ran}D$, they coincide if and only if they have the same dimension (and the latter has dimension k). $\square$

Lemma 2. *Let (C, D) be a strategy profile with C linear, and let $A := DC \in \mathbb{R}^{n\times n}$. Then S and R, respectively, have expected payoff*

$$\mathbb{E}\left[\mathbf{u}_S\right] = \text{Tr}\left[\left(2SA - A^\top SA\right)V\right] \text{ and } \mathbb{E}\left[\mathbf{u}_R\right] = \text{Tr}\left[\left(2RA - A^\top RA\right)V\right].$$

Hence, if A is a V^{-1}-projection, then

$$\mathbb{E}\left[\mathbf{u}_S\right] = \text{Tr}\left(SAV\right) \text{ and } \mathbb{E}\left[\mathbf{u}_R\right] = \text{Tr}\left(RAV\right).$$

Proof. First, any $M \in \mathbb{R}^{n\times n}$ has

$$\mathbb{E}\left[\boldsymbol{\omega}^\top M\boldsymbol{\omega}\right] = \mathbb{E}\left[\text{Tr}\left(M\boldsymbol{\omega}\boldsymbol{\omega}^\top\right)\right] = \text{Tr}\left(M\mathbb{E}\left[\boldsymbol{\omega}\boldsymbol{\omega}^\top\right]\right) = \text{Tr}\left(MV\right).$$

Meanwhile, S's payoff takes the form

$$\begin{aligned}
\mathbf{u}_S &= \boldsymbol{\omega}^\top S\boldsymbol{\omega} - (\boldsymbol{\omega} - A\boldsymbol{\omega})^\top S(\boldsymbol{\omega} - A\boldsymbol{\omega}) \\
&= \boldsymbol{\omega}^\top\left[S - (I - A^\top)S(I - A)\right]\boldsymbol{\omega} \\
&= \boldsymbol{\omega}^\top\left(A^\top S + SA - A^\top SA\right)\boldsymbol{\omega} \\
&= \boldsymbol{\omega}^\top\left(2SA - A^\top SA\right)\boldsymbol{\omega}^\top.
\end{aligned}$$

The formula for $\mathbb{E}\left[\mathbf{u}_S\right]$ comes from combining these two calculations. Specializing now to the case that A is a V^{-1}-projection, we get

$$\text{Tr}\left(A^\top SAV\right) = \text{Tr}\left[SAV\left(A^\top V^{-1}\right)V\right] = \text{Tr}\left[SAV\left(V^{-1}A\right)V\right] = \text{Tr}\left[SA^2V\right] = \text{Tr}\left[SAV\right],$$

yielding the simplified payoff formula. Identical calculations (replacing S with R) give the payoff formulae for $\mathbb{E}\left[\mathbf{u}_R\right]$. $\square$

Lemma 3. *Let (C, D) be a strategy profile with C linear, and let $A := DC \in \mathbb{R}^{n\times n}$. Then D is an R best response to C if and only if A is a V^{-1}-projection. Moreover, a unique R best response exists for any rank-k linear S strategy.*

Proof. Define the differentiable function $G : \mathbb{R}^{n\times n} \to \mathbb{R}$ by letting

$$G(\tilde{A}) := \operatorname{Tr}\left[\left(2R\tilde{A} - \tilde{A}^\top R\tilde{A}\right)V\right] = 2\operatorname{Tr}(R\tilde{A}V) - \operatorname{Tr}\left(\tilde{A}^\top R\tilde{A}V\right).$$

We begin by noting the following conditions are equivalent for an R strategy D:

(i) D is a best response for R.

(ii) $D \in \operatorname{argmax}_{\tilde{D}\in\mathbb{R}^{n\times k} \text{ full rank}} G(\tilde{D}C)$.

(iii) $\frac{\partial}{\partial\epsilon}\Big|_{\varepsilon=0} G\left([D+\varepsilon M]\,C\right) = 0$ for every $M \in \mathbb{R}^{n\times k}$.

(iv) $D \in \operatorname{argmax}_{\tilde{D}\in\mathbb{R}^{n\times k}} G(\tilde{D}C)$.

First, Lemma 2 tells us (i) and (ii) are equivalent, and (iv) obviously implies (ii). Next, (ii) implies (iii) because the set of full-rank matrices is open in $\mathbb{R}^{n\times k}$, and so $D+\varepsilon M$ has full rank for any $M \in \mathbb{R}^{n\times k}$ and all $\varepsilon \in \mathbb{R}$ close enough to zero. Then, to see (iii) implies (iv), it suffices to see that $\tilde{D} \mapsto G(\tilde{D}C)$ is concave. To that end, note that $\tilde{D} \mapsto \tilde{D}C$ and $\tilde{A} \mapsto 2\operatorname{Tr}(R\tilde{A}V)$ are linear, and the function $\tilde{A} \mapsto \operatorname{Tr}\left(\tilde{A}^\top R\tilde{A}V\right)$ is (as a quadratic and nonnegative function) convex—nonnegativity follows because any $\tilde{A}$ has $\operatorname{Tr}\left(\tilde{A}^\top R\tilde{A}V\right) = \operatorname{Tr}\left[\left(\sqrt{R}\tilde{A}\sqrt{V}\right)^\top\left(\sqrt{R}\tilde{A}\sqrt{V}\right)\right]$ and a positive semidefinite matrix has nonnegative trace. Given that the four conditions are equivalent, we work with condition (iii) below.

Letting $A := DC$, any $M \in \mathbb{R}^{n\times k}$ has

$$\begin{aligned}
G\left([D+\varepsilon M]\,C\right) &= 2\operatorname{Tr}(RAV) - \operatorname{Tr}\left(A^\top RA\right) - \varepsilon^2\operatorname{Tr}\left(C^\top M^\top RMCV\right)\\
&\quad +2\varepsilon\operatorname{Tr}\left(RMCV\right) - 2\varepsilon\operatorname{Tr}\left(A^\top RMCV\right)\\
\implies \frac{1}{2}\frac{\partial}{\partial\epsilon}\Big|_{\varepsilon=0} G\left([D+\varepsilon M]\,C\right) &= \operatorname{Tr}(RMCV) - \operatorname{Tr}(A^\top RMCV)\\
&= \operatorname{Tr}(CVRM) - \operatorname{Tr}(CVA^\top RM)\\
&= \operatorname{Tr}\left[CV(I-A^\top)RM\right]
\end{aligned}$$

So D is a best response for R if and only if $CV(I-A^\top)RM$ has zero trace for every $M \in \mathbb{R}^{n\times k}$, or equivalently if and only if $CV(I-A^\top)R = 0$. Right multiplying this equation by the invertible matrix R^{-1} and left multiplying by the full-column-rank matrix D, we can equivalently express this condition as $AV(I-A^\top) = 0$. Transposing, D is a best response for R if and only if $VA^\top = AVA^\top$. This condition is equivalent to $A^\top$ being a V-projection, which in turn is equivalent to A being a V^{-1}-projection.

Let us next see that any full-rank linear S strategy C admits an R best response. Given the above argument, we need to find an R strategy D such that DC is a V^{-1}-projection. And indeed, by the rank-nullity theorem, left multiplication by C bijectively maps the V^{-1}-orthogonal complement $\mathcal{X}$ of $\ker C$ to $\mathrm{ran} C = \mathbb{R}^k$, so we can choose D so that left multiplication by it is the inverse of this map. Then, left multiplication by DC fixes $\mathcal{X}$ and annihilates $\ker C$, making it the V^{-1}-projection onto $\mathcal{X}$.

Finally, let us observe that the above R best response is unique. The reason is that any R strategy has full column rank, and so the composition of any R strategy with a given linear S strategy has the same kernel. Uniqueness then follows from an orthogonal projection being determined by its kernel. ◻

Lemma 4. *For $A \in \mathbb{R}^{n\times n}$, letting $E := [\mathbf{1}_{i=j\le k}]_{i,j=1}^n$, the following are equivalent:*

1. *The matrix A has rank k and is both a S-projection and a V^{-1}-projection.*
2. *Some S-perspective P exists such that $A = (P^\top)^{-1} E P^\top$.*

Proof. First, suppose $A = (P^\top)^{-1} E P^\top$ for some S-perspective $P = [p_1 \cdots p_n]$. Let us now see A is a rank-k map that is both a S-projection and a V^{-1}-projection. First, P (and hence its transpose) having full rank means A has rank equal to $\mathrm{rank} E = k$. Second, A is idempotent because the similar matrix E is. Third, that P is a perspective means $P^\top V P = I$ and so $A = VPEP^\top$. Therefore, $V^{-1}A = PEP^\top = A^\top V^{-1}$, making A a V^{-1}-projection. Finally, P being an S-perspective means $S = P\Lambda P^\top$ for some diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$. Therefore,

$$A^\top S - SA = PEP^{-1}P\Lambda P^\top - P\Lambda P^\top (P^\top)^{-1} E P^\top = P\left(E\Lambda - \Lambda E\right)P^\top,$$

which is zero because Λ and E are both diagonal. Hence, A is an S-projection.

Conversely, suppose A has rank k and is both a S-projection and a V^{-1}-projection. In what follows, define the rank-k matrix $\tilde{A} := \sqrt{V}^{-1} A \sqrt{V}$ and the positive definite matrix $\tilde{S} := \sqrt{V} S \sqrt{V}$. That A is a V^{-1}-projection tells us $\tilde{A}$ is an orthogonal projection (with respect to the standard inner product); and that A is an S-projection tells us

$$0 = \sqrt{V}(SA - A^\top S)\sqrt{V} = \tilde{S}\tilde{A} - \tilde{A}^\top \tilde{S} = \tilde{S}\tilde{A} - \tilde{A}\tilde{S}.$$

The symmetric matrices $\tilde{S}$ and $\tilde{A}$ commute, meaning they are simultaneously diagonalizable by an orthonormal (according to the standard inner product) basis. Since

$\tilde{A}$ is an orthogonal projection matrix, this means there exists an orthonormal (again, according to the standard inner product) basis $(\tilde{p}_i)_{i=1}^k$ for ran$\tilde{A}$ and an orthonormal basis $(\tilde{p}_i)_{i=k+1}^n$ for $\ker \tilde{A}$ such that all of $\tilde{p}_1, \ldots \tilde{p}_n$ are eigenvectors for $\tilde{S}$. Now, define $\tilde{P} := [\tilde{p}_1 \cdots \tilde{p}_n] \in \mathbb{R}^{n\times n}$ and $P := \sqrt{V}^{-1}\tilde{P} \in \mathbb{R}^{n\times n}$. Let us see P witnesses the desired property for A. First, because $\tilde{P}$ is orthogonal by construction, P is a perspective. Next, the orthogonal matrix $\tilde{P}$ has $\tilde{S}$-eigenvectors as its columns, implying $\tilde{S} = \tilde{P}\Lambda\tilde{P}^\top$ for some diagonal $\Lambda \in \mathbb{R}^{n\times n}$. Therefore,

$$P\Lambda P^\top = \left(\sqrt{V}^{-1}\tilde{P}\right)\left(\tilde{P}^\top\tilde{S}\tilde{P}\right)\left(\tilde{P}^\top\sqrt{V}^{-1}\right) = \sqrt{V}^{-1}\tilde{S}\sqrt{V}^{-1} = S.$$

Hence, P is an S-perspective. Finally, observe that $P^\top = \tilde{P}^\top\sqrt{V}^{-1}$ which has inverse $\sqrt{V}\tilde{P}$, and that

$$\tilde{P}E\tilde{P}^\top = \tilde{P}\left(\sum_{i=1}^k e_i e_i^\top\right)\tilde{P}^\top = \sum_{i=1}^k \tilde{p}_i\tilde{p}_i^\top = \tilde{A}.$$

Thus, $\left(P^\top\right)^{-1}EP^\top = \sqrt{V}\tilde{P}E\tilde{P}^\top\sqrt{V}^{-1} = \sqrt{V}\tilde{A}\sqrt{V}^{-1} = A$, as required. □

Proof of Proposition 1. By Lemma 1 and Lemma 3, a given strategy profile (C, D) is an equilibrium if and only if C is linear and $A := DC \in \mathbb{R}^{n\times n}$ is a rank-k matrix that is both a S-projection and a V^{-1}-projection. Meanwhile, for the strategy profile (C, D) that communicates perspective $P = [p_1 \cdots p_n]$, we know $C \in \mathbb{R}^{k\times n}$ is a linear strategy, and

$$C = \begin{pmatrix} I_k & 0_{k,n-k}\end{pmatrix} P^\top \text{ and } D = \left(P^\top\right)^{-1}\begin{pmatrix} I_k \\ 0_{n-k,k}\end{pmatrix}.$$

Letting $E := \left[\mathbf{1}_{i=j\le k}\right]_{i,j} \in \mathbb{R}^{n\times n}$ then gives $DC = \left(P^\top\right)^{-1}EP^\top$. The desired equivalence then follows from Lemma 4. □

Notation 2. *For any perspectives P and Q, define their **correlation** matrix $\Re(P,Q) \in \mathbb{R}^{n\times n}$, whose ij-entry is the correlation $\mathbb{E}\left[\boldsymbol{\omega}_i^P\boldsymbol{\omega}_j^Q\right]$. Algebraically, the correlation matrix takes the form $\Re(P,Q) = P^\top VQ$.*

Lemma 5. *Let Q be an R-perspective with associated weights $\rho = (\rho_i)_{i=1}^n \in \mathbb{R}^n$. If a strategy profile (C, D) communicates perspective P, then it yields expected R payoff*

$$\mathbb{E}\left[\mathbf{u}_R\right] = \sum_{i=1}^k e_i^\top\Re^2(P,Q)\rho = \sum_{j=1}^n \rho_j\sum_{i=1}^k\left(\mathbb{E}[\boldsymbol{\omega}_i^P\boldsymbol{\omega}_j^Q]\right)^2.$$

If P is an S-perspective with associated weights $(\sigma_i)_{i=1}^n$, then the strategy profile yields expected S payoff $\mathbb{E}[\mathbf{u}_S] = \sum_{i=1}^k \sigma_i$.

Proof. Recall that P being a perspective means $P^\top VP = I$; and note that $E := [\mathbf{1}_{i=j\leq k}]_{i,j} \in \mathbb{R}^{n\times n}$ has $E^\top E = E$. So letting $M := E\Re(P,Q)$, matrix $A := DC$ has

$$Q^\top AVQ = Q^\top (P^\top)^{-1}(E)P^\top VQ = Q^\top (VP)(E^\top E)P^\top VQ = M^\top M.$$

Any $j \in [n]$ therefore has[16]

$$\mathrm{Tr}(Qe_je_j^\top Q^\top AV) = e_j^\top Q^\top AVQe_j = (Me_j)^\top (Me_j) = \sum_{i=1}^n M_{ij}^2 = \sum_{i=1}^k \Re^2(P,Q)_{ij}.$$

Now, because Q is an R-perspective with weights $(\rho_i)_{i=1}^n$, we can write

$$R = \sum_{j=1}^n \rho_j Qe_je_j^\top Q^\top.$$

Recalling from Lemma 2 that $\mathbb{E}[\mathbf{u}_R] = \mathrm{Tr}(RAV)$, we therefore have

$$\mathbb{E}[\mathbf{u}_R] = \sum_{j=1}^n \rho_j \,\mathrm{Tr}(Qe_je_j^\top Q^\top AV) = \sum_{j=1}^n \rho_j \sum_{i=1}^k \Re^2(P,Q)_{ij} = \sum_{i=1}^k e_i^\top \Re^2(P,Q)\rho.$$

Finally, in the case P is an S-perspective with associated weights $(\sigma_i)_{i=1}^n$, identical calculations show that $\mathbb{E}[\mathbf{u}_S] = \sum_{i=1}^k e_i^\top \Re^2(P,P)\sigma = \sum_{i=1}^k e_i^\top I\sigma = \sum_{i=1}^k \sigma_i$. □

Proof of Proposition 2. Apply Lemma 5 to equilibria. □

Definition 2. An **R-extremal S-perspective** is an S-perspective P such that, for any $k \in \{1,\ldots,n-1\}$ and S-perspective $\tilde{P}$, letting (s,r) be the payoff pair from communicating perspective $\tilde{P}$ in the capacity-k game, there exist S-perspectives $\bar{P}$ and $\underline{P}$ that are issue-equivalent to P such that the payoff pairs $(\bar{s},\bar{r})$ and $(\underline{s},\underline{r})$ induced by communicating $\bar{P}$ and $\underline{P}$, respectively, have $\underline{s} = s = \bar{s}$ and $\underline{r} \leq r \leq \bar{r}$.

Lemma 6. *Let P be an S-perspective with associated weights $(\sigma_i)_{i=1}^n$, and define the partition $\mathfrak{I}$ of $[n]$ with $i,j \in [n]$ in the same cell of $\mathfrak{I}$ if and only if $\sigma_i = \sigma_j$. Then, for*

$$\mathcal{M} := \left\{M \in \mathbb{R}^{n\times n} : \ M \text{ orthogonal, } M_{ij} = 0 \text{ for } i \text{ and } j \text{ in distinct } \mathfrak{I}\text{-cells}\right\},$$

[16]Recall, any $m \in \mathbb{Z}_+$ has $[m] = \{1,\ldots,m\} = \{i \in \mathbb{N} : \ i \leq m\}$.

the set of S-perspectives is equal to $\left\{PM^{\top}\Pi : \ M \in \mathcal{M}, \ \Pi \in \mathbb{R}^{n\times n} \text{ permutation matrix}\right\}$. *Moreover, S attains the same the expected payoff from communicating perspective $PM^{\top}\Pi$ or perspective $P\Pi$, for any $M \in \mathcal{M}$ and permutation matrix Π. Finally, some S-perspective and R-perspective exist.*

Proof. Recall, $P = [p_1 \cdots p_n]$ is a perspective if and only if $p_1, \ldots, p_n$ is a $\langle\cdot,\cdot\rangle_V$-orthonormal basis. Moreover, left multiplication by $\tilde{S} := SV$ is a self-adjoint operator with respect to $\langle\cdot,\cdot\rangle_V$, and P is an S-perspective if and only if the $\langle\cdot,\cdot\rangle_V$-orthonormal basis $p_1, \ldots, p_n$ consists of eigenvectors of $\tilde{S}$. In this case, the weights $(\sigma_i)_{i=1}^n$ associated with perspective P are exactly the eigenvalues associated with the eigenbasis:

$$\langle x, \tilde{S}x\rangle_V = \sum_{i=1}^{n} \sigma_i \langle x, p_i\rangle_V^2 \text{ for all } x \in \mathbb{R}^n.$$

The result now follows from the spectral theorem. The $\tilde{S}$-eigenspaces are pairwise $\langle\cdot,\cdot\rangle_V$-orthogonal and span $\mathbb{R}^n$, and P is an S-perspective and only if its columns consist of orthonormal bases of these eigenspaces. Each $J \in \mathfrak{J}$ corresponds to an eigenspace $\operatorname{span}\{p_j\}_{j\in J}$ whose associated eigenvalue we call (in a mild abuse) σ_J. Then some $\langle\cdot,\cdot\rangle_V$-orthonormal basis $\{p_i : \ i \in J\}$ exists for the σ_J eigenspace, and the set of all $\langle\cdot,\cdot\rangle_V$-orthonormal basese for it is the set of orthogonal linear combinations of $\{p_i : \ i \in J\}$. Hence, an S-perspective P exists, and the set of all S-perspectives is the set of all column permutations of eigenspace-by-eigenspace orthogonal transformations of P. Moreover, communicating *any* perspective k_J of whose first k columns belong to the σ_J eigenspace for each $J \in \mathfrak{J}$ yields (given Lemma 5) S expected payoff $\sum_{J\in\mathfrak{J}} k_J \sigma_J$. Therefore, the expected payoff from communicating $PM^{\top}\Pi$ for $M \in \mathcal{M}$ and permutation matrix Π is invariant to M.

Finally, the existence of an R-perspective follows from applying the above analysis to R instead of S. □

Lemma 7. *Given the n weights associated with some S-perspective, a unique S-perspective exists if and only if these n weights are distinct.*

Proof. This lemma follows directly from Lemma 6. If the weights are all distinct, then the set $\mathcal{M}$ (as described in the statement of that lemma) consists of all diagonal matrices with diagonal entries in $\{1, -1\}$; and if the weights are not all distinct, then $\mathcal{M}$ is infinite. □

Lemma 8. *Some R-extremal S-perspective exists.*

Proof. Given Lemma 6, some S-perspective P and R-perspective Q exist. Let $\mathfrak{J}$ and $\mathcal{M}$ be as defined in the statement of that lemma for the set of S-perspectives. Observe, a matrix $M \in \mathbb{R}^{n\times n}$ belongs to $\mathcal{M}$ if and only if there exist $(M_J)_{J\in\mathfrak{J}}$ such that M is block diagonal with $J \times J$ block M_J for each $J \in J$.

Let us set up some useful notation. Let $\rho = (\rho_i)_{i=1}^n \in \mathbb{R}^n$ be the weights associated with R-perspective Q, and let $\tilde{R} \in \mathbb{R}^{n\times n}$ denote the diagonal matrix with diagonal ρ. Let $G := \Re(P, Q) = P^\top V Q$, which is orthogonal. For any $J \in \mathfrak{J}$, let $G_J \in \mathbb{R}^{J\times n}$ be the submatrix of G consisting of its J rows, and define the symmetric matrix $\tilde{R}_J := G_J \tilde{R} G_J^\top \in \mathbb{R}^{J\times J}$. The spectral theorem delivers orthogonal $\tilde{M}_J \in \mathbb{R}^{J\times J}$ such that $\tilde{M}_J \tilde{R}_J \tilde{M}_J^\top$ is diagonal with diagonal vector $\tilde{x}_J \in \mathbb{R}^J$. Let $\tilde{x} = (\tilde{x}_J)_{J\in\mathfrak{J}} \in \mathbb{R}^n$, let $\tilde{M} \in \mathcal{M}$ have its $J\times J$ block equal to $\tilde{M}_J$ for each $J \in \mathfrak{J}$, and define the S-perspective $\tilde{P} := P\tilde{M}^\top$. We will show that the $\tilde{P}$ is R-extremal.

Consider any $J \in \mathfrak{J}$ and $M \in \mathcal{M}$ with $J \times J$ block M_J. By construction, the J rows of MG are given by $M_J G_J$. Therefore, any $i \in J$ has

$$
\begin{aligned}
\left[\Re^2(PM^\top, Q)\rho\right]_i &= \left[\Re(PM^\top, Q)\ \tilde{R}\ \Re(PM^\top, Q)^\top\right]_{ii} \\
&= \left[MG\tilde{R}(MG)^\top\right]_{ii} \\
&= \left[M_J G_J \tilde{R}(M_J G_J)^\top\right]_{ii} \\
&= \left[(M_J \tilde{M}_J^\top)(\tilde{M}_J \tilde{R}_J \tilde{M}_J^\top)(M_J \tilde{M}_J^\top)^\top\right]_{ii} \\
&= (B_J \tilde{x})_i,
\end{aligned}
$$

where $B_J \in \mathbb{R}^{J\times J}$ is the entrywise square of the orthogonal matrix $M_J \tilde{M}_J^\top$. Hence, $[\Re^2(PM^\top, Q)\rho]_J = B_J \tilde{x}_J$ is majorized by $\tilde{x}_J$ by the Schur-Horn theorem (Theorem 2.B.6, Marshall et al., 2011), and specializing this calculation to $M_J = \tilde{M}_J$ tells us $[\Re^2(\tilde{P}, Q)\rho]_J = \tilde{x}_J$. So $[\Re^2(\tilde{P}, Q)\rho]_J$ majorizes $[\Re^2(PM^\top, Q)\rho]_J$ for every $M \in \mathcal{M}$.

Now, take any $\hat{M} \in \mathcal{M}$ and permutation matrix $\hat{\Pi}$, and consider the strategy profile (for capacity k) that communicates perspective $P\hat{M}^\top\hat{\Pi}$—Lemma 6 tells us any S-perspective is of this form. We want to find some S-perspective, issue equivalent to $P\tilde{M}^\top\tilde{\Pi}$, such that communicating it yields the same S payoff as $P\hat{M}^\top\hat{\Pi}$ and a weakly higher [resp. lower] R payoff than $P\hat{M}^\top\hat{\Pi}$. To that end, let $\hat{P} := P\hat{M}^\top$ and $\hat{x} := \Re^2(\hat{P}, Q)\rho$, and let K be the set of all $i \in [n]$ such that some $j \in [k]$ has $\hat{\Pi}_{ij} = 1$. Let $\Pi \in \mathcal{M}$ be a permutation matrix such that, for each $J \in \mathfrak{J}$, every entry of $(\Pi\tilde{x})_{J\cap K}$ is weakly higher [resp. lower] than every entry of $(\Pi\tilde{x})_{J\setminus K}$; and let $\hat{\Pi} := \Pi\tilde{\Pi}$. Then,

$$\sum_{i=1}^{k} e_i^\top \left(\tilde{\Pi}^\top \tilde{x} - \hat{\Pi}^\top \hat{x}\right) = \sum_{i=1}^{k} (\hat{\Pi} e_i)^\top (\Pi\tilde{x} - \hat{x}) = \sum_{J\in\mathfrak{J}} \sum_{i\in J\cap K} (\Pi\tilde{x} - \hat{x})_i ,$$

which is nonnegative [resp. nonpositive] because $\tilde{x}_J$ majorizes $\hat{x}_J$ for every $J \in \mathfrak{J}$.

Finally, let us see that $\tilde{P}\tilde{\Pi}$ is an S-perspective that satisfies the desired payoff comparisons. First, observe $\tilde{P}\tilde{\Pi} = (P\tilde{M}^\top)(\Pi^\top\hat{\Pi}) = P(\Pi\tilde{M})^\top\hat{\Pi}$ and $\Pi\tilde{M} \in \mathcal{M}$. Hence, Lemma 6 tells us $\tilde{P}\tilde{\Pi}$ is an S-perspective, and that communicating it yields the same expected S payoff as communicating $\hat{P}\hat{\Pi} = P\hat{M}^\top\hat{\Pi}$. Toward the R payoff comparison, Lemma 5 says that sharing perspective $\hat{P}\hat{\Pi}$ yields an expected R payoff of

$$\sum_{i=1}^{k} e_i^\top \mathfrak{R}^2(\hat{P}\hat{\Pi}, Q)\rho = \sum_{i=1}^{k} e_i^\top \hat{\Pi}^\top \mathfrak{R}^2(\hat{P}, Q)\rho = \sum_{i=1}^{k} e_i^\top \hat{\Pi}^\top \hat{x},$$

and an identical calculation says sharing perspective $\tilde{P}\tilde{\Pi}$ yields an expected R payoff of $\sum_{i=1}^{k} e_i^\top \tilde{\Pi}^\top \tilde{x}$. The previous paragraph establishes the desired ranking of these two quantities, and so $\tilde{P}\tilde{\Pi}$ is as required. Therefore, $\tilde{P}$ is R-extremal. $\square$

Lemma 9. *Fix a vector $\mu \in \mathbb{R}^n$ and a perspective Q. Given $x \in \mathbb{R}^n$, vector μ majorizes x if and only if a perspective P exists such that $\mathfrak{R}^2(P, Q)\mu = x$.*

Proof. Let us observe that the set of possible $\mathfrak{R}^2(P, Q)$ matrices, as we range over all perspectives P, coincides with the set of all orthostochastic matrices—that is, the entrywise squares of orthogonal matrices. Indeed, defining the orthogonal matrix $\tilde{Q} := \sqrt{V}Q$, the former set is

$$\begin{aligned}
\{\mathfrak{R}(P, Q) :\ P \text{ a perspective}\} &= \left\{P^\top V Q :\ P \in \mathbb{R}^{n\times n} \text{ has } P^\top V P = I\right\} \\
&= \left\{P^\top \sqrt{V}\tilde{Q} :\ P \in \mathbb{R}^{n\times n} \text{ has } \sqrt{V}P \text{ orthogonal}\right\} \\
&= \left\{\tilde{P}^\top \tilde{Q} :\ \tilde{P} \in \mathbb{R}^{n\times n} \text{ orthogonal}\right\},
\end{aligned}$$

which is equal to the set of orthogonal matrices because the latter forms a group under multiplication. So $\{\mathfrak{R}^2(P, Q) :\ P \text{ a perspective}\}$ is the set of orthostochastic matrices.

Therefore, x admits a perspective P with $\mathfrak{R}^2(P, Q)\mu = x$ if and only if $x = M\mu$ for some orthostochastic matrix M. But then the Schur-Horn theorem (Theorem 2.B.6, Marshall et al., 2011) shows the latter condition is equivalent to μ majorizing x. □

Lemma 10. *Fix an R loss matrix R, and let $\rho = (\rho_i)_{i=1}^n$ be the weights associated with some R-perspective. Given $\mathcal{E} \subseteq \mathbb{R}$, the following are equivalent:*

(i) Some perspective P exists such that $\mathcal{E}$ is the set of R payoffs associated with strategy profiles communicating perspectives that are issue-equivalent to P. (Hence, $\mathcal{E}$ is finite and nonempty.)

(ii) Condition (i) holds, and is witnessed by a perspective that is the unique S-perspective (which is therefore R-extremal) for some S loss S.

(iii) Some $v \in \mathbb{R}^n$ majorized by ρ has $\mathcal{E} = \left\{\sum_{i\in J} v_i : \; J \subseteq [n] \text{ has } |J| = k\right\}$.

Moreover, given any S loss matrix S: the highest, lowest Pareto-optimal, and lowest equilibrium R payoffs all come from some set $\mathcal{E}$ satisfying the above equivalent conditions.

Proof. First, let us see (i) is equivalent to its strengthening (ii). We need only verify that any perspective P can be the unique S-perspective for some S loss S. And indeed, the S loss matrix given by $S = \sum_{j=1}^n jPe_je_j^\top P^\top$ admits P as an S-perspective with associated weights $(j)_{j=1}^n$. Thus, by Lemma 7, this S-perspective is unique, hence R-extremal.

Next, let us observe (i) is equivalent to (iii). To see it, let Q be the R-perspective with associated weights $(\rho_i)_{i=1}^n$. Lemma 5 tells us (i) is equivalent to the existence of some perspective P such that $v = \Re^2(P,Q)\rho$ has $\mathcal{E} = \left\{\sum_{i\in J} v_i : \; J \subseteq [n] \text{ has } |J| = k\right\}$. This condition is then equivalent to (iii) by Lemma 9.

Having established the three-way equivalence, we need only prove the last assertion. To that end, fix some S loss S. Let P be some R-extremal S-perspective (which exists by Lemma 8). Proposition 1 tells us all profiles that communicate a perspective issue-equivalent to P are equilibria. Proposition 1 also says every equilibrium communicates some S-perspective, and so weakly Pareto dominates one of these equilibria and so is weakly Pareto dominated by one of these equilibria. The claim follows. □

Proof of Proposition 3. Recall, for loss matrix R, the vector $(\rho_i)_{i=1}^n \in \mathbb{R}^n$ records the weights on the n issues according an R-perspective Q. Permuting columns of Q, we may assume $\rho_1 \geq \cdots \geq \rho_n$. Given Lemma 10, condition (i) holds if and only if some $v \in \mathbb{R}^n$ majorized by ρ has

$$\max_{J \subseteq [n]:\ |J|=k} \sum_{i \in J} v_i = \bar{r} \quad \text{and} \quad \min_{J \subseteq [n]:\ |J|=k} \sum_{i \in J} v_i = \underline{r}. \tag{4}$$

This condition is invariant to permuting the entries of v, so we can without loss of generality focus on v with $v_1 \geq \cdots \geq v_n$. In this case, the majorization constraint on v says $\sum_{i=1}^{m}(\rho_i - v_i)$ is nonnegative for every $m \in [n]$ and zero for $m = n$; and condition (4) reduces to $\bar{r} = \sum_{i=1}^{k} v_i$ and $\underline{r} = \sum_{i=n-k+1}^{n} v_i$. Now, let $\tilde{k} := \min\{k, n-k\}$. Replacing the first $\tilde{k}$ entries of v with their average, replacing the last $\tilde{k}$ entries of v with their average, and (if $k \neq \frac{n}{2}$) replacing all other entries with their average preserves the majorization constraint (and the property that $v_1 \geq \cdots \geq v_n$) and does not affect condition (4). We can therefore restrict attention to v that are weakly decreasing, and constant on each of these three sets of indices. Moreover, note that a v of this form is majorized by ρ if and only if it satisfies

$$\sum_{i=1}^{n} v_i = \sum_{i=1}^{n} \rho_i \text{ and } \sum_{i=1}^{m} v_i \leq \sum_{i=1}^{m} \rho_i \text{ for each } m \in \{k, n-k\}. \tag{5}$$

We can further simplify (5) by substituting in (4). To that end, let $\tilde{r}$ denote k times the average value of all of the v entries other than the first $\tilde{k}$ and last $\tilde{k}$ if $k \neq \frac{n}{2}$, and let $\tilde{r}$ be arbitrary otherwise. Then the equality constraint (4) is equivalent to $\underline{r}+\bar{r}+\frac{n-2k}{k}\tilde{r} = \frac{n}{k}r_M$, and (given this equality constraint) the two inequality constraints are equivalent to $\underline{r} \geq r_L$ and $\bar{r} \leq r_H$.[17]

Summarizing the conclusion of the previous paragraph, we have shown a pair $(\bar{r}, \underline{r}) \in \mathbb{R}^2$ satisfies (i) if and only if some $\tilde{r}$ has $r_L \leq \underline{r} \leq \tilde{r} \leq \bar{r} \leq r_H$ and $\underline{r} + \bar{r} + \frac{n-2k}{k}\tilde{r} = \frac{n}{k}r_M$. But now, a given $\underline{r} \leq \bar{r}$ admit some $\tilde{r}$ between them such that $\underline{r} + \bar{r} + \frac{n-2k}{k}\tilde{r} = \frac{n}{k}r_M$ if and only if $\frac{n}{k}r_M$ is between $\underline{r} + \bar{r} + \frac{n-2k}{k}\underline{r}$ and $\underline{r} + \bar{r} + \frac{n-2k}{k}\bar{r}$. Rearranging, this condition is the same as r_M lying between $\frac{k}{n}\underline{r} + \left(1 - \frac{k}{n}\right)\bar{r}$ and $\frac{k}{n}\bar{r} + \left(1 - \frac{k}{n}\right)\underline{r}$—which also requires that r_M lie between $\underline{r}$ and $\bar{r}$. The proposition follows. □

Proof of Corollary 2. The equivalence of the first two conditions holds because the

[17] To see this form of the equality constraint, observe that $\frac{n}{k}r_M$ is the sum of all entries of ρ and $\underline{r} + \bar{r}$ is the sum of the first k and last k entries of v. The term $\frac{n-2k}{k}\tilde{r}$ can be any sign, and adjusts the sum of v entries for the $|n-2k|$ entries in the middle that are uncounted or double-counted.

range of β specified in Proposition 3 belongs to $(0, 1)$.

Now, for any perspective P, the vector $\Re^2(P, Q)\rho$ has its entries averaging to $\frac{1}{k} r_M$ because $\Re^2(P, Q)$ is stochastic. The equivalence of the third item and (say) the first therefore follows directly from Proposition 2. $\square$

Definition 3. Given a subset X of a Euclidean space, say a subset of X is **generic** in X if it contains an open and dense subset X, whose complement in the affine span of X is Lebesgue null in that affine span.

Notation 3. *Let $\mathbb{L}$ denote the set of positive definite matrices, a convex open set of the Euclidean space of pairs of symmetric $n \times n$ matrices. Say a property of matrix R or S [resp. pairs (S, R)] holds* ***generically*** *if the set of matrices [resp. pairs] for which it holds is generic in $\mathbb{L}$ [resp. $\mathbb{L}^2$].*

Remark 1. *Note that the definition of genericity in Definition 3 is stronger than measure theoretic genericity (i.e., its complement being Lebesgue null in the ambient affine space) and stronger than topological genericity (containing a countable intersection of open and dense sets). It follows that all our results that invoke generic parameter choices will continue to hold if one replaces our definition with either of these notions of genericity.*

Lemma 11. *If X is a semialgebraic subset of some Euclidean space, then any semialgebraic dense subset of X is generic in X. In particular, any semialgebraic dense subset of $\mathbb{L}$ or $\mathbb{L}^2$ is generic.*

Proof. Let X_1 be any dense semialgebraic set in X, and let $X_0 := X \setminus X_1$. Throughout this proof, we refer to definitions and results from Van den Dries (1998)—hereafter vdD98. By cell decomposition (Proposition 3.2.11 of vdD98, which applies by Corollary 2.2.11 of vdD98), each of X_1 and X_0 can be represented as a disjoint union of finitely many cells (as in Definition 2.2.3 of vdD98) defined by polynomial functions. As noted in vdD98 immediately after this definition, a semialgebraic set has the same dimension (in the sense of Definition 4.1.1 from vdD98) as the underlying ambient linear space if and only if it contains a full-dimensional cell. Hence, because X_1 is dense in X and semialgebraic, it must be that all of the cells comprising X_0 have strictly lower dimension than X. Each low-dimensional cell in the decomposition therefore has low-dimensional closure too by vdD98's Theorem 4.1.8, and so is nowhere dense (because it has empty interior and is open in its closure) and Lebesgue null (Lemma

7.25 Rudin, 1987). The union of all full-dimensional cells in the decomposition is then a subset of X_1 as required for the first assertion.

To see the second assertion, it then remains to note the convex set $\mathbb{L}$ (hence also $\mathbb{L}^2$) is semialgebraic. And indeed, S is positive definite matrix if and only if $x^\top S x > 0$ for every nonzero $x \in \mathbb{R}^n$. $\square$

Lemma 12. *For a generic loss matrix S, the n weights associated with some S-perspective are distinct.*

Proof. Let $\mathbb{L}_1 \subseteq \mathbb{L}$ denote the set of $S \in \mathbb{L}$ with distinct weights. Because $\mathbb{L}_1$ is semialgebraic, Lemma 11 tells us it suffices to see an arbitrary $S \in \mathbb{L}$ is a limit of members of $\mathbb{L}_1$. To that end, take some S-perspective P (which exists by Lemma 6), witnessed by weights $\sigma_1, \dots, \sigma_n$. We can then write $S = \sum_{j=1}^n \sigma_j P e_j e_j^\top P^\top$. For $\varepsilon > 0$, define the positive definite matrix

$$S^\varepsilon = \sum_{j=1}^n (\sigma_j + j\varepsilon) P e_j e_j^\top P^\top,$$

which converges to S as $\varepsilon \to 0$. By construction, P is a S-perspective with weights $(\sigma_j + j\varepsilon)_{j=1}^n$. Meanwhile, whenever $\varepsilon > 0$ is small enough, the n numbers $(\sigma_j + j\varepsilon)_{j=1}^n$ are all distinct, so that $S^\varepsilon \in \mathbb{L}_1$. $\square$

Proof of Corollary 3. Let Q is an R-perspective with weights $(\rho_i)_{i=1}^n$, and let r_L, r_M, r_H be as defined in the preamble to Proposition 3. Say a pair $(\bar{r}, \underline{r})$ is payoff feasible if it satisfies the two equivalent conditions of that proposition—or equivalently, some S loss S exists such that $\bar{r}$ and $\underline{r}$ are R's best and worst equilibrium payoffs, respectively. In what follows, we detail the precise conditions for each of (i) and (ii) to hold, and then we argue that both conditions hold for generic R.[18]

First, if $k = \frac{1}{2}n$, then Proposition 3 tells us any payoff-feasible $(\bar{r}, \underline{r})$ has $\frac{1}{2}\bar{r} + \frac{1}{2}\underline{r} = r_M$, and so no two payoff-feasible pairs are coordinate-wise ranked, meaning neither (i) nor (ii) holds.

Now, using Proposition 3, let us see (i) is equivalent to having $k \neq \frac{n}{2}$ and ρ not being a constant vector. Given the previous paragraph, we can focus on the case of $k \neq \frac{1}{2}n$. If ρ is a constant vector, then $(r_M, r_M) = (r_H, r_L)$ is the unique payoff-feasible pair, so (i) fails. Conversely, if ρ is not a constant vector, then $r_L < r_M < r_H$, and

[18]The proof of the genericity result can be simplified by not characterizing the exact conditions for each of (i) and (ii). We present the current version for completeness.

$\mathrm{co}\left\{\frac{k}{n}, 1-\frac{k}{n}\right\}$ contains a neighborhood of $\frac{1}{2}$. So take any $\varepsilon > 0$ small enough that $\varepsilon < r_M - r_L$ and $\varepsilon < r_H - r_M$. Then, any payoff pair in a small enough neighborhood of $(r_M + \varepsilon, r_M - \varepsilon)$ is payoff feasible, delivering (i).

Next, we argue that (ii) is equivalent to either having $k < \frac{n}{2}$ and the last $n-k$ entries of ρ all coinciding, or having $k > \frac{n}{2}$ and the first k entries of ρ all coinciding. Given the paragraph before the previous one, we can focus on the case of $k \neq \frac{1}{2}n$. Now, let $\tilde{k} := \min\{k, n-k\}$. Then, by Proposition 3, a given $\underline{r} \in \mathbb{R}$ is such that $(r_H, \underline{r})$ is payoff feasible if and only if $\underline{r} \in [r_L, r_M]$ and some $\beta \in \left[\frac{\tilde{k}}{n}, 1-\frac{\tilde{k}}{n}\right]$ has $r_M = (1-\beta)\underline{r} + \beta r_H$. So (i) holds if and only if some $\underline{r} > r_L$ satisfies these conditions. We know (r_H, r_L) is payoff feasible because (Proposition 2) the best and worst equilibrium payoffs for sender $S = R$ are r_H and r_L. We can then raise r_L and maintain payoff feasibility if and only if β can be correspondingly lowered to maintain $r_M = (1-\beta)\underline{r} + \beta r_H$. So (i) holds if and only if

$$r_M > \left(1 - \frac{\tilde{k}}{n}\right) r_L + \frac{\tilde{k}}{n} r_H.$$

If $k < \frac{n}{2}$ [resp. $k > \frac{n}{2}$] then this inequality rearranges to say that the average of the last $n-k$ [resp. first k] entries of ρ is strictly greater than the average of the last k [resp. first $n-k$] entries of ρ; this condition is in turn equivalent to the last $n-k$ [resp. first k] entries of ρ all coinciding.

Finally, we turn to the genericity claim. If $k \neq \frac{n}{2}$, then $\max\{k, n-k\} \geq 2$. Hence, given the above characterizations, failure of either of (i) or (ii) requires ρ to have at least two entries in common. The corollary then follows from applying Lemma 12 to R. □

Proof of Proposition 4. In what follows, let $M_1 := \mathfrak{R}^2(P_1, Q)$ and $M_2 := \mathfrak{R}^2(P_2, Q)$; and let $\mathcal{K} := \{\kappa \in \{0,1\}^n : \sum_{i=1}^n \kappa_i = k\}$, the set of weight vectors that put weight 1 on k components and 0 on $n-k$.

By assumption, for either $\ell \in \{1, 2\}$, the matrix S_ℓ admits a unique perspective, meaning any perspective is obtained from P_ℓ by permuting columns and multiplying some by -1. Multiplying a column of perspective P by -1 leaves $\mathfrak{R}^2(P, Q)$ unchanged, whereas permuting columns of P permutes the rows of $\mathfrak{R}^2(P, Q)$ in the same way. Therefore, in light of Proposition 1 and Lemma 5, if the weights associated with R-perspective Q are $(\rho_i)_{i=1}^n \in \mathbb{R}^n$, then $\bar{u}_R^k(S_\ell, R) = \max_{\kappa \in \mathcal{K}}\left\{\kappa^\top M_\ell \rho\right\}$ and $\underline{u}_R^k(S_\ell, R) = \min_{\kappa \in \mathcal{K}}\left\{\kappa^\top M_\ell \rho\right\} = -\max_{\kappa \in \mathcal{K}}\left\{-\kappa^\top M_\ell \rho\right\}$.

Now, say $S_1 \succsim_+ S_2$ if every R with perspective Q has $\bar{u}^k_R(S_1, R) \geq \bar{u}^k_R(S_2, R)$; and say $S_1 \succsim_- S_2$ if every R with perspective Q has $\underline{u}^k_R(S_1, R) \leq \underline{u}^k_R(S_2, R)$. Observe,

$$\begin{aligned}
S_1 \succsim_\pm S_2 \quad &\Longleftrightarrow \quad \text{every } \rho \in \mathbb{R}^n_{++} \text{ has } \max_{\kappa \in \mathcal{K}} \left\{ \pm \kappa^\top M_1 \rho \right\} \geq \max_{\lambda \in \mathcal{K}} \left\{ \pm \lambda^\top M_2 \rho \right\} \\
&\Longleftrightarrow \quad \text{every } \lambda \in \mathcal{K} \text{ has } \inf_{\rho \in \mathbb{R}^n_{++}} \max_{\kappa \in \mathcal{K}} \left\{ \pm \left(\kappa^\top M_1 - \lambda^\top M_2 \right) \rho \right\} \geq 0 \\
&\Longleftrightarrow \quad \text{every } \lambda \in \mathcal{K} \text{ has } \max_{\kappa \in \mathrm{co}\mathcal{K}} \inf_{\rho \in \mathbb{R}^n_{++}} \left\{ \pm \left(\kappa^\top M_1 - \lambda^\top M_2 \right) \rho \right\} \geq 0 \\
&\Longleftrightarrow \quad \text{every } \lambda \in \mathcal{K} \text{ admits } \kappa \in \mathrm{co}\mathcal{K} \text{ with } \pm \left(\kappa^\top M_1 - \lambda^\top M_2 \right) \geq 0,
\end{aligned}$$

where the third equivalence follows from Sion's minimax theorem. But next note that the matrices M_1, M_2 are bistochastic, hence row stochastic. Therefore, letting $e \in \mathbb{R}^n$ be the vector with all entries 1, any $\kappa, \lambda \in \mathrm{co}\mathcal{K}$ have

$$\pm \left(\kappa^\top M_1 - \lambda^\top M_2 \right) e = \pm \left(\kappa^\top e - \lambda^\top e \right) = \pm(k - k) = 0.$$

Thus, the only way the vector inequality $\pm \left(\kappa^\top M_1 - \lambda^\top M_2 \right) \geq 0$ can hold is if the left-hand side is the zero vector, that is, $\kappa^\top M_1 = \lambda^\top M_2$.

So we have shown conditions $S_1 \succsim_+ S_2$ and $S_1 \succsim_- S_2$ are each equivalent to the condition that every $\lambda \in \mathcal{K}$ admits $\kappa \in \mathrm{co}\mathcal{K}$ with $\kappa^\top M_1 = \lambda^\top M_2$, as required. □

Lemma 13. *Suppose Q is an R-perspective with associated weights $\rho = (\rho_i)_{i=1}^n$, perspective P is an R-extremal S-perspective such that $v := \mathfrak{R}^2(P, Q)\rho$ have $v_1 \geq \cdots \geq v_n$, and $\hat{v} := \frac{1}{n} \sum_{i=1}^n v_i = \frac{1}{n} \sum_{i=1}^n \rho_i$.*

(i) Under R-best equilibrium selection: capacity n is optimal if and only if $\gamma \leq v_n$, capacity 0 is optimal if and only if $\gamma \geq v_1$, and capacity $k \in \{1, \ldots, n-1\}$ is optimal if and only if $v_k \geq \gamma \geq v_{k+1}$.

(ii) Under R-worst equilibrium selection: capacity n is optimal if and only if $\gamma \leq \hat{v}$, capacity 0 is optimal if and only if $\gamma \geq \hat{v}$, and capacity $k \in \{1, \ldots, n-1\}$ is optimal if and only if $\gamma = v_1 = \cdots = v_n = \hat{v}$.

Proof. For any given $k \in \{0, \ldots, n\}$, Proposition 1 and P being R-extremal tell us a best and worst equilibrium for R take the form of communicating whichever perspective R most prefers among those that are issue-equivalent to P. By Lemma 5, then, the best and worst R equilibrium payoff given capacity k are $\sum_{i=1}^k v_i$ and $\sum_{i=n-k+1}^n v_i$, respectively.[19] Because $v_1 \geq \cdots \geq v_n$, these sums are weakly concave in k and weakly convex in k, respectively, and the latter is not affine in k unless v is a constant vector. Because the investment cost is linear in k, the observations of the previous sentence apply identically to the best and worst R equilibrium payoffs net of investment costs.

Consider the case of R-best equilibrium selection. Because R's net payoff is weakly concave in investment, an investment level is optimal if and only if R cannot profitably increase or decrease investment by 1. Meanwhile, the net payoff difference between investment level $\tilde{k} \in [n]$ and investment level $\tilde{k}-1$ is $v_{\tilde{k}} - \gamma$, so the optimality characterization in the lemma follows.

Turn now to the case of R-worst equilibrium selection. Because R's net payoff is weakly convex in investment, one of the extreme investment levels $\{0, n\}$ is optimal. The payoff difference between investment levels n and 0 is $n(\hat{v} - \gamma)$. Hence, the stated characterization of when each of $\{0, n\}$ is optimal follows. Finally, given that the objective is weakly convex on its scalar domain, no interior investment level is optimal unless this objective is constant (in which case all are optimal). But the latter condition holds if and only if $\gamma = v_1 = \cdots = v_n$, in which case this quantity also coincides with $\hat{v}$. The lemma follows. □

Lemma 14. *Given R, take any R-perspective Q with associated weights $(\rho_i)_{i=1}^n$. Suppose R is not proportional to V^{-1} (equivalently $\{\rho_i\}_{i=1}^n$ do not all coincide). Then, generic $S \in \mathbb{L}$ has a unique S-perspective P, which has all entries of $v = \Re^2(P, Q)\rho$ distinct.*

Proof. Let $\mathcal{P} \subseteq \mathbb{R}^{n\times n}$ denote the set of all perspectives. For any distinct $i, j \in [n]$, define $\varphi_{ij} := (e_i - e_j)^\top \Re^2(\cdot, Q)\rho : \mathcal{P} \to \mathbb{R}$.

Below, we will establish that φ_{ij} is not zero on any nonempty neighborhood for any distinct $i, j \in [n]$. Let us first see how this would deliver the lemma. In this case, $\mathcal{P} \setminus \varphi_{ij}^{-1}(0)$ is dense in $\mathcal{P}$, and so every $S \in \mathbb{L}$ with unique perspective is in the closure of the set of such loss matrices with perspective in $\mathcal{P} \setminus \varphi_{ij}^{-1}(0)$.[20] But then, Lemma 7 and Lemma 12 imply the matrices in $\mathbb{L}$ whose unique perspective lives in $\mathcal{P} \setminus \varphi_{ij}^{-1}(0)$ is dense in $\mathbb{L}$, and so (since it is semialgebraic) generic by Lemma 11. But

[19] Although that lemma and its inputs assume a capacity in $\{1, \ldots, n-1\}$, the proofs apply identically to the cases of $\{0, n\}$. Alternatively, the payoff formulas are trivial to prove for these special cases, given the equilibrium characterizations described in Section 2.

[20] Holding the n distinct weights fixed and perturbing the perspective amounts to a small perturbation of the loss matrix.

a finite intersection of generic sets is generic; hence, for a generic set of matrices in $\mathbb{L}$, the unique perspective is in $\mathcal{P} \setminus \bigcup_{i,j\in[n]:\ i\neq j} \varphi_{ij}^{-1}(0)$. But any P in this set is such that $\Re^2(P,Q)\rho$ has all entries distinct, by definition.

So fixing distinct $i, j \in [n]$, it remains to see $\varphi = \varphi_{ij}$ is not zero on any neighborhood. To that end, let $\mathcal{P}_0$ be the interior of $\varphi^{-1}(0)$ in $\mathcal{P}$, which is open by construction; and let $\mathcal{P}_+$ comprise the positive-determinant members of $\mathcal{P}$.

Now, we argue $\mathcal{P}_0$ is closed in $\mathcal{P}$. Consider any P in its closure. Because $\mathcal{P}$ is the zero set of the polynomial $\tilde{P} \mapsto \tilde{P}^\top V \tilde{P} - I$ on $\mathbb{R}^{n\times n}$, the real-analytic implicit function theorem (Theorem 2.3.5, Krantz and Parks, 2002) yields an open neighborhood $\mathcal{N}$ of a Euclidean space and an analytic map from said neighborhood onto a neighborhood of P in $\mathcal{P}$—without loss, say $\mathcal{N}$ is convex. The composition of this analytic map with the polynomial φ is therefore analytic too (Proposition 2.2.8, Krantz and Parks, 2002). But because P is in the closure of $\mathcal{P}_0$, every element of the convex set $\mathcal{N}$ is on some line that intersects the open neighborhood in $\mathcal{N}$ on which this composed function is zero. Applying the identity theorem (Corollary 1.2.7 Krantz and Parks, 2002), it follows that the composed function is globally zero on $\mathcal{N}$. Therefore, φ is zero in a neighborhood of P, meaning $P \in \mathcal{P}_0$ too.

Next, let us see that $\mathcal{P}_+$ is either disjoint from $\mathcal{P}_0$ or contained in it. Indeed, because the positive definite matrix $\sqrt{V}$ has strictly positive determinant, and the self-map on $\mathbb{R}^{n\times n}$ given by left-multiplication by $\sqrt{V}$ is an invertible linear map that maps $\mathcal{P}$ onto the orthogonal group $O(n,\mathbb{R})$, it follows that this linear map is a homeomorphism between $\mathcal{P}_+$ and the special orthogonal group. Because the latter is connected (Example 9.14, Baker, 2003), it follows that $\mathcal{P}_+$ is. But in light of the previous paragraph (and the definition), $\mathcal{P}_0$ is clopen, implying $\mathcal{P}_0 \cap \mathcal{P}_+$ is clopen in $\mathcal{P}_+$. As the only clopen sets in $\mathcal{P}_+$ are $\mathcal{P}_+$ itself and the empty set, the claim follows.

But now, observe that the self-map $\mathcal{P}$ that multiplies the first column by -1 is a homeomorphism that maps the open set $\mathcal{P}_+$ onto its open complement, and that preserves the value of φ. Therefore, $\mathcal{P}_0$ is invariant under this self-map. Hence, the conclusion of the previous paragraph implies $\mathcal{P}_0$ is either the empty set or all of $\mathcal{P}$.

Given the above paragraph, we will know $\mathcal{P}_0$ is empty if we determine it is not equal to $\mathcal{P}$. To see this fact, note that for any permutation matrix $\Pi \in \mathbb{R}^{n\times n}$, the perspective $Q\Pi$ has $\varphi(Q\Pi) = (e_i - e_j)^\top \Pi^\top \Re^2(Q,Q)\rho = [\Pi(e_i - e_j)]^\top \rho$. But then, because ρ is not a constant vector, some appropriate choice of this permutation yields $\varphi(Q\Pi) \neq 0$. Hence, $Q\Pi \in \mathcal{P} \setminus \mathcal{P}_0$, delivering the lemma. $\square$

Proof of Proposition 5. Letting $v, \hat{v}$ be as given in the statement of Lemma 13, that lemma tells us the cutoffs $(\gamma_H, \gamma_M, \gamma_L) := (v_1, \hat{v}, v_n)$ satisfy all three items in the list. Moreover, $\gamma_H \geq \gamma_M \geq \gamma_L > 0$ in general, with all inequalities strict if $v_1 > v_n$.

It remains to see that, generically, $v_1 > v_n$ for some R-extremal S-perspective. First, Lemma 12 tells us a dense set of $R \in \mathbb{L}$ admit an R-perspective which has associated weights $\rho = (\rho_i)_{i=1}^n$ that are all distinct. Second, by Lemma 14, for any R of this form, a generic set of $S \in \mathbb{L}$ admit a unique S-perspective (which is then R-extremal) such that $\Re^2(P, Q)\rho$ is not a constant vector. In particular, the set of pairs $(S, R) \in \mathbb{L}^2$ each with distinct weights and yielding the desired $v_1 > v_n$ property is dense in $\mathbb{L}^2$. Being semialgebraic by definition, it is generic by Lemma 11 as desired. $\square$

Proof of Proposition 6. Let $M_1 := \Re^2(P_1, Q)$ and $M_2 := \Re^2(P_2, Q)$. Given Proposition 1 and Proposition 2, the R-best equilibrium value differs for S losses S_1 and S_2 for some capacity if and only if some $k \in [n-1]$ exists such that the sum of the k highest entries of $M_1\rho$ and $M_2\rho$ are different. Because M_1 and M_2 being bistochastic means these two vectors have the same sum of entries ($\sum_{i=1}^n \rho_i$), this condition is equivalent to $M_1\rho$ and $M_2\rho$ not being rearrangements of one another. So we must show that generic $\rho \in \mathbb{R}^n_{++}$ have $M_1\rho \neq \Pi M_2\rho$ for every permutation matrix $\Pi \in \mathbb{R}^{n\times n}$, and that all such ρ admit the desired thresholds.

First, let us establish the genericity property. For any permutation matrix $\Pi \in \mathbb{R}^{n\times n}$, note that $\ker(M_1 - \Pi M_2)$ is a linear subspace of $\mathbb{R}^n$, and is a proper subspace because $P_1 \not\sim_Q P_2$. But then, because $\mathbb{R}^n_{++}$ is a convex open subset of $\mathbb{R}^n$, it follows that $\mathbb{R}^n_{++} \setminus \ker(M_1 - \Pi M_2)$ is open and dense in $\mathbb{R}^n_{++}$ with Lebesgue-null complement in it—so, generic in it. Therefore, the intersection of this finite collection of sets (ranging over the finitely many permutation matrices) is also generic in $\mathbb{R}^n_{++}$. In what follows, fix some ρ belonging to this generic set.

Let v^1 and v^2 be the rearrangements of $M_1\rho$ and $M_2\rho$, respectively, with $v^1_1 \geq \cdots \geq v^1_n$ and $v^2_1 \geq \cdots \geq v^2_n$. We have $v^1 \neq v^2$ by definition of ρ, and v^1 majorizes v^2 because $P_1 \succsim_Q P_2$. Therefore, every $k \in [n-1]$ has $\sum_{i=1}^k (v^1_i - v^2_i) \geq 0$, but the inequality cannot hold with equality for every k. So let $\underline{k}$ and $\bar{k}_0$ be the first and last index, respectively, in $[n-1]$ for which the inequality holds strictly, and let $\bar{k} := 1 + \bar{k}_0 \in [n]$. Now, let $(\gamma_{HH}, \gamma_H, \gamma_L, \gamma_{LL}) := (v^1_{\underline{k}}, v^2_{\underline{k}}, v^2_{\bar{k}}, v^1_{\bar{k}})$. The definitions of $\underline{k}$ and $\bar{k}$ tell us every $k \in [n]$ with $k < \underline{k}$ or $k > \bar{k}$ has $v^1_{\underline{k}} = v^2_{\underline{k}}$, and hence $\gamma_{HH} \neq \gamma_H \geq \gamma_L \neq \gamma_{LL}$. It then follows from the majorization inequalities that $\gamma_{HH} > \gamma_H \geq \gamma_L > \gamma_{LL}$. We can now apply Lemma 13 in each of these three regions. Costs in (γ_{LL}, γ_L) yield $k_2^* \geq \bar{k} > k_1^*$; costs in (γ_H, γ_{HH}) yield $k_2^* < \underline{k} \leq k_1^*$; and for costs outside of $[\gamma_{LL}, \gamma_{HH}]$, optimal R-best-equilibrium investment is determined by the exact same first-order conditions for both S_1 and S_2. $\square$

B Proofs for Discussion Section

Lemma 15. *Every S-projection of rank k is equal to DC for some R strategy D and S best response C; and every such strategy profile yields such an S-projection.*

Proof. The second assertion is proven in Lemma 1. Toward the first assertion, consider any S-projection A with rank k. Because it has rank k, some $D \in \mathbb{R}^{n\times k}$ exists with the same range—let the columns be any basis for said range. Lemma 1 tells us some S best response C exists, and this map is a rank-k linear map such that DC is an S-projection. But then, $DC = A$ as required, because an orthogonal projection on any inner-product space is determined by its range. $\square$

Proof of Proposition 7. In what follows, call our main model *the original game*, and let *the common-interest game* refer to a modification of our main model in which R too has loss matrix S.

Below, we will establish that some S-preferred strategy profile exists; first, let us see the equivalence assuming existence. Indeed, in this case, a pure strategy profile is S-preferred if and only if it is a best pure-strategy Nash equilibrium of the common-interest game (in normal form). But both in the common-interest game and in the original game, a strategy profile is a Nash equilibrium of the normal form if and only if it is outcome equivalent to a Bayes Nash equilibrium. The equivalence then follows from Corollary 1.

All that remains is existence of an S-preferred strategy profile. By Lemma 1, every R strategy admits an S-best response. So we can without loss show optimality only among strategy profiles in which S is best responding. Given Lemma 15 and Lemma 2, we need only see an optimum exists to the program

$$\max_{A \text{ an } S\text{-projection of rank } k} \operatorname{Tr}\left[\left(2SA - A^\top SA\right) V\right].$$

Since the domain is compact and the objective continuous, some optimum exists. $\square$

Lemma 16. *Some optimum exists in the delegation problem, and every optimum has linear S strategy. If (C, D) is an optimum, and $A = DC \in \mathbb{R}^{n\times n}$, then*

$$A^\top \left[R(I-A)VS + SV(I-A^\top)R\right](1-A) = 0.$$

Hence, if A is a V^{-1}-projection, then it is also an R-projection.

Proof. Any feasible profile, hence any optimum, in the delegation problem has linear S strategy by Lemma 1. Given Lemma 15, a committed R can induce $\mathbf{a}$ if and only if $\mathbf{a} = A\boldsymbol{\omega}$ for some S-projection A of rank k. Her payoff from such a strategy profile is $\mathrm{Tr}\left[\left(2RA - A^\top RA\right)V\right]$ by Lemma 2. She therefore solves the program

$$\max_{A \text{ an } S\text{-projection of rank } k} \mathrm{Tr}\left[\left(2RA - A^\top RA\right)V\right].$$

Since the domain is compact and the objective continuous, some optimum exists.

Now, define

$$(\tilde{A}, \tilde{V}, \tilde{R}) = (\sqrt{S}A\sqrt{S}^{-1},\ \sqrt{S}V\sqrt{S},\ \sqrt{S}^{-1}R\sqrt{S}^{-1}).$$

The matrices $\tilde{V}$ and $\tilde{R}$ are symmetric, and A is S-orthogonal if and only if $\tilde{A}$ is an orthogonal matrix. Hence,

$$\begin{aligned}
\mathrm{Tr}\left[\left(2RA - A^\top RA\right)V\right] &= \mathrm{Tr}\left[\sqrt{S}V\left(2RA - A^\top RA\right)\sqrt{S}^{-1}\right] \\
&= \mathrm{Tr}\left(2\tilde{V}\tilde{R}\tilde{A} - \tilde{V}\tilde{A}\tilde{R}\tilde{A}\right) = \mathrm{Tr}\left(2\tilde{V}\tilde{R}\tilde{A} - \tilde{V}\tilde{A}\tilde{R}\tilde{A}^\top\right)
\end{aligned}$$

Because any orthogonal matrix can be conjugated by $\sqrt{S}$ to get an S-projection, any optimum for the delegation problem then maximizes $\mathrm{Tr}\left(2\tilde{V}\tilde{R}\tilde{A} - \tilde{V}\tilde{A}\tilde{R}\tilde{A}^\top\right)$ over all rank-k orthogonal projection matrices. Now, for any skew-symmetric matrix $Z \in \mathbb{R}^{n\times n}$, the matrix exponential e^Z is orthogonal, and so too is $e^{-\beta Z}\tilde{A}e^{\beta Z}$ for each $\beta \in \mathbb{R}$. Optimality then requires $\mathrm{Tr}\left(2\tilde{V}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z} - \tilde{V}e^{\beta Z}\tilde{A}e^{\beta Z}\tilde{R}e^{-\beta Z}\tilde{A}e^{-\beta Z}\right)$ to be maximized at $\beta = 0$. In particular, it requires the first-order condition

$$\left.\frac{\partial}{\partial\beta}\right|_{\beta=0} \mathrm{Tr}\left(2\tilde{V}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z} - \tilde{V}e^{-\beta Z}\tilde{A}e^{\beta Z}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z}\right) = 0 \tag{6}$$

to hold. To write (6) more explicitly, note $e^{0Z} = I$ and $\frac{\partial}{\partial\beta}\Big|_{\beta=0} e^{\beta Z} = Z$, so that

$$\frac{\partial}{\partial\beta}\Big|_{\beta=0} \left[e^{-\beta Z}\tilde{A}e^{\beta Z}\right] = I\tilde{A}Z + (-Z)\tilde{A}I = [\tilde{A}, Z].$$

Therefore,

$$\begin{aligned} & \frac{\partial}{\partial\beta}\Big|_{\beta=0} \left[2\tilde{V}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z} - \tilde{V}e^{-\beta Z}\tilde{A}e^{\beta Z}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z}\right] \\ = \;& 2\tilde{V}\tilde{R}[\tilde{A}, Z] - \tilde{V}\left\{(\tilde{A})\tilde{R}[\tilde{A}, Z] + [\tilde{A}, Z]\tilde{R}(\tilde{A})\right\} \\ = \;& 2\tilde{V}\tilde{R}(\tilde{A}Z - Z\tilde{A}) - \tilde{V}\tilde{A}\tilde{R}(\tilde{A}Z - Z\tilde{A}) - \tilde{V}(\tilde{A}Z - Z\tilde{A})\tilde{R}\tilde{A}, \end{aligned}$$

and hence skew symmetry of Z implies

$$\begin{aligned} & \frac{\partial}{\partial\beta}\Big|_{\beta=0} \mathrm{Tr}\left(2\tilde{V}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z} - \tilde{V}e^{-\beta Z}\tilde{A}e^{\beta Z}\tilde{R}e^{-\beta Z}\tilde{A}e^{\beta Z}\right) \\ = \;& \mathrm{Tr}\left[2\tilde{V}\tilde{R}(\tilde{A}Z - Z\tilde{A}) - \tilde{V}\tilde{A}\tilde{R}(\tilde{A}Z - Z\tilde{A}) - \tilde{V}(\tilde{A}Z - Z\tilde{A})\tilde{R}\tilde{A}\right] \\ = \;& \mathrm{Tr}\left[\left(2\tilde{V}\tilde{R}\tilde{A} - 2\tilde{A}\tilde{V}\tilde{R} - \tilde{V}\tilde{A}\tilde{R}\tilde{A} + \tilde{A}\tilde{V}\tilde{A}\tilde{R} - \tilde{R}\tilde{A}\tilde{V}\tilde{A} + \tilde{A}\tilde{R}\tilde{A}\tilde{V}\right)Z\right] \\ = \;& \mathrm{Tr}\left[\left(-2\tilde{A}\tilde{R}\tilde{V} - 2\tilde{A}\tilde{V}\tilde{R} + \tilde{A}\tilde{R}\tilde{A}\tilde{V} + \tilde{A}\tilde{V}\tilde{A}\tilde{R} + \tilde{A}\tilde{V}\tilde{A}\tilde{R} + \tilde{A}\tilde{R}\tilde{A}\tilde{V}\right)Z\right] \\ = \;& -2\,\mathrm{Tr}\left\{\tilde{A}\left[\tilde{R}(I - \tilde{A})\tilde{V} + \tilde{V}(I - \tilde{A})\tilde{R}\right]Z\right\}. \end{aligned}$$

Requiring (6) to hold for every skew-symmetric matrix Z then amounts to requiring that matrix $\tilde{A}M$ be symmetric, where we define the symmetric matrix

$$M := \tilde{R}(I - \tilde{A})\tilde{V} + \tilde{V}(I - \tilde{A})\tilde{R}.$$

But then symmetry of $\tilde{A}M$ equivalently says M commutes with the orthogonal projection $\tilde{A}$, so that $\tilde{A}M(I - \tilde{A}) = M\tilde{A}(I - \tilde{A}) = 0$. Hence,

$$\begin{aligned} 0 \;=\; & \sqrt{S}\tilde{A}M(I - \tilde{A})\sqrt{S} \\ =\; & \sqrt{S}\tilde{A}\left[\tilde{R}(I - \tilde{A})\tilde{V} + \tilde{V}(I - \tilde{A})\tilde{R}\right](I - \tilde{A})\sqrt{S} \\ =\; & SAS^{-1}\left[R(I - A)VS + SV(I - SAS^{-1})R\right](I - A) \\ =\; & A^\top\left[R(I - A)VS + SV(I - A^\top)R\right](I - A), \end{aligned}$$

where the last equality follows from A being an S-projection.

Finally, toward the last assertion, suppose A is a V^{-1}-projection in addition to being an S-projection, and is optimal in the delegation problem. In this case, $AVS = VA^\top S = VSA$, so that

$$\begin{aligned} 0 &= A^\top \left[R(I-A)VS + SV(I-A^\top)R\right](I-A) \\ &= A^\top RVS(I-A)(I-A) + A^\top(I-A^\top)SVR(I-A) \\ &= A^\top RVS(I-A) + 0SVR(I-A) && \text{(since } A \text{ is idempotent)} \\ &= A^\top R(I-A)VS + 0. \end{aligned}$$

But then VS being invertible means $A^\top R(I-A) = 0$. So A is an R-projection. □

Proof of Proposition 8. We begin with some useful notation. Let $\mathbb{L}_1$ be the set of all $S \in \mathbb{L}$ such that an S-perspective is witnessed by n distinct weights. Define the correspondence $\mathcal{A} : \mathbb{L} \rightrightarrows \mathbb{R}^{n\times n}$ by letting $\mathcal{A}(S)$ be the set of rank-k matrices that are simultaneously S-projections and V^{-1}-projections. Define $\varphi : (\mathbb{R}^{n\times n})^2 \to \mathbb{R}^{n\times n}$ by letting $\varphi(R, A) := A^\top RA - RA = (I-A)^\top RA$. By Lemma 1 and Lemma 3, every equilibrium (whatever R is) has a linear S strategy and yields $\mathbf{a} = A\boldsymbol{\omega}$ for some $A \in \mathcal{A}(S)$; and Lemma 16 tells us such an equilibrium can also be optimal in the delegation problem only if $\varphi(R, A)$ zero. So the proposition will follow if we establish that generic $(S, R) \in \mathbb{L}^2$ have $\varphi(R, A)$ nonzero for every $A \in \mathcal{A}(S)$.

Below, we will show that every $S \in \mathbb{L}_1$ and $A \in \mathcal{A}(S)$ admit some $R \in \mathbb{L}$ such that $\varphi(R, A) \neq 0$. Let us first see how the proposition would follow. Given $S \in \mathbb{L}_1$ and $A \in \mathcal{A}(S)$, because $\varphi(\cdot, A)$ is linear, it would follow that a dense set of $R \in \mathbb{L}$ have $\varphi(R, A) \neq 0$—indeed, any proper convex combination R of $R^0, R^1 \in \mathbb{L}$ with $\varphi(R^0, A) = 0 \neq \varphi(R^1, A)$ has $\varphi(R, A) \neq 0$. But then (again given such S and A), because the set of such $R \in \mathbb{L}$ is semialgebraic, it is generic by Lemma 11. Now, for every $S \in \mathbb{L}_1$, it follows from Lemma 4 that $\mathcal{A}(S)$ is finite, and so (a finite intersection of generic sets being generic) generic $R \in \mathbb{L}$ have $\{\varphi(S, R, A)\}_{A\in\mathcal{A}(S)}$ all nonzero. In particular, because $\mathbb{L}_1$ is dense in $\mathbb{L}$ by Lemma 12, a dense set of $(S, R) \in \mathbb{L}^2$ has $\varphi(R, A)$ nonzero for every $A \in \mathcal{A}(S)$. Finally, this set is generic by Lemma 11.

So now fix $S \in \mathbb{L}_1$ and $A \in \mathcal{A}(S)$. It remains to see some $R \in \mathbb{L}$ has $\varphi(R, A) \neq 0$. And indeed, because the idempotent matrix $A^\top$ has rank $k \in \{1, \ldots, n-1\}$, it has some eigenvectors $p, q \in \mathbb{R}^n$ with respective eigenvalues 1 and 0. So consider the loss matrix $R := S + (p+q)(p+q)^\top$. That A is an S-perspective then tells us

$$\begin{aligned} \varphi(R, A) &= \varphi(S, A) + \varphi\left((p+q)(p+q)^\top, A\right) \\ &= 0 + \left[(I - A^\top)(p+q)\right]\left[A^\top(p+q)\right]^\top \\ &= qp^\top, \end{aligned}$$

which is nonzero because p and q are. $\square$