Unrestricted Information Acquisition*

Tommaso Denti†

November 21, 2015

Abstract

In a game, when there is uncertainty about the state of fundamentals, the players' behavior depends not only on their information about the state, but also crucially on their information about each other's information. Before taking actions, what do the players choose to know about what the others know? In this paper, I propose a tractable model of information acquisition to address this question in a systematic way. To unmask the primitive incentives to acquire information, the model does not impose restrictions on the information choice technology: the players can acquire information not only about the state, but also about each other's information in a flexible way. In coordination games, I show that the players have a strong incentive to know what the others know. In investment games, this leads to risk-dominance as the unique solution. In linear-quadratic games, this generates nonfundamental volatility. I further show that this incentive weakens as the game gets large and players small. In large investment games, multiple equilibria arise where the players focus on information about the state. In linear-quadratic games, nonfundamental volatility vanishes if no player is central in the game.

1 Introduction

In a game, when there is uncertainty about the state of fundamentals, the players have an incentive to acquire information not only about the state, but also about

*I am grateful to Muhamet Yildiz and Juuso Toikka for their guidance and support. For helpful discussions, I thank Alessandro Bonatti, Giulia Brancaccio, Drew Fudenberg, Massimo Marinacci, Jean Tirole, Alex Wolitzky, and, especially, Glenn Ellison.

†Department of Economics, MIT. Email: tdenti@mit.edu.

each other's information. So far, the literature has focused on what the players want to know about the state. Information choice, represented by signals, is assumed to be "independent": the players can observe only signals conditionally independent given the state (e.g., Persico [2000], Bergemann and Välimäki [2002], Tirole [2015]).[1] However, also what the players know about what the others know is crucial for the strategic interaction and the outcome of the game, for instance, in coordination games such as currency attacks (e.g., Morris and Shin [1998]).

Recently, in coordination games, a few models have been proposed to study what the players want to know about each other's information. In most of these models, information choice is assumed to be "rigid": the players can modify the noise of their signal only up to some parameter (e.g., Hellwig and Veldkamp [2009], Myatt and Wallace [2012], Pavan [2014]). In contrast, Yang [2015] allows information choice to be "flexible," keeping the independence assumption: the players can acquire any information they want about the state, but not about each other's information. Both independence and rigidity are natural starting points. Nevertheless, we do not know to what extent our predictions depend on primitive incentives to acquire information, or on exogenous restrictions on the information choice technology.

In this paper, I drop both independence and rigidity, and propose a model of "unrestricted" information choice: the players can acquire information not only about the state, but also about each other's information in a flexible way. The model is tractable under broad conditions on the cost of information, and provides a systematic framework for studying information choice in games. My analysis highlights two main patterns. First, in coordination games, the players have a strong incentive to learn what the others know: for instance, I show that this can explain the onset of phenomena such as bank runs and currency crises. Second, this incentive weakens as the game gets large and players small: for instance, I show that this leads nonfundamental volatility to vanish in canonical linear-quadratic games.

The model has two key features. First, the players can arbitrarily modify the noise of their signal, and achieve any desired correlation (in the widest sense of this word) with the state and the signals of the others. Second, the players face a tradeoff between learning the state and each other's information. Their cost of information depends not only on the correlation of their signal with the state, but also on the

[1] What I call independent information choice is sometimes named "private" information choice in the literature (e.g., Hellwig et al. [2012]).

correlation with the signals of the others. It represents the players' effort to learn the state and each other's information. The cost of information is assumed to satisfy only two broad monotonicity conditions: it is increasing with respect to the order of Blackwell [1951] both in the players' own signal and in the signals of the opponents. These conditions are satisfied, for instance, by mutual information (Shannon [1948]) and its generalization based on f-divergences (Csiszár [1974]).[2] The model extends existing models of independent or rigid information choice. It allows the players to acquire information also about each other's information, and in a flexible way.

The tractability of the model relies on a version of the revelation principle: even if the players can choose any type of signal, without loss of generality we can assume they observe "direct signals" to study the information they acquire and the actions they take. Direct signals not only convey information, but also "directly" tell the players what to do with that information.[3] By merging information and action, direct signals make transparent how the information choice relates to the primitives incentives to acquire information, which are driven by the action choice. For instance, I show that if a player's utility depends only on some statistic of the state and the others' actions, then, at the optimum, that statistic is sufficient to explain the dependence of the player's direct signal on the state and the others' direct signals.

Mutual information provides a tractable functional form for the cost of information and a natural starting point for the analysis of the model. For the case of mutual information, I provide a general equilibrium characterization for potential games. In potential games (Monderer and Shapley [1996]), the players' incentives to take actions can be described by a single global function called "potential." The equilibrium characterization I provide distinguishes the quality from the quantity of information acquired by the players. On one hand, the quality of information is summarized by the potential. On the other hand, the quantity of information can be studied in isolation through an auxiliary complete-information potential game. Examples of potential games are investment games and linear-quadratic games (on an undirected network), which are core applications I study in this paper (with and without mutual

[2]In economics, as a measure of the cost of information, mutual information has been popularized by Sims [2003] and the rational inattention literature (see, e.g., Wiederholt [2010] for an overview). Recently, f-divergences have been used by Maccheroni et al. [2006] in ambiguity aversion, and by Hébert [2015] in security design.

[3]Direct signals, sometimes called "recommendation strategies," are common in the literature on flexible information choice.

information).

My first main application of the model is to investment games. Investment games with common knowledge of fundamentals have been the traditional multiple-equilibrium model in economics (e.g., Diamond and Dybvig [1983], Katz and Shapiro [1986], Obstfeld [1996]). The return on investment is increasing in the state and the share of players who decide to invest. This complementarity generates a coordination problem and multiple equilibria when information is complete. In their pathbreaking work on global games, Carlsson and Van Damme [1993] showed that this multiplicity is not an intrinsic feature of investment games, but an artifact of the knife-edge common knowledge assumption. Their work has prompted a vast rethinking of investment games from a global-games perspective (e.g., Morris and Shin [1998], Rochet and Vives [2004], Goldstein and Pauzner [2005]).[4]

In many situations, the players do not decide whether to invest on the basis of some fixed prior information. Instead, the choice of what to know is a key component of the strategic interaction. In investment games with information acquisition, a compelling intuition suggests that the coordination problem in the investment decision could translate into a coordination problem in the information choice, restoring multiple equilibria. The players could coordinate on learning a threshold event such as "the state is positive," and invest when the event realizes, and not invest otherwise. Different thresholds would correspond to different equilibria, leading to multiplicity. This intuition has been remarkably formalized by Yang [2015] in a model of independent and flexible information choice, with mutual information as the cost of information.

The intuition for multiplicity, however, does not take into account that the players have also an incentive to know what the others know. The players are not interested in the event "the state is positive" *per se*. Their primitive incentive is to acquire information about the event "the return on investment is positive." Even for small cost of information, these two events are different, since the players' actions are not perfectly correlated with the state. In particular, for intermediate values of the state, the positivity of the return is overwhelmingly determined by the players' actions, which reflect their information. This creates a strong incentive for the players to know what the others know. This incentive is shut down if information choice is

[4]The literature on global games goes well beyond investment games. See, e.g., Morris and Shin [2003] for an early survey, or Frankel et al. [2003], Weinstein and Yildiz [2007], and Morris et al. [2015] for recent theoretical developments. The pitfalls of multiple-equilibrium models are discussed by Morris and Shin [2000].

assumed to be independent, and the players can acquire information only about the state.

Dropping independence and keeping mutual information, I show that, indeed, the players have a strong incentive to know what the others know, and this leads to risk-dominance as the unique solution, as in global games. To match each other's investment decision, as the cost of information becomes small, the players learn more and more about each other's information, at the expense of learning about the state. Furthermore, it is relatively more valuable to learn if the other players choose to play the risk-dominant action, since, by definition, it is the riskier action to mismatch. This channel fosters coordination on the risk-dominant action and drives equilibrium uniqueness. The formal analysis relies on the potential structure of investment games. In fact, this limit uniqueness result extends to all potential games: when the cost of information is negligible, for every state, the players select the action profile maximizing the potential.[5]

I further show that the players' incentive to learn what the others know weakens as the investment game gets large and players small. In fact, in large investment games with unrestricted information acquisition, I show that multiple equilibria arise where the players endogenously acquire information only about the state. The distinction between finite and large investment games is not simply a theoretical curiosity, since most investment games studied in applications feature a large number of small players.

I investigate more deeply the relation between structure of the game and information choice in my second main application of the model to linear-quadratic games. In linear-quadratic games, the players want to minimize the squared distance between their action and an unknown target, a linear combination of the state and the opponents' actions. A network matrix summarizes the impact of the players' actions on each other's targets. In linear-quadratic games, as customary in the literature, I focus on equilibria where the players observe normally distributed signals (but all deviations are allowed). Without functional form assumptions on the cost of information, I show that what two players know about each other's information is determined in equilibrium not only by the strength of their link, but by the entire network of relations. This happens because they value each other's information also for what it tells them about the information of the other players.[6] I further show that nonfundamental

[5]Provided that there are dominance regions (as it is characteristically assumed in global games). In investment games, the profile of risk-dominant actions is the maximizer of the potential.

[6]Calvó-Armengol et al. [2015] make a related point in a model of costly communication, which

volatility arises as an average of the players' Katz-Bonacich centralities in the network (Katz [1953], Bonacich [1987]). The centrality of a player can be seen as an overall measure of her opponents' incentive to learn what she knows. An implication of this result is that, in large networks without central players, nonfundamental volatility vanishes and the players behave as if information acquisition was independent.

A recurring theme of my analysis is that, with information acquisition, the size of the game matters for equilibrium predictions. The size of the game affects the impact of the players' actions on each other's utilities, and therefore their incentive to know what the others know. In particular, as the game gets large and players small, the players' incentive to acquire information about each other's information weakens. I explore this issue in my last main application of the model to large games. I consider general large games which extend both large investment games and linear-quadratic games with many players. I ask whether in all equilibria of the game the players focus on information about the state and behave as if information acquisition was independent. With mutual information as the cost of information, I show that the answer to this question depends on the strategic motives for actions. When actions are strategic substitutes, in any equilibrium the players choose conditionally independent signals. In particular, nonfundamental volatility vanishes. In contrast, when actions are strategic complements, there are equilibria where the players coordinate on acquiring information about each other's information, generating correlation between their signals that is unexplained by the state.[7]

Outline of the Paper. The rest of the paper is organized as follows. Section 2 collects standard definitions and notations I adopt throughout. Section 3 describes the model of unrestricted information acquisition and discuss the assumptions on the cost of information. Section 4 introduces direct signals and the revelation principle. Section 5 presents the main tools to analyze the model with mutual information as the cost of information. In Sections 6, 7, and 8, I apply the model to investment games, linear-quadratic games, and large games, respectively. In Section 9, I relate the model to the literature. Section 10 concludes. All proofs are in the appendix.

---

can be reinterpreted as a model of rigid information choice.

[7]In beauty contests, which are examples of large linear-quadratic games, Hellwig and Veldkamp [2009] first point out a connection between the strategic motives for actions and information choice.

2 Preliminary Definitions and Notations

In this section, I introduce standard definitions and notations I adopt throughout. It should be noted that I use bold lower-case letters for random variables (e.g., $\boldsymbol{x}$). Keeping that in mind, the reader can safely skip this section and come back to it when necessary.

Random Variables. In this paper, a random variable is a measurable function from a probability space $(\Omega, \mathcal{F}, P)$ into a Polish space.[8] Bold lower-case letters are used for random variables, upper-case letters for sets, and lower-case letters for deterministic variables: e.g., the random variable $\boldsymbol{x}$ takes values in the set X with typical element x. The distribution of the random variable $\boldsymbol{x}$, which is a probability measure on X, is denoted by $P_{\boldsymbol{x}}$. The set of probability measures on X is denoted by $\Delta(X)$, with typical element P_X. If $\boldsymbol{x}$ is real valued, then the symbol $E[\boldsymbol{x}]$ denotes its expected value. If $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is a profile of random variables, the sub-profile $(\boldsymbol{x}_j : j \neq i)$ is denoted by $\boldsymbol{x}_{-i}$.

Conditional Probabilities. Write $\boldsymbol{x}|\boldsymbol{w} \sim \boldsymbol{x}'|\boldsymbol{w}'$ if the conditional distribution of $\boldsymbol{x}$ given $\boldsymbol{w}$ is equal to the the conditional distribution of $\boldsymbol{x}'$ given $\boldsymbol{w}'$. Abusing notation, the symbol $\sim$ is also used to specify distributions: e.g., I use $\boldsymbol{x}|\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{w}, 1)$ to indicate that the conditional distribution of $\boldsymbol{x}$ given $\boldsymbol{w}$ is normal with mean $\boldsymbol{w}$ and unit variance. The symbol $(\boldsymbol{x} \perp \boldsymbol{x}')|\boldsymbol{w}$ indicates that $\boldsymbol{x}$ and $\boldsymbol{x}'$ are conditionally independent given $\boldsymbol{w}$.

Mutual Information. Given a pair of random variables $\boldsymbol{x}$ and $\boldsymbol{w}$, the *mutual information* of $\boldsymbol{x}$ and $\boldsymbol{w}$ is given by $I(\boldsymbol{x}; \boldsymbol{w})$ such that

$$I(\boldsymbol{x}; \boldsymbol{w}) = \int \log \left(\frac{dP_{(\boldsymbol{x},\boldsymbol{w})}}{d(P_{\boldsymbol{x}} \times P_{\boldsymbol{w}})} \right) dP_{(\boldsymbol{x},\boldsymbol{w})},$$

where $\frac{dP_{(\boldsymbol{x},\boldsymbol{w})}}{d(P_{\boldsymbol{x}} \times P_{\boldsymbol{w}})}$ is the density of the joint distribution of $\boldsymbol{x}$ and $\boldsymbol{w}$ with respect to the product of their marginals. If that density does not exist, set $I(\boldsymbol{x}; \boldsymbol{w}) = \infty$. To illustrate, let $\boldsymbol{x}$ and $\boldsymbol{w}$ be discrete with probability mass function p. The mutual

[8] A Polish space is a separable completely metrizable topological space, endowed with the Borel sigma-algebra. Examples are closed and open subsets of finite-dimensional vector spaces.

information of $\boldsymbol{x}$ and $\boldsymbol{w}$ is

$$I(\boldsymbol{x};\boldsymbol{w}) = E\left[\log \frac{p(\boldsymbol{w}|\boldsymbol{x})}{p(\boldsymbol{w})}\right] = E\left[\sum_{w\in W} p(w|\boldsymbol{x}) \log p(w|\boldsymbol{x})\right] - \sum_{w\in W} p(w) \log p(w).$$

Mutual information can be read as the expected value of the log-likelihood ratio of the posterior and prior of $\boldsymbol{w}$ given $\boldsymbol{x}$ (first equality), or as the expected reduction in the entropy of $\boldsymbol{w}$ from observing $\boldsymbol{x}$ (second equality).

Sufficient Statistics. Given a random variable $\boldsymbol{x}$, a *statistic* $f(\boldsymbol{x})$ of $\boldsymbol{x}$ is a random variable which is measurable with respect to $\boldsymbol{x}$. If $\boldsymbol{w}$ is another random variable, the statistic $f(\boldsymbol{x})$ is *sufficient* for $\boldsymbol{w}$ if one of the following equivalent conditions hold: (i) $\boldsymbol{x}|(\boldsymbol{w}, f(\boldsymbol{x})) \sim \boldsymbol{x}|f(\boldsymbol{x})$, (ii) $(\boldsymbol{x} \perp \boldsymbol{w})|f(\boldsymbol{x})$, or (iii) $\boldsymbol{w}|\boldsymbol{x} \sim \boldsymbol{w}|f(\boldsymbol{x})$.

3 Model

In this section, I present the model of unrestricted information acquisition. In the model, the players' information is represented by signals, that is, random variables. The players can arbitrarily modify the correlation between their signal, the state, and the signals of the others. The players face a tradeoff between learning the state and the others' information: the players' cost of information depends not only on the correlation of their signal with the state, but also with the signals of the others. Only broad conditions are imposed on the cost of information. The conditions are satisfied, for instance, by mutual information.

3.1 Game and Equilibrium Notion

A probability space $(\Omega, \mathcal{F}, P)$ is fixed. There is a countable set N of *players*, with typical element i and cardinality n.[9] The *state* of fundamentals is denoted by θ. Uncertainty about the state is represented by a random variable $\boldsymbol{\theta} : \Omega \to \Theta$, where Θ is a Polish space. In the game, the players first acquire information, then take actions. The players do not observe each other's information choice before taking actions. It is easier first to describe the game as if it happened in two subsequent

[9] Therefore, n can be finite or infinite.

phases: the information acquisition phase and the action phase. Then, I lay out its strategic form, which I study throughout the paper.

In the information acquisition phase, the players choose signals. For every player i, a Polish space X_i is fixed. Player i's *signal* is a random variable $\boldsymbol{x}_i : \Omega \to X_i$. Denote by $\boldsymbol{X}_i$ the set of signals available to player i. I assume that there are no restrictions on the information acquisition technology, that is, I assume that:

- For all players i, if $P_{X\times\Theta} \in \Delta(X \times \Theta)$ and $(\boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim P_{X_{-i}\times\Theta}$ for some $\boldsymbol{x}_{-i} \in \boldsymbol{X}_{-i}$, then there is $\boldsymbol{x}_i \in \boldsymbol{X}_i$ such that $(\boldsymbol{x}, \boldsymbol{\theta}) \sim P_{X\times\Theta}$.

- If $P_{X\times\Theta} \in \Delta(X\times\Theta)$ and $\boldsymbol{\theta} \sim P_\Theta$, then there is $\boldsymbol{x} \in \boldsymbol{X}$ such that $(\boldsymbol{x}, \boldsymbol{\theta}) \sim P_{X\times\Theta}$.

- For all players i, if $f : X_i \to X_i$ measurable and $\boldsymbol{x}_i \in \boldsymbol{X}_i$, then $f(\boldsymbol{x}_i) \in \boldsymbol{X}_i$.

Put differently, given any $\boldsymbol{x}_{-i}$, player i can arbitrarily modify the conditional distribution of $\boldsymbol{x}_i$ given $\boldsymbol{x}_{-i}$ and $\boldsymbol{\theta}$. Information choice, therefore, is fully flexible. Moreover, any joint distribution of signals and state can be achieved. Finally, the players can always choose to observe only functions of their signal.[10]

Given signal profile $\boldsymbol{x}$, player i's *cost of information* is given by $C_i(\boldsymbol{x}, \boldsymbol{\theta}) \in [0, \infty]$. The cost of information $C_i(\boldsymbol{x}, \boldsymbol{\theta})$ measures the correlation (in the widest sense of this word) between $\boldsymbol{x}_i$ and $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$, and it represents i's effort to learn the state and her opponents' information. I assume that $C_i(\boldsymbol{x}, \boldsymbol{\theta})$ satisfies the following properties:

- It depends only on the joint distribution of signals and state.

- It is increasing in Blackwell's order both in i's signal and her opponents' signals.

For instance, the cost of information can be proportional to the mutual information of $\boldsymbol{x}_i$ and $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$: $C_i(\boldsymbol{x}, \boldsymbol{\theta}) = \lambda_i I(\boldsymbol{x}_i; \boldsymbol{x}_{-i}, \boldsymbol{\theta})$, where $\lambda_i > 0$ is a scale factor. (See Section 3.2 for formal definitions and more examples, and Section 3.3 for discussion of the assumptions).

After the information acquisition phase, the players take actions. Denote by A_i player i's Polish space of *actions*, with typical element a_i. I assume that X_i is large enough such that X_i includes A_i (possibly up to an isomorphism). If the action profile a is selected, player i gets measurable *utility* $u_i(a, \theta) \in \mathbb{R}$, depending also on the state θ. The players take actions after observing the realization of their signal.

[10] See Appendix B for an example of probability space and available signals that satisfy the above assumptions.

Denote by $s_i(x_i)$ the action taken by player i if the realization of her signal is x_i. Call the measurable function $s_i : X_i \to A_i$ *contingency plan*. Denote by S_i the set of all possible contingency plans.

Combining the two phases, the game can be written in strategic form:

- The set of players is N.
- Player i's *strategy* consists of a signal $\boldsymbol{x}_i \in \boldsymbol{X}_i$ and a contingency plan $s_i \in S_i$.
- Player i's *payoff* is $E[u_i(s(\boldsymbol{x}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}, \boldsymbol{\theta})$.

To have well-defined expectations, assume $E[\sup_{a \in A} u_i(a, \boldsymbol{\theta})] < \infty$ for all players i.

As solution concept, I adopt pure-strategy Nash equilibrium, with the additional harmless requirement that the players' payoffs are finite in equilibrium:

Definition 1. A strategy profile $(\boldsymbol{x}, s)$ is an *equilibrium* if, for all players i and strategies $(\boldsymbol{x}'_i, s'_i)$,

$$E[u_i(s(\boldsymbol{x}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}, \boldsymbol{\theta}) \geq E[u_i(s'_i(\boldsymbol{x}'_i), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}'_i, \boldsymbol{x}_{-i}, \boldsymbol{\theta}),$$

and the left-hand side is finite.

3.2 Conditions on the Cost of Information

In the game, I assume that the players' cost of information, a function of the joint distribution of signals and state, is increasing in Blackwell's order both in the players' own signal and the signals of the others. Here, to formally present these monotonicity conditions, I fix player i, abstract away from the strategic setting, and consider the cost of information $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i) \in [0, \infty]$, where $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ are arbitrary random variables living in Ω. In the game, $\boldsymbol{x}_i \in \boldsymbol{X}_i$ and $\boldsymbol{w}_i = (\boldsymbol{x}_{-i}, \boldsymbol{\theta})$. Again, $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$ depends only on the joint distribution of $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$. In Section 3.3, I discuss how to interpret the assumptions on the cost of information in the game.

Given $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i) < \infty$, the monotonicity conditions read as follows:

Condition 1. If $f(\boldsymbol{x}_i)$ is a statistic of $\boldsymbol{x}_i$, then $C_i(f(\boldsymbol{x}_i), \boldsymbol{w}_i) \leq C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$, with equality if and only if $f(\boldsymbol{x}_i)$ is sufficient for $\boldsymbol{w}_i$.

Condition 2. If $f(\boldsymbol{w}_i)$ is a statistic of $\boldsymbol{w}_i$, then $C_i(\boldsymbol{x}_i, f(\boldsymbol{w}_i)) \leq C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$, with equality if and only if $f(\boldsymbol{w}_i)$ is sufficient for $\boldsymbol{x}_i$.

Conditions 1 and 2 say that $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$ is increasing both in $\boldsymbol{x}_i$ and in $\boldsymbol{w}_i$ with respect to the order of Blackwell [1951]. The quantity $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$ can be seen as a measure of dependence between $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$. Condition 1 says that $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ are more "correlated" than $f(\boldsymbol{x}_i)$ and $\boldsymbol{w}_i$, while Condition 2 that they are are more "correlated" than $\boldsymbol{x}_i$ and $f(\boldsymbol{w}_i)$.[11] For instance, if $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ are unidimensional and jointly normal, then Conditions 1 and 2 are satisfied if and only if $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$ depends only on and is increasing in $|Cor(\boldsymbol{x}_i, \boldsymbol{w}_i)|$, the correlation coefficient of $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ (in absolute value).[12] Notice also that Conditions 1 and 2 are ordinal: new cost functions arise by taking monotone transformations of existing ones.

There is a wide range of cost functions that satisfy Conditions 1 and 2. Some possible specifications follow:

No Information. Under Conditions 1 and 2, $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$ is minimal when $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ are independent: if $\boldsymbol{x}_i \perp \boldsymbol{w}_i$, then $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i) \leq C_i(\boldsymbol{x}'_i, \boldsymbol{w}'_i)$ for any other pair of random variables $\boldsymbol{x}'_i$ and $\boldsymbol{w}'_i$. When $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ are independent, no information is contained in $\boldsymbol{x}_i$ about $\boldsymbol{w}_i$. The case where no information is available can be represented by choosing C_i such that $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i) = 0$ if $\boldsymbol{x}_i \perp \boldsymbol{w}_i$, and $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i) = \infty$ else. Such C_i is a simple example of cost function that satisfies Conditions 1 and 2.

Mutual Information. Mutual Information (Shannon [1948]) is a standard measure of dependence between random variables, and satisfies Conditions 1 and 2, that is, if $C_i(\boldsymbol{x}_i, \boldsymbol{w}_i) = I(\boldsymbol{x}_i; \boldsymbol{w}_i)$, then C_i satisfies Conditions 1 and 2.[13] In economics, as a measure of the cost of information, mutual information has been popularized by Sims [2003] and the rational inattention literature.

f-Divergences. Mutual information $I(\boldsymbol{x}_i; \boldsymbol{w}_i)$ is the K-L divergence of the joint distribution of $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ from the product of their marginals (Kullback and Leibler [1951]). By considering f-divergences (Csiszár [1967]), which are generalizations of

[11] See also Rényi [1959] for different but related conditions on measures of dependence.

[12] Results from Hansen and Torgersen [1974] can be used to extend this characterization to the multidimensional case.

[13] The monotonicity of mutual information with respect to Blackwell's order is a consequence of the data processing inequality (e.g., Cover and Thomas [2006, pp. 35-37]).

the K-L divergence, it is possible to construct a spectrum of cost functions satisfying Conditions 1 and 2.[14]

3.3 Discussion

In the game, Condition 1 says that the more information a player acquires, the higher her cost of information. This is a common assumption in models of information choice. On the other hand, Condition 2 is distinctive to unrestricted information acquisition. It says that the player's cost of information is increasing also in the information the others acquire. Broadly speaking, Condition 2 reflects the idea that it is hard to learn exactly what others know, and the more they know, the harder to do so. For the game, its key implication is that the players face a tradeoff between learning the state and each other's information.

The model can arise from an underlying model of information choice where the players' cost of information depends only on their own signal (as a random variable), and not on its correlation with the signals of the others. From this perspective, Conditions 1 and 2 can be interpreted as underlying richness assumptions on the set of available signals. I postpone to Section 9 the discussion of other settings to which the model may more primarily relate.

Tradeoff. For the game, Condition 2 implies that the players face a tradeoff between learning the state and each other's information. To see this, think of the conditional distributions of $\boldsymbol{\theta}$ and $\boldsymbol{x}_{-i}$ given $\boldsymbol{x}_i$ as player i's information about the state and her opponents' information, respectively. Given Condition 2, if she is willing to disregard some information about her opponents' information, player i has a signal less expensive than $\boldsymbol{x}_i$ to acquire the same information about the state. Indeed, instead of $\boldsymbol{x}_i$, she can observe signal $\boldsymbol{x}_i'$ such that $\boldsymbol{x}_i'|(\boldsymbol{x}_{-i},\boldsymbol{\theta}) \sim \boldsymbol{x}_i|\boldsymbol{\theta}$. Clearly, we have

[14]Take $f:(0,\infty)\to\mathbb{R}$ strictly convex such that $\lim_{t\to\infty} f(t)/t=\infty$. Define $I_f(\boldsymbol{x}_i;\boldsymbol{w}_i)$ such that

$$I_f(\boldsymbol{x}_i;\boldsymbol{w}_i) = \int f\left(\frac{dP_{(\boldsymbol{x}_i,\boldsymbol{w}_i)}}{d(P_{\boldsymbol{x}_i}\times P_{\boldsymbol{w}_i})}\right) d(P_{\boldsymbol{x}_i}\times P_{\boldsymbol{w}_i}),$$

where $\frac{dP_{(\boldsymbol{x}_i,\boldsymbol{w}_i)}}{d(P_{\boldsymbol{x}_i}\times P_{\boldsymbol{w}_i})}$ is the density of the joint distribution of $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ with respect to the product of their marginals. If that density does not exist, set $I_f(\boldsymbol{x}_i;\boldsymbol{w}_i)=\infty$. The quantity $I_f(\boldsymbol{x}_i;\boldsymbol{w}_i)$ is the f-divergence of the joint distribution of $\boldsymbol{x}_i$ and $\boldsymbol{w}_i$ from the product of their marginals. Mutual information corresponds to the case $f(t)=t\log t$ for all $t\in(0,\infty)$. Applying Liese and Vajda [1987, Corollary 1.29], if $C_i(\boldsymbol{x}_i,\boldsymbol{w}_i)=I_f(\boldsymbol{x}_i;\boldsymbol{w}_i)$, then C_i meets Conditions 1 and 2.

$\boldsymbol{\theta}|\boldsymbol{x}_i' \sim \boldsymbol{\theta}|\boldsymbol{x}_i$, and therefore i's information about the state is the same given $\boldsymbol{x}_i'$ or $\boldsymbol{x}_i$. Put differently, in the two situations, she exerts the same level of effort to learn the state. However, since $\boldsymbol{x}_i'|(\boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim \boldsymbol{x}_i'|\boldsymbol{\theta}$, the state $\boldsymbol{\theta}$ as a statistic of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}_i'$. This reflects the idea that, choosing $\boldsymbol{x}_i'$, player i focuses on the state, and exerts less effort to learn the others' information. Overall, using Condition 2,

$$C_i(\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{\theta}) = C_i(\boldsymbol{x}_i', \boldsymbol{\theta}) = C_i(\boldsymbol{x}_i, \boldsymbol{\theta}) \leq C_i(\boldsymbol{x}, \boldsymbol{\theta}),$$

which implies that $\boldsymbol{x}_i'$ is indeed less expensive than $\boldsymbol{x}_i$.

Richness. In the model, the players' cost of information is a function of the joint distribution of signals and state. This structure for the cost of information can arise from an underlying model of information choice where the players' cost of information depends only on their own signal. Conditions 1 and 2 can then be seen as underlying richness assumptions on the set of available signals. Informally, assume player i's signal $\boldsymbol{x}_i$ has an underlying cost $C_i^*(\boldsymbol{x}_i)$. The cost depends on $\boldsymbol{x}_i$ as a random variable. Let $\boldsymbol{X}$ satisfies the following richness conditions:

- Condition 1*: Given signal profile $\boldsymbol{x} \in \boldsymbol{X}$, for every measurable $f : X_i \to X_i$, player i has another signal $\boldsymbol{x}_i' \in \boldsymbol{X}_i$ such that

  $$(\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim (f(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \quad \text{and} \quad C_i^*(\boldsymbol{x}_i') \leq C_i^*(\boldsymbol{x}_i),$$

  with equality if and only if $f(\boldsymbol{x}_i)$ as a statistic of $\boldsymbol{x}_i$ is sufficient for $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$.

- Condition 2*: Given signal profile $\boldsymbol{x} \in \boldsymbol{X}$, for every statistic $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$, player i has another signal $\boldsymbol{x}_i' \in \boldsymbol{X}_i$ such that

  $$(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})) \sim (\boldsymbol{x}_i, f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})) \quad \text{and} \quad C_i^*(\boldsymbol{x}_i') \leq C_i^*(\boldsymbol{x}_i),$$

  with equality if and only if $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ as a statistic of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}_i$.

Condition 1* says that player i can throw away additional information about $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ contained in $\boldsymbol{x}_i$ that is not carried by $f(\boldsymbol{x}_i)$. Condition 2*, instead, means that player i can disregard additional information about $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ that is not relevant for $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$. Conditions 1* and 2* intuitively correspond to Conditions 1 and 2. In fact, starting from C_i^*, it is possible to derive a cost function C_i that satisfies Conditions 1 and 2:

$$C_i(\boldsymbol{x}, \boldsymbol{\theta}) = \min\{C_i^*(\boldsymbol{x}_i') : (\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim (\boldsymbol{x}, \boldsymbol{\theta})\}.$$

Given $\boldsymbol{x}_{-i}$, $C_i(\boldsymbol{x}, \boldsymbol{\theta})$ is the cost under C_i^* of the least expensive signal that achieves the target distribution $P_{(\boldsymbol{x},\boldsymbol{\theta})}$. To study the game, assuming C_i^* as the cost information is equivalent to assuming C_i, since the focus is on pure-strategy equilibria.

4 Direct Signals

In the information acquisition phase, the players can choose any type of signal. Nevertheless, in the analysis of the game, we can always assume they pick a direct signal, as I show in this section. For player i, a *direct signal* $\boldsymbol{x}_i$ is a random variable that takes values in her action space.[15] It is paired with the identity function as contingency plan.[16] Direct signals not only convey information to the players, but also directly tell them what to do with that information. In this section, I show that it is without loss of generality to focus on equilibria in direct signals (revelation principle): any equilibrium of the game can be replicated by an "equivalent" equilibrium in direct signals. For the analysis of the game, direct signals have also the advantage of merging information and action. This makes more transparent how the information choice relates to the primitive incentives to acquire information, which are driven by the action choice. An example is given by the following characterization of optimal direct signals in terms of sufficient statistics, which concludes the section: if a player's utility depends only on some statistic of the state and the others' actions, then, at the optimum, that statistic is sufficient to explain the dependence of the player's direct signal on the state and the others' direct signals.

4.1 Revelation Principle

To state the revelation principle, I need a notion of "equivalence" for equilibria of the game. It is natural to think that two equivalent equilibria $(\boldsymbol{x}, s)$ and $(\boldsymbol{x}', s')$ should induce the same distribution over actions and states: $(s(\boldsymbol{x}), \boldsymbol{\theta}) \sim (s'(\boldsymbol{x}'), \boldsymbol{\theta})$. In games with information choice, however, also the players' information is an equilibrium out-

[15]That is, the image of $\boldsymbol{x}_i$ is included in A_i. Recall that, in Section 3.1, I have assumed that X_i includes A_i (possibly up to an isomorphism).

[16]That is, it is always paired with $s_i : X_i \to A_i$ such that $s_i|_{A_i} = id_{A_i}$.

come of interest. To determine whether two equilibria are informationally equivalent, I use the ranking of signal profiles introduced by Bergemann and Morris [2015]:

Definition 2. The signal profile $\boldsymbol{x}$ is *individually sufficient* for the signal profile $\boldsymbol{x}'$ if, for all players i, the signal $\boldsymbol{x}_i$ as a statistic of $(\boldsymbol{x}_i, \boldsymbol{x}'_i)$ is sufficient for $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$.

Individual sufficiency is a many-player extension of Blackwell's order. It intuitively captures when one signal profile contains more information than another: if $\boldsymbol{x}$ is individually sufficient for $\boldsymbol{x}'$, then signal $\boldsymbol{x}'_i$ does not provide new information about $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ relative to $\boldsymbol{x}_i$.

The next proposition is a version of the revelation principle for the model:

Proposition 1. *If $(\boldsymbol{x}, s)$ is an equilibrium, then the direct signals $s(\boldsymbol{x})$ form an equilibrium. Moreover, $\boldsymbol{x}$ and $s(\boldsymbol{x})$ are individually sufficient for each other.*

Proposition 1 says that, for every equilibrium $(\boldsymbol{x}, s)$ of the game, there exists an "equivalent" equilibrium $s(\boldsymbol{x})$ in direct signals. First, the two equilibria generate the same joint distributions of actions and state. Second, they are individually sufficient for each other, i.e., informationally equivalent in the sense of Bergemann and Morris [2015]. Therefore, it is without loss of generality to study the one in direct signals.

The revelation principle of Proposition 1 reflects the idea that the players do not value information per se, but only to take better actions. For an intuition, take the perspective of player i. First, any information provided by $\boldsymbol{x}_i$ that is not contained in $s_i(\boldsymbol{x}_i)$ is superfluous for player i. Analogously, any information contained in $\boldsymbol{x}_{-i}$ that does not affect $s_{-i}(\boldsymbol{x}_{-i})$ is irrelevant for her. Moreover, observing $s_i(\boldsymbol{x}_i)$ is less expensive than observing $\boldsymbol{x}_i$ (Condition 1), and it is less costly to learn $s_{-i}(\boldsymbol{x}_{-i})$ rather than $\boldsymbol{x}_{-i}$ (Condition 2). Overall, this means that, in equilibrium, direct signal $s_i(\boldsymbol{x}_i)$ is as informative as $\boldsymbol{x}_i$, and not only a best reply to $(\boldsymbol{x}_{-i}, s_{-i})$, but also a best reply to direct signals $s_{-i}(\boldsymbol{x}_{-i})$. Indeed, in the first place player i focuses on $s_{-i}(\boldsymbol{x}_{-i})$ rather than $\boldsymbol{x}_{-i}$.

Direct signals, sometimes called "recommendation strategies" (e.g., Ravid [2015]), are common in the literature on flexible information choice. In fact, several existing models feature results similar to Proposition 1.[17] In these models, the cost of information is independent of the others' signals. Hence, versions of Condition 1 are sufficient for the revelation principle to hold. Here, instead, also Condition 2 is needed.

[17] Even if the informational equivalence is often omitted.

4.2 Best Replies and Sufficient Statistics

In applications, the players' utilities often depend only on some statistic of the state and their opponents' actions. Taking the perspective of player i, there is a measurable function f defined on $A_{-i} \times \Theta$ such that, for all $a_i \in A_i$,

$$f(a_{-i}, \theta) = f(a'_{-i}, \theta') \quad \Rightarrow \quad u_i(a_i, a_{-i}, \theta) = u_i(a_i, a'_{-i}, \theta'). \tag{1}$$

Abusing notation, write $u_i(a_i, f(a_{-i}, \theta))$ for $u_i(a, \theta)$.

Assuming (1), the next lemma characterizes optimal direct signals in term of sufficient statistics:

Lemma 1. *Fix player i and assume there is a function f such that (1) holds. Let $\boldsymbol{x}$ be a profile of direct signals such that i's payoff is finite at $\boldsymbol{x}$. Direct signal $\boldsymbol{x}_i$ is a best reply to $\boldsymbol{x}_{-i}$ if and only if the following two conditions hold: (i) $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ as a statistic of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}_i$, and (ii) $\boldsymbol{x}_i$ solves the optimization problem*

$$\textit{maximize } E[u_i(\boldsymbol{x}'_i, f(\boldsymbol{x}_{-i}, \boldsymbol{\theta}))] - C_i(\boldsymbol{x}'_i, f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})) \textit{ over direct signals } \boldsymbol{x}'_i.$$

Lemma 1 says that, at the optimum, the statistic $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient to explain the dependence of $\boldsymbol{x}_i$ on $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$. Therefore, to study the information player i acquires about $\boldsymbol{x}_{-i}$ and $\boldsymbol{\theta}$, it is enough to consider the information she wants to acquire about $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$. Lemma 1 is key to the analysis of linear-quadratic games in Section 7.

Lemma 1 reflects the same idea that underlies the revelation principle of Proposition 1: for the players, information is only instrumental in taking better actions. Intuitively, since player i's action interacts with the others' actions and the state only through the statistic f, any information contained in $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ which does not affect $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is irrelevant. Moreover, it is less costly to learn $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ rather than $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ (Condition 2). As a result, the player focuses on $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ in information acquisition, rather than $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$, as Lemma 1 states.

5 Mutual Information

Mutual information provides a tractable functional form for the cost of information and a natural starting point for the analysis of the model. In this section, I introduce the main tools to study equilibria in direct signals with cost of information given by mutual information. In general games, best replies with mutual information take the form of a Bayesian multinomial-logit model (Csiszár [1974], Matějka and McKay [2015]). In potential games (Monderer and Shapley [1996]), I use this result to separately characterize the quality and quantity of information the players acquire in equilibrium. On one hand, the quality of information is summarized by the potential. On the other hand, the quantity of information can be studied in isolation through an auxiliary complete-information potential game. Examples of potential games are investment games and linear-quadratic games, which I study in the next sections (with and without mutual information).

5.1 Best Replies with Mutual Information

With mutual information as cost of information, player i best replies to her opponents' direct signals by solving the following the optimization problem:

$$\text{maximize } E[u_i(\boldsymbol{x},\boldsymbol{\theta})] - \lambda_i I(\boldsymbol{x}_i;\boldsymbol{x}_{-i},\boldsymbol{\theta}) \text{ over direct signals } \boldsymbol{x}_i.$$

Notice that the objective function depends only on the joint distribution of $\boldsymbol{x}$ and $\boldsymbol{\theta}$. Furthermore, since information acquisition is flexible, player i can arbitrarily modify the dependence of $\boldsymbol{x}_i$ on $\boldsymbol{x}_{-i}$ and $\boldsymbol{\theta}$. Therefore, best replying means choosing $\boldsymbol{x}_i|(\boldsymbol{x}_{-i},\boldsymbol{\theta})$ to maximize $E[u_i(\boldsymbol{x},\boldsymbol{\theta})] - \lambda_i I(\boldsymbol{x}_i;\boldsymbol{x}_{-i},\boldsymbol{\theta})$, taking $P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}$ as given.

Csiszár [1974] studies an analogous optimization problem in rate-distortion theory, a branch of information theory that analyzes lossy data-compression. The following lemma, an immediate implication of his work, provides optimality conditions for the best-reply problem with mutual information:

Lemma 2. *Fix player i and assume $C_i = \lambda_i I$. Let $\boldsymbol{x}$ be a profile of direct signals such that i's payoff is finite at $\boldsymbol{x}$. Direct signal $\boldsymbol{x}_i$ is a best reply to $\boldsymbol{x}_{-i}$ if and only if the following two conditions hold:*

$$\boldsymbol{x}_i|(\boldsymbol{x}_{-i},\boldsymbol{\theta}) \sim \frac{e^{u_i(\cdot,\boldsymbol{x}_{-i},\boldsymbol{\theta})/\lambda_i}}{\int_{A_i} e^{u_i(a_i',\boldsymbol{x}_{-i},\boldsymbol{\theta})/\lambda_i} dP_{\boldsymbol{x}_i}(a_i')} dP_{\boldsymbol{x}_i}, \tag{2}$$

and, for all $a_i \in A_i$,

$$\int_{A_{-i}\times\Theta} \frac{e^{u_i(a,\theta)/\lambda_i}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)/\lambda_i} dP_{\boldsymbol{x}_i}(a_i')} dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta) \leq 1. \tag{3}$$

In the context of rational inattention, Matějka and McKay [2015] interpret (2) as a Bayesian version of the multinomial logit model (in settings with only one player). Condition (3) disciplines the marginal distribution $P_{\boldsymbol{x}_i}$. For instance, if $\boldsymbol{x}_{-i}$ and $\boldsymbol{\theta}$ are degenerate and concentrated on some a_{-i} and θ, condition (2) is satisfied by any direct signal $\boldsymbol{x}_i$ that is degenerate. Condition (3), instead, is satisfied when i's degenerate signal is concentrated on some best reply to a_{-i} in the complete information game corresponding to θ.

The conditions of Lemma 2 describe two different aspects of information choice: quality and quantity of information. On one hand, condition (2) provides a formula for the density of $P_{(\boldsymbol{x},\boldsymbol{\theta})}$ with respect to $P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}$. The density is given by $f_i : A \times \Theta \to \mathbb{R}_+$ such that

$$f_i(a,\theta) = \frac{e^{u_i(a,\theta)/\lambda_i}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)/\lambda_i} dP_{\boldsymbol{x}_i}(a_i')}.$$

The density captures the dependence of i's signal on the state and the signals of i's opponents. It represents the quality of i's information about the state and the others' information. On the other hand, condition (3) pins down the the marginal distribution of i's signal, which can be seen as an overall measure of the quantity of i's information. Intuitively, a diffuse marginal corresponds to a variable and informative signal. At the other extreme, when the marginal is concentrated on some action, the player acquires no information at all.

5.2 Potential Games with Mutual Information

For the rest of this section, let $n < \infty$ and $v : A \times \Theta \to \mathbb{R}$ be a *potential*: for all $i \in N$ and $a_i, a_i' \in A_i$,

$$u_i(a_i, a_{-i}, \theta) - u_i(a_i', a_{-i}, \theta) = v(a_i, a_{-i}, \theta) - v(a_i', a_{-i}, \theta), \quad \forall a_{-i} \in A_{-i} \text{ and } \forall \theta \in \Theta.$$

Building on Lemma 2, the next lemma characterizes equilibrium direct signals for potential games, with mutual information as cost of information.

Lemma 3. *Consider a potential game. Assume $C_i = \lambda I$ for all $i \in N$. The profile of direct signals $\boldsymbol{x}$ is an equilibrium if and only if the following two conditions hold:*

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(P_{\boldsymbol{\theta}} \times_{i\in N} P_{\boldsymbol{x}_i})}(a,\theta) = \frac{e^{v(a,\theta)/\lambda}}{\int_A e^{v(a',\theta)/\lambda} d(\times_{i\in N} P_{\boldsymbol{x}_i})(a')}, \quad a.s., \tag{4}$$

and, for all players i and action a_i,

$$\int_\Theta \frac{\int_{A_{-i}} e^{v(a_i,a'_{-i},\theta)/\lambda} d(\times_{j\neq i} P_{\boldsymbol{x}_j})(a'_{-i})}{\int_A e^{v(a',\theta)/\lambda} d(\times_{i\in N} P_{\boldsymbol{x}_i})(a')} dP_{\boldsymbol{\theta}}(\theta) \leq 1. \tag{5}$$

Condition (4) and (5) provide a separate description of the quality and the quantity of information the players acquire in equilibrium.[18]

On one hand, Condition (4) gives the density of the joint distribution of signals and state with respect to the product of their marginal distributions. The density captures the dependence of the players' signals on each other and the state. It represents the quality of the players' information and is summarized by the potential. Therefore, it provides an immediate connection between primitive properties of the game and information in equilibrium. For instance, if every A_i and Θ are totally ordered spaces, then signals and state are affiliated (Milgrom and Weber [1982]) whenever the potential is supermodular, i.e., whenever there are strategic complementarities.[19]

Condition (5), instead, depends only on the marginal distributions of signals and state. As argued above after Lemma 2, the marginal distributions can seen as overall measures of the quantity of the players' information. Condition (5) disciplines these quantities. For instance, if $\boldsymbol{\theta}$ is degenerate and concentrated on some θ, condition (4) is satisfied by any profile $\boldsymbol{x}$ that is degenerate. Condition (5), instead, is satisfied when the degenerate signals are concentrated on a pure-strategy Nash equilibrium of the complete information game corresponding to θ.

A first implication of Lemma 3 is that, for negligible cost of information, state-by-state, the players select the maximizer of the potential:

Proposition 2. *Fix action profile a such that every component a_i has a dominance region, i.e., there is a set of states Θ_{a_i} with $P(\boldsymbol{\theta} \in \Theta_{a_i}) > 0$ such that*

$$\inf_{\theta\in\Theta_{a_i}} \inf_{a'_i\neq a_i} \inf_{a_{-i}\in A_{-i}} u_i(a_i, a_{-i}, \theta) - u_i(a'_i, a_{-i}, \theta) > 0.$$

[18] Observe that both (4) and (5) do not depend on the choice of the potential, since the potential is unique up to an additive constant that depends on the state, but not on the actions.

[19] If every A_i and Θ are totally ordered spaces, the potential is supermodular if and only if, for all $i \in N$, utility u_i has non-decreasing differences in (a_i, θ) and (a_i, a_j), for all players $j \neq i$.

Consider a potential game. Assume $C_i = \lambda I$ for all $i \in N$. For every λ, select an equilibrium in direct signals $\boldsymbol{x}_\lambda$. Then, $P_{\boldsymbol{\theta}}$-almost surely,

$$v(a, \theta) > \sup_{a' \neq a} v(a', \theta) \quad \Rightarrow \quad \lim_{\lambda \to 0} P(\boldsymbol{x}_\lambda = a | \theta) = 1.$$

Proposition 2 says that, for every state, the players select the action profile maximizing the potential, when the cost of information is negligible. As the cost of information gets smaller, for every state, the density in (4) puts more weight on the action profile maximizing the potential. With dominance regions, the product of the signals' marginal distributions assigns positive probability to the maximizer. Hence, in the limit, signals are concentrated on the maximizer of the potential, conditioned on the state. In investment games, Proposition 2 implies a risk-dominance selection result, which I discuss in depth in Section 6.

Another implication of Lemma 3 is that all equilibria in direct signals are parametrized by the marginal distributions of the players' signals: if two profiles of equilibrium direct signals have the same marginal distributions, then they are equivalent in the sense that they induce the same distributions over actions and states. The proposition illustrates what marginal distributions arise in equilibrium. To state the result, denote by $\alpha = (\alpha_i \in \Delta(A_i) : i \in N)$ a generic profile of marginal distributions. Abusing notation, write also

$$\alpha = \times_{i \in N} \alpha_i \in \Delta(A) \quad \text{and} \quad \alpha_{-i} = \times_{j \neq i} \alpha_j \in \Delta(A_{-i}), \quad \forall i \in N.$$

Proposition 3. *Consider a potential game. Assume $C_i = \lambda I$ for all $i \in N$. Fix a profile of marginal distributions α. The following statements are equivalent:*

(i) There is an equilibrium in direct signals $\boldsymbol{x}$ with $\boldsymbol{x}_i \sim \alpha_i$ for all players i.

(ii) The profile α is a pure-strategy equilibrium of the auxiliary complete-information potential game defined by $V : \Delta(A) \to \mathbb{R}$ such that

$$V(\alpha) = \int_\Theta \log \left(\int_A e^{v(a,\theta)/\lambda} d\alpha(a) \right) dP_{\boldsymbol{\theta}}(\theta).$$

(iii) For all players i and action a_i,

$$\int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i,a'_{-i},\theta)/\lambda} d\alpha_{-i}(a'_{-i})}{\int_{A} e^{v(a',\theta)/\lambda} d\alpha(a')} dP_{\boldsymbol{\theta}}(\theta) \leq 1.$$

Proposition 3 says that the players choose the marginal distributions of their signals as if they played an auxiliary complete-information potential game defined by V (equivalence of (i) and (ii)). Combined with Lemma 3, this result implies that the overall characterization of equilibria in direct signals boils down to the study of the auxiliary potential game. Statement (iii) provides the optimality conditions for the best-reply problem in the auxiliary game. The optimality conditions correspond to the inequalities in (5).

The auxiliary potential games can be studied to determine existence and uniqueness of equilibria in direct signals. When V has a maximizer, there is an equilibrium in direct signals. When V is strictly concave, the equilibrium in direct signals is unique. I conclude the section by providing a broad existence result for potential games, with mutual information as cost of information:

Corollary 1. *Consider a potential game. Assume $C_i = \lambda I$ for all $i \in N$. Then there exists an equilibrium in direct signals whenever the following conditions hold:*

- *For all players i, the action space A_i is compact.*
- *For every $\theta \in \Theta$, the function $v(\cdot, \theta)$ is upper semi-continuous.*
- *The expected value $E\left[\sup_{a\in A} |v(a, \boldsymbol{\theta})|\right]$ is finite.*

6 Investment Games

Starting from this section, I apply the model and tools so far developed to specific games. Here, I focus on investment games (e.g., Morris and Shin [1998]), canonical examples of coordination games. The main question I address is the following: is information choice a source of equilibrium indeterminacy in coordination games?

When information acquisition is assumed to be independent, investment games have multiple equilibria, with mutual information as cost of information (Yang [2015]): the coordination problem in the investment decision translates into a coordination problem in the choice of information about the state.

Dropping independence and keeping mutual information, I show that the coordination problem is broken by the players' incentive to acquire information about each other's information. When there are finitely many players and the cost of information is negligible, risk-dominance is the unique solution, as in global games (Carlsson and Van Damme [1993]). However, I also show that, when the game is large, multiple equilibria arise where the players behave as if information acquisition was independent.

The distinction between finite and large games emerges because the number of players in the game affects the impact of the players' action on each other's utilities, and therefore their incentive to learn what the others know. This channel is not present when information is exogenously given and the size of the game is mostly chosen on the basis of tractability considerations. Finally, for the finite-player case, I provide a general equilibrium characterization for any level of cost of information, relying on the potential structure of investment games.

6.1 Setup

Every player has two actions: $\forall i \in N$, $A_i = \{0, 1\}$ where one stands for *invest*, and zero for *not invest*. For player i, the *return on investment* $\rho(\bar{a}_{-i}, \theta) \in \mathbb{R}$ is non-decreasing in the state $\theta \in \mathbb{R}$ and the share of opponents who decide to invest $\bar{a}_{-i} = \frac{1}{n-1}\sum_{j \neq i} a_j$.[20] The utility of not invest is normalized to zero. Overall, player i's utility is $u_i(a, \theta) = a_i \rho(\bar{a}_{-i}, \theta)$ for all $a \in A$ and $\theta \in \Theta$. I assume that $P(\rho(1, \boldsymbol{\theta}) < 0) > 0$ and $P(\rho(0, \boldsymbol{\theta}) > 0) > 0$, i.e., both invest and not invest have dominance regions.

Risk Dominance. Action a_i is *risk-dominant* at state θ if, in the complete information game corresponding to θ, action a_i is a strict best reply to uniform conjecture over the share of opponents who decide to invest (Harsanyi and Selten [1988], Morris and Shin [2003]). In particular, if $n < \infty$, invest is risk dominant at θ if

$$\frac{1}{n}\sum_{m=0}^{n-1} \rho\left(\frac{m}{n-1}, \theta\right) > 0.$$

[20] If $n = \infty$, set $\bar{a}_{-i} = \bar{a} = \limsup_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} a_i$.

If the inequality holds in the other direction, the risk-dominant action is not invest. State θ is a *risk-dominance threshold* if $\frac{1}{n}\sum_{m=0}^{n-1}\rho\left(\frac{m}{n-1},\theta\right) = 0$. The risk-dominant action can also be interpreted as the action that is riskier to mismatch. For instance, if there are two players, invest is risk-dominant at θ when $-\rho(1,\theta) < \rho(0,\theta)$, i.e., when the loss from not investing when the opponent invests is larger than the loss from investing when the the opponent does not invest.

Potential. Investment games with finitely many players are potential games. For instance, a potential is the function $v : A \times \Theta \to \mathbb{R}$ such that

$$v(a,\theta) = \sum_{m=0}^{|a|-1} \rho\left(\frac{m}{n-1},\theta\right), \quad \text{with } |a| = \sum_{i\in N} a_i.$$

Linear Return. While my analysis applies to general return functions, I sometimes consider the special case of *linear return* (Morris and Shin [2000]) as a concrete example: for all $i \in N$, $a_{-i} \in A_{-i}$, and $\theta \in \Theta$,

$$\rho(\bar{a}_{-i},\theta) = \theta - r(1-\bar{a}_{-i}), \quad \text{with } r > 0.$$

The scalar r parametrizes the degree of strategic complementarity in actions. If $\theta > r$ ($\theta < 0$, resp.), the action invest (not invest, resp.) is strictly dominant in the corresponding complete information game. On the other hand, if $\theta \in [0,r]$, both "all invest" and "all not invest" are equilibria. The risk-dominance threshold is $r/2$.

6.2 Risk-Dominance Selection for $n < \infty$

With finitely many players, I show that coordination on the risk-dominant action is achieved for negligible cost of information, with mutual information as cost of information. This happens because, for intermediate values of the state, the players want to correlate their direct signals with the direct signals of the others, in order to match their actions. Furthermore, correlation is more valuable when the direct signals of the others suggest them to play the risk-dominant action, since the risk-dominant action is the riskier to mismatch. This channel is not present when information choice is independent and the players can acquire only information about the state. Indeed, if so, multiple equilibria arise (Yang [2015]). The risk-dominance selection result with unrestricted information acquisition parallels the one in global games (Carlsson and Van Damme [1993]), but under very different informational assumptions.

Assume $n < \infty$. To formally prove risk-dominance selection, the key observation is that the risk-dominant action is the maximizer of the potential: given $t \in \{0, 1\}$, if action t is risk-dominant at θ, then

$$v(t, \dots, t, \theta) > v(a', \theta), \quad \forall a' \neq (t, \dots, t).$$

Given $C_i = \lambda I$ for all players i, since both invest and not invest have dominance regions, the hypothesis of Proposition 2 is satisfied, and therefore the following risk-dominance selection result holds:

Corollary 2. *Consider an investment game with $n < \infty$. Assume $C_i = \lambda I$ for all $i \in N$. For every λ, select an equilibrium $\boldsymbol{x}_\lambda$ in direct signals. Then, for all $i \in N$,*

$$\lim_{\lambda \to 0} P(\boldsymbol{x}_{i,\lambda} = 1 | \boldsymbol{\theta}) = \begin{cases} 1 & \textit{if invest is risk dominant at } \boldsymbol{\theta}, \\ 0 & \textit{if not invest is risk dominant at } \boldsymbol{\theta}. \end{cases}$$

Corollary 2 says that the players coordinate on the risk-dominant action, when the game is finite and the cost of information is negligible. While the proof of Corollary 2 relies on mutual information and the potential structure of the game, a more primitive intuition can be given. For simplicity, consider the case of two players i (she) and j (he). Take the prospective of player i. For intermediate values of the state, player i has a strong incentive to correlate her direct signal with j's direct signal, in order to match his action. Furthermore, recall that it is riskier to mismatch the risk-dominant action rather than then the risk-dominated one, under complete information. For negligible cost of information, this implies that player i has a stronger incentive to correlate her signal with j's signal when his signal suggests him to play the risk-dominant action. For instance, if invest is risk dominant at θ, this means that

$$\frac{P(\boldsymbol{x}_i \neq \boldsymbol{x}_j | \boldsymbol{x}_j = 1, \theta)}{P(\boldsymbol{x}_i = \boldsymbol{x}_j | \boldsymbol{x}_j = 1, \theta)} < \frac{P(\boldsymbol{x}_i \neq \boldsymbol{x}_j | \boldsymbol{x}_j = 0, \theta)}{P(\boldsymbol{x}_i = \boldsymbol{x}_j | \boldsymbol{x}_j = 0, \theta)}.$$

But then player j realizes that he is more likely to be matched if he plays the risk-dominant action. This channel fosters coordination on the risk-dominant action, underlying Corollary 2.[21]

The driving force of Corollary 2 is the players' incentive to learn what the others know. In the informal argument provided above, this is reflected by player i trying to correlate her direct signal with j's direct signal, state by state. When information choice is independent, i's choice variable is $\boldsymbol{x}_i|\boldsymbol{\theta}$, and not $\boldsymbol{x}_i|(\boldsymbol{x}_j, \boldsymbol{\theta})$. Therefore, this force is mute, if the players can choose only conditionally independent signals given the state. In fact, with independent and flexible information acquisition, Yang [2015] shows that investment games have multiple equilibria, when the cost of information is sufficiently small. In his analysis, the coordination problem in the investment decision translates into a coordination in the choice of information about the state, generating multiplicity. Yang [2015] adopts mutual information as cost of information, as I do in this application of the model. In his formulation, the cost of signal $\boldsymbol{x}_i$ is given by $\lambda I(\boldsymbol{x}_i; \boldsymbol{\theta})$. When signals are conditionally independent given the state, $I(\boldsymbol{x}_i; \boldsymbol{\theta}) = I(\boldsymbol{x}_i; \boldsymbol{x}_{-i}, \boldsymbol{\theta})$. Hence, dropping the independence assumption, unrestricted information acquisition can be seen as a generalization of the model of Yang [2015]. Therefore, the difference between the limit uniqueness of Corollary 2 and the multiplicity of Yang [2015] cannot be ascribed to different functional form assumptions on the cost of information.

Corollary 2 mimics the risk-dominance selection of Carlsson and Van Damme [1993] for global games, but under very different informational assumptions. The global game approach provides a natural perturbation of complete information that selects a unique equilibrium in investment games. With complete information, investment games have multiple equilibria for intermediate values of the state: both "all invest" and "all not invest" are strict equilibria. In global games, the players' information is incomplete and exogenously given: in the canonical formulation, the players observe the state plus some noise, independent of the state and across players. As the noise becomes small, the players coordinate on the risk-dominant action. With information acquisition, the case of negligible cost of information can be seen as an

[21]For this argument to go through, it is necessary that, for every state, the players both invest and do not invest with some probability. With mutual information, this happens for two reasons: dominance regions, and absolute continuity of the joint distribution of signals and state with respect to the product of their marginals, in equilibrium (Lemma 3). This implication of mutual information is shared by the class of cost functions based on f-divergences I presented in Section 3.2. These cost functions, therefore, provide a natural starting point to extend Proposition 2 beyond mutual information.

alternative perturbation of complete information.[22] With unrestricted information acquisition, for the case of finitely many players, Corollary 2 also selects a unique equilibrium, and the selection coincides with the one in global games.

6.3 Multiplicity for $n = \infty$

With infinitely many players, I show that multiple equilibria arise. Since the game is large, the players' incentive to acquire information about each other's information weaken. The players may coordinate on acquiring information about different aspects of the state, generating multiplicity. In fact, if $n = \infty$, there are multiple equilibria where the players behave as if information acquisition was independent, with mutual information as cost of information, as in Yang [2015].

When there are infinitely many players, the coordination problem in the investment decision translates into a coordination problem in the choice of information about the state. Hence, multiple equilibria arise. For an intuition, assume the players believe that nonfundamental volatility vanishes, i.e., $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) = 0$. Since the return on investment depends on the individual actions only through their average, the players are happy with overlooking information about each other's information, if they believe that the average can be inferred from the state. Hence, they choose conditionally independent signals given the state, and, by the law of large numbers, nonfundamental volatility does vanish, and their belief is correct in equilibrium. Overall, this means that it is possible to sustain equilibria where the players behave as if information acquisition was independent, when $n = \infty$. In these equilibria, the player focus on information about the state, and may coordinate on acquiring information about different threshold events $\{\boldsymbol{\theta} > \hat{\theta}\}$, and invest when the event realizes, and do not invest otherwise. Different thresholds correspond to different equilibria, generating multiplicity.

The next proposition formalizes this observation:

Proposition 4. *Consider an investment game with $n = \infty$ and linear return. Assume $C_i = \lambda I$ for all $i \in N$. Moreover, suppose that:*

- *$P_{\boldsymbol{\theta}}$ is absolutely continuous with respect to the Lebesgue measure.*

[22]This analogy is suggestive but not literally correct. For instance, according to Proposition 2, in the limit of negligible cost of information, by observing their own signal, the players can tell if invest or not invest is risk dominant, but they cannot determine what is the realized state, as under complete information.

- *$E[e^{-\rho(1,\boldsymbol{\theta})/\lambda}] > 1$ and $E[e^{\rho(0,\boldsymbol{\theta})/\lambda}] > 1$.*

If $r > 4\lambda$, then there exist infinitely many equilibria in direct signals that are conditionally independent given the state. In particular, for every threshold $\hat{\theta} \in [0, r]$, there are equilibria $\{\boldsymbol{x}_\lambda\}$ in direct signals such that, for all players i, $(\boldsymbol{x}_{i,\lambda} \perp \boldsymbol{x}_{-i,\lambda})|\boldsymbol{\theta}$ and

$$\lim_{\lambda \to 0} P(\boldsymbol{x}_{i,\lambda} = 1|\boldsymbol{\theta}) = \begin{cases} 1 & \text{if } \boldsymbol{\theta} > \hat{\theta}, \\ 0 & \text{if } \boldsymbol{\theta} \leq \hat{\theta}. \end{cases}$$

Proposition 4 says that multiple equilibria emerge, when there are infinitely many players, and the cost of information is sufficiently small. The proposition is an immediate consequence of the analysis of Yang [2015]. When $n = \infty$, the equilibria he found under independence can be replicated with unrestricted information acquisition. Following Yang [2015], I focus on the case of linear return, and assume that distribution of the state is absolutely continuous with respect to the Lebesgue measure.[23] When n is finite and the cost of information is negligible, Corollary 2 says that the players invest when the state is above the risk-dominance threshold, and do not invest otherwise. When $n = \infty$ and the cost of information is negligible, Proposition 4 says that, for every threshold outside the dominance regions, there is an equilibrium where the players invest when the state is above the threshold, and do not invest otherwise.

In global games, the risk-dominance selection result does not depend on the number of players in the game. Therefore, while Corollary 2 parallels it, Proposition 4 provides a sharp contrast. This difference emerges because, with information acquisition, the size of the game modify the players' incentives to acquire information, and therefore the players' information in equilibrium. If, instead, information is exogenously given as in global games, this channel is not present, and there is no connection between the size of the game and the players' information.

The distinction between finite and infinite investment games is not simply a theoretical curiosity. In fact, most of the applications of investment games have a large set of atomless players. When information is exogenously given, the choice of the size of the game is mostly driven by tractability considerations. Corollary 2 and Proposition 4 show that, with information acquisition, the size of the game matters also for equilibrium predictions.

[23] The inequalities $E[e^{-\rho(1,\boldsymbol{\theta})/\lambda}] > 1$ and $E[e^{\rho(0,\boldsymbol{\theta})/\lambda}] > 1$ rule out degenerate equilibria.

6.4 Equilibrium Characterization for $n < \infty$

I conclude the analysis of investment games by providing a general equilibrium characterization for finitely many players and positive cost of information. The characterization shows that all equilibria are symmetric and parametrized by the marginal probability of invest. Moreover, signals and state are affiliated. Biased versions of the risk-dominance thresholds arise to determine of the investment decision. I also illustrate these facts in the concrete example of two players and linear return.

Investment games with finitely many players are potential games. Therefore, with mutual information as cost of information, the results presented in Section 5.2 apply, and can be used to provide the following characterization of equilibria in direct signals:

Proposition 5. *Consider an investment game with $n < \infty$. Assume $C_i = \lambda I$ for all $i \in N$. Furthermore, suppose that $E[e^{-\rho(1,\boldsymbol{\theta})/\lambda}] > 1$ and $E[e^{\rho(0,\boldsymbol{\theta})/\lambda}] > 1$. Direct signals $\boldsymbol{x}$ arise in equilibrium if and only if there is $p \in (0,1)$ such that the following two conditions hold: for all $a \in A$,*

$$P(\boldsymbol{x} = a|\theta) = \frac{\prod_{l=0}^{|a|-1} e^{\rho(\frac{l}{n-1},\theta)/\lambda}\binom{n}{|a|}p^{|a|}(1-p)^{n-|a|}}{\sum_{m=0}^{n}\prod_{l=0}^{m-1} e^{\rho(\frac{l}{n-1},\theta)/\lambda}\binom{n}{m}p^{m}(1-p)^{m-n}}, \quad a.s., \tag{6}$$

and

$$\int_{\Theta} \frac{\sum_{m=0}^{n-1}\prod_{l=0}^{m-1} e^{\rho(\frac{l}{n-1},\theta)/\lambda}\binom{n-1}{m}p^{m}(1-p)^{m-n}}{\sum_{m=0}^{n}\prod_{l=0}^{m-1} e^{\rho(\frac{l}{n-1},\theta)/\lambda}\binom{n}{m}p^{m}(1-p)^{m-n}} dP_{\boldsymbol{\theta}}(\theta) = 1. \tag{7}$$

In particular, an equilibrium in direct signals exists. Moreover:

(i) $P(\boldsymbol{x}_i = 1) = p$ for all $i \in N$, and signals are exchangeable.

(ii) Signals and state are affiliated.

(iii) $P(\{\boldsymbol{x}_i = 1, \forall i \in N\}|\{\boldsymbol{x}_i = \boldsymbol{x}_j, \forall i,j \in N\}, \theta)$ is non-decreasing in θ and equal to one half for all states $\hat{\theta}$ such that

$$\frac{1}{n}\sum_{l=0}^{n-1}\rho(l,\hat{\theta}) = \lambda \log \frac{1-p}{p}. \tag{8}$$

Proposition 5 characterizes equilibrium behavior in terms of p, the marginal probability of invest. Different marginal probabilities of invest correspond to different equilibria. Multiple equilibria arise when equation (7) has multiple solutions in p. The inequalities $E[e^{-\rho(1,\boldsymbol{\theta})/\lambda}] > 1$ and $E[e^{\rho(0,\boldsymbol{\theta})/\lambda}] > 1$ rule out degenerate equilibria.[24]

In equilibrium, all the players invest with the same marginal probability. More broadly, signals are exchangeable: this means that equilibrium behavior is symmetric. The intuition for symmetry parallels the one under complete information: players who invest more have weaker incentives to invest than players who invest less, since the opponents of the former invest less than the opponents of the latter. Hence, the players must invest with the same probability, and equilibrium behavior must be symmetric.

Proposition 5 also says that, in equilibrium, signals and state are affiliated, i.e., they positively depend on each other (Milgrom and Weber [1982]). Since the return on investment is non-decreasing in the state and the share of opponents who decide to invest, actions are strategic complements, and incentives to invest are monotone in the state. These complementarities translate into affiliation of the signal structure.

By affiliation, the conditional probability of "all invest" given "coordination" is monotone in the state, i.e., $P(\{\boldsymbol{x}_i = 1, \forall i \in N\} | \{\boldsymbol{x}_i = \boldsymbol{x}_j, \forall i, j \in N\}, \theta)$ is non-decreasing in θ. In addition, Proposition 5 says that, given coordination, the players are equally likely to invest and not invest at states $\hat{\theta}$ satisfying (8). For instance, with linear return,

$$\hat{\theta} = \frac{r}{2} + \lambda \log \frac{1-p}{p}.$$

These states can be interpreted as biased risk-dominance thresholds. The bias is positive when $p < 1/2$, and negative otherwise. The size of the bias is decreasing in the cost of information λ.

Proposition 5 can be used to further investigate equilibrium behavior under different specification of the return function. For instance, for the case of two players and linear return, Figure 1 depicts the conditional probability of invest $P(\boldsymbol{x}_i = 1|\theta)$ and the conditional probability of coordination $P(\boldsymbol{x}_1 = \boldsymbol{x}_2|\theta)$ as a function of the state θ.

Figure 1 can be interpreted as follows. In equilibrium, the players follow a threshold strategy corresponding to $\hat{\theta}$. The probability of invest (solid line) is increasing in θ and steepest at $\hat{\theta}$: the players acquire information about the event $\{\boldsymbol{\theta} \geq \hat{\theta}\}$, and invest when the event is realized, and do not invest otherwise. Coordination is

[24] From Lemma 3, there is an equilibrium such that all the players invest for sure if and only if $E[e^{-\rho(1,\boldsymbol{\theta})/\lambda}] \leq 1$. Analogously, there is an equilibrium such that all the players do not invest for sure if and only if $E[e^{\rho(0,\boldsymbol{\theta})/\lambda}] \leq 1$.

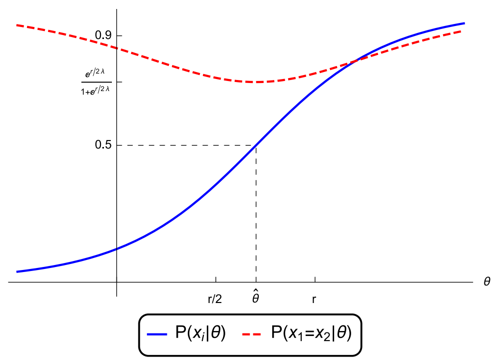

Figure 1: Equilibrium for two players and linear return.

therefore harder for values of θ around the threshold, while it is easier to coordinate on tail events: the probability of coordination (dashed line) is minimized at $\hat{\theta}$, and increases as we move away from the threshold. Nevertheless, the amount of coordination the players achieve is bounded by $e^{r/2\lambda}/(1+e^{r/2\lambda})$. The bound is increasing in r, the degree of strategic complementarities in actions, and decreasing in λ, the cost of information. Furthermore, the bound is tight, since it goes to 1 whenever r goes to infinity or λ goes to zero.

7 Linear-Quadratic Games

In the previous section, the analysis of investment games provides a sharp contrast between information choice in finite and large games: a risk-dominance selection result holds in finite investment games, while in large investment games multiple equilibria emerge. This happens because the number of players in the game affects the impact of the players' actions on each other's utilities, and therefore their incentive to learn what the others know. In this section, to investigate more deeply the relation between structure of the game and information choice, I consider linear-quadratic games, where the impact of the players' actions on each other's utilities can be summarized by a network matrix.[25]

I characterize the information the players in equilibrium acquire about each other's information in terms of the underlying network. In particular, I show that nonfundamental volatility arise as an average of the players' centralities in the network,

[25]See, e.g., Bergemann and Morris [2013, Section 2.3] for an overview of the many economic applications of linear-quadratic games, when information is exogenously given.

overall measures of the players' incentive to learn what the others know. An implication of this result is that, in large networks without central players, nonfundamental volatility vanishes and the players behave as if information choice was independent. I derive these conclusions for general cost of information (only Conditions 1 and 2), focusing on equilibria where the players' direct signals are normally distributed (but of course all deviations are allowed). Finally, to provide a concrete example of large linear-quadratic games, I consider the special case of beauty contests (Morris and Shin [2002]) with mutual information as cost of information. I show that beauty contests with mutual information have a unique Gaussian equilibrium, and I use this uniqueness result to carry out comparative statistic analysis.

7.1 Setup

Let $n < \infty$ and $A_i = \mathbb{R}$ for all $i \in N$. Assume $\Theta = \mathbb{R}^m$ for some positive integer m, and $\boldsymbol{\theta} \sim \mathcal{N}(\mu_{\boldsymbol{\theta}}, \Sigma_{\boldsymbol{\theta}})$. The (undirected) network linking the players is represented by a symmetric $n \times n$ matrix Γ such that $\gamma_{ii} = 0$ for all $i \in N$. The matrix Γ is a contraction: $\sum_{j \neq i} |\gamma_{ij}| < 1$ for all $i \in N$.[26] Player i's utility: for all $a \in A$ and $\theta \in \Theta$,

$$u_i(a, \theta) = -\frac{1}{2}(a_i - w_i)^2, \quad \text{with } w_i = \sum_{j \neq i} \gamma_{ij} a_j + \sum_{l=1}^{m} \delta_{il} \theta_l.$$

The variable w_i is player i's target. The weight δ_{il} is some arbitrary real number, possibly zero.

Centralities. Denote by b_i the *Katz-Bonacich centrality* of player i in network Γ (Katz [1953], Bonacich [1987]):

$$b_i = \frac{1}{n} \sum_{j \in N} \left(\sum_{k=0}^{\infty} \Gamma^k \right)_{ij}.$$

The number $\left(\sum_{k=0}^{\infty} \Gamma^k\right)_{ij}$ counts the total number of walks from player i to player j in network Γ. The Katz-Bonacich centrality b_i counts the total number of walks starting from i (which I scale by the number of players in the game).[27] Henceforth, I

[26] Γ is a contraction when $\mathbb{R}^n$ is endowed with the L^1-norm, and $\mathbb{R}^{n \times n}$ with the corresponding operator norm.

[27] The centrality is well defined since Γ is a contraction.

refer to b_i simply as the centrality of player i.[28]

Potential. Linear-quadratic games on an undirected network are potential games. For instance, a potential is the function $v : A \times \Theta \to \mathbb{R}$ such that

$$v(a, \theta) = \sum_{i \in N} \sum_{l=1}^{m} \delta_{il} \theta_l a_i - \frac{1}{2} \sum_{i \in N} a_i^2 + \frac{1}{2} \sum_{i \in N} \sum_{j \neq i} \gamma_{ij} a_i a_j.$$

7.2 Information Choice and the Network Structure

I characterize the information the players in equilibrium acquire about each other's information in terms of the underlying network. The characterization shows that what player i know about what player j knows depends not only on the link between i and j, but on the entire network. This happens because i values j's information also for what she can learn from j's information about the other players' information. I also show that nonfundamental volatility is an average of the players' centralities in the network, which are overall measures of the players' incentive to learn what the others know.

In the analysis of linear-quadratic games, I focus on Gaussian equilibria:[29] direct signals are *Gaussian* if the vector $(\boldsymbol{x}, \boldsymbol{\theta})$ has the multivariate normal distribution.[30] The Gaussianity assumption is customary in the literature on linear-quadratic games, and this common ground facilitates the comparison between my results and existing findings. Here, of course, I allow all possible deviations.[31]

The next proposition provides an equilibrium characterization of the the players' information about each other's information in terms of the underlying network:

Proposition 6. *Consider a linear-quadratic game. Let* $\boldsymbol{x}$ *be a Gaussian equilibrium*

[28] See, e.g., Jackson [2008] for more background on this and related definitions of centrality.

[29] It is well-known that, with mutual information as cost of information, Gaussian signals are (individually) optimal when utility is linear-quadratic and uncertainty is normally distributed. In fact, exploiting the potential structure of the game, the results of Section 5.2 can be used to verify that Gaussian equilibria exist in this setting, at least with mutual information as cost of information.

[30] It should be noted that the covariance matrix of the vector $(\boldsymbol{x}, \boldsymbol{\theta})$ may be singular.

[31] Nevertheless, the results of this section could have been derived only by allowing deviations in Gaussian signals. Such restrictions would make easier to prove equilibrium existence for general cost of information.

in direct signals. Denote by $\tilde{\Gamma}$ *the (directed) network such that, for all* $i, j \in N$,

$$\tilde{\gamma}_{ij} = \gamma_{ij} \frac{Var(\boldsymbol{x}_j) - Var(\boldsymbol{x}_j | \boldsymbol{x}_{-j}, \boldsymbol{\theta})}{Var(\boldsymbol{x}_j)}, \quad \textit{if } Var(\boldsymbol{x}_j) > 0.$$

If $Var(\boldsymbol{x}_j) = 0$, *set* $\tilde{\gamma}_{ij} = 0$. *Then, for all players* i *and* j,

$$Cov(\boldsymbol{x}_i, \boldsymbol{x}_j | \boldsymbol{\theta}) = Var(\boldsymbol{x}_i | \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \left(\sum_{l=0}^{\infty} \tilde{\Gamma}^l \right)_{ij}.$$

In particular, given $\bar{\boldsymbol{x}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x}_i$,

$$Var(\bar{\boldsymbol{x}} | \boldsymbol{\theta}) = \frac{1}{n} \sum_{i \in N} Var(\boldsymbol{x}_i | \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \tilde{b}_i,$$

where $\tilde{b}_i$ is i's centrality in $\tilde{\Gamma}$.[32]

Proposition 6 says that, in equilibrium, player i's information about player j's information is given by the sum of all walks from player i to player j in the "adjusted network" $\tilde{\Gamma}$ (scaled by $Var(\boldsymbol{x}_i | \boldsymbol{x}_{-i}, \boldsymbol{\theta})$). In particular, nonfundamental volatility (measured by $Var(\bar{\boldsymbol{x}} | \boldsymbol{\theta})$) is given by the average of the players' centralities in $\tilde{\Gamma}$ (scaled by $Var(\boldsymbol{x}_i | \boldsymbol{x}_{-i}, \boldsymbol{\theta})$). In $\tilde{\Gamma}$, the primitive link from i to j is adjusted by the share of the volatility of j's signal that cannot be reduced by acquiring information about the state and about information held by players other than j. The adjustment term can also be interpreted as the precision of j's signal, as measured by the correlation between j's signal and j's target. In fact, in equilibrium, $(\boldsymbol{x}_j \perp (\boldsymbol{x}_{-j}, \boldsymbol{\theta})) | \boldsymbol{w}_j$ (Lemma 1). Therefore,

$$\frac{Var(\boldsymbol{x}_j) - Var(\boldsymbol{x}_j | \boldsymbol{x}_{-j}, \boldsymbol{\theta})}{Var(\boldsymbol{x}_j)} = Cor(\boldsymbol{x}_j, \boldsymbol{w}_j)^2.$$

Proposition 6 shows that, in equilibrium, player i's information about player j's information is determined not only by the link between i and j, but by the entire network. This happens because player i has two two different incentives to learn what player j knows. On one hand, j's action affects i's target. Therefore, to match her own target, i values j's information in proportion to the strength of their link. On the other hand, j's information can also be used to infer what the other players know.[33]

[32] Since Γ is contraction, $\tilde{\Gamma}$ is a contraction. Hence, $\tilde{b}_i$ is well defined.

[33] In principle, player i has also a third incentive to learn what player j knows: she can use j's

Hence, i values j's information also in proportion to (i) the strength of her link with players other than j, and (ii) j's information about the opponents shared by i and j. However, j's information about others' information also depends on the strength of all j's links, and so forth. In conclusion, in equilibrium, the entire network structure matters for the players' information about each other's information.[34]

When information is exogenously given, the literature on linear-quadratic games has emphasized public information as key source of nonfundamental volatility (e.g., Morris and Shin [2002], Angeletos and Pavan [2007]).[35] Proposition 6 highlights an alternative source of nonfundamental volatility: the players' incentive to acquire information about each other's information. In fact, player i's centrality can be seen as an aggregate measure of her opponents' incentive to learn what she knows.

7.3 Bound on Nonfundamental Volatility

When no player is central in the network, I show that all the players focus on information about the state and nonfundamental volatility vanishes. This happens, for instance, in large networks of atomless players: an information aggregation result cancels the noise of the players' targets and weaken their incentive to acquire information about each other's information.

The next corollary of Proposition 6 provides a bound for the information the players acquire about each other's information, and therefore for nonfundamental volatility in equilibrium:

Corollary 3. *Consider a linear-quadratic game. Let $\boldsymbol{x}$ be a Gaussian equilibrium in direct signals. Denote by $\hat{\Gamma}$ the (undirected) network such that, for all $i, j \in N$,*

$$\hat{\gamma}_{ij} = |\gamma_{ij}|.$$

information to make inference about the state. Here, since all statements are "given the state," this incentive does not appear.

[34] A similar point has recently been made by Calvó-Armengol et al. [2015] in a model of costly communication, which can be reinterpreted as a model of rigid information choice.

[35] The interplay of independent information choice and public information is studied by Colombo et al. [2014].

Then, for all players i and j,

$$|Cov(\boldsymbol{x}_i, \boldsymbol{x}_j|\boldsymbol{\theta})| \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|)^2} \left(\sum_{k=0}^{\infty} \hat{\Gamma}^k \right)_{ij},$$

where $\boldsymbol{\theta}_i = \sum_{l=1}^{m} \delta_{il} \boldsymbol{\theta}_l$. In particular, given $\bar{\boldsymbol{x}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x}_i$,

$$Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|)^2} \frac{1}{n} \sum_{i=1}^{n} \hat{b}_i,$$

where $\hat{b}_i$ is i's centrality in $\hat{\Gamma}$.[36]

Corollary 3 says that, in equilibrium, player i's information about player j's information is bounded by the sum of all walks from player i to player j in the original network Γ (in absolute value), modulo a constant factor. This implies that nonfundamental volatility (measured by $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta})$) is bounded by the average of the players' centralities in Γ (in absolute value), modulo a constant factor.

Corollary 3 shows that, if no player is central in the network, all the players focus on information about the state and nonfundamental volatility vanishes. This happens, of course, when $\Gamma = 0$ and there is no strategic interaction. More interestingly, this happens also when the network is large and contain many atomless players. To see this, consider the case of identical players: for all players i and j with $i \neq j$,

$$\gamma_{ij} = r \frac{1}{n-1}, \quad r \in (-1, 1).$$

When players are identical,

$$\max_{j \in N} \sum_{i \in N} |\gamma_{ij}| = |r| \quad \text{and} \quad \left(\sum_{k=0}^{\infty} \hat{\Gamma}^k \right)_{ij} \leq 1(i = j) + \frac{|r|}{n - 1 - |r|}.$$

Hence, Corollary 3 implies that, for all players i and j with $i \neq j$,

$$|Cov(\boldsymbol{x}_i, \boldsymbol{x}_j|\boldsymbol{\theta})| \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - |r|)^2} \frac{|r|}{n - 1 - |r|}.$$

[36] Since Γ is contraction, $\hat{\Gamma}$ is a contraction. Hence, $\hat{b}_i$ is well defined.

Moreover,

$$Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1-|r|)^2} \left(\frac{1}{n} + \frac{|r|}{n-1-|r|} \right)$$

As the number of players increases, the upper bounds become closer and closer to zero. Therefore, the players acquire less and less information about each other's information, and focus more and more on the state. In particular, nonfundamental volatility vanishes.

Corollary 3 follows almost immediately from Proposition 6.[37] However, in the case of identical players, a more primitive intuition can be given to understand why nonfundamental volatility vanishes as the game gets large. The driving force is an information aggregation result, which cancels the noise of the players' targets, and make the players willing to acquire information only about the state. To provide more intuition, here I sketch an informal argument:

"Proof". For simplicity, I give the argument "in the limit," i.e., for infinitely many identical players. In addition, assume that the state is unidimensional and the players' common target is $w = r_0\theta + r\bar{a}$, where $r_0 \in \mathbb{R}$ and $\bar{a}$ is the players' average action. I wish to show that $Var(\boldsymbol{w}|\boldsymbol{\theta}) = 0$, i.e., the noise of the target vanishes. Sufficiently, I argue that $Cor(\bar{\boldsymbol{x}}, \boldsymbol{\theta})^2 = 1$, i.e., the average action aggregates the players' individual information and is perfectly correlated with the state.

First, for all players i, observe that utility u_i depends on (a, θ) only through w. Hence, $\boldsymbol{w}$ is sufficient to explain the dependence of $\boldsymbol{x}_i$ on $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ (Lemma 1). In particular, this implies that the players' signals are conditionally independent given the target. Furthermore, since $|r| < 1$ (contraction), it can be shown that $Var(\boldsymbol{x}_i)$ is uniformly bounded, and therefore also $Var(\boldsymbol{x}_i|\boldsymbol{w})$. As a consequence, a law of large numbers kicks in and gives $\bar{\boldsymbol{x}} = E[\bar{\boldsymbol{x}}|\boldsymbol{w}]$. By Gaussianity, $E[\bar{\boldsymbol{x}}|\boldsymbol{w}]$ is linear in $\boldsymbol{w}$, and therefore also linear in $\boldsymbol{\theta}$ and $\bar{\boldsymbol{x}}$. Hence, overall, the equality $\bar{\boldsymbol{x}} = E[\bar{\boldsymbol{x}}|\boldsymbol{w}]$ implies that $\boldsymbol{\theta}$ and $\bar{\boldsymbol{x}}$ are linearly related, i.e., $Cor(\bar{\boldsymbol{x}}, \boldsymbol{\theta})^2 = 1$, as wanted. □

7.4 Beauty Contests with Mutual Information

In beauty contests, I characterize the unique Gaussian equilibrium in direct signals, with mutual information as cost of information. Beauty contests are linear-quadratic

[37] The fact that Γ is a contraction is used to uniformly bound $Var(\boldsymbol{x}_i)$, and therefore $Var(\boldsymbol{x}_i|\boldsymbol{x}_{-i}, \boldsymbol{\theta})$.

games with infinitely many identical players. In beauty contests, no players is central in the network. Hence, in equilibrium, nonfundamental volatility vanishes and the players behave as if information acquisition was independent. In the concrete example of mutual information as cost of information, I use this result to provide a closed-form equilibrium characterization, and carry out comparative statistic analysis. In particular, the players' choice of information about the state inherits the strategic motives for actions, as in Hellwig and Veldkamp [2009].

For the rest of this section, assume $n = \infty$. Let again $A_i = \mathbb{R}$ for all $i \in N$, and suppose the state is unidimensional and normally distributed with mean $\mu_{\boldsymbol{\theta}}$ and variance $\sigma^2_{\boldsymbol{\theta}}$. Player i's utility is

$$u_i(a, \theta) = -\frac{1}{2}(a_i - w)^2, \quad \text{with } w = r_0\theta + r\bar{a}.$$

The players' common target is w, a linear combination of the state and the average action $\bar{a}$. The weights r_0 and r are real numbers with $|r| < 1$ (contraction).

With infinitely many players, I say that direct signals are *Gaussian* if the following two conditions hold:

(i) For all $i \in N$, the vector $(\theta, \boldsymbol{x}_1, \dots, \boldsymbol{x}_i)$ has the multivariate normal distribution.

(ii) The sequence $\left\{\frac{1}{m}\sum_{i=1}^m \boldsymbol{x}_i\right\}$ converges in L_2 to the average action $\bar{\boldsymbol{x}}$.

Notice that the definition extends the one previously given for finitely many players. Moreover, it implies that also the vector $(\boldsymbol{\theta}, \boldsymbol{x}_1, \dots, \boldsymbol{x}_i, \bar{\boldsymbol{x}})$ has the multivariate normal distribution, for all $i \in N$.

The next proposition characterizes the unique Gaussian equilibrium in beauty contests, with mutual information as cost of information:

Proposition 7. *Consider a beauty contest. Assume $C_i = \lambda I$ for all players i, with $\lambda < \sigma^2_{\boldsymbol{\theta}} r_0^2$. Then there exists a unique Gaussian equilibrium $\boldsymbol{x}$ in direct signals. To describe the equilibrium, let $\{\boldsymbol{\epsilon}_i\}$ be a sequence of standard normal random variables, independent of each other and of the state. Then, for all players i,*

$$\boldsymbol{x}_i = \frac{r_0}{1-r}(\mu_{\boldsymbol{\theta}} + \alpha(\boldsymbol{\theta} - \mu_{\boldsymbol{\theta}}) + \beta\boldsymbol{\epsilon}_i),$$

where the coefficients α *and* β *are given by*

$$\alpha = \sqrt{\left(\frac{1}{2r} - 1\right)^2 + \frac{(1-r)}{r}\left(1 - \frac{\lambda}{\sigma_{\boldsymbol{\theta}}^2 r_0^2}\right)} - \left(\frac{1}{2r} - 1\right),$$
$$\beta = \sigma_{\boldsymbol{\theta}}\sqrt{\alpha(1-\alpha)(1-r)}.$$

Proposition 7 says that, in equilibrium, the players' signals are conditionally independent given the state. In particular, this implies that nonfundamental volatility vanishes (i.e., $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) = 0$). This result is a limit version of Proposition 3, and it can be shown also without assuming mutual information as cost of information. Mutual information is used to uniquely pin down equilibrium behavior (in the class of Gaussian equilibria).

Hellwig and Veldkamp [2009] and Myatt and Wallace [2012] analyze beauty contests with rigid information acquisition, and point out that multiple Gaussian equilibria emerge, with mutual information as cost of information.[38] In both models, due to the rigidity in information choice, nonfundamental volatility does not vanish in equilibrium. For instance, in the specification of Myatt and Wallace [2012], the signals share some common noise that the players' parametrized information choice cannot affect (see Section 9.4 for a review of their model). Dropping the rigidity assumption, Proposition 7 shows that in equilibrium the players behave as if information acquisition was independent, and nonfundamental volatility does vanish. For this reason, under unrestricted information acquisition, beauty contests have a unique Gaussian equilibrium, with mutual information as cost of information.[39]

The characterization provided by Proposition 7 can be used to carry out comparative static analysis. Here, I illustrate how the equilibrium is affected by changes in r, the degree of strategic interaction.

Figure 2 shows that the more complementary the actions are, the more information the players acquire about the state: $|Cor(\boldsymbol{x}_i, \boldsymbol{\theta})|$ is increasing in r. Figure 2 also shows that the more complementary the actions are, the more correlated the signals

[38]In the model of Hellwig and Veldkamp [2009], multiplicity arises for general cost of information, unless independence is assumed on top of rigidity. In the model of Myatt and Wallace [2012], multiplicity arises when the cost of information does not satisfy a convexity assumption, which is not met by mutual information.

[39]In general, independent information choice does not guarantee equilibrium uniqueness, as shown by Yang [2015] in investment games. See also Section 6.3 on large investment games with unrestricted information acquisition.

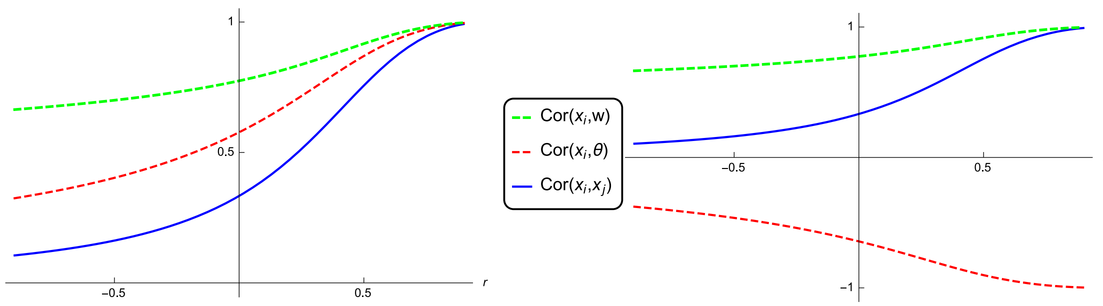

Figure 2: Effect of r on information choice. Left panel $r_0 > 0$, right panel $r_0 < 0$.

are with each other and the target: both $|Cor(\boldsymbol{x}_i, \boldsymbol{x}_j)|$ and $|Cor(\boldsymbol{x}_i, \boldsymbol{w})|$ are increasing in r. This happens because, under independence, the signals' correlation structure is explained by the state (i.e., $Cor(\boldsymbol{x}_i, \boldsymbol{x}_j) = Cor(\boldsymbol{x}_i, \boldsymbol{\theta})Cor(\boldsymbol{x}_j, \boldsymbol{\theta})$). Therefore, the more information the players acquire about the state, the more correlated their signals are.

Figure 2 connects the strategic motives for actions to information choice. In beauty contests, Hellwig and Veldkamp [2009] are the first to point out this connection. As they argue, the underlying channel is the impact of the degree of strategic interaction on the variance of the target, which is illustrated in Figure 3.

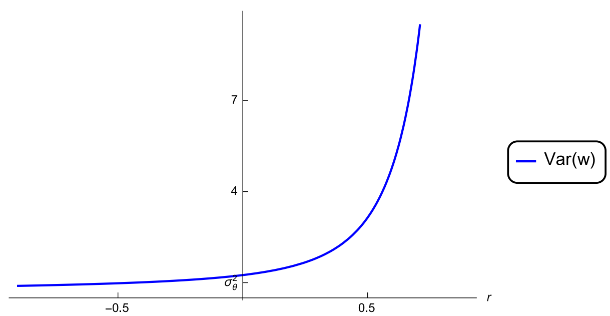

Figure 3: Effect of r on the target's variance.

Figure 3 shows that the variance of the target is increasing in the degree of strategic interaction. The variance of the target is a measure of the players' overall incentive to acquire information, and is given by

$$Var(\boldsymbol{w}) = r_0^2\sigma_{\boldsymbol{\theta}}^2 + r^2 Var(\bar{\boldsymbol{x}}) + 2r_0 r Cor(\bar{\boldsymbol{x}}, \boldsymbol{\theta}) Std(\bar{\boldsymbol{x}})\sigma_{\boldsymbol{\theta}}.$$

Since $Cor(\bar{\boldsymbol{x}}, \boldsymbol{\theta})$ has the same sign of r_0 (see Figure 2), their product is always positive.

As a result, when r is positive, the more information about the state the players acquire (i.e., the higher $|Cor(\bar{\boldsymbol{x}}, \boldsymbol{\theta})|$ is), the higher $Var(\boldsymbol{w})$ is, and therefore the higher the incentives to acquire information are. Conversely, when r is negative, the more information about the state the players acquire, the lower $Var(\boldsymbol{w})$ is, and therefore the lower the incentives to acquire information are. In this sense, the choice information about the state inherits the strategic motives for actions.

The connection between information choice and strategic motives for actions is also reflected by the impact of r on individual volatility $Var(\boldsymbol{x}_i)$, aggregate volatility $Var(\bar{\boldsymbol{x}})$, and dispersion $Var(\boldsymbol{x}_i - \bar{\boldsymbol{x}})$, as Figure 4 illustrates.

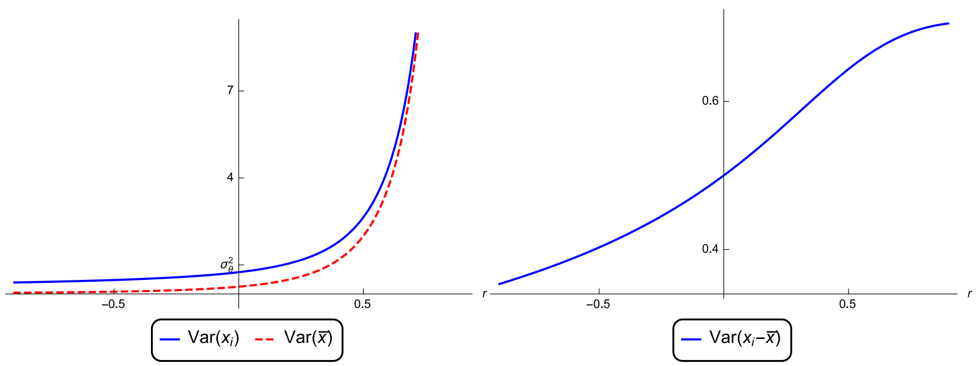

Figure 4: Effect of r on individual and aggregate volatility, and dispersion.

Figure 4 shows that individual volatility, aggregate volatility, and dispersion are increasing in r. In equilibrium, individual volatility $Var(\boldsymbol{x}_i)$ is the product of $Cor(\boldsymbol{x}_i, \boldsymbol{w})^2$ and $Var(\boldsymbol{w})$, and therefore it is increasing in r. Furthermore, aggregate volatility $Var(\bar{\boldsymbol{x}})$ is equal to $Cov(\boldsymbol{x}_i, \boldsymbol{x}_j)$, which is increasing in r as both $Var(\boldsymbol{x}_i)$ and $Cor(\boldsymbol{x}_i, \boldsymbol{x}_j)$ are. Dispersion $Var(\boldsymbol{x}_i - \bar{\boldsymbol{x}})$, instead, is the product of $Var(\boldsymbol{x}_i)$ and $1 - Cor(\boldsymbol{x}_i, \boldsymbol{x}_j)$, and therefore, ex ante, the effect of r is ambiguous: $Var(\boldsymbol{x}_i)$ increases, but $1 - Cor(\boldsymbol{x}_i, \boldsymbol{x}_j)$ decreases. Here, the positive effect on the individual volatility dominates.

8 Large Games

For large linear-quadratic games, the analysis in the previous section implies a stark conclusion: in Gaussian equilibria, nonfundamental volatility vanishes and independence in information choice can be assumed without loss of generality. In this section, I ask whether this conclusion extends to more general large games and beyond the class of Gaussian equilibria.[40]

With mutual information as cost of information, I show that the answer depends on the strategic motives for actions. When actions are strategic substitutes, nonfundamental volatility does vanish, and in equilibrium information choice is independent. If, instead, actions are strategic complements, nonfundamental volatility may not vanish, and in equilibrium the players may choose signals not conditionally independent given the state. The results are derived for large games where the players' utilities depend on each other's action only through a (possibly nonlinear) average of their actions. The setup generalizes both large investment games (Section 6.3) and beauty contests (Section 7.4).

8.1 Setup

Let $n = \infty$ and $A_i \subseteq \mathbb{R}$ for all players i. Associate a non-decreasing measurable function $f_i : A_i \to \mathbb{R}$ to each player i. Given action profile a, define the *average action* $\bar{a}$ *by*

$$\bar{a} = \limsup_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} f_i(a_i).$$

The value $f_i(a_i)$ measures the individual contribution of player i to the average action. Since f_i is non-decreasing, player i's contribution to the average action is positive. Overall, the players' utilities depend on the others' actions only through the average action: for all players i,

$$u_i(a, \theta) = u_i(a_i, \bar{a}, \theta), \quad \forall a \in A \text{ and } \forall \theta \in \Theta.$$

Utilities have increasing/decreasing differences in individual and average action. Individual and average action are *strategic complements* if, for all players i, $a_i \geq a_i'$ and $\bar{a} \geq \bar{a}'$ implies

$$u_i(a_i, \bar{a}, \theta) - u_i(a_i', \bar{a}, \theta) \geq u_i(a_i, \bar{a}', \theta) - u_i(a_i', \bar{a}', \theta), \quad \forall \theta \in \Theta.$$

[40] Large games (e.g., Kalai [2004]) are extensively used to represent economic environments with many atomless players. In this environments, individual deviations have small or null strategic relevance.

If the displayed inequality holds in the opposite direction, then individual and average action are *strategic substitutes*. Note that, since f_i is non-decreasing in a_i, also individual actions are strategic complements (substitutes, resp.) when individual and average action are strategic complements (substitutes, resp.).

8.2 Substitutability

When individual and average action are strategic substitutes, with mutual information as cost of information, I show that nonfundamental volatility vanishes, and the players behave as if information acquisition was independent. This happens because the strategic motives for actions are inherited by the players' choice of information about the others' information. From player i's perspective, it means that the more information i's opponents gather about each other's information, the less information player i wants to gather about her opponents' information.

The next proposition states that nonfundamental volatility vanishes, when actions are strategic substitutes:

Proposition 8. *Consider a large game. Assume $C_i = \lambda_i I$ for all $i \in N$. Let $\boldsymbol{x}$ be equilibrium direct signals such that $\sup_{i \in N} Var(f_i(\boldsymbol{x}_i)) < \infty$. If individual and average action are strategic substitutes, then $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) = 0$ and $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|\boldsymbol{\theta}$ for all players i.*

When individual and average action are strategic substitutes, Proposition 8 says that nonfundamental volatility $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta})$ vanishes. Since the players' utilities depend on the others' actions only through the average action, this also implies that signals are conditionally independent given the state. Indeed, by Lemma 1, the average action and the state are sufficient to explain the dependence of the players' signals on each other, i.e., $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|(\bar{\boldsymbol{x}}, \boldsymbol{\theta})$. Hence, since the average action is conditionally degenerate given the state, in equilibrium the state is sufficient to explain the dependence of the players' signals on each other, i.e., $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|\boldsymbol{\theta}$. The regularity condition $\sup_{i \in N} Var(f_i(\boldsymbol{x}_i)) < \infty$ is trivially satisfied when the functions f_i are uniformly bounded. In general, contraction properties of the game can also be used to established that $\sup_{i \in N} Var(f_i(\boldsymbol{x}_i)) < \infty$ (e.g., in beauty contests).

The proof of Proposition relies on mutual information and his implications for information choice (Lemma 2). However, a more primitive intuition can be given. The main idea is that nonfundamental volatility vanishes because the the strategic motives for actions are inherited by the players' choice of information about the others' information. For an intuition, take the perspective of player i. The key is the relation between $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta})$ and $Cov(\boldsymbol{x}_j, \boldsymbol{x}_k|\boldsymbol{\theta})$, where j and k represent i's opponents. Since i's utility depends on the others' actions only through the average action, player i's incentives to acquire information about her opponents' information can be measured by $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta})$. If i's opponents acquire more information about each other's information, then their signals depend more on each other, and therefore $|Cov(\boldsymbol{x}_j, \boldsymbol{x}_k|\boldsymbol{\theta})|$ increases. Moreover, since actions are strategic substitutes, the players want to "mismatch" each other's actions, and therefore $Cov(\boldsymbol{x}_j, \boldsymbol{x}_k|\boldsymbol{\theta})$ is negative. Overall, hence, if i's opponents acquire more information about each other's information, $Cov(\boldsymbol{x}_j, \boldsymbol{x}_k|\boldsymbol{\theta})$ decreases. However, $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta})$ is increasing in $Cov(\boldsymbol{x}_j, \boldsymbol{x}_k|\boldsymbol{\theta})$. As a consequence, in conclusion, if i's opponents gather more information about each other's information, then i's incentives to acquire information about her opponents' information become weaker. This means that the only equilibrium outcome is $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) = 0$, as Proposition 8 establishes.

8.3 Complementarity

When individual and average action are strategic complements, nonfundamental volatility may not vanish, and the players' signals may not be conditionally independent given the state. In an example, I illustrate that the players may be willing to buy correlation devices, due to the coordination motives for actions.

I use a simple example to provide a "converse" to Proposition 8. The example is a large investment game with state-independent utilities. Examples with nontrivial incentives to acquire information about the state can also be easily constructed. However, the case of state-independent utilities make the intuition more transparent.

Example 1. Consider a large investment game ($n = \infty$) with return on investment

$$\rho(\bar{a}, \theta) = \log\left(\frac{\bar{a}}{1 - \bar{a}}\right).$$

Note that the return is (strictly) increasing in the average action, and it does not depend on the state. Therefore, individual and average action are strategic complements (but not strategic substitutes), and the players do not have incentives to acquire information about the state. For the cost of information, assume $C_i = I$ for all players i.

Now I construct equilibria where the players coordinate on buying a "correlation device." Fix $\gamma \in [1/2, 1)$, and let $\boldsymbol{z}$ be a binary random variable, independent of the state, taking values either γ or $(1-\gamma)$ with equal probabilities. Using Lemma 2, it is easy to verify that there is an equilibrium in direct signals $\boldsymbol{x}$ such that

$$P(\boldsymbol{x}_i = 1 | \boldsymbol{x}_{-i}, \boldsymbol{\theta}) = P(\boldsymbol{x}_i = 1 | \boldsymbol{z}) = \boldsymbol{z}, \quad \forall i \in N.$$

In the equilibrium, $\boldsymbol{z} = \bar{\boldsymbol{x}}$ and therefore

$$Var(\bar{\boldsymbol{x}} | \boldsymbol{\theta}) = Var(\boldsymbol{z} | \boldsymbol{\theta}) = Var(\boldsymbol{z}) = 1/4 - \gamma(1-\gamma).$$

Moreover, the joint distribution of the signals of players i and j is given by

$P_{(\boldsymbol{x}_i, \boldsymbol{x}_j)}$	1	0
1	$1/2 - \gamma(1-\gamma)$	$\gamma(1-\gamma)$
0	$\gamma(1-\gamma)$	$1/2 - \gamma(1-\gamma)$

Hence, γ parametrizes the dependence of i's signal on j's signal. □

In Example 1, the random variable $\boldsymbol{z}$ is unrelated to fundamentals. Nevertheless, in equilibrium, all players choose to acquire information about $\boldsymbol{z}$, even if it is costly. The random variable $\boldsymbol{z}$, in fact, corresponds to the share of players who decide to invest, i.e., $\boldsymbol{z} = \bar{\boldsymbol{x}}$. When $\boldsymbol{z}$ is equal to $1-\gamma$, the average action is low, and it is optimal not to invest; when $\boldsymbol{z}$ is equal to γ, instead, the average action is high, and it optimal to invest. This generates an incentive to acquire information about $\boldsymbol{z}$, which can be interpreted as an expensive correlation device.

The coefficient γ parametrizes the variance of $\boldsymbol{z}$, and, contextually, the players' incentives to acquire information about each other's information. When γ is high, the players have high incentives to acquire information about $\boldsymbol{z}$, and their signals are strongly correlated. The opposite is true when γ is low. In the extreme case $\gamma = 1/2$, the players' signals are independent, and nonfundamental volatility vanishes.

The equilibrium construction relies on the equality $\boldsymbol{z} = \bar{\boldsymbol{x}}$, which is not an equilibrium outcome when actions are strategic substitutes. Intuitively, by contradiction, assume that $\boldsymbol{z} = \bar{\boldsymbol{x}}$ and actions are strategic substitutes. With substitutability, when

$\bar{\boldsymbol{x}}$ is high, the players do not want to invest. Hence, when $\boldsymbol{z}$ is high, it is less likely that the players invest. As a result, also the average action is low, and therefore $\bar{\boldsymbol{x}}$ is low when $\boldsymbol{z}$ is high, contradicting $\boldsymbol{z} = \bar{\boldsymbol{x}}$.

9 Related Literature

In this section, I discuss how the model can be related to more structured settings of information choice, and how it compares to existing models in the literature.[41]

The cost-acquisition technology I consider in this paper applies quite naturally to any setting where agents must purchase information from third (unmodeled) parties. In markets for information, the buyers' cost of information depends on each other's demands, hence on each other's information choice. A more structured foundation for the model would be provided by a model of information choice where the players repeatedly acquire information about the state and each other's information, before taking actions.

So far, the literature has focused on the players' choice of information about the state. In most of the existing models, the players can choose only signals that are conditionally independent given the state. Unrestricted information acquisition can be seen as a generalization of these models: when information choice is unrestricted, the players can acquire information not only about the state, but also about each other's information. Recently, a few works have dropped the independence assumption in models of rigid information choice. In these settings, the players can modify the signals' noise only up to some parameter. Unrestricted information acquisition can also be seen as a generalization of these models: when information choice is unrestricted, the players can acquire information both about the state and each other's information in a flexible way.

9.1 Markets for Information

A tradeoff between learning the state and the others' information is faced by the buyers in many markets for information. The tradeoff emerges from the dependence of the price of information on the buyers' demands. For instance, consider a market

[41] I focus on static games with information acquisition. See, e.g., Liu [2011], Kim and Lee [2014], and Ravid [2015], for examples of information choice in dynamic games. See, e.g., Veldkamp [2011] for a broader perspective on information choice in economics.

for information where experts are sellers. The players are the buyers. To acquire information, the players choose what experts to consult. Their cost of information summarizes the prices they pay to the experts. The more players decide to consult the same expert, the higher the price charged by the expert (e.g., if the expert's marginal productivity is decreasing). In this situation, the players need to consult the same expert to correlate their signals, i.e., to acquire information about each other's information. However, this comes with a higher cost. Therefore, the players face a tradeoff between learning the state and each others' information.

9.2 Information Acquisition in Extensive Form

The model could also be viewed as the normal form of an extensive-form game where the players repeatedly acquire information, before taking actions. In the extensive form, there may be several stages of information choice. First, the players acquire information about the state. Then, they acquire information about each other's first-stage information, and so forth. This could happen by spying on each other or by engaging in costly communication. In this setting, it is natural to think that, in each stage, the players' cost of information depends on the information acquired by the other players in the previous stages. An extensive-form foundation for the model could be an interesting direction for future research.

9.3 Independent Information Acquisition

In models of independent information acquisition, signals are conditionally independent given the state, and the cost of information $C_i(\boldsymbol{x}_i; \boldsymbol{\theta})$ depends only on the relation between signal and state. See, e.g., Persico [2000], Bergemann and Välimäki [2002], Yang [2015], and Tirole [2015]. Unrestricted information acquisition can be seen as a generalization of these models. To illustrate, observe that, when $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|\boldsymbol{\theta}$, the state $\boldsymbol{\theta}$ as a statistics of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}_i$. As a result, if Condition 2 is satisfied, then $C_i(\boldsymbol{x}_i; \boldsymbol{x}_{-i}, \boldsymbol{\theta}) = C_i(\boldsymbol{x}_i; \boldsymbol{\theta})$. The equality says that, if we add the constraint $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|\boldsymbol{\theta}$ to unrestricted information acquisition, then we go back to independent information acquisition. Overall, with independent information acquisition, the players can acquire information only about the state. Instead, with unrestricted information acquisition, the players can acquire information also about each other's information.

Investment games (Section 6) exemplified the different implications of independent and unrestricted information choice. In finite investment games, with flexible and independent information acquisition, multiple equilibria emerge (Yang [2015]). Instead, dropping the independence assumption, I showed that the players' incentive to learn what the others know come into play, and lead to risk-dominance as the unique solution.

With unrestricted information acquisition, independence may endogenously arise as an equilibrium outcome. For instance, this happens in large games, where the players' incentive to learn what the others know weaken. I explored this theme in large investment games (Section 6.3), large linear-quadratic games (Sections 7.3 and 7.4), and general large games (Section 8).

9.4 Rigid Information Acquisition

In models of rigid information acquisition, the players can modify the signals' noise only up to some parameter. See, e.g., Hellwig and Veldkamp [2009], Myatt and Wallace [2012], and Pavan [2014]. To illustrate the relation with unrestricted information acquisition, I focus on the model used by Myatt and Wallace [2012] and Pavan [2014] in beauty contests. In this model, information choice is rigid but not independent:

Example 2. Let the state $\boldsymbol{\theta}$ be a univariate normal random variable. There are m information sources $\boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$ such that, for all $l = 1, \ldots, m$,

$$\boldsymbol{z}_l = \boldsymbol{\theta} + \sigma_l \boldsymbol{\eta}_l, \quad \text{with } \sigma_l \in [0, \infty] \text{ and } \boldsymbol{\eta}_l \sim \mathcal{N}(0, 1).$$

The players choose how much to learn about the information source. Player i's signal is a vector $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{im})$ such that, for all $l = 1, \ldots, m$,

$$\boldsymbol{x}_{il} = \boldsymbol{z}_l + \tau_{il} \boldsymbol{\epsilon}_{il}, \quad \text{with } \tau_{il} \in [0, \infty] \text{ and } \boldsymbol{\epsilon}_{il} \sim \mathcal{N}(0, 1).$$

The noises $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_m)$ and $\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}_{i1}, \ldots, \boldsymbol{\epsilon}_{im})$ all independent of each other and of the state. Player i's information choice is parametrized by $\tau_i = (\tau_{i1}, \ldots, \tau_{im})$. The cost of information is some decreasing function of τ_i. □

In Example 2, information choice is not independent due to the common noise $\boldsymbol{\eta}$, and it is rigid because the player i can modify the signals' noise only up to the parameter τ_i. Unrestricted information acquisition can be seen as a generalization of this model. Intuitively, in Example 2, the cost of signal $\boldsymbol{x}_i$ depends on how much effort the player puts in learning $\boldsymbol{z}$, and can be written as $C_i(\boldsymbol{x}_i; \boldsymbol{z})$. Furthermore, the random variable $\boldsymbol{z}$ as a statistic of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta}, \boldsymbol{z})$ is sufficient for $\boldsymbol{x}_i$. Therefore, if Condition 2 is satisfied, then $C_i(\boldsymbol{x}_i; \boldsymbol{z}) = C_i(\boldsymbol{x}_i; \boldsymbol{x}_{-i}, \boldsymbol{\theta}, \boldsymbol{z})$. This is consistent with the idea that player i acquires information about the state and each other's information only through the vector $\boldsymbol{z}$ of information sources. Now, when we relax the rigidity in information choice, player i is better off by focusing on the state and her opponents' information, and disregarding $\boldsymbol{z}$. Under Condition 2, she chooses $\boldsymbol{x}_i$ such that $\boldsymbol{x}_i | (\boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim \boldsymbol{x}_i | (\boldsymbol{x}_{-i}, \boldsymbol{\theta}, \boldsymbol{z})$, and therefore $C_i(\boldsymbol{x}_i; \boldsymbol{x}_{-i}, \boldsymbol{\theta}, \boldsymbol{z}) = C_i(\boldsymbol{x}_i; \boldsymbol{x}_{-i}, \boldsymbol{\theta})$, as in unrestricted information acquisition. Overall, unrestricted information acquisition generalizes rigid information choice by expanding the set of signals available to the players.

Beauty contests (Section 7.4) provided an example of the different implications of rigid and unrestricted information acquisition. In Myatt and Wallace [2012], beauty contests with mutual information have multiple Gaussian equilibria. On the other hand, when information choice is unrestricted, beauty contests with mutual information have a unique Gaussian equilibrium. This happens because, with unrestricted information acquisition, in equilibrium, nonfundamental volatility vanishes and the players choose signals conditionally independent given the state. In Myatt and Wallace [2012], this cannot happen, since the players' signals share the common noise $\boldsymbol{\eta}$ and the exposure to this noise, parametrized by $\sigma = (\sigma_1, \ldots, \sigma_m)$, is not a choice variable.

10 Conclusion

Traditional economic analysis has taken the players' information exogenously given. In many settings, however, the players do not interact on the basis of some fixed prior information. Instead, what they choose to know is a central part of the game they play (Arrow [1996]). Learning about the underlying state of fundamentals is necessary, but not sufficient to choose the best course of actions. The players understand that the actions the others take depend on the information they have. Hence, to best reply to the opponents' behavior, the players have also an incentive to learn what the others know.

In this paper, I proposed a tractable model to study information choice in general games where fundamentals are not known. The model does not impose technological restrictions on the players' information choice. Therefore, it exposes the players' primitive incentives to acquire information about the state and each other's information. An essential feature of the model is that the players face a tradeoff between learning the state and each other's information. Such tradeoff has an intuitive appeal: it reflects the idea that it is hard to learn exactly what others know. Nevertheless, a microfoundation of the model can help us better understand its nature and sources. Mutual information is a natural benchmark for the cost of information. While the model assumes only broad monotonicity conditions on the cost of information, developing tractable alternatives to mutual information seems necessary for applications.

From my analysis of the model, two main patterns for strategic information choice emerged. First, in coordination games, the players have a strong incentive to know what the others know. Not taking that into account crucially affects our predictions. For instance, in investment games, if information acquisition is assumed to be independent and the players can acquire information only about the state, multiple equilibria emerge. Instead, when the players can acquire information also about each other's information, the incentive to know what the others know comes into play and leads to risk-dominance as the unique solution.

Second, the size of the game matters with information acquisition. In large games with small players, the players' incentive to know what the others know weakens. In large investment games, this restores equilibrium multiplicity. In linear-quadratic games without central players, nonfundamental volatility vanishes. The relevance of the size of the game is distinctive to information choice. When information is exogenously given, the size of the game often does not affect predictions and is chosen on the basis of tractability considerations. My analysis of general large game with mutual information is a first step to better understand strategic information choice in many-player settings.

A Proofs

A.1 Proof of Proposition 1

The proof is divided into two steps. The first step verifies that, if $(\boldsymbol{x}, s)$ is an equilibrium, then $s(\boldsymbol{x})$ is an equilibrium in direct signals. The second step checks that, if $(\boldsymbol{x}, s)$ is an equilibrium, then $\boldsymbol{x}$ and $s(\boldsymbol{x})$ are individually sufficient for each other.

Step 1. If $(\boldsymbol{x}, s)$ is an equilibrium, then $s(\boldsymbol{x})$ is an equilibrium in direct signals.

Proof of the Step. Let $(\boldsymbol{x}, s)$ be an equilibrium. Fix player i. We wish to show that direct signal $s_i(\boldsymbol{x}_i)$ is a best reply to direct signals $s_{-i}(\boldsymbol{x}_{-i})$. Notice first that, since $s_i(\boldsymbol{x}_i)$ and $s_{-i}(\boldsymbol{x}_{-i})$ are statistics of $\boldsymbol{x}_i$ and $\boldsymbol{x}_{-i}$, respectively, Conditions 1 and 2 imply that

$$C_i(s(\boldsymbol{x}), \boldsymbol{\theta}) \leq C_i(s_i(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \leq C_i(\boldsymbol{x}, \boldsymbol{\theta}).$$

As a consequence, for $s_i(\boldsymbol{x}_i)$ to be a best reply to $s_{-i}(\boldsymbol{x}_{-i})$, it is enough to check that, for all deviations $(\boldsymbol{x}'_i, s'_i)$,

$$E[u_i(s(\boldsymbol{x}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}, \boldsymbol{\theta}) \geq E[u_i(s'_i(\boldsymbol{x}'_i), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}'_i, s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}). \tag{9}$$

To see that (9) holds, choose a direct signal $\boldsymbol{x}''_i$ such that $\boldsymbol{x}''_i|(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is equal to $s'_i(\boldsymbol{x}'_i)|(s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})$. Notice first that $\boldsymbol{x}''_i|(\boldsymbol{x}_{-i}, \boldsymbol{\theta}) = \boldsymbol{x}''_i|(s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})$, i.e., $(s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})$ as a statistic of $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}''_i$. By Condition 2

$$C_i(\boldsymbol{x}''_i, \boldsymbol{x}_{-i}, \boldsymbol{\theta}) = C_i(\boldsymbol{x}''_i, s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}). \tag{10}$$

Next, observe that $(\boldsymbol{x}''_i, s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}) \sim (s'_i(\boldsymbol{x}'_i), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})$ implies

$$E[u_i(\boldsymbol{x}''_i, s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] = E[u_i(s_i(\boldsymbol{x}'_i), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})], \tag{11}$$

$$C_i(\boldsymbol{x}''_i, s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}) = C_i(s'_i(\boldsymbol{x}'_i), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}). \tag{12}$$

Finally, since $s_i(\boldsymbol{x}'_i)$ is a statistics of $\boldsymbol{x}'_i$, by Condition 1

$$C_i(s'_i(\boldsymbol{x}'_i), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}) \leq C_i(\boldsymbol{x}'_i, s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}) \tag{13}$$

In conclusion, since $(\boldsymbol{x}_i, s_i)$ is a best reply to $(\boldsymbol{x}_{-i}, s_{-i})$,

$$
\begin{aligned}
E[u_i(s(\boldsymbol{x}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}, \boldsymbol{\theta}) \geq & E[u_i(\boldsymbol{x}_i'', s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}_i'', \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \\
= & E[u_i(\boldsymbol{x}_i'', s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}_i'', s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}) \\
= & E[u_i(s_i(\boldsymbol{x}_i'), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] - C_i(s_i(\boldsymbol{x}_i'), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}) \\
\geq & E[u_i(s_i(\boldsymbol{x}_i'), s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}_i', s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}),
\end{aligned}
$$

where the first equality holds by (10), the second equality by (11) and (12), and the last inequality by (13). This chain of inequalities proves (9), and therefore that $s_i(\boldsymbol{x}_i)$ is a best reply to $s_{-i}(\boldsymbol{x}_{-i})$. Since this is true for all players, $s(\boldsymbol{x})$ is an equilibrium in direct signals, as desired. □

Step 2. If $(\boldsymbol{x}, s)$ is an equilibrium, then $\boldsymbol{x}$ and $s(\boldsymbol{x})$ are individually sufficient for each other.

Proof of the Step. Let $(\boldsymbol{x}, s)$ be an equilibrium. Fix player i. Since $s_i(\boldsymbol{x}_i)$ is a statistic of $\boldsymbol{x}_i$, $\boldsymbol{x}_i$ as a statistic of $(\boldsymbol{x}_i, s_i(\boldsymbol{x}_i))$ is sufficient for $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$. For the opposite direction, notice that, since $(\boldsymbol{x}_i, s_i)$ is a best reply to $(\boldsymbol{x}_{-i}, s_{-i})$,

$$
E[u_i(s(\boldsymbol{x}), \boldsymbol{\theta})] - C_i(\boldsymbol{x}, \boldsymbol{\theta}) \geq E[u_i(s(\boldsymbol{x}), \boldsymbol{\theta})] - C_i(s_i(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta}),
$$

which implies that $C_i(\boldsymbol{x}, \boldsymbol{\theta}) \leq C_i(s_i(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta})$ since expected utility is finite. Moreover, since $s_i(\boldsymbol{x}_i)$ is a statistic of $\boldsymbol{x}_i$, by Condition 1 we have that $C_i(\boldsymbol{x}, \boldsymbol{\theta})$ is bigger than $C_i(s_i(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta})$. Hence $C_i(\boldsymbol{x}, \boldsymbol{\theta})$ equals $C_i(s_i(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta})$: since $C_i(\boldsymbol{x}, \boldsymbol{\theta})$ is finite, by Condition 1 we must have that $s_i(\boldsymbol{x}_i)$ is sufficient for $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$, that is, $(\boldsymbol{x}_i \perp (\boldsymbol{x}_{-i}, \boldsymbol{\theta}))|s_i(\boldsymbol{x}_i)$. But then $(\boldsymbol{x}_i \perp (s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta}))|s_i(\boldsymbol{x}_i)$, which says that $s_i(\boldsymbol{x}_i)$ is sufficient for $(s_{-i}(\boldsymbol{x}_{-i}), \boldsymbol{\theta})$ as a statistic of $(\boldsymbol{x}_i, s_i(\boldsymbol{x}_i))$. Since this is true for all players, $\boldsymbol{x}$ and $s(\boldsymbol{x})$ are individually sufficient for each other, as desired. □

A.2 Proof of Lemma 1

Sufficiently, I show that for any strategy $(\boldsymbol{x}_i, s_i)$ there is a direct signal $\boldsymbol{x}_i'$ such that (i) $f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}_i'$, and (ii)

$$
\begin{aligned}
E[u_i(s_i(\boldsymbol{x}_i), \boldsymbol{x}_{-i}, \boldsymbol{\theta})] - C_i(\boldsymbol{x}, \boldsymbol{\theta}) &\leq E[u_i(\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{\theta})] - C_i(\boldsymbol{x}_i', \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \\
&= E[u_i(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i}, \boldsymbol{\theta}))] - C_i(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i}, \boldsymbol{\theta})),
\end{aligned}
$$

with strict inequality if the left-hand side is finite and $f(\boldsymbol{x}_{-i},\boldsymbol{\theta})$ is not sufficient for $s_i(\boldsymbol{x}_i)$. Hence, pick direct signal $\boldsymbol{x}_i'$ such that $\boldsymbol{x}_i'|(\boldsymbol{x}_{-i},\boldsymbol{\theta}) \sim s_i(\boldsymbol{x}_i)|f(\boldsymbol{x}_{-i},\boldsymbol{\theta})$. Since $\boldsymbol{x}_i'|f(\boldsymbol{x}_{-i},\boldsymbol{\theta}) \sim \boldsymbol{x}_i'|(\boldsymbol{x}_{-i},\boldsymbol{\theta})$, the statistic $f(\boldsymbol{x}_{-i},\boldsymbol{\theta})$ is sufficient for $\boldsymbol{x}_i$, and therefore (i) holds. To see that also (ii) holds, first notice that $(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i},\boldsymbol{\theta})) \sim (s_i(\boldsymbol{x}_i), f(\boldsymbol{x}_{-i},\boldsymbol{\theta}))$, which implies

$$\begin{aligned} E[u_i(s_i(\boldsymbol{x}_i),\boldsymbol{x}_{-i},\boldsymbol{\theta})] &= E[u_i(s_i(\boldsymbol{x}_i), f(\boldsymbol{x}_{-i},\boldsymbol{\theta}))] = E[u_i(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i},\boldsymbol{\theta}))] \\ &= E[u_i(\boldsymbol{x}_i',\boldsymbol{x}_{-i},\boldsymbol{\theta})]. \end{aligned} \tag{14}$$

Moving to the cost of information, observe that

$$\begin{aligned} C_i(\boldsymbol{x},\boldsymbol{\theta}) \geq C_i(s_i(\boldsymbol{x}_i),\boldsymbol{x}_{-i},\boldsymbol{\theta}) &\geq C_i(s_i(\boldsymbol{x}_i), f(\boldsymbol{x}_{-i},\boldsymbol{\theta})) \\ &= C_i(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i},\boldsymbol{\theta})) = C_i(\boldsymbol{x}_i',\boldsymbol{x}_{-i},\boldsymbol{\theta}) \end{aligned} \tag{15}$$

where the first inequality holds by Condition 1, the second inequality by Condition 2, the first equality by $(\boldsymbol{x}_i', f(\boldsymbol{x}_{-i},\boldsymbol{\theta})) \sim (s_i(\boldsymbol{x}_i), f(\boldsymbol{x}_{-i},\boldsymbol{\theta}))$, and the second equality by Condition 2 and $f(\boldsymbol{x}_{-i},\boldsymbol{\theta})$ sufficient for $\boldsymbol{x}_i'$. Note that, whenever $C_i(\boldsymbol{x},\boldsymbol{\theta})$ is finite, the second inequality is strict if $f(\boldsymbol{x}_{-i},\boldsymbol{\theta})$ is not sufficient for $s_i(\boldsymbol{x}_i)$, again by Condition 2. Therefore, combining (14) and (15), we obtain (ii), concluding the proof.

A.3 Proof of Lemma 3

To ease the exposition, the proof is divided in two parts. In first part, I prove necessity. In the second part, I prove sufficiency. Without loss of generality, assume $\lambda = 1$. Throughout, given M nonempty subset of N, define $A_M = \times_{i\in M} A_i$, $a_M = (a_i : i \in M)$, $\boldsymbol{x}_M = (\boldsymbol{x}_i : i \in M)$, and $Q_{\boldsymbol{x}_M} = \times_{i\in M} P_{\boldsymbol{x}_i}$. If $M = N$ or $M = N \setminus \{i\}$, write also $Q_{\boldsymbol{x}}$ and $Q_{\boldsymbol{x}_{-i}}$ instead of $Q_{\boldsymbol{x}_M}$, respectively.

Necessity

Let $\boldsymbol{x}$ be an equilibrium in direct signals. Here, I show that the following two conditions hold:

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}})}(a,\theta) = \frac{e^{v(a,\theta)}}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')}, \qquad Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}} - a.s.. \tag{16}$$

$$1 \geq \int_\Theta \frac{\int_{A_{-i}} e^{v(a_i,a'_{-i},\theta)} dQ_{\boldsymbol{x}_{-i}}(a'_{-i})}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')} dP_{\boldsymbol{\theta}}(\theta), \qquad \forall i \in N \text{ and } \forall a_i \in A_i. \tag{17}$$

The proof proceeds by steps, and the last step concludes.

Step 1. For all players i, the following two conditions hold:

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})})}(a,\theta) = \frac{e^{v(a,\theta)}}{\int_{A_i} e^{v(a'_i,a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a'_i)}, \qquad P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})} - a.s..$$

$$1 \geq \int_{A_{-i}\times\Theta} \frac{e^{v(a_i,a_{-i},\theta)}}{\int_{A_i} e^{v(a'_i,a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a'_i)} dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta), \qquad \forall a_i \in A_i.$$

Proof of the Step. Fix player i. Since $\boldsymbol{x}_i$ is a best reply to $\boldsymbol{x}_{-i}$, by Lemma 2, the following two conditions hold:

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})})}(a,\theta) = \frac{e^{u_i(a,\theta)}}{\int_{A_i} e^{u_i(a'_i,a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a'_i)}, \qquad P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})} - a.s.. \tag{18}$$

$$1 \geq \int_{A_{-i}\times\Theta} \frac{e^{u_i(a_i,a_{-i},\theta)}}{\int_{A_i} e^{u_i(a'_i,a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a'_i)} dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta), \qquad \forall a_i \in A_i. \tag{19}$$

Now, notice that for all $a_i, a'_i \in A_i$, $a_{-i} \in A_{-i}$, and $\theta \in \Theta$,

$$\begin{aligned}\frac{e^{u_i(a,\theta)}}{\int_{A_i} e^{u_i(a'_i,a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a'_i)} &= \frac{1}{\int_{A_i} e^{u_i(a'_i,a_{-i},\theta)-u_i(a,\theta)} dP_{\boldsymbol{x}_i}(a'_i)} \\ &= \frac{1}{\int_{A_i} e^{v(a'_i,a_{-i},\theta)-v(a,\theta)} dP_{\boldsymbol{x}_i}(a'_i)} \\ &= \frac{e^{v(a,\theta)}}{\int_{A_i} e^{v(a'_i,a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a'_i)},\end{aligned} \tag{20}$$

where the second equality holds since v is a potential. The proof of the step is concluded by plugging (20) into (18) and (19). □

Step 2. Let M be a nonempty subset of N. For all $i \in M$,

$$P_{(\boldsymbol{x}_M,\boldsymbol{\theta})} \ll P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})}.$$

Proof of the Step. The proof proceeds by induction on the cardinality of M. The basis step, when $M = N$, follows from Step 1. For the inductive step, take player $i \in M$. Let B be a measurable subset of $A_M \times \Theta$ such that $(P_{\boldsymbol{x}_i} \times P_{\boldsymbol{x}_{M\setminus\{i\}}})(B) = 0$: the goal is to show that also $P_{(\boldsymbol{x}_M,\boldsymbol{\theta})}(B) = 0$. For every $a_i \in A_i$, write

$$B_{a_i} = \{a_{M\setminus\{i\}} \in A_{M\setminus\{i\}} : (a_{M\setminus\{i\}}, a_i) \in B\}.$$

Now, take player $j \notin M$ and define $M' = M \cup \{j\}$. Notice that

$$\begin{aligned}
P_{(\boldsymbol{x}_M,\boldsymbol{\theta})}(B) &= P_{(\boldsymbol{x}_{M'},\boldsymbol{\theta})}(B \times A_j) \\
&= \int_{B\times A_j} \frac{dP_{(\boldsymbol{x}_{M'},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})})}(a_{M'},\theta)d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})})(a_{M'}) \\
&= \int_{A_i}\int_{B_{a_i}\times A_j} \frac{dP_{(\boldsymbol{x}_{M'},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})})}(a_{M'},\theta)dP_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})}(a_{M'\setminus\{i\}})dP_{\boldsymbol{x}_i}(a_i) \\
&= \int_{B}\int_{A_j} \frac{dP_{(\boldsymbol{x}_{M'},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})})}(a_{M'},\theta)dP_{\boldsymbol{x}_j}(a_j|a_{M\setminus\{i\}})d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})})(a_M),
\end{aligned}$$

where the second equality holds by inductive hypothesis. Since $P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})}$ assigns probability zero to B, therefore, we conclude that $P_{(\boldsymbol{x}_M,\boldsymbol{\theta})}(B) = 0$, as desired. □

Step 3. Let M be a nonempty subset of N. For all $i \in M$, almost surely with respect to $P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})}$,

$$\frac{dP_{(\boldsymbol{x}_M,\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})})}(a_M,\theta) = \frac{\int_{A_{N\setminus M}} e^{v(a'_{N\setminus M},a_M,\theta)}dQ_{\boldsymbol{x}_{N\setminus M}}(a'_{N\setminus M})}{\int_{A_i}\int_{A_{N\setminus M}} e^{v(a'_{N\setminus M},a'_i,a_{M\setminus\{i\}},\theta)}dQ_{\boldsymbol{x}_{N\setminus M}}(a'_{N\setminus M})dP_{\boldsymbol{x}_i}(a'_i)}.$$

Proof of the Step. The proof proceeds by induction on the cardinality of M. The basis step, when $M = N$, follows from Step 1. For the inductive step, take players $i \in M$. Moreover, pick player $j \notin M$, and define $M' = M \cup \{j\}$. Notice that, almost surely with respect to $P_{\boldsymbol{x}_i} \times P_{\boldsymbol{x}_j} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})}$,

$$\frac{dP_{(\boldsymbol{x}_{M'},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})})}\frac{dP_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_j} \times P_{(\boldsymbol{x}_M,\boldsymbol{\theta})})} = \frac{dP_{(\boldsymbol{x}_{M'},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_j} \times P_{(\boldsymbol{x}_M,\boldsymbol{\theta})})}\frac{dP_{(\boldsymbol{x}_M,\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})})},$$

where the densities exist by Step 2. Hence, by inductive hypothesis, almost surely with respect to $P_{\boldsymbol{x}_i} \times P_{\boldsymbol{x}_j} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})}$,

$$\frac{dP_{(\boldsymbol{x}_{M'\setminus\{i\}},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_j} \times P_{(\boldsymbol{x}_M,\boldsymbol{\theta})})}(a_{M'\setminus\{i\}},\theta) = \frac{\int_{A_i}\int_{A_{N\setminus M'}} e^{v(a'_{N\setminus M'},a_{M'\setminus\{i\}},a'_i,\theta)} dQ_{\boldsymbol{x}_{N\setminus M'}}(a'_{N\setminus M'}) dP_{\boldsymbol{x}_i}(a'_i)}{\int_{A_{N\setminus M}} e^{v(a'_{N\setminus M},a_M,\theta)} dQ_{\boldsymbol{x}_{N\setminus M}}(a'_{N\setminus M})} \times \frac{dP_{(\boldsymbol{x}_M,\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})})}(a_M,\theta).$$

Integrating both sides with respect to $P_{\boldsymbol{x}_j}$, almost surely with respect to $P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})}$,

$$1 = \frac{\int_{A_i}\int_{A_{N\setminus M}} e^{v(a'_{N\setminus M},a_{M\setminus\{i\}},a'_i,\theta)} dQ_{\boldsymbol{x}_{N\setminus M}}(a'_{N\setminus M}) dP_{\boldsymbol{x}_i}(a'_i)}{\int_{A_{N\setminus M}} e^{v(a'_{N\setminus M},a_M,\theta)} dQ_{\boldsymbol{x}_{N\setminus M}}(a'_{N\setminus M})} \frac{dP_{(\boldsymbol{x}_M,\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{M\setminus\{i\}},\boldsymbol{\theta})})}(a_M,\theta),$$

as desired. □

Step 4 (Necessity)**.** Conditions (16) and (17) hold.

Proof of the Step. First, observe that by Step 3, almost surely with respect to $Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}}$,

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}})}(a,\theta) = \prod_{i=1}^{n} \frac{dP_{(\boldsymbol{x}_i,\ldots,\boldsymbol{x}_n,\boldsymbol{\theta})}}{dP_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{i+1},\ldots,\boldsymbol{x}_n,\theta)}}(a_i,\ldots,a_n,\theta) = \frac{e^{v(a,\theta)}}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')},$$

which gives (16). In particular, for all players i,

$$\frac{dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})}(a_{-i},\theta) = \frac{\int_{A_i} e^{v(a'_i,a_{-i},\theta)}}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')}, \quad Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}} - a.s..$$

Hence, for all players i and $a_i \in A_i$,

$$
\begin{aligned}
1 &\geq \int_{A_{-i}\times\Theta} \frac{e^{v(a_i,a_{-i},\theta)}}{\int_{A_i} e^{v(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')} dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta) \\
&= \int_{A_{-i}\times\Theta} \frac{e^{v(a_i,a_{-i},\theta)}}{\int_{A_i} e^{v(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')} \frac{dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})} d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})(a_{-i},\theta) \\
&= \int_{A_{-i}\times\Theta} \frac{e^{v(a_i,a_{-i},\theta)}}{\int_{A} e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')} d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})(a_{-i},\theta) \\
&= \int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i,a_{-i}',\theta)} dQ_{\boldsymbol{x}_{-i}}(a_{-i}')}{\int_{A} e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')} dP_{\boldsymbol{\theta}}(\theta),
\end{aligned}
$$

where the first inequality holds by Step 1. Therefore, also (17) holds, as desired. □

Sufficiency

Let $\boldsymbol{x}$ be a profile of direct signals such that the following two conditions hold:

$$
\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}})}(a,\theta) = \frac{e^{v(a,\theta)}}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')}, \qquad Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}} - a.s.. \tag{21}
$$

$$
1 \geq \int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i,a_{-i}',\theta)} dQ_{\boldsymbol{x}_{-i}}(a_{-i}')}{\int_{A} e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')} dP_{\boldsymbol{\theta}}(\theta), \qquad \forall i \in N \text{ and } \forall a_i \in A_i. \tag{22}
$$

Here, I show that $\boldsymbol{x}$ is an equilibrium, i.e., by Lemma 2, for all players i,

$$
\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})})}(a,\theta) = \frac{e^{u_i(a,\theta)}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')}, \qquad P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})} - a.s.. \tag{23}
$$

$$
1 \geq \int_{A_{-i}\times\Theta} \frac{e^{u_i(a_i,a_{-i},\theta)}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')} dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta), \qquad \forall a_i \in A_i. \tag{24}
$$

Fix player i. We make two preliminary observations. First, notice that, by (21),

$$
\frac{dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})}(a_{-i},\theta) = \frac{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')}{\int_{A_i} e^{v(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i)}, \quad Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}} - a.s..
$$

Moreover, observe that for all $a \in A$,

$$\frac{e^{v(a,\theta)}}{\int_{A_i} e^{v(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')} = \frac{1}{\int_{A_i} e^{v(a_i',a_{-i},\theta)-v(a,\theta)} dP_{\boldsymbol{x}_i}(a_i')} = \frac{1}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)-u_i(a,\theta)} dP_{\boldsymbol{x}_i}(a_i')}$$
$$= \frac{e^{u_i(a,\theta)}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')},$$

where the second equality holds since v is a potential. Now, condition (23) comes from the following chain of equality: almost surely with respect to $P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}$,

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(P_{\boldsymbol{x}_i} \times P_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})})}(a,\theta) = \frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(Q_{\boldsymbol{x}} \times P_{\boldsymbol{\theta}})}(a,\theta) \frac{d(Q_{\boldsymbol{x}_{-i}} \times P_\theta)}{dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}}(a_{-i},\theta)$$
$$= \frac{e^{v(a,\theta)}}{\int_{A_i} e^{v(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')} = \frac{e^{u_i(a,\theta)}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')}.$$

Finally, condition (24) comes from the following chain of inequalities: for all $a_i \in A_i$,

$$1 \geq \int_{A_{-i}\times\Theta} \frac{e^{v(a_i,a_{-i},\theta)}}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')} d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})(a_{-i},\theta)$$
$$= \int_{A_{-i}\times\Theta} \frac{e^{v(a_i,a_{-i},\theta)}}{\int_A e^{v(a',\theta)} dQ_{\boldsymbol{x}}(a')} \frac{d(Q_{\boldsymbol{x}_{-i}} \times P_{\boldsymbol{\theta}})}{dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}}(a_{-i},\theta) dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta)$$
$$= \int_{A_{-i}\times\Theta} \frac{e^{u_i(a_i,a_{-i},\theta)}}{\int_{A_i} e^{u_i(a_i',a_{-i},\theta)} dP_{\boldsymbol{x}_i}(a_i')} dP_{(\boldsymbol{x}_{-i},\boldsymbol{\theta})}(a_{-i},\theta),$$

where the first inequality is given by (24). Since the choice of player i was arbitrary, $\boldsymbol{x}$ is an equilibrium, as desired.

A.4 Proof of Proposition 2

For every λ, fix an equilibrium in direct signals $\boldsymbol{x}_\lambda$. The proof proceeds in steps, and the last step concludes.

Step 1. For all players i, if a_i has dominance regions, then

$$\liminf_{\lambda \to 0} P(\boldsymbol{x}_{i,\lambda} = a_i) > 0.$$

Proof of the Step. Fix player i and a_i with dominance region Θ_{a_i}. Set

$$t = \inf_{\theta \in \Theta_{a_i}} \inf_{a_i' \neq a_i} \inf_{a_{-i} \in A_{-i}} v(a_i, a_{-i}, \theta) - v(a_i', a_{-i}, \theta) > 0.$$

Fix $\lambda > 0$. By Lemma 3,

$$1 \geq \int_{A_{-i} \times \Theta} \frac{e^{v(a_i, a_{-i}, \theta)/\lambda}}{\int_{A_i} e^{v(a_i', a_{-i}, \theta)/\lambda} P_{\boldsymbol{x}_{i,\lambda}}(da_i')} dP_{(\boldsymbol{x}_{-i,\lambda}, \boldsymbol{\theta})}(a_{-i}, \theta).$$

Moreover, observe that

$$\int_{A_{-i} \times \Theta} \frac{e^{v(a_i, a_{-i}, \theta)/\lambda}}{\int_{A_i} e^{v(a_i', a_{-i}, \theta)/\lambda} P_{\boldsymbol{x}_{i,\lambda}}(da_i')} dP_{(\boldsymbol{x}_{-i,\lambda}, \boldsymbol{\theta})}(a_{-i}, \theta) =$$
$$\int_{A_{-i} \times \Theta} \frac{1}{\int_{A_i \setminus \{a_i\}} e^{(v(a_i', a_{-i}, \theta) - v(a_i, a_{-i}, \theta))/\lambda} P_{\boldsymbol{x}_{i,\lambda}}(da_i') + P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\})} dP_{(\boldsymbol{x}_{-i,\lambda}, \boldsymbol{\theta})}(a_{-i}, \theta) \geq$$
$$\frac{P_{\boldsymbol{\theta}}(\Theta_{a_i})}{e^{-t/\lambda} + P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\})},$$

where the last inequality holds by Markov's inequality. Hence,

$$P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\}) \geq P_{\boldsymbol{\theta}}(\Theta_{a_i}) - e^{-t/\lambda}.$$

Since the choice of λ was arbitrary,

$$\liminf_{\lambda \to 0} P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\}) \geq P_{\boldsymbol{\theta}}(\Theta_{a_i}) > 0,$$

as desired. □

Step 2 (Proposition 2)**.** If a has dominance regions, then, a.s. with respect to $P_{\boldsymbol{\theta}}$,

$$v(a, \theta) > \sup_{a' \neq a} v(a', \theta) \quad \Rightarrow \quad \lim_{\lambda \to 0} P(\boldsymbol{x}_\lambda = a | \theta) = 1.$$

Proof of the Step. Assume that a has dominance regions. For all $\lambda > 0$, by Lemma

3, almost surely with respect to $P_{\boldsymbol{\theta}}$,

$$P(\boldsymbol{x}_\lambda = a|\theta) = \frac{e^{v(a,\theta)/\lambda}\prod_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\})}{\int_A e^{v(a',\theta)/\lambda} d(\times_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}})(a')}$$
$$= \frac{\prod_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\})}{\int_{A\setminus\{a\}} e^{(v(a',\theta)-v(a,\theta))/\lambda} d(\times_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}})(a') + \prod_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\})}.$$

Since a has dominance regions, by Step 1

$$\liminf_{\lambda\to 0}\prod_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}}(\{a_i\}) > 0.$$

Moreover, if $v(a,\theta) > \sup_{a'\neq a} v(a',\theta)$, then

$$\limsup_{\lambda\to 0}\int_{A\setminus\{a\}} e^{(v(a',\theta)-v(a,\theta))/\lambda} d(\times_{i=1}^n P_{\boldsymbol{x}_{i,\lambda}})(a') \leq \limsup_{\lambda\to 0} e^{(\sup_{a'\neq a} v(a',\theta)-v(a,\theta))/\lambda} = 0.$$

Therefore, almost surely with respect to $P_{\boldsymbol{\theta}}$,

$$v(a,\theta) > \sup_{a'\neq a} v(a',\theta) \quad\Rightarrow\quad \lim_{\lambda\to 0} P(\boldsymbol{x}_\lambda = a|\theta) = 1,$$

as desired. □

A.5 Proof of Proposition 3

Without loss of generality, let $\lambda = 1$. From Lemma 3, it is clear that (i) implies (iii). The rest of the proof is organized as follows. First, I prove that (iii) implies (i). Next, I show that (ii) implies (iii). Finally, I show that (iii) implies (ii). The arguments for the last two implications are inspired by Csiszár [1974], in particular by the proof of his Lemma 1.4.

(iii) implies (i)

Assume that (iii) holds. Take a profile of direct signals $\boldsymbol{x}$ such that

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(\alpha\times P_{\boldsymbol{\theta}})}(a,\theta) = \frac{e^{v(a,\theta)}}{\int_A e^{v(a',\theta)} d\alpha(a')}, \quad a.s.. \tag{25}$$

By Lemma 3, to prove that $\boldsymbol{x}$ is an equilibrium, we only need to show $P_{\boldsymbol{x}_i} = Q_i$ for all players i. Fix player i. Notice that

$$\int_{A_i} \int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i, a_{-i}, \theta)} d\alpha_{-i}(a_{-i})}{\int_A e^{v(a,\theta)} d\alpha(a)} dP_{\boldsymbol{\theta}}(\theta) d\alpha_i(a_i) = \int_{\Theta} \frac{\int_A e^{v(a,\theta)} d\alpha(a)}{\int_A e^{v(a,\theta)} d\alpha(a)} dP_{\boldsymbol{\theta}}(\theta) = 1.$$

Hence, by (iii),

$$\int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i, a'_{-i}, \theta)} d\alpha_{-i}(a'_{-i})}{\int_A e^{v(a',\theta)} d\alpha(a')} dP_{\boldsymbol{\theta}}(\theta) = 1, \quad \alpha_i - a.s..$$

Moreover, by (25),

$$\frac{dP_{\boldsymbol{x}_i}}{d\alpha_i}(a_i) = \int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i, a'_{-i}, \theta)} d\alpha_{-i}(a'_{-i})}{\int_A e^{v(a',\theta)} d\alpha(a')} dP_{\boldsymbol{\theta}}(\theta), \quad \alpha_i - a.s..$$

In conclusion, $dP_{\boldsymbol{x}_i}/d\alpha_i = 1$ almost surely with respect to α_i, and therefore $P_{\boldsymbol{x}_i} = \alpha_i$, as desired.

(ii) implies (iii)

Assume that (ii) holds. Fix player i and action a_i. Define the function $f : [0,1] \to \mathbb{R}$ such that, for all $t \in [0,1]$,

$$f(t) = \int_{\Theta} \log \left((1-t) \int_A e^{v(a',\theta)} d\alpha(a') + t \int_{A_{-i}} e^{v(a_i, a'_{-i}, \theta)} d\alpha_{-i}(a'_{-i}) \right) dP_{\boldsymbol{\theta}}(\theta).$$

Notice that, for all $t \in [0,1]$,

$$f(t) = \int_{\Theta} \log \left(\int_{A_i} \int_{A_{-i}} e^{v(a',\theta)} d\alpha_{-i}(a'_{-i}) d\alpha_{i,t}(a'_i) \right) dP_{\boldsymbol{\theta}}(\theta),$$

where $\alpha_{i,t} = (1-t)\alpha_i + t\alpha'_i$ and $\alpha'_i(\{a_i\}) = 1$. By (ii), α_i is a best reply to α_{-i} in the auxiliary game V. Therefore, the function f is maximized at $t = 0$, i.e.,

$$0 \geq f'(0) = \int_{\Theta} \frac{\int_{A_{-i}} e^{v(a_i, a'_{-i}, \theta)} d\alpha_{-i}(a'_{-i})}{\int_A e^{v(a',\theta)} d\alpha(a')} dP_{\boldsymbol{\theta}}(\theta) - 1.$$

Hence, (iii) holds, as desired.

(iii) implies (ii)

Assume that (iii) holds. Fix player i. Define

$$f(a_i,\theta)=\int_{A_{-i}} e^{v(a_i,a'_{-i},\theta)}d\alpha_{-i}(a'_{-i}),\quad \forall a_i\in A_i \text{ and } \forall\theta\in\Theta.$$

To check that α_i is a best reply to α_{-i} in auxiliary game G, it is enough to show that, for all $\alpha'_i\in\Delta(A_i)$,

$$\int_\Theta \log\frac{\int_{A_i} f(a_i,\theta)d\alpha'_i(a_i)}{\int_{A_i} f(a_i,\theta)d\alpha_i(a_i)}dP_{\boldsymbol\theta}(\theta)\le 0.$$

Fix $\alpha'_i\in\Delta(A_i)$, and consider the probability $\tilde\alpha_i\in\Delta(A_i\times\Theta)$ such that

$$\frac{d\tilde\alpha_i}{d(\alpha'_i\times P_{\boldsymbol\theta})}(a_i,\theta)=\frac{f(a_i,\theta)}{\int_{A_i} f(a'_i,\theta)d\alpha'_i(a'_i)},\quad \alpha'_i\times P_{\boldsymbol\theta}-a.s..$$

Observe that

$$\begin{aligned}\int_\Theta \log\frac{\int_{A_i} f(a_i,\theta)d\alpha'_i(a_i)}{\int_{A_i} f(a_i,\theta)d\alpha_i(a_i)}dP_{\boldsymbol\theta}(\theta)&=\int_{A_i\times\Theta}\log\frac{\int_{A_i} f(a_i,\theta)d\alpha'_i(a_i)}{\int_{A_i} f(a_i,\theta)d\alpha_i(a_i)}d\tilde\alpha_i(a'_i,\theta)\\&\le\log\int_{A_i}\int_\Theta\frac{f(a'_i,\theta)}{\int_{A_i} f(a_i,\theta)d\alpha_i(a_i)}dP_{\boldsymbol\theta}(\theta)d\alpha'_i(a'_i)\\&\le 0,\end{aligned}$$

where the last inequality holds by (iii). Since the choice of α'_i was arbitrary, α_i is a best reply to α_{-i} in the auxiliary game. Since the choice of player i was arbitrary, (ii) holds, as desired.

A.6 Proof of Corollary 1

Without loss of generality, assume $\lambda=1$. By Proposition 3, to prove that equilibrium existence, it is enough to check that the V has a maximizer. Since A_i is a compact Polish space for all players i, the product $\times_{i\in N}\Delta(A_i)$ is a compact Polish space (in the weak topology). Hence, V has a maximizer whenever it is upper semi-continuous. To verify upper semi-continuity, for every player i, take a sequence $\{\alpha_{i,m}: m\in\mathbb{N}\}$

such that $\alpha_{i,m} \to \alpha_i$. Then $\alpha_m \to \alpha$ and

$$\begin{aligned}\limsup_{m\to\infty} \int_\Theta \log\left(\int_A e^{v(a,\theta)} \alpha_m(da)\right) P_{\boldsymbol{\theta}}(d\theta) &\leq \int_\Theta \limsup_{m\to\infty} \log\left(\int_A e^{v(a,\theta)} \alpha_m(da)\right) P_{\boldsymbol{\theta}}(d\theta) \\ &\leq \int_\Theta \log\left(\limsup_{m\to\infty} \int_A e^{v(a,\theta)} \alpha_m(da)\right) P_{\boldsymbol{\theta}}(d\theta) \\ &\leq \int_\Theta \log\left(\int_A e^{v(a,\theta)} \alpha(da)\right) P_{\boldsymbol{\theta}}(d\theta),\end{aligned}$$

where the first inequality holds by the (reverse) Fatou lemma (applicable because the expected value $E\left[\sup_{a\in A} |v(a,\boldsymbol{\theta})|\right]$ is finite), and the last inequality by upper semi-continuity of $v(a,\theta)$ in a. Hence, V is upper semi-continuous, as desired.

A.7 Proof of Proposition 5

Without loss of generality, assume $\lambda = 1$. First, notice that an equilibrium in direct signals exists by Corollary 1. Moreover, observe that, if there is $p \in (0,1)$ such that (6) and (7) hold, then, by Lemma 3, the corresponding direct signals form an equilibrium. For the rest of the proof, fix equilibrium direct signals $\boldsymbol{x}$. Write

$$\begin{aligned} Q &= \times_{i\in N} P_{\boldsymbol{x}_i} \in \Delta(A). \\ Q_{-i} &= \times_{j\neq i} P_{\boldsymbol{x}_j} \in \Delta(A_{-i}), \quad \forall i \in N. \\ Q_{-ij} &= \times_{k\neq j,i} P_{\boldsymbol{x}_k} \in \Delta(A_{-ij}), \quad \forall i,j \in N \text{ with } i \neq j. \end{aligned}$$

By Lemma 3, the following two conditions hold:

$$\frac{dP_{(\boldsymbol{x},\boldsymbol{\theta})}}{d(Q\times P_{\boldsymbol{\theta}})}(a,\theta) = \frac{e^{v(a,\theta)}}{\int_A e^{v(a',\theta)} dQ(a')}, \quad Q \times P_{\boldsymbol{\theta}} - a.s.. \tag{26}$$

$$1 \geq \int_\Theta \frac{\int_{A_{-i}} e^{v(a_i, a'_{-i},\theta)} dQ_{-i}(a'_{-i})}{\int_A e^{v(a',\theta)} dQ(a')} dP_{\boldsymbol{\theta}}(\theta), \quad \forall i \in N \text{ and } \forall a_i \in A_i. \tag{27}$$

Since $E\left[e^{-\rho(1,\boldsymbol{\theta})/\lambda}\right] > 1$ and $E\left[e^{\rho(0,\boldsymbol{\theta})/\lambda}\right] > 1$, it is easy to check that all signals must be not degenerate, otherwise (27) does not hold. Moreover, since there are only two actions both played with positive marginal probability, (27) must hold with equality. Hence, to complete the proof, it is enough to show that statements (i), (ii), and (iii) in Lemma 3 hold for $\boldsymbol{x}$. The rest of the proof proceeds by steps. Step 3 reports a result showed by Yang [2015] in the proof of his Proposition 2.

Step 1 ((ii) holds)**.** The random variables $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_n$, and $\boldsymbol{\theta}$, are affiliated.

Proof of the Step. By (26), it is enough to check that the density

$$f(a,\theta) = \frac{e^{v(a,\theta)}}{\int_A e^{v(a',\theta)} dQ(a')}, \quad \forall a \in A \text{ and } \forall \theta \in \Theta,$$

is log-supermodular. Take $a, a' \in A$ and $\theta, \theta' \in \Theta$. Without loss of generality, let $\theta \geq \theta'$. Since ρ is non-decreasing, v is supermodular, and therefore e^v is log-supermodular:

$$e^{v(a \vee a',\theta)} e^{v(a \wedge a',\theta')} \geq e^{v(a,\theta)} e^{v(a',\theta')}.$$

Dividing both sides by $\int_A e^{v(a'',\theta)} dQ(a'') \int_A e^{v(a'',\theta')} dQ(a'')$, we get

$$f(a \vee a', \theta) f(a \wedge a', \theta') \geq f(a,\theta) f(a',\theta'),$$

as desired. □

Step 2. For all players i,

$$P\left(P(\boldsymbol{x}_i = 1|\boldsymbol{\theta}) \neq P(\boldsymbol{x}_i = 1)\right) > 0.$$

Proof of the Step. Fix player i. By (26),

$$\log \frac{P(\boldsymbol{x}_i = 1|\theta)}{1 - P(\boldsymbol{x}_i = 1|\theta)} = \log \frac{P(\boldsymbol{x}_i = 1)}{1 - P(\boldsymbol{x}_i = 1)} + \log \frac{\int_{A_{-i}} e^{v(1,a_{-i},\theta)} dQ_{-i}(a_{-i})}{\int_{A_{-i}} e^{v(0,a_{-i},\theta)} dQ_{-i}(a_{-i})}, \quad P_{\boldsymbol{\theta}} - a.s.. \tag{28}$$

By $E\left[e^{-\rho(1,\boldsymbol{\theta})/\lambda}\right] > 1$,

$$P\left(\max_{l=0,\ldots,n-1} \rho\left(\frac{l}{n-1}, \boldsymbol{\theta}\right) \leq \rho(1,\boldsymbol{\theta}) < 0\right) > 0.$$

Since $v(1,a_{-i},\theta) - v(0,a_{-i},\theta) = \rho(|a_{-i}|,\theta)$ for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, this implies that

$$P\left(e^{v(1,a_{-i},\boldsymbol{\theta})} < e^{v(0,a_{-i},\boldsymbol{\theta})},\ \forall a_{-i} \in A_{-i}\right) > 0.$$

Hence,

$$P\left(\int_{A_{-i}} e^{v(1,a_{-i},\boldsymbol{\theta})} Q_{-i}(da_{-i}) < P(\boldsymbol{x}_i = 1)\right) > 0,$$

as desired. $\square$

Step 3 (Yang [2015])**.** Take $f, g : \Theta \to (0,1)$ measurable such that

$$\log \frac{f(\theta)}{1-f(\theta)} - \log \frac{g(\theta)}{1-g(\theta)} \leq \log \frac{\int f dP_{\boldsymbol{\theta}}}{1-\int f dP_{\boldsymbol{\theta}}} - \log \frac{\int g dP_{\boldsymbol{\theta}}}{1-\int g dP_{\boldsymbol{\theta}}}, \quad P_{\boldsymbol{\theta}} - a.s..$$

If $P_{\boldsymbol{\theta}}(\theta : g(\theta) \neq \int g dP_{\boldsymbol{\theta}}) > 0$, then $\int g dP_{\boldsymbol{\theta}} \geq \int f dP_{\boldsymbol{\theta}}$.

Step 4 ((i) holds)**.** There is $p \in (0,1)$ such that $P(\boldsymbol{x}_i = 1) = p$ for all players i. Hence, signals are exchangeable.

Proof of the Step. For all players i, define $p_i = P(\boldsymbol{x}_i = 1)$. By (26),

$$\log \frac{P(\boldsymbol{x}_i = 1|\theta)}{1 - P(\boldsymbol{x}_i = 1|\theta)} = \log \frac{p_i}{1-p_i} + \log \frac{\int_{A_{-i}} e^{v(1,a_{-i},\theta)} dQ_{-i}(a_{-i})}{\int_{A_{-i}} e^{v(0,a_{-i},\theta)} dQ_{-i}(a_{-i})}, \quad P_{\boldsymbol{\theta}} - a.s.. \tag{29}$$

Fix players i and j with $i \neq j$. Define, for all $t \in (0,1)$ and $\theta \in \Theta$,

$$f(t,\theta) = \frac{t \int_{A_{-ij}} e^{v(1,1,a_{-i},\theta)} dQ_{-ij}(a_{-ij}) + (1-t) \int_{A_{-ij}} e^{v(1,0,a_{-i},\theta)} dQ_{-ij}(a_{-ij})}{t \int_{A_{-ij}} e^{v(0,1,a_{-ij},\theta)} dQ_{-ij}(a_{-ij}) + (1-t) \int_{A_{-ij}} e^{v(0,0,a_{-ij},\theta)} dQ_{-ij}(a_{-ij})}.$$

In particular, notice that

$$f(p_j, \theta) = \frac{\int_{A_{-i}} e^{v(1,a_{-i},\theta)} dQ_{-i}(a_{-i})}{\int_{A_{-i}} e^{v(0,a_{-i},\theta)} dQ_{-i}(a_{-i})}, \quad \forall \theta \in \Theta.$$
$$f(p_i, \theta) = \frac{\int_{A_{-j}} e^{v(1,a_{-j},\theta)} dQ_{-j}(a_{-j})}{\int_{A_{-j}} e^{v(0,a_{-j},\theta)} dQ_{-j}(a_{-j})}, \quad \forall \theta \in \Theta.$$

Differentiating f with respect to t, it is easy to see that f is non-decreasing in t whenever, for all $\theta \in \Theta$,

$$\frac{\int_{A_{-ij}} e^{v(1,1,a_{-i},\theta)} dQ_{-ij}(a_{-ij})}{\int_{A_{-ij}} e^{v(1,0,a_{-i},\theta)} dQ_{-ij}(a_{-ij})} \geq \frac{\int_{A_{-ij}} e^{v(0,1,a_{-ij},\theta)} dQ_{-ij}(a_{-ij})}{\int_{A_{-ij}} e^{v(0,0,a_{-ij},\theta)} dQ_{-ij}(a_{-ij})}. \tag{30}$$

Since ρ is non-decreasing, v is supermodular, and therefore e^v is log-supermodular. Hence, by Karlin and Rinott [1980], for all $\theta \in \Theta$, the function

$$(a_i, a_j) \to \int_{A_{-ij}} e^{v(a_i,a_j,a_{-ij},\theta)} dQ_{-ij}(a_{-ij})$$

is log-supermodular. Therefore, the inequality (30) is satisfied, and f is non-decreasing in t. Assume now that $p_i \geq p_j$. From (29),

$$\log \frac{P(\boldsymbol{x}_i = 1|\theta)}{1 - P(\boldsymbol{x}_i = 1|\theta)} - \log \frac{P(\boldsymbol{x}_j = 1|\theta)}{1 - P(\boldsymbol{x}_j = 1|\theta)} = \log \frac{p_i}{1 - p_i} - \log \frac{p_j}{1 - p_j} + \log \frac{f(p_j, \theta)}{f(p_i, \theta)}, \quad Q_{\boldsymbol{\theta}} - a.s..$$

Since f is non-decreasing in t and $p_i \geq p_j$,

$$\log \frac{P(\boldsymbol{x}_i = 1|\theta)}{1 - P(\boldsymbol{x}_i = 1|\theta)} - \log \frac{P(\boldsymbol{x}_j = 1|\theta)}{1 - P(\boldsymbol{x}_j = 1|\theta)} \leq \log \frac{p_i}{1 - p_i} - \log \frac{p_j}{1 - p_j}, \quad Q_{\boldsymbol{\theta}} - a.s..$$

By Step 2, the hypothesis of Step 3 is satisfied. Hence $p_j \geq p_i$, which implies $p_i = p_j$, as desired. □

Step 5 ((iii) holds)**.** $P(\{\boldsymbol{x}_i = 1, \forall i \in N\}|\{\boldsymbol{x}_i = \boldsymbol{x}_j, \forall i, j \in N\}, \theta) = \frac{1}{2}$ whenever

$$\frac{1}{n} \sum_{l=0}^{n-1} \rho(l, \theta) = \log \frac{1 - p}{p}, \quad \text{with } p = P(x_i = 1), \, \forall i \in N. \tag{31}$$

Proof of the Step. Let $B = \{(0, \dots 0), (1, \dots 1)\}$. By (26) and Step 4,

$$P(\boldsymbol{x} = (1, \dots 1)|\boldsymbol{x} \in B, \theta) = \frac{e^{v(1,\dots,1,\theta)} p^n}{e^{v(1,\dots,1,\theta)} p^n + (1 - p)^n}, \quad P_{\boldsymbol{\theta}} - a.s..$$

Hence, when θ satisfies (31), $P(\boldsymbol{x} = (1, \dots 1)|\boldsymbol{x} \in B, \theta) = 1/2$, as desired. □

A.8 Proof of Proposition 6

Let $\boldsymbol{x}$ be a Gaussian equilibrium in direct signals. The proof is based on the following lemma:

Lemma 4. *For all players i, $Var(\boldsymbol{x}_i) = Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)$ and $(\boldsymbol{x}_i \perp (\boldsymbol{x}_{-i}, \boldsymbol{\theta}))|\boldsymbol{w}_i$.*

Proof of the Lemma. Fix player i. Since u_i depends on (a_{-i}, θ) only through the target w_i, by Lemma 1 we have $(\boldsymbol{x}_i \perp (\boldsymbol{x}_{-i}, \boldsymbol{\theta}))|\boldsymbol{w}_i$. Moreover, $\boldsymbol{x}_i$ is solves the following optimization problem:

$$\text{maximize } -\frac{1}{2}(\boldsymbol{x}'_i - \boldsymbol{w}_i)^2 - C_i(\boldsymbol{x}'_i; \boldsymbol{w}_i) \text{ over direct signals } \boldsymbol{x}'_i. \tag{32}$$

Now I use (32) to show $Var(\boldsymbol{x}_i) = Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)$, and conclude the proof of the step.

By contradiction, suppose that $Var(\boldsymbol{x}_i) \neq Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)$. Consider the alternative direct signal

$$\boldsymbol{x}_i' = \frac{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)}{Var(\boldsymbol{x}_i)}(\boldsymbol{x}_i - E[\boldsymbol{x}_i]) + E[\boldsymbol{x}_i].$$

By Condition 1, $C_i(\boldsymbol{x}_i'; \boldsymbol{w}_i) = C_i(\boldsymbol{x}_i, \boldsymbol{w}_i)$. Moreover,

$$E[-(\boldsymbol{x}_i - \boldsymbol{w}_i)^2] - E[-(\boldsymbol{x}_i' - \boldsymbol{w}_i)^2] = -\frac{1}{Var(\boldsymbol{x}_i)}(Var(\boldsymbol{x}_i) - Cov(\boldsymbol{x}_i, \boldsymbol{w}_i))^2 < 0,$$

where the last inequality holds by the assumption $Var(\boldsymbol{x}_i) \neq Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)$. In conclusion,

$$E[-(\boldsymbol{x}_i - \boldsymbol{w}_i)^2] - C_i(\boldsymbol{x}_i; \boldsymbol{w}_i) < E[-(\boldsymbol{x}_i' - \boldsymbol{w}_i)^2] - C_i(\boldsymbol{x}_i'; \boldsymbol{w}_i),$$

and therefore $\boldsymbol{x}_i$ cannot be a solution of (32): contradiction. As a result, $Var(\boldsymbol{x}_i)$ is equal to $Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)$ in the first place, as desired. □

To ease the exposition, assume that $Var(\boldsymbol{x}_i) > 0$ for all players i. (If not, just focus on the set of players whose signals are not degenerate). The proof proceeds in steps, and the last step concludes.

Step 1. For all $i, j \in N$ with $i \neq j$,

$$\frac{Var(\boldsymbol{w}_i)}{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)} Cov(\boldsymbol{x}_j, \boldsymbol{x}_i | \boldsymbol{\theta}) - \sum_{k \neq i} \gamma_{ik} Cov(\boldsymbol{x}_j, \boldsymbol{x}_k | \boldsymbol{\theta}) = 0.$$

Proof of the Step. Fix $i, j \in N$ with $i \neq j$. Define $\tilde{\boldsymbol{x}}_i = \frac{Var(\boldsymbol{w}_i)}{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)} \boldsymbol{x}_i$. To prove the claim, it is enough to show

$$Cov(\boldsymbol{x}_j, \tilde{\boldsymbol{x}}_i | \boldsymbol{\theta}) = Cov(\boldsymbol{x}_j, \boldsymbol{w}_i | \boldsymbol{\theta}).$$

The above equality immediately comes from

$$\begin{aligned} E[(\boldsymbol{x}_j - E[\boldsymbol{x}_j])(\tilde{\boldsymbol{x}}_i - E[\tilde{\boldsymbol{x}}_i]) | \boldsymbol{\theta}] &= E[E[(\boldsymbol{x}_j - E[\boldsymbol{x}_j])(\tilde{\boldsymbol{x}}_i - E[\tilde{\boldsymbol{x}}_i]) | \boldsymbol{\theta}, \boldsymbol{w}_i, \boldsymbol{x}_j] | \boldsymbol{\theta}] \\ &= E[(\boldsymbol{x}_j - E[\boldsymbol{x}_j]) E[\tilde{\boldsymbol{x}}_i - E[\tilde{\boldsymbol{x}}_i] | \boldsymbol{w}_i] | \boldsymbol{\theta}] \\ &= E[(\boldsymbol{x}_j - E[\boldsymbol{x}_j])(\boldsymbol{w}_i - E[\boldsymbol{w}_i]) | \boldsymbol{\theta}], \end{aligned}$$

where the second equality holds by $(\boldsymbol{x}_i \perp (\boldsymbol{x}_j, \boldsymbol{\theta})) | \boldsymbol{w}_i$ (Step 4). □

Step 2. For all $i \in N$,

$$\frac{Var(\boldsymbol{w}_i)}{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)} Var(\boldsymbol{x}_i|\boldsymbol{\theta}) - \sum_{k \neq i} \gamma_{ik} Cov(\boldsymbol{x}_i, \boldsymbol{x}_k|\boldsymbol{\theta}) = \frac{Var(\boldsymbol{w}_i)}{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)} Var(\boldsymbol{x}_i|\boldsymbol{w}_i).$$

Proof of the Step. Fix $i, j \in N$ with $i \neq j$. Define $\tilde{\boldsymbol{x}}_i = \frac{Var(\boldsymbol{w}_i)}{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)} \boldsymbol{x}_i$. To prove the claim, it is enough to show

$$Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{\theta}) - Cov(\tilde{\boldsymbol{x}}_i, \boldsymbol{w}_i|\boldsymbol{\theta}) = Var(\boldsymbol{x}_i|\boldsymbol{w}_i).$$

First, notice that

$$\begin{aligned}
E[(\tilde{\boldsymbol{x}}_i - E[\tilde{\boldsymbol{x}}_i])(\boldsymbol{w}_i - E[\boldsymbol{w}_i])|\boldsymbol{\theta}] &= E[E[(\tilde{\boldsymbol{x}}_i - E[\tilde{\boldsymbol{x}}_i])(\boldsymbol{w}_i - E[\boldsymbol{w}_i])|\boldsymbol{w}_i, \boldsymbol{\theta}]|\boldsymbol{\theta}] \\
&= E[E[(\tilde{\boldsymbol{x}}_i - E[\tilde{\boldsymbol{x}}_i])|\boldsymbol{w}_i](\boldsymbol{w}_i - E[\boldsymbol{w}_i])|\boldsymbol{\theta}] \\
&= E[(\boldsymbol{w}_i - E[\boldsymbol{w}_i])^2|\boldsymbol{\theta}],
\end{aligned}$$

where the second equality holds by $(\boldsymbol{x}_i \perp \boldsymbol{\theta})|\boldsymbol{w}_i$ (Step 4). Moreover, observe that

$$\begin{aligned}
Var(\boldsymbol{w}_i|\boldsymbol{\theta}) &= Var(\tilde{\boldsymbol{x}}_i - (\tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i)|\boldsymbol{\theta}) \\
&= Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{\theta}) + Var(\tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i|\boldsymbol{\theta}) - 2Cov(\tilde{\boldsymbol{x}}_i, \tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i|\boldsymbol{\theta}) \\
&= Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{\theta}) + Var(\tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i|\boldsymbol{\theta}) - 2Cov(\tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i, \tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i|\boldsymbol{\theta}) \\
&= Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{\theta}) - Var(\tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i|\boldsymbol{\theta}) = Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{\theta}) - Var(\tilde{\boldsymbol{x}}_i - \boldsymbol{w}_i) \\
&= Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{\theta}) - Var(\tilde{\boldsymbol{x}}_i|\boldsymbol{w}_i),
\end{aligned}$$

where the third and fifth equalities hold by $(\boldsymbol{x}_i \perp \boldsymbol{\theta})|\boldsymbol{w}_i$ (Step 4). Combining the two chains of equalities, we obtain the desired result. □

Step 3. For all players i,

$$\frac{Var(\boldsymbol{w}_i)}{Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)} = \frac{Var(\boldsymbol{x}_i)}{Var(\boldsymbol{x}_i) - Var(\boldsymbol{x}_i|\boldsymbol{x}_{-i}, \boldsymbol{\theta})}.$$

Proof of the Step. It is enough to observe that (i) $Var(\boldsymbol{x}_i|\boldsymbol{x}_{-i}, \boldsymbol{\theta}) = Var(\boldsymbol{x}_i|\boldsymbol{w}_i)$ since $(\boldsymbol{x}_i \perp (\boldsymbol{x}_{-i}, \boldsymbol{\theta}))|\boldsymbol{w}_i$ (Step 4), and (ii)

$$Cov(\boldsymbol{x}_i, \boldsymbol{w}_i) = Var(\boldsymbol{w}_i) Cor(\boldsymbol{x}_i, \boldsymbol{w}_i)^2,$$

since $Var(\boldsymbol{x}_i) = Cov(\boldsymbol{x}_i, \boldsymbol{w}_i)$ (Step 4). $\square$

Step 4 (Proposition 6)**.** For all players i and j,

$$Cov(\boldsymbol{x}_i, \boldsymbol{x}_j|\boldsymbol{\theta}) = Var(\boldsymbol{x}_i|\boldsymbol{x}_{-i}, \boldsymbol{\theta}) \left(\sum_{k=0}^{\infty} \tilde{\Gamma}^k \right)_{ij}.$$

Proof of the Step. Write $\Sigma_{\boldsymbol{x}|\boldsymbol{\theta}}$ for the conditional covariance matrix of $\boldsymbol{x}$ given $\boldsymbol{\theta}$. Define auxiliary matrices M and M' such that, for all $i, j \in N$,

$$\begin{aligned} m_{ij} &= \frac{Var(\boldsymbol{x}_i)}{Var(\boldsymbol{x}_i) - Var(\boldsymbol{x}_i|\boldsymbol{x}_{-i}, \boldsymbol{\theta})} 1(i = j), \\ m'_{ij} &= Var(\boldsymbol{x}_i|\boldsymbol{x}_{-i}, \boldsymbol{\theta}) 1(i = j). \end{aligned}$$

Since $(\boldsymbol{x}_i \perp (\boldsymbol{x}_{-i}, \boldsymbol{\theta}))|\boldsymbol{w}_i$ for all players i (Step 4), then $m'_{ii} = Var(\boldsymbol{x}_i|\boldsymbol{w}_i)$ for all players i. Hence, By Steps 1, 2, and 3,

$$\Sigma_{\boldsymbol{x}|\boldsymbol{\theta}}(M - \Gamma) = M'M.$$

Since $\tilde{\Gamma} = \Gamma M^{-1}$, we get

$$\Sigma_{\boldsymbol{x}|\boldsymbol{\theta}}(\mathcal{I} - \tilde{\Gamma}) = M',$$

where $\mathcal{I}$ is the identity matrix. Since $\tilde{\Gamma}$ is a contraction, the matrix $\mathcal{I} - \tilde{\Gamma}$ is invertible. Hence, $\Sigma_{\boldsymbol{x}|\boldsymbol{\theta}} = M'(\mathcal{I} - \tilde{\Gamma})^{-1}$, as desired. $\square$

A.9 Proof of Corollary 3

Let $\boldsymbol{x}$ be a Gaussian equilibrium in direct signals. The proof proceeds in steps, and the last step concludes.

Step 1. For all players i,

$$Var(\boldsymbol{x}_i) \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|)^2}.$$

Proof of the Step. Pick some player i such that $Var(\boldsymbol{x}_i) \geq Var(\boldsymbol{x}_j)$ for all $j \in N$.

Notice that

$$\begin{aligned} Var(\boldsymbol{x}_i) = Cov(\boldsymbol{x}_i, \boldsymbol{w}_i) &= Cov(\boldsymbol{x}_i, \boldsymbol{\theta}_i) + \sum_{j \neq i} Cov(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ &\leq Std(\boldsymbol{x}_i) Std(\boldsymbol{\theta}_i) + Var(\boldsymbol{x}_i) \sum_{j \neq i} |\gamma_{ij}|, \\ &\leq Std(\boldsymbol{x}_i) Std(\boldsymbol{\theta}_i) + Var(\boldsymbol{x}_i) \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|, \end{aligned}$$

where the first equality holds by Lemma 4 (see Section A.8). As a consequence,

$$\frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|)^2} \geq Var(\boldsymbol{x}_i) \geq Var(\boldsymbol{x}_j), \quad \forall j \in N,$$

as desired. □

Step 2 (Corollary 3)**.** For all players i and j,

$$|Cov(\boldsymbol{x}_i, \boldsymbol{x}_j | \boldsymbol{\theta})| \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|)^2} \left(\sum_{k=0}^{\infty} \hat{\Gamma}^k \right)_{ij},$$

Proof of the Step. By Proposition 6,

$$|Cov(\boldsymbol{x}_i, \boldsymbol{x}_j | \boldsymbol{\theta})| = Var(\boldsymbol{x}_i | \boldsymbol{x}_{-i}, \boldsymbol{\theta}) \left| \left(\sum_{k=0}^{\infty} \tilde{\Gamma}^k \right)_{ij} \right|$$

Therefore, from Step 1,

$$|Cov(\boldsymbol{x}_i, \boldsymbol{x}_j | \boldsymbol{\theta})| \leq \frac{\max_{i \in N} \sigma^2_{\boldsymbol{\theta}_i}}{(1 - \max_{j \in N} \sum_{i \in N} |\gamma_{ij}|)^2} \left| \left(\sum_{k=0}^{\infty} \tilde{\Gamma}^k \right)_{ij} \right|$$

Finally, since

$$|\tilde{\gamma}_{ij}| = |\gamma_{ij}| \frac{Var(\boldsymbol{x}_j) - Var(\boldsymbol{x}_j | \boldsymbol{x}_{-j}, \boldsymbol{\theta})}{Var(\boldsymbol{x}_j)} \leq |\gamma_{ij}|, \quad \forall i, j \in N,$$

then $\left| \left(\sum_{k=0}^{\infty} \tilde{\Gamma}^k \right)_{ij} \right| \leq \left(\sum_{k=0}^{\infty} \hat{\Gamma}^k \right)_{ij}$. The inequality follows. □

A.10 Proof of Proposition 7

Let $\boldsymbol{x}$ be a profile of Gaussian direct signals. From Lemmas 1 and 2, $\boldsymbol{x}$ is an equilibrium if and only if, for all players i and j with $i \neq j$,

$$E[\boldsymbol{x}_i] = E[\boldsymbol{w}], \tag{33}$$

$$Var(\boldsymbol{x}_i) = Cov(\boldsymbol{x}_i, \boldsymbol{w}) = \max\{Var(\boldsymbol{w}) - \lambda, 0\}, \tag{34}$$

$$Var(\boldsymbol{w})Cov(\boldsymbol{x}_i, \boldsymbol{x}_j) = Cov(\boldsymbol{x}_i, \boldsymbol{w})Cov(\boldsymbol{x}_j, \boldsymbol{w}), \tag{35}$$

$$Var(\boldsymbol{w})Cov(\boldsymbol{x}_i, \boldsymbol{\theta}) = Cov(\boldsymbol{x}_i, \boldsymbol{w})Cov(\boldsymbol{\theta}, \boldsymbol{w}). \tag{36}$$

In particular, (33)-(36) imply that signals are exchangeable, which, from now, I assume without loss of generality.

First, I show that, in equilibrium, $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|\boldsymbol{\theta}$ for all players i. If $Var(\boldsymbol{x}_i) = 0$ for all players i, there is nothing to do. Hence, assume that $Var(\boldsymbol{x}_i) > 0$ for all players i. In particular, by (34), this implies that $Var(\boldsymbol{w}) > 0$ and $Cov(\boldsymbol{x}_i, \boldsymbol{w}) > 0$ for all players i. From (35) and (36), averaging across players and taking the limit,

$$Var(\boldsymbol{w})Var(\bar{\boldsymbol{x}}) = Cov(\bar{\boldsymbol{x}}, \boldsymbol{w})^2, \quad \text{and} \quad Var(\boldsymbol{w})Cov(\bar{\boldsymbol{x}}, \boldsymbol{\theta}) = Cov(\bar{\boldsymbol{x}}, \boldsymbol{w})Cov(\boldsymbol{\theta}, \boldsymbol{w}).$$

Combining the two equalities:

$$Cov(\bar{\boldsymbol{x}}, \boldsymbol{\theta})Cov(\bar{\boldsymbol{x}}, \boldsymbol{w}) = Var(\bar{\boldsymbol{x}})Cov(\boldsymbol{\theta}, \boldsymbol{w}).$$

Since $\boldsymbol{w} = r_0\boldsymbol{\theta} + r\bar{\boldsymbol{x}}$, this implies

$$r_0 Cov(\bar{\boldsymbol{x}}, \boldsymbol{\theta})^2 = r_0 Var(\bar{\boldsymbol{x}})Var(\boldsymbol{\theta}).$$

Since $r_0 \neq 0$ and $Var(\boldsymbol{\theta}) > 0$, it must be that $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) = 0$. Hence, $Var(\boldsymbol{w}|\boldsymbol{\theta}) = 0$ and $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|\boldsymbol{\theta}$ for all players i, by (35).

Now, since, in equilibrium, signals are exchangeable and conditionally independent given the state, without loss of generality, for all players i, assume

$$\boldsymbol{x}_i = \tilde{\alpha}_0 + \tilde{\alpha}\boldsymbol{\theta} + \tilde{\beta}\boldsymbol{\epsilon}_i, \tag{37}$$

where $\tilde{\alpha}_0$, $\tilde{\alpha}$ and $\tilde{\beta}$ are arbitrary scalars. Notice that $\bar{\boldsymbol{x}} = \tilde{\alpha}_0 + \tilde{\alpha}\boldsymbol{\theta}$. Hence

$$\boldsymbol{w} = (r_0 + \tilde{\alpha} r)\boldsymbol{\theta} + r\tilde{\alpha}_0. \tag{38}$$

Clearly, (35) and (36) are satisfied. Now combine the expressions (37) and (38) with the conditions (33) and (34) to obtain equilibrium equations only in terms of the parameters $\tilde{\alpha}_0$, $\tilde{\alpha}$ and $\tilde{\beta}$: $\boldsymbol{x}$ is an equilibrium if and only if (37) hold and

$$\begin{aligned}
\tilde{\alpha}_0 + \tilde{\alpha}\mu_{\boldsymbol{\theta}} &= (r_0 + \tilde{\alpha} r)\mu_{\boldsymbol{\theta}} + r\tilde{\alpha}_0, \\
\tilde{\alpha}^2\sigma^2_{\boldsymbol{\theta}} + \tilde{\beta}^2 &= \tilde{\alpha}(r_0 + \tilde{\alpha} r)\sigma^2_{\boldsymbol{\theta}}, \\
\tilde{\alpha}(r_0 + \tilde{\alpha} r)\sigma^2_{\boldsymbol{\theta}} &= \max\{(r_0 + \tilde{\alpha} r)^2\sigma^2_{\boldsymbol{\theta}} - \lambda, 0\}.
\end{aligned}$$

Given the assumption $\lambda < r_0^2\sigma^2_{\boldsymbol{\theta}}$, it is easy to check that this system has a unique solution in $\tilde{\alpha}_0$, $\tilde{\alpha}$ and $\tilde{\beta}^2$ such that

$$\begin{aligned}
\tilde{\alpha}_0 &= \frac{r_0}{1-r}\mu_{\boldsymbol{\theta}} - \tilde{\alpha}\mu_{\boldsymbol{\theta}}, \\
\tilde{\alpha} &= \frac{r_0}{1-r}\frac{\sqrt{(1-2r)^2 + 4r(1-r)(1-\frac{\lambda}{\sigma^2_{\boldsymbol{\theta}} r_0^2})} - (1-2r)}{2r}, \\
\tilde{\beta}^2 &= \tilde{\alpha}(r_0 + \tilde{\alpha} r - \tilde{\alpha})\sigma^2_{\boldsymbol{\theta}}.
\end{aligned}$$

The choice of the sign of β does not affect the distribution of signals and state, and therefore we can take $\tilde{\beta} \geq 0$ without loss of generality. Plugging $\tilde{\alpha}_0$, $\tilde{\alpha}$ and $\tilde{\beta}$ in (37) and rearranging, it is easy to see that the representation in the proposition in terms of α and β holds, as desired.

A.11 Proof of Proposition 8

Fix an equilibrium $\boldsymbol{x}$ in direct signals such that $\sup_{i \in N} Var(f_i(\boldsymbol{x}_i)) < \infty$. The proof proceeds by steps, and the last step proves Proposition 8. Throughout, for all $\bar{a} \in \bar{A}$ and $\theta \in \Theta$, define

$$g_i(a_i|\bar{a}, \theta) = \frac{e^{u_i(a_i, \bar{a}, \theta)/\lambda_i}}{\int_{A_i} e^{u_i(a_i, \bar{a}, \theta)/\lambda_i} dP_{\boldsymbol{x}_i}(a_i)}, \quad \forall i \in N \text{ and } \forall a_i \in A_i.$$

Whenever the integral exists, define also

$$h_i(\bar{a}, \theta) = \int_{A_i} f_i(a_i) g_i(a_i|\bar{a}, \theta) dP_{\boldsymbol{x}_i}(a_i).$$

Step 1. With probability one,

$$\bar{\boldsymbol{x}} = \limsup_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} h_i(\bar{\boldsymbol{x}}, \boldsymbol{\theta}).$$

Proof of the Step. Define

$$\boldsymbol{z}_i = f_i(\boldsymbol{x}_i) - E[f_i(\boldsymbol{x}_i)|\bar{\boldsymbol{x}}, \boldsymbol{\theta}], \quad \forall i \in N.$$

Since the players' utilities depend on the others' actions only through the average action, $(\boldsymbol{x}_i \perp \boldsymbol{x}_{-i})|(\boldsymbol{\theta}, \bar{\boldsymbol{x}})$ for all players i, by Lemma 1. Therefore, the family of random variables $\{\boldsymbol{z}_i : i \in N\}$ is independent. Moreover, they have mean zero and $\sup_{i\in N} Var(\boldsymbol{z}_i) < \infty$, since $\sup_{i\in N} Var(f_i(\boldsymbol{x}_i)) < \infty$ by assumption. Hence, by the strong law of large numbers, with probability one,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} \boldsymbol{z}_i = 0. \tag{39}$$

By definition, $f_i(\boldsymbol{x}_i) = E[f_i(\boldsymbol{x}_i)|\bar{\boldsymbol{x}}, \boldsymbol{\theta}] - \boldsymbol{z}_i$, for all players i. Hence, with probability one,

$$\begin{aligned} \bar{\boldsymbol{x}} &= \limsup_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} (E[f_i(\boldsymbol{x}_i)|\bar{\boldsymbol{x}}, \boldsymbol{\theta}] - \boldsymbol{z}_i) \\ &= \limsup_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} E[f_i(\boldsymbol{x}_i)|\bar{\boldsymbol{x}}, \boldsymbol{\theta}], \\ &= \limsup_{m\to\infty} \frac{1}{m} \sum_{i=1}^{m} h_i(\bar{\boldsymbol{x}}, \boldsymbol{\theta}) \end{aligned}$$

where the second equality holds by (39), and the third inequality, which concludes the proof of the step, by Lemma 2. □

Step 2. Fix $\theta \in \Theta$ and $i \in N$. For all $\bar{a}, \bar{a}' \in \bar{A}$,

$$\bar{a} \geq \bar{a}' \quad \Rightarrow \quad h_i(\bar{a}, \theta) \leq h_i(\bar{a}', \theta).$$

Proof of the Step. Assume $\bar{a} \geq \bar{a}'$. Since individual and average action are strategic substitutes,

$$a_i \geq a_i' \quad \Rightarrow \quad e^{u_i(a_i,\bar{a},\theta)/\lambda_i} e^{u_i(a_i',\bar{a},\theta)/\lambda_i} \leq e^{u_i(a_i,\bar{a}',\theta)/\lambda_i} e^{u_i(a_i',\bar{a}',\theta)/\lambda_i}.$$

Therefore

$$a_i \geq a_i' \quad \Rightarrow \quad \frac{g_i(a_i|\bar{a}', \theta)}{g_i(a_i|\bar{a}, \theta)} \geq \frac{g_i(a_i'|\bar{a}', \theta)}{g_i(a_i'|\bar{a}, \theta)}. \tag{40}$$

Define the pair of probabilities Q and Q' on A_i such that

$$\frac{dQ}{dP_{\boldsymbol{x}_i}}(a_i) = g_i(a_i|\bar{a}, \theta), \quad a.s.,$$
$$\frac{dQ'}{dP_{\boldsymbol{x}_i}}(a_i) = g_i(a_i|\bar{a}', \theta), \quad a.s..$$

By (40), Q' (first-order) stochastically dominates Q. Hence, since f_i is non-decreasing by assumption,

$$h_i(\bar{a}, \theta) = \int_{A_i} f_i(a_i) dQ(a_i) \leq \int_{A_i} f_i(a_i) dQ'(a_i) = h_i(\bar{a}', \theta),$$

as desired. □

Step 3. For all $\theta \in \Theta$, the set

$$\bar{A}_\theta = \left\{ \bar{a} \in \bar{A} : \bar{a} = \limsup_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} h_i(\bar{a}, \theta) \right\}.$$

contains at most one element.

Proof of the Step. Take $\bar{a}, \bar{a}' \in \bar{A}_\theta$: we wish to show that $\bar{a} = \bar{a}'$. Assume $\bar{a} \geq \bar{a}'$, without loss of generality. By Step 2,

$$h_i(\bar{a}, \theta) \leq h_i(\bar{a}', \theta), \quad \forall i \in N.$$

As a result, by monotonicity of the limit superior,

$$\bar{a} = \limsup_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} h_i(\bar{a},\theta) \le \limsup_{m\to\infty} \frac{1}{m}\sum_{i=1}^{m} h_i(\bar{a}',\theta) = \bar{a}'.$$

Hence $\bar{a} \le \bar{a}'$. Since we assumed that $\bar{a} \ge \bar{a}'$, we obtain $\bar{a} = \bar{a}'$, as desired. □

Step 4 (Proposition 8)**.** The conditional distribution of $\bar{\boldsymbol{x}}$ is degenerate given $\boldsymbol{\theta}$, and therefore $Var(\bar{\boldsymbol{x}}|\boldsymbol{\theta}) = 0$.

Proof of the Step. By Step 1, the distribution $P_{(\bar{\boldsymbol{x}},\boldsymbol{\theta})}$ assigns probability one to the set

$$\{(\bar{a},\theta) \in \bar{A} \times \Theta : \bar{a} \in \bar{A}_\theta\},$$

where $\bar{A}_\theta$ is defined as in Step 3. This means that

$$P(\bar{\boldsymbol{x}} \in \bar{A}_{\boldsymbol{\theta}}|\boldsymbol{\theta}) = 1. \tag{41}$$

By Step 3, for every θ, the set $\bar{A}_\theta$ contains at most one element. Hence, (41) implies that the conditional distribution of $\bar{\boldsymbol{x}}$ is degenerate given $\boldsymbol{\theta}$, as desired. □

B Example

In this section, I provide an example of probability space and available signals that satisfy the assumptions of Section 3.1 when $n < \infty$. Throughout, fix a probability $P_\Theta \in \Delta(\Theta)$ representing uncertainty about the state.

To construct $(\Omega, \mathcal{F}, P)$, first define $\Omega_0 = \Theta$ and $\Omega_t = [0,1]$ for all $t = 1, 2, \ldots$. Next, set $\Omega = \prod_{t=0}^{\infty} \Omega_t$, and let $\mathcal{F}$ be the product sigma-algebra. Now, define $\boldsymbol{\theta} : \Omega \to \Theta$ such that $\boldsymbol{\theta}(\omega) = \omega_0$. Furthermore, for every $t = 1, 2 \ldots$, define $\boldsymbol{z}_t : \Omega \to [0,1]$ such that $\boldsymbol{z}_t(\omega) = \omega_t$. Finally, choose $P \in \Delta(\Omega)$ such that:

- The random variables $\boldsymbol{\theta}, \boldsymbol{z}_1, \boldsymbol{z}_2, \ldots$ are independent.
- $\boldsymbol{\theta} \sim P_\Theta$ and, for every $t = 1, 2, \ldots$, $\boldsymbol{z}_t$ is uniformly distributed.

For every player i, let $\boldsymbol{X}_i$ be the set of random variables $\boldsymbol{x}_i : \Omega \to X_i$ such that, $\boldsymbol{x}_i$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_t)$ for some choice of t.

Now, I verify that the assumptions of Section 3.1 are satisfied. First, fix player i, and take $P_{X\times\Theta} \in \Delta(X \times \Theta)$ such that $(\boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim P_{X_{-i}\times\Theta}$ for some $\boldsymbol{x}_{-i} \in \boldsymbol{X}_{-i}$. I want to find $\boldsymbol{x}_i \in \boldsymbol{X}_i$ such that $(\boldsymbol{x}, \boldsymbol{\theta}) \sim P_{X\times\Theta}$.

We begin with the case of X_i being a compact subset of $\mathbb{R}$. For every $x_{-i} \in X_{-i}$ and $\theta \in \Theta$, pick the distribution function $F(\cdot|x_{-i}, \theta) : X_i \to [0, 1]$ corresponding to the conditional probability over X_i given (x_{-i}, θ), according to $P_{X\times\Theta}$. Write $F^{-1}(\cdot|x_{-i}, \theta) : [0, 1] \to X_i$ for the generalized inverse distribution function.[42] Pick t large enough such that $\boldsymbol{x}_{-i}$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_{t-1})$. Define $\boldsymbol{x}_i : \Omega \to X_i$ such that

$$\boldsymbol{x}_i(\omega) = F^{-1}(\boldsymbol{z}_t(\omega)|\boldsymbol{x}_{-i}(\omega), \boldsymbol{\theta}(\omega)), \quad \forall \omega \in \Omega.$$

Notice that $\boldsymbol{x}_i$ is measurable with respect to $\mathcal{F}$, since $F^{-1} : [0, 1] \times X_{-i} \times \Theta \to X_i$ is jointly measurable.[43] Since $\boldsymbol{x}_{-i}$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_{t-1})$, $\boldsymbol{x}_i$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_t)$. Hence, $\boldsymbol{x}_i \in \boldsymbol{X}_i$. Finally, since

$$\boldsymbol{x}_i|(\boldsymbol{x}_{-i}, \boldsymbol{\theta}) \sim F(\cdot|\boldsymbol{x}_{-i}, \boldsymbol{\theta}),$$

then $(\boldsymbol{x}, \boldsymbol{\theta}) \sim P_{X\times\Theta}$, as desired.

Consider now the general case where X_i is an arbitrary Polish space. Let X_i' be a compact subset of $\mathbb{R}$ that is Borel isomorphic to X_i, that is, there is a bijection $\phi : X_i \to X_i'$ such that both ϕ and ϕ^{-1} are measurable (Parthasaraty [1967, Ch. 1]). Denote by $P_{X\times\Theta} \circ \phi^{-1} \in \Delta(X_i' \times X_{-i} \times \Theta)$ the pushforward of $P_{X\times\Theta}$ given the map $(x, \theta) \mapsto (\phi(x_i), x_{-i}, \theta)$. From above, we can find a random variable $\boldsymbol{x}_i' : \Omega \to X_i'$ with the following two properties. First, the joint distribution of $\boldsymbol{x}_i'$ and $(\boldsymbol{x}_{-i}, \boldsymbol{\theta})$ is $P_{X\times\Theta} \circ \phi^{-1}$. Second, $\boldsymbol{x}_i'$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_t)$ for some choice of t. Then, define $\boldsymbol{x}_i : \Omega \to X_i$ such that $\boldsymbol{x}_i(\omega) = \phi^{-1}(\boldsymbol{x}_i'(\omega))$ for all $\omega \in \Omega$. Clearly, $\boldsymbol{x}_i$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_t)$ and $(\boldsymbol{x}, \boldsymbol{\theta}) \sim P_{X\times\Theta}$, as desired.

Applying the same technique, one can verify that, if $P_{X\times\Theta} \in \Delta(X \times \Theta)$ and $\boldsymbol{\theta} \sim P_\Theta$, then there is $\boldsymbol{x} \in \boldsymbol{X}$ such that $(\boldsymbol{x}, \boldsymbol{\theta}) \sim P_{X\times\Theta}$. Finally, if $\boldsymbol{x}_i \in \boldsymbol{X}_i$ and $f : X_i \to X_i$ measurable, then $f(\boldsymbol{x}_i) \in \boldsymbol{X}_i$, since $f(\boldsymbol{x}_i)$ is measurable with respect to $\boldsymbol{x}_i$, and $\boldsymbol{x}_i$ is measurable with respect to $(\boldsymbol{\theta}, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_t)$ for some choice of t. In conclusion, all the assumptions of Section 3.1 are satisfied, as desired.

[42]That is, $F^{-1}(p|x_{-i}, \theta) = \inf\{x_i \in X_i : F(x_i|x_{-i}, \theta) \geq p\}$ for all $p \in [0, 1]$.

[43]Because F^{-1} is measurable in $(x_{-i}, \theta) \in X_{-i} \times \Theta$ and left continuous in $p \in [0, 1]$.

References

Angeletos, G.-M. and A. Pavan (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75, 1103–1142.

Arrow, K. J. (1996): "The Economics of Information: An Exposition," *Empirica*, 23, 119–128.

Bergemann, D. and S. Morris (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

——— (2015): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, Forthcoming.

Bergemann, D. and J. Välimäki (2002): "Information Acquisition and Efficient Mechanism Design," *Econometrica*, 70, 1007–1033.

Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 93–102.

Bonacich, P. (1987): "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 92, 1170–1182.

Calvó-Armengol, A., J. de Martí, and A. Prat (2015): "Communication and Influence," *Theoretical Economics*, 10, 649–690.

Carlsson, H. and E. Van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

Colombo, L., G. Femminis, and A. Pavan (2014): "Information Acquisition and Welfare," *The Review of Economic Studies*, 81, 1438–1483.

Cover, T. M. and J. A. Thomas (2006): *Elements of Information Theory*, Wiley-Interscience, 2nd ed.

Csiszár, I. (1967): "Information-Type Measures of Difference of Probability Distributions and Indirect Observations," *Studia Scientiarum Mathematicarum Hungarica*, 2, 299–318.

——— (1974): "On an Extremum Problem of Information Theory," *Studia Scientiarum Mathematicarum Hungarica*, 9, 57–71.

Diamond, D. W. and P. H. Dybvig (1983): "Bank Runs, Deposit Insurance and Liquidity," *Journal of Political Economy*, 90, 881–894.

Frankel, D. M., S. Morris, and A. Pauzner (2003): "Equilibrium Selection in Global Games with Strategic Complementarities," *Journal of Economic Theory*, 108, 1–44.

Goldstein, I. and A. Pauzner (2005): "Demand–Deposit Contracts and the Probability of Bank Runs," *The Journal of Finance*, 60, 1293–1327.

Hansen, O. H. and E. N. Torgersen (1974): "Comparison of Linear Normal Experiments," *The Annals of Statistics*, 2, 367–373.

Harsanyi, J. C. and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*, MIT Press.

Hébert, B. (2015): "Moral Hazard and the Optimality of Debt," Mimeo.

Hellwig, C., S. Kohls, and L. Veldkamp (2012): "Information Choice Technologies," *The American Economic Review: Papers and Proceedings*, 102, 35–40.

Hellwig, C. and L. Veldkamp (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *The Review of Economic Studies*, 76, 223–251.

Jackson, M. O. (2008): *Social and Economic Networks*, Princeton University Press.

Kalai, E. (2004): "Large Robust Games," *Econometrica*, 72, 1631–1665.

Karlin, S. and Y. Rinott (1980): "Classes of Orderings of Measures and Related Correlation Inequalities. I. Multivariate Totally Positive Distributions," *Journal of Multivariate Analysis*, 10, 467–498.

Katz, L. (1953): "A New Status Index Derived from Sociometric Analysis," *Psychometrika*, 18, 39–43.

Katz, M. L. and C. Shapiro (1986): "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, 94, 882–841.

Kim, K. and F. Z. X. Lee (2014): "Information Acquisition in a War of Attrition," *American Economic Journal: Microeconomics*, 6, 37–78.

Kullback, S. and R. A. Leibler (1951): "On Information and Sufficiency," *The Annals of Mathematical Statistics*, 22, 79–86.

Liese, F. and I. Vajda (1987): *Convex Statistical Distances*, Teubner.

Liu, Q. (2011): "Information Acquisition and Reputation Dynamics," *The Review of Economic Studies*, 78, 1400–1425.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica.*

Matějka, F. and A. McKay (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *The American Economic Review*, 105, 272–298.

Milgrom, P. R. and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

Monderer, D. and L. S. Shapley (1996): "Potential Games," *Games and Economic Behavior*, 14, 124–143.

Morris, S. and H. S. Shin (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *The American Economic Review*, 88, 587–597.

——— (2000): "Rethinking Multiple Equilibria in Macroeconomic Modeling," *NBER Macroeconomics Annual*, 15, 139–191.

——— (2002): "Social Value of Public Information," *The American Economic Review*, 92, 1521–1534.

——— (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics (Proceedings of the Eight World Congress of the Econometric Society)*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, Cambridge University Press, 56–114.

Morris, S., H. S. Shin, and M. Yildiz (2015): "Common Belief Foundations of Global Games," Princeton University - William S. Dietrich II Economic Theory Center, Working Paper 069_2015.

Myatt, D. P. and C. Wallace (2012): "Endogenous Information Acquisition in Coordination Games," *The Review of Economic Studies*, 79, 340–374.

Obstfeld, M. (1996): "Models of Currency Crises with Self-Fulfilling Features," *European Economic Review*, 40, 1037–1047.

Parthasaraty, K. R. (1967): *Probability Measures on Metric Spaces*, Academic Press.

Pavan, A. (2014): "Attention, Coordination, and Bounded Recall," Mimeo.

Persico, N. (2000): "Information Acquisition in Auctions," *Econometrica*, 68, 135–148.

Ravid, D. (2015): "Bargaining with Rational Inattention," Mimeo.

Rényi, A. (1959): "On Measures of Dependence," *Acta Mathematica Hungarica*, 10, 441–451.

Rochet, J.-C. and X. Vives (2004): "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association*, 2, 1116–1147.

Shannon, C. E. (1948): "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 379–423.

Sims, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

Tirole, J. (2015): "Cognitive Games and Cognitive Traps," Mimeo.

Veldkamp, L. (2011): *Information Choice in Macroeconomics and Finance*, Princeton University Press.

Weinstein, J. and M. Yildiz (2007): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75, 365–400.

Wiederholt, M. (2010): "Rational Inattention," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Palgrave Macmillan, online ed.

Yang, M. (2015): "Coordination with Flexible Information Acquisition," *Journal of Economic Theory*, 158, 721–738.