Robust Predictions in Games with Rational Inattention*

Tommaso Denti[†]
NYU Stern

Doron Ravid[‡]
University of Chicago

July 14, 2023

Abstract

We derive robust predictions in games involving flexible information acquisition, also known as *rational inattention* (Sims, 2003). These predictions remain accurate regardless of the exact specification of players' learning abilities. Compared to scenarios where information is predetermined, rational inattention reduces welfare and introduces additional constraints on behavior. We show these constraints generically do not bind; the two knowledge regimes are behaviorally indistinguishable in most environments. Yet, we demonstrate the welfare difference they generate is robust: optimal policy depends on whether one assumes information is given or acquired. We provide the necessary tools for policy analysis in this context.

1. Introduction

The economics discipline has long recognized that information matters for incentives. Many studies have also noted the reverse, namely, that incentives shape information. The theory of rational inattention, initiated by Sims (2003), is a case in point. This theory suggests that economic agents effectively use their limited ability to gather

*We thank Alex Bloedel, Benjamin Brooks, Modibo Camara, Marina Halac, Emir Kamenica, Ilia Krasikov, Elliot Lipnowski, Andrew McClellan, Stephen Morris, Luciano Pomatto, Daniel Rappoport, Phil Reny, and Joe Root for helpful discussions.

[†] td838@nyu.edu

[‡] dravid@uchicago.edu

and process information: faced with uncertainty, they flexibly acquire and use information in an optimal fashion, trading off costs and benefits. The rational inattention approach has been very successful, seeing a broad array of applications.[1]

Despite their success, most rational inattention models come with a caveat: they require one to make non-trivial assumptions on the information acquisition environment, and these assumptions can be challenging to verify. One reason is that information costs and constraints are often affected by factors that are difficult to directly measure, such as time, effort, and cognitive resources. A few tests of rational inattention have been carried out in laboratory settings, and the results have been concerning. For example, Dean and Neligh (forthcoming) provide experimental evidence against the ubiquitous entropy-reduction cost (Matějka and McKay, 2015). Overall, these considerations raise the question of what predictions of rational inattention are robust to the exact specification of agents' learning abilities.

In this paper, we address these concerns by developing a framework for predictions that are valid across all information acquisition technologies that represent rational inattention. Using this framework, we identify the features that distinguish rational inattention from traditional models where information is given. We demonstrate that in most settings, there is no substantive difference in terms of behavioral outcomes, yet a robust distinction emerges when considering welfare consequences.

Our findings carry two key implications. Firstly, they indicate that studies that utilize rational inattention as a behavioral model need to consider and account for the subtle details of the subjects' attention limitations, either implicitly or explicitly; otherwise, the analysis reverts back to standard information economics. Secondly, our paper highlights a significant array of situations where mistakenly assuming information is given, when it is actually acquired, can lead to misleading welfare conclusions. We provide tools for identifying such situations and for conducting robust welfare analysis in economies with rationally inattentive agents.

Our model begins by considering a scenario in which strategic agents, referred to as players, are faced with the task of simultaneously selecting from a set of alternative actions. The utility of these actions is contingent upon an unknown state of nature. This scenario is commonly known as a *base game*. Our objective is to generate predictions regarding the players' behavior and welfare within this base game. Given the

[1]See Maćkowiak, Matějka, and Wiederholt (2023) for a recent review.

inherent uncertainty surrounding the true payoff state, which directly influences the attractiveness of various actions, these predictions rely on two factors: the information that players are capable of acquiring and the associated costs of obtaining that information.

Standard rational inattention models obtain predictions by combining a base game with what we term an *information technology*, which specifies players' learning capabilities regarding the payoff state. In early studies, this technology assumes that information costs are proportional to entropy reduction, and that players can flexibly gather information, provided that it is independent of the other agents' information (e.g., Maćkowiak and Wiederholt, 2009; Yang, 2015). Recent papers allow for different cost functions and potentially correlated signals.[2]

Together, a base game and an information technology define *an information acquisition game*. This game begins with players covertly choosing what costly information to acquire. Subsequently, each player privately uses the acquired knowledge to take an action. Predictions are derived by solving for Nash equilibrium.

We depart from the standard approach, and do not pair the base game with a fixed information technology. Instead, we simultaneously find all the predictions one can obtain with a technology that is consistent with rational inattention. Specifically, we consider information technologies that satisfy three properties. First, no information comes at zero cost. Second, (strictly) more informative signals (in the sense of Blackwell, 1951, 1953) cost (strictly) more. And third, information choice is flexible. These properties are satisfied by virtually all applications of rational inattention. We are agnostic whether signals are independent or correlated across players.

Theorem 1 characterizes all behavioral and welfare implications of rational inattention in a given base game. The behavior of the players is summarized by the joint distribution of the actions they take and the payoff state drawn by nature. We call this distribution the *outcome* of the game. The welfare aspect is determined by the *value* that each player obtains from the game, which corresponds to their payoff net of any costs they incurred for acquiring information.

We show rational inattention can generate any outcome that satisfies two constraints: *obedience* and *separation*. Each constraint comes from a different optimality condition: obedience represents optimal behavior given a fixed signal structure,

[2]See, e.g., Morris and Yang (2022); Ravid, Roesler, and Szentes (2022); Denti (forthcoming); Hebert and La'O (forthcoming)

whereas separation stems from players' incentives to collect information. We follow the literature (Bergemann and Morris, 2013, 2016), and use the term *Bayes correlated equilibrium* (BCE) to refer to obedient outcomes. Thus, Theorem 1 says an outcome can be generated by rational inattention if and only if it is a separated BCE (sBCE).

Theorem 1 also establishes precise bounds on the net payoff that each player can achieve in a given equilibrium outcome. The upper bound corresponds to the scenario in which the player acquires information at essentially no cost, and we refer to this bound as the outcome's *gross value*. The lower bound represents the situation where the player gathers information that is so expensive that they would have been equally satisfied without acquiring any information at all. We name this lower bound the outcome's *uninformed value*. Theorem 1 demonstrates that, given a separated BCE, each player has the potential to achieve any payoff that surpasses their uninformed value but remains below their gross value.

Theorem 1 allows us to study whether rationally inattentive agents exhibit special behavioral properties, and the impact information costs have on welfare. For this purpose, we compare the predictions made by rational inattention to those generated by models where information is an exogenous variable. Under exogenous information, players do not choose what they know; instead, they are endowed with a fixed signal structure. Bergemann and Morris (2013, 2016) show that an outcome can be induced by some signal structure if and only if it is a BCE (i.e., it is obedient). Welfare is given by the outcome's gross value, since no learning costs are paid when information is predetermined.

Whereas the separation constraint can dramatically shrink the BCE set, such situations turn out to be rare. Specifically, we show in Theorem 2 that, for generic preferences, one can approximate every obedient outcome with outcomes that are both obedient and separated. To put it simply, by introducing small perturbations to the payoffs of a base game, we can transform the game into a scenario where nearly all BCEs become separated, even if the original set of non-separated BCEs was extensive. Thus, unless one commits to a highly structured economic setting or makes specialized assumptions on agents' information technologies, rational inattention is behaviorally indistinguishable from exogenous information.

Our analysis also reveals that rational inattention and exogenous information, despite yielding the same behavioral predictions, can have distinct welfare implications. Intuitively, when information is exogenous, the value of information must go to the players. By contrast, in the worst case under rational inattention, players must pay this value to purchase their information. Theorem 3 explores the impact of this intuition on the set of attainable payoffs. The theorem demonstrates that, for generic preferences, rational inattention expands the set of achievable payoffs when compared to exogenous information. This expansion is not trivial, as it often results in a strict inclusion of additional payoffs within the set.

The welfare difference between rational inattention and exogenous information can have concrete repercussions for public policy. We give a proof of concept in Section 6, where we show how to use our results to conduct robust welfare in economies where agents are rationally inattentive. Specifically, we consider a utilitarian social planner who understands the economic environment (i.e., the base game), but does not know the source of players' information (i.e., the information technology). The planner takes a worst-case approach, as is common in robust mechanism design.[3] First, we characterize the set of binary-action symmetric games in which this planner's welfare evaluations depend on whether she assumes players' information is given or acquired. We then apply this characterization to a regime change game, in which several investors choose whether or not to attack a distressed financial institution. We show that the planner may choose to bolster the institution's fundamentals if she thinks information is exogenous, but she should never do so if she believes the investors are rationally inattentive.

Since many settings of interest are non-generic (e.g., auctions), we conclude the paper by studying environments for which the separation constraint binds. We provide a tight characterization, and show that separation is an all-or-nothing refinement of BCE: the sBCE set is either dense or nowhere dense in the set of BCEs. Thus, whenever the separation constraint binds, it dramatically reduces the set of attainable outcomes.

Related literature. Our paper is related to several strands of literature. First, we contribute to the study of robustness in game-theoretic predictions. Especially pertinent is the work that uses the BCE solution concept to obtain robust predictions in games with incomplete information (e.g., Bergemann, Brooks, and Morris, 2017;

[3]See Carroll (2019) for a recent survey of this literature.

Brooks and Du, 2021). In particular, we show the robustness criterion of such studies extend to include endogenously determined information à la rational inattention, as long as the focus is only on the moments of the action-state distribution (e.g., the seller's revenue in an auction), and the separation constraint is not binding. Even when separation binds, our genericity result means such constraint should be taken seriously only if the analyst is confident about the non-generic features of the economic environment. For studies that focus on welfare (e.g., Bergemann, Brooks, and Morris, 2015), our framework provides a road-map for understanding whether their conclusions continue to hold once one accounts for the cost of information.

To best of our knowledge, our paper is the first to conduct a robust welfare analysis under rational inattention. Several studies have considered the effects of costly information acquisition on public policy and efficiency (e.g., Colombo, Femminis, and Pavan, 2014; Ravid, Roesler, and Szentes, 2022; Angeletos and Sastry, 2023; Hebert and La'O, forthcoming). A distinctive feature of our work is that we evaluate welfare in a manner that is independent of the specific details of agents' learning abilities. The tools we develop are portable and can help policy makers when estimating information costs is challenging or infeasible.

More broadly, we contribute to the literature on rational inattention in games (e.g., Yang, 2015; Ravid, 2020; Morris and Yang, 2022; Denti, forthcoming). Within this literature, the closely related work of Denti (2021) is the first to consider the question of robustness. Denti (2021) studies a two-player signaling game where the receiver needs to pay a cost to monitor the sender's action. He shows that, when costs are strictly monotone, off-path beliefs play no role in equilibrium. He then characterizes the behavioral predictions that are consistent with some strictly monotone cost function. As part of this characterization, he obtains what is essentially the single-player version of our Theorem 1. Our Theorem 1 expands on Denti's (2021) result by allowing for multiple players, and by studying welfare (Denti, 2021, does not discuss payoffs). In addition, our analysis of generic games and the comparison to exogenous information have no parallel in Denti (2021).

One can interpret some of our results as providing a test for rational inattention in a strategic settings: simply check whether the observed outcome satisfies obedience and separation. Moreover, our genericity result suggests that testing separation requires a highly structured and controlled environment. Thus, our paper can be seen as adding to the growing literature on the testable implications of rational inattention (e.g., Caplin and Dean, 2015; Caplin, Dean, and Leahy, 2022; Denti, 2022; Lipnowski and Ravid, 2022). Within this literature, the most closely related paper is Caplin and Dean (2015).[4] They develop a test for whether a single agent's choices in multiple menus are consistent with costly information acquisition, but do not require costs to be strictly monotone in information. Their characterization includes obedience, as well as another condition called *no improving attention cycles* (NIAC), which restricts the agent's behavior across decision problems. Since we consider a fixed base game, the NIAC restriction does not apply in our setting. By contrast, Caplin and Dean's (2015) characterization does not require separation, because they allow for costs that are not strictly monotone. Hence, their characterization differs from ours in that it only considers a single agent, does not require costs to be strictly monotone, and accounts for multiple decision problems.

Our work also expands on the uses of correlated equilibrium and its cousins for spanning the set of predictions attainable across various ways of extending a base game (e.g., Aumann, 1974; Myerson, 1982; Forges, 1993; Bergemann and Morris, 2016; Doval and Ely, 2020). An early and closely related paper within this literature is Lipman and Srivastava (1990). They consider base games without payoff uncertainty, and ask which correlated equilibria can be attained by extending the game to allow players to acquire costly information about a common payoff-irrelevant state. They maintain two assumptions on players' information technology: costs are (ordinally) symmetric across players, and players must use partitional information. They show an obedient outcome can be generated from an information technology satisfying their assumptions if and only if it satisfies a cyclical monotonicity condition across players that is similar to Caplin and Dean's (2015) NIAC. Our work differs from theirs in that we allow for payoff uncertainty, asymmetric costs, and non-partitional information, and that we require costs to be strictly increasing in informativeness. In addition, Lipman and Srivastava (1990) have no analog of our payoff bounds.

[4]Another related paper is de Clippel and Rozen (2022). They consider a two-staged game where the first stage player chooses how much to obfuscate the state, and the second stage player chooses what to learn about that state whenever it is obfuscated. In their setting, the second player essentially faces a single agent decision problem. Using this fact, de Clippel and Rozen (2022) apply Caplin and Dean's (2015) results to obtain testable prediction that are valid across all cost functions. They then test these predictions in a lab experiment.

2. Setup

We consider a game involving a finite number of players where the players face uncertain payoffs, call it the *base game*. Before playing this base game, the players can covertly acquire costly information to reduce the uncertainty they face. We call *information technology* the description of learning resources. Together, a base game and an information technology define an *information acquisition game*, our main object of study. We focus on two equilibrium predictions: the induced *outcome* of the base game, and the net *value* each player obtains from information acquisition. In this section, we precisely define all these objects and provide the associated notation.

Base Game. Let I be a finite set of players, with typical element i. Each player i has to choose an action a_i from a finite set A_i. As usual, we define $A_{-i} = \prod_{j \neq i} A_j$ and $A = A_i \times A_{-i}$. Accordingly, we use $a_{-i} = (a_j)_{j \neq i}$ to denote the action profile of all players other than i, and $a = (a_i, a_{-i})$ to denote the entire action profile. Throughout the paper we adopt the same notational conventions for all Cartesian products indexed by $I \setminus \{i\}$ and I.

Players are expected utility maximizers who care about each other's actions as well as an exogenous variable θ, which is drawn from a finite set Θ according to a full-support probability measure $\pi \in \Delta(\Theta)$. We denote by $u_i : A \times \Theta \to \mathbb{R}$ player i's von Neumann-Morgenstern utility function.

Embracing an established terminology (see, e.g., Bergemann and Morris, 2016), we call a tuple $\mathcal{G} = (I, \Theta, \pi, (A_i, u_i)_{i \in I})$ a **base game**.

Information Technologies. Before taking an action in the base game, each player has the opportunity to acquire information about θ as well as other exogenous quantities of potential interests (e.g., sunspots, noisy public information). We succinctly represent them by a single variable z taking values in a finite set Z; a Markov kernel $\zeta : \Theta \to \Delta(Z)$ details the conditional distribution of z given θ.

Following Blackwell (1951), we model the acquisition of information using experiments. An experiment for player i is a Markov kernel $\xi_i : Z \times \Theta \to \Delta(X_i)$, where X_i is a finite space of signal realizations privately observable by player i. The functions ξ_i details how the distribution of i's signal x_i depends on θ and z.

By construction, the players' signals are conditionally independent given θ and

z. However, they may be correlated given θ only. Thus, our framework incorporates a form of correlated information acquisition, as in, among others, Hellwig and Veldkamp (2009), Myatt and Wallace (2012), Hebert and La'O (forthcoming), and Denti (forthcoming). Indeed, one can simply view z as a modeling device for situations in which players have access to information sources with correlated noise (e.g., newspapers with similar slants, consultants from the same firm).

The acquisition of information faces two kinds of frictions. First, each player is constrained in the kind of experiments she can use: player i can only choose experiments that lie in a given set $\mathcal{E}_i$. Second, experiments come at a cost, where $C_i : \mathcal{E}_i \to \mathbb{R}_+$ denotes i's cost function. As a normalization, we assume the existence of an experiment that costs zero; that is, $C_i(\xi_i) = 0$ for some $\xi_i \in \mathcal{E}_i$.

We call a tuple $\mathcal{T} = (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i \in I})$ an **information technology**.

Information Acquisition Games. Together, a base game $\mathcal{G}$ and an information technology $\mathcal{T}$ define an **information acquisition game**. The game begins with the realization of the state of nature $\omega = (z, \theta)$. Without observing the state, the players simultaneously and covertly choose experiments, and pay their costs. Then, each player privately observes the outcome of their own experiment and takes an action. We use $\sigma_i : X_i \to \Delta(A_i)$ to denote player i's action plan in this game, and let Σ_i be the set of i's action plans.

The solution concept we adopt is Nash equilibrium. A strategy for player i consists of an experiment $\xi_i \in \mathcal{E}_i$ and an action plan $\sigma_i \in \Sigma$. A strategy profile $(\xi_i^*, \sigma_i^*)_{i \in I}$ is an equilibrium if for all players i, (ξ_i^*, σ_i^*) maximizes

$$\left[\sum_{a,x,z,\theta} u_i(a,\theta)\sigma_i(a_i|x_i)\xi_i(x_i|z,\theta)\prod_{j\neq i}\sigma_j^*(a_j|x_j)\xi_j^*(x_j|z,\theta)\zeta(z|\theta)\pi(\theta)\right] - C_i(\xi_i).$$

over all $\xi_i \in \mathcal{E}_i$ and $\sigma_i \in \Sigma_i$. The objective function consists of two terms: the benefit of information (in square brackets) and the cost of information. As common in applications, benefit and cost are additively separable.

Equilibrium Predictions. We summarize the equilibria of information acquisition games using two statistics: the induced outcome of the base game, and the net value each player obtains from information acquisition.

The induced **outcome** of the base game is the joint distribution $p \in \Delta(A \times \Theta)$ of the players' actions and the payoff-relevant state. Note that the marginal distribution of θ must be the prior π; we denote by $\Delta_\pi(A \times \Theta)$ the set of probability measures over $A \times \Theta$ whose marginal on Θ is π.

The net **value** player i obtains from information acquisition is simply i's expected payoff, counting both benefits and costs of information. We denote by $v = (v_i)_{i \in I} \in \mathbb{R}^I$ the vector of such values. Since v_i includes player i's information acquisition costs, it may differ from the expectation of u_i under p.

3. Rational Inattention: A Definition

Our aim is to study the behavioral outcomes and welfare values that can be generated in equilibrium of information acquisition games as we fix the base game and range over all information technologies that represent rational inattention. Consistently with the literature, we interpret rational inattention as information technologies where the set of feasible experiments is flexible, and where it is costly to acquire more information.

To give a precise definition of "more information," we build on the classic ranking of experiments due to Blackwell (1951, 1953). Given a pair of experiments ξ_i and ξ_i', we say ξ_i **Blackwell dominates** ξ_i' (denoted $\xi_i \succsim \xi_i'$) if there exists a Markov kernel $g : X_i \to \Delta(X_i)$ such that for every $x_i \in X_i$, $\theta \in \Theta$, and $z \in Z$ with $\zeta(z|\theta) > 0$,

$$\xi_i'(x_i|z, \theta) = \sum_{x_i' \in X_i} g(x_i|x_i')\xi(x_i'|z, \theta).$$

Thus, ξ_i Blackwell dominates ξ_i' if one can generate ξ_i' by garbling the output of ξ_i. As shown by Blackwell, $\xi_i \succsim \xi_i'$ if and only if player i is better off observing the output of ξ_i rather than the output of ξ_i' (holding fixed other players' behavior). In this sense, ξ_i is more informative than ξ_i'. We write $\xi_i \succ \xi_i'$ whenever $\xi_i \succsim \xi_i'$ and $\xi_i' \not\succsim \xi_i$.

We say a set of feasible experiments $\mathcal{E}_i$ is **flexible** if, whenever a given amount of information is feasible, a lower amount of information is feasible as well: whenever $\xi_i \in \mathcal{E}_i$ and $\xi_i \succsim \xi_i'$, then $\xi_i' \in \mathcal{E}_i$. We say a cost function C_i is **monotone** if less informative experiments are cheaper to acquire: for all $\xi_i, \xi_i' \in \mathcal{E}_i$ such that $\xi_i \succsim \xi_i'$ (resp., $\xi_i \succ \xi_i'$), we have $C_i(\xi_i) \geq C_i(\xi_i')$ (resp., $C_i(\xi_i) > C_i(\xi_i')$). We say a technology $\mathcal{T}$ represents **rational inattention** if for every player, the set of feasible experiments is flexible and the cost function is monotone.

Many authors regard flexibility as the key difference between rational inattention and traditional information-acquisition models (see, e.g., Maćkowiak, Matějka, and Wiederholt, 2023, Section 2). Applications of rational inattention often assume all experiments are feasible: in the language of this paper, $\mathcal{E}_i = \Delta(X_i)^{Z\times\Theta}$. Our results are unchanged if we adopt this more extreme definition of flexibility. The reason is that there are no observable differences between experiments that are unfeasible and experiments that are excessively costly.

By pairing flexibility with monotonicity, we postulate that players can save on costs by only acquiring the information they actually use in making decisions. Matějka and McKay (2015) provide a standard example of monotone cost function: $C_i(\xi_i)$ equals the expected reduction in uncertainty about θ and z, as measured by Shannon entropy. More broadly, one could substitute Shannon's entropy with any other strictly concave measure of uncertainty (Caplin, Dean, and Leahy, 2022). One could also consider increasing transformations of these costs (Denti, 2022; Zhong, 2022), or any differentiable cost function whose derivative is strictly convex (Lipnowski and Ravid, 2022).[5]

4. Robust Predictions

This section presents our first main result: a robust characterization of the outcomes and values that can be generated via rational inattention. To provide a benchmark, we first review the case of exogenous information.

The class of information acquisition games includes situations in which the players' information is predetermined. One can obtain them by considering technologies in which every player has only one feasible experiment (whose cost is zero by our normalization). Each such technology can be identified with a tuple $\mathcal{S} = (Z, \zeta, (X_i, \xi_i)_{i\in I})$. Of course, an information acquisition game $(\mathcal{G}, \mathcal{S})$ is just a game of incomplete information à la Harsanyi, where $\mathcal{S}$ is the predetermined information structure.

Among other results, Bergemann and Morris (2016) characterize the equilibrium outcomes that can arise in a game of incomplete information $(\mathcal{G}, \mathcal{S})$ as we fix the base game $\mathcal{G}$ and range over all information structures $\mathcal{S}$: they call them Bayes correlated equilibria. A **Bayes correlated equilibrium** (BCE) of a base game $\mathcal{G}$ is an outcome

[5] In each information acquisition game, prior beliefs are exogenously determined by π and ζ. Thus, the issue of experiment-based vs. posterior-based information costs that sometimes arise in applications of rational inattention (see, e.g., Ravid, 2020; Denti, Marinacci, and Rustichini, 2022) is irrelevant here.

$p \in \Delta_\pi(A \times \Theta)$ such that for all $i \in I$ and $a_i, b_i \in A_i$,

$$\sum_{a_{-i},\theta} \left(u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta)\right) p(a_i, a_{-i}, \theta) \geq 0. \tag{1}$$

Following standard terminology, we name (1) the **obedience constraint**.[6] The standard way of viewing this constraint is through the lens of a mediator who generates p by observing the state and stochastically sending an action recommendation to each player. The players are willing to follow these recommendations if and only if the obedience constraint is satisfied. Even if our framework does not include an actual mediator, throughout the paper we find it helpful to use the mediator's metaphor as expositional device.

Calculating players' values under exogenous information is straightforward. If an outcome p arises under exogenous information in a base game $\mathcal{G}$, player i's expected payoff is

$$\bar{v}_i(p) = \sum_{a,\theta} u_i(a, \theta) p(a, \theta).$$

We call $\bar{v}_i(p)$ the **gross value** for player i, since it ignores information costs. Let $\bar{v}(p) = (\bar{v}_i(p))_{i \in I}$ be the vector of gross values.

What changes when information is endogenous? Our analysis highlights two main differences between rational inattention and exogenous information.

The first difference is that outcomes generated by costly information acquisition must satisfy an additional constraint, which we name the *separation constraint*. To present this constraint, we require a few definitions. Given an outcome $p \in \Delta_\pi(A \times \Theta)$, a player $i \in I$, and an action $a_i \in A_i$, let $p(a_i) = \sum_{a_{-i},\theta} p(a_i, a_{-i}, \theta)$ be the probability of player i taking action a_i under p, and let

$$\text{supp}_i(p) = \{a_i \in A_i : p(a_i) > 0\}$$

be the set of i's actions that have positive probability. For each $a_i \in \text{supp}_i(p)$, let $p_{a_i} \in \Delta(A_{-i} \times \Theta)$ be the conditional distribution of the actions of the players other

[6] Our definition of BCE corresponds to the specialization of Bergemann and Morris's (2016) definition to the case in which players' original type spaces are degenerate. Forges (1993) refers to such outcomes as *universal Bayesian solutions*.

than i and the payoff-relevant state: for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$p_{a_i}(a_{-i}, \theta) = \frac{p(a_i, a_{-i}, \theta)}{p(a_i)}.$$

We also denote the set of player i's best responses to p_{a_i} via

$$BR(p_{a_i}) = \underset{b_i \in A_i}{\operatorname{argmax}} \sum_{a_{-i}, \theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

Finally, we say an outcome $p \in \Delta_\pi(A \times \Theta)$ satisfies the **separation constraint** if for all $i \in I$ and $a_i, b_i \in \operatorname{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \text{implies} \quad BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing. \tag{2}$$

In other terms, an outcome satisfies the separation constraint if distinct beliefs have separate best responses. We refer to a BCE that satisfies the separation constraint as a **separated BCE** (sBCE).

The second difference between rational inattention and exogenous information is in the set of payoffs players obtain from a given outcome. When information is given, each player i's value is completely determined by the outcome p. By contrast, under rational inattention the same outcome can arise under multiple cost functions, and so is consistent with a set of values. As we demonstrate next, this set of values is convex, with an upper bound given by the gross value $\bar{v}_i(p)$. The lower bound is given by (what we call) the outcome's *uninformed value*, which is i's maximal value if she has to act before receiving the mediator's recommendation. Formally, the **uninformed value** of an outcome p in a base game $\mathcal{G}$ for player i is the quantity

$$\underline{v}_i(p) = \max_{b_i \in A_i} \sum_{a, \theta} u_i(b_i, a_{-i}, \theta) p(a, \theta).$$

Let $\underline{v}(p) = (\underline{v}_i(p))_{i \in I}$ be the vector of uninformed values.[7]

The next result summarizes our characterization of rational inattention in games:

[7] The uninformed value is related to the notion of *coarse correlated equilibrium* (see, e.g., Roughgarden, 2016, Definition 13.5). Coarse correlated equilibrium relaxes the obedience constraint, requiring instead that no player can benefit from deviating to a single action from all of the mediator's recommendation; that is, $\bar{v}_i(p) \geq \underline{v}_i(p)$ for every i.

Theorem 1. *Fix a base game $\mathcal{G}$. A rational-inattention technology $\mathcal{T}$ exists that induces the outcome-value pair (p, v) in an equilibrium of $(\mathcal{G}, \mathcal{T})$ if and only if*

(i) the outcome p is a separated BCE, and

(ii) for every $i \in I$, $v_i = \underline{v}_i(p) = \bar{v}_i(p)$ or $v_i \in [\underline{v}_i(p), \bar{v}_i(p))$.

Thus, an outcome-value pair is consistent with rational inattention if and only if two conditions hold. First, the outcome satisfies both the obedience constraint and the separation constraint. And second, each player's value is weakly above her uninformed value, but strictly below her gross value.

We now explain why the theorem's conditions are necessary. That every outcome generated by rational inattention must be obedient follows from the necessity of the obedience constraint under exogenous information. The reason is that any outcome one can attain when players choose their information must also be attainable if players where exogenously endowed with their chosen experiment. To understand why the separation constraint is necessary, consider a player i who takes with positive probability a pair of actions a_i and b_i such that $p_{a_i} \neq p_{b_i}$. As in BCE, we can interpret a_i and b_i as signals. With rational inattention, informative signals are costly. To save on information costs, the player could substitute a_i and b_i with a single action recommendation c_i. For this substitution not to be profitable, it must be that either $c_i \notin BR(p_{a_i})$ or $c_i \notin BR(p_{b_i})$. Since the choice of c_i is arbitrary, it must be that $BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing$.

Next, we explain the necessity of Theorem 1's payoff bounds. To understand the lower bound, suppose we have an information technology and an equilibrium that induces an outcome p. By assumption, player i always has the option of remaining uninformed at no cost. Therefore, i's optimal payoff must be higher than $\underline{v}_i(p)$. As for the upper bound, recall that information costs are non-negative. Consequently, player i's payoff from an equilibrium that induces p must be below her payoffs if her information was free; that is, $v_i \leq \bar{v}_i(p)$. Moreover, her information must come at a strictly positive cost whenever $\underline{v}_i(p) \neq \bar{v}_i(p)$: to generate p in this case player i must acquire *some* information. Hence, player i can actually attain her gross value from an outcome only if that value coincides with the uninformed value.

We now briefly review our proof that conditions (i) and (ii) of Theorem 1 are sufficient. The proof is constructive, and is based on a result by Denti (2021). Denti (2021) studies a two-player signaling game where the receiver pays a cost to monitor the sender's action. Among other results, Denti (2021) shows that, when the sender takes every action with positive probability, any receiver behavior satisfying (what we call in this paper) obedience and separation can be justified via rational inattention. To prove the "if" direction of Theorem 1, we extend Denti's (2021) single-agent construction so that it applies to settings in which multiple agents simultaneously acquire information. In addition, we show that any payoff between the uninformed and the gross value can be generated in equilibrium—Denti (2021) focuses on equilibrium outcomes and does not discuss achievable payoffs.

Theorem 1 shows that any BCE that arises from rational inattention must satisfy the separation constraint. Next, we record that the separation constraint never eliminates all of a game's Bayes correlated equilibria; that is, the set of separated BCEs is non-empty.

Corollary 1. *Every base game $\mathcal{G}$ admits a separated BCE.*

The proof is straightforward (details omitted): A technology where all feasible experiments are free and uninformative is flexible and monotone. The corresponding game with information acquisition admits an equilibrium by standard arguments (information is de facto exogenous). It follows from Theorem 1 that the outcome of such an equilibrium is a separated BCE.[8]

Next, to help the reader familiarize themselves with the concepts we introduced in this section, we show Theorem 1's implications in a simple coordination game.

Example 1. There are two players, $I = \{1, 2\}$, each of which chooses a binary action. The players get a payoff of 1 if their actions match, -1 otherwise. Specifically, $A_i = \{-1, 1\}$ and $u_i(a_1, a_2) = a_1 a_2$ for each $i \in I$. The payoff state is degenerate, hence omitted.

We first give an example of a non-separated BCE. Let $p \in \Delta(A)$ be the outcome such that assigns $(a_1, a_2) = (1, 1)$ a probability of $1/2$; each other pair of actions occurs with probability $1/6$. It is easy to see p is obedient: conditional on any action recommendation, the probability the other player takes the same action is at least $1/2$. However, this BCE is not separated: $p_{a_1=1}(a_2 = 1) = 3/4 \neq 1/2 = p_{a_1=-1}(a_2 = 1)$, but $a_1 = 1$ is a best response to both action recommendations. Thus,

[8] Another immediate consequence of this argument is that in the special case in which Θ is a singleton, all (pure or mixed) Nash equilibria of the base game are separated BCEs.

the recommendations $a_1 = 1$ and $a_1 = -1$ have a common best response, but lead to different beliefs.

For an example of a separated BCE, take $q \in \Delta(A)$ to be the uniform distribution over the two action profiles $(a_1, a_2) = (1, 1)$ and $(a_1, a_2) = (-1, -1)$. Because the players are perfectly coordinating their actions, taking the action that the mediator recommends is a strict best response. It follows q is a separated BCE.

We now demonstrate the difference between gross and uninformed value. Consider the sBCE q. Since the players perfectly coordinate their actions, the gross value from this BCE is $\bar{v}_1(q) = \bar{v}_2(q) = 1$. The uninformed value, however, is lower: a player who does not see the mediator's recommendation gets a payoff of 0 no matter what action she takes, given that the co-player takes both actions with equal probability. Therefore, $\underline{v}_1(q) = \underline{v}_2(q) = 0$.

We conclude the section with a brief discussion of what happens if we allow for non-flexible and non-monotone technologies. To accommodate such technologies, we need to adjust Theorem 1's statement in two ways. First, the separation constraint is no longer necessary. And second, players can attain their gross value even when it differs from their uninformed value. We refer the reader to Online Appendix D for the precise details.

5. Generic Environments

Theorem 1 shows rational inattention differs from exogenous information in two ways. First, rational inattention reduces the set of equilibrium outcomes: an additional separation constraint must be satisfied. Second, rational inattention expands the set of achievable payoffs for any given equilibrium outcome: every payoff between the uniformed value and the gross value can be obtained. In this section, we prove that only the second difference is meaningful in generic environments.

We adopt the following notion of genericity: We fix a finite set of players I, a finite set of payoff states Θ, a full-support prior $\pi \in \Delta(\Theta)$, and a finite set of actions A_i for each player $i \in I$. To obtain a base game, it remains to specify a profile of utility function $u = (u_i)_{i \in I}$. We identify u with an element of the Euclidean space $\mathbb{R}^{I \times A \times \Theta}$, and say a statement is true for **generic** u if the closure of the subset in $\mathbb{R}^{I \times A \times \Theta}$ for which it is false has Lebesgue measure zero. We denote by $BCE(u)$ and $sBCE(u)$ the sets of BCEs and separated BCEs for the base game corresponding to u.

As long as the sets Θ, I, and A are non-trivial, one can find many examples of $u \in \mathbb{R}^{I \times A \times \Theta}$ for which the separation constraint imposes sharp restrictions on behavior, so that most BCE are not separated—we discuss a few simple cases in Online Appendix E. However, the next theorem shows that all such examples are non-generic.

Theorem 2. *For generic u, the set $sBCE(u)$ is dense in the set $BCE(u)$.*

Thus, the environments in which rational inattention predict different outcomes than incomplete information are knife edge.[9] An important caveat to this result is the notion of genericity we use: it is the most common for the static games we study in this paper, but also the most permissive. For example, according to this notion of genericity, many important economic applications—such as auctions—are non-generic. The notion of genericity is also not appropriate if the base game is not actually static but represents the strategic form of a primitive dynamic game. We discuss sBCE in non-generic environments in Section 7.

One might be tempted to think that Theorem 2 follows from indifferences being fragile, so to speak. This intuition works, but only for the single agent case. When I has one element, the set of *strict* BCEs—i.e., the set of outcomes p such that $BR(p_{a_i}) = \{a_i\}$ for every player i and $a_i \in \text{supp}_i(p)$—is generically dense in the BCE set. Theorem 2 then follows from noting that every strict BCE is separated.[10]

The situation radically changes when there are at least two players. The reason is that, with two or more players, indifferences emerge in equilibrium in generic fashion: when I has more than one element, there exists an open set of games where *no* BCE is strict. For example, consider games in a neighborhood of Matching Pennies: they have a unique correlated equilibrium (the fully-mixed Nash equilibrium) in which both players are indifferent between both actions. In sum, to prove Theorem 2 beyond the single-agent case, one cannot hope to show that indifferences are somewhat fragile.

Our proof of Theorem 2 combines two independent lemmas. The first lemma shows that for any BCE p and utility profile u, there exists a perturbation of u that makes p separated. In other words, every BCE of a given game is the limit of separated BCEs of nearby games.

[9]It is easy to construct generic examples where the set of separated BCEs is not close. In particular, the statement "$sBCE(u) = BCE(u)$ for generic u" is false. It is also not true that "$\text{cl}(sBCE(u)) = BCE(u)$ for all u."

[10]See Online Appendix F for a detailed argument.

The second lemma shows that for generic games, any BCE that is a limit of separated BCE in nearby games is also a separated BCE when the game is held fixed. Specifically, the second lemma shows that the correspondence $u \mapsto \mathrm{cl}(sBCE(u))$, which takes utilities to the closure of the sBCE set, generically is upper hemicontinuous (in fact, continuous). To prove it, we employ ideas from Blume and Zame (1994), who study the algebraic geometry of Nash, perfect, and sequential equilibria. In particular, we use the Tarsky-Siedenberg Theorem to show that $u \mapsto \mathrm{cl}(sBCE(u))$ has semi-algebraic graph, and so must be continuous for all utility profiles outside a closed low-dimensional manifold.

Theorem 2 implies that, in generic games, rational inattention and exogenous information are outcome equivalent. Next, we show this equivalence does not extend to players' welfare. In other words, even though the two knowledge regimes generically yield the same behavioral predictions, they can have distinct welfare implications.

For $u \in \mathbb{R}^{I \times A \times \Theta}$, let $V_R(u)$ be the closure of the set of attainable value vectors under rational inattention. By Theorem 1, these are the value vectors that lie between the uninformed value and the gross value of some limit of separated BCEs:[11]

$$V_R(u) = \{v \in [\underline{v}(p,u), \bar{v}(p,u)] : p \in \mathrm{cl}(sBCE(u))\}, \tag{3}$$

where $\underline{v}(p,u)$ and $\bar{v}(p,u)$ make explicit the dependence of uninformed and gross values on the players' utility functions.

We also denote by $V_I(u)$ the set of attainable value vectors under exogenous information. By Bergemann and Morris (2016), this is the set of gross value vectors attainable in some BCE,

$$V_I(u) = \{\bar{v}(p,u) : p \in BCE(u)\}. \tag{4}$$

For an arbitrary u, there is no simple relationship between $V_R(u)$ and $V_I(u)$: $sBCE(u)$ is a subset of $BCE(u)$, but $\bar{v}(p,u)$ is an element of $[\underline{v}(p,u), \bar{v}(p,u)]$, so one cannot easily conclude that $V_R(u)$ contains or is contained by $V_I(u)$, or neither.

In generic environments, the comparison is simpler:

[11]For two vectors of real numbers $\lambda, \gamma \in \mathbb{R}^n$ such that $\lambda \leq \gamma$, we use the box, or ordered interval notation $[\lambda, \gamma] = \{\eta \in \mathbb{R}^n : \lambda \leq \eta \leq \gamma\}$ to denote the set of all vectors between λ and γ.

Theorem 3. *For generic u, $V_I(u) \subseteq V_R(u)$. In addition, if $|I| \geq 2$, $|\Theta| \geq 2$, and $|A_i| \geq 2$ for at least two distinct players i, then the set of u for which $V_I(u) \subset V_R(u)$ has non-empty interior. If instead $|I| = 1$, then $V_I(u) = V_R(u)$ for generic u.*

Thus, in generic environments, rational inattention expands the set of achievable payoffs, and it does so in a non-trivial way (i.e., with strict inclusion) for a class of environments of positive measure. As we demonstrate in Section 6, the difference between $V_I(u)$ and $V_R(u)$ has substantial implications for welfare analysis: if a utilitarian social planner mistakenly assumes that information is given rather than acquired, she may end up choosing a sub-optimal policy, even in situations in which the two knowledge regimes are outcome equivalent. Our results give the planner the tools to avoid such mistakes.

Theorem 3 distinguishes between single-agent and many-player settings. In generic single-agent settings, rational inattention and exogenous information are not only outcome equivalent (as Theorem 2 states), but also welfare equivalent: $BCE(u) = sBCE(u)$ and $V_I(u) = V_R(u)$ for generic u. In generic many-player environments, rational inattention and exogenous information are only outcome equivalent: $BCE(u) = sBCE(u)$ for generic u, but $V_I(u) \subset V_R(u)$ for a set of utility profiles u with non-empty interior.

We now clarify the difference between the case of a single agent and the case of multiple players. Suppose first there is only one agent, and let p^0 be a BCE in which the agent receives no information. Since the agent's payoffs depend only on the correlation between her action and the state, and because the state's marginal distribution is constant across all BCEs, the agent's uninformed value is also constant across all BCEs, and equals her expected utility under no information; that is, every BCE p has $\underline{v}(p, u) = \bar{v}(p^0, u)$. Moreover, since more information always helps the agent, the BCE p^0 minimizes the agent's gross value across all BCEs. In other words, the minimal value of $V_R(u)$ and $V_I(u)$ is the same, and equals $\bar{v}(p^0, u)$. The single-agent result then follows from noting that, for generic u, $V_R(u)$ and $V_I(u)$ are convex sets that have the same maximum value (i.e., complete information).

Suppose now there are multiple players. In this case, each player cares not only about her action and the state, but also about the actions of others. Consequently, it is possible that to minimize one player's uninformed value, a BCE must change the distribution of another player's actions. If such a change requires giving players valuable information, one can obtain a gap between the minimal uninformed value and the lowest gross value for a given player. The proof of Theorem 3 constructs a game with this property, and shows that, for this game, this property is preserved under small perturbations.

6. Application: Robust Welfare Analysis

In this section we demonstrate how one can use our results to conduct robust welfare analysis in an economy where agents exhibit rational inattention. The proofs for the results in this section are in the online appendix.

We consider an economy that consists of a fixed set of agents, I, who play an information acquisition game, $(\mathcal{G}, \mathcal{T})$. The structure of the game depends on the policy enacted by a utilitarian social planner. The planner has a good understanding of the policy's material implications, i.e., she knows a given policy leads to a given $\mathcal{G}$. However, the planner is unsure about the accompanying information technology $\mathcal{T}$.

We focus on two cases regarding the source of the agents' information. The planner either postulates that information is exogenously given, or that it is generated by rational inattention. In both cases, the planner identifies a policy with the corresponding base game $\mathcal{G}$ and employs a robust criterion that evaluates it according to the worst-case utilitarian welfare across all relevant information technologies $\mathcal{T}$ and ensuing equilibria.

For the exogenous information case, Bergemann and Morris's (2016) results imply one can find the social value of a policy $\mathcal{G}$ by minimizing the sum of the agents' gross payoffs across all BCEs. Specifically, let $\bar{w}(p)$ be the utilitarian welfare of an outcome p assuming the players' payoffs are given by their gross value,

$$\bar{w}(p) = \sum_i \bar{v}_i(p).$$

Then the planner's value of a base game $\mathcal{G}$ under exogenous information is

$$\bar{w} = \min_{p \in BCE} \bar{w}(p),$$

where BCE is the set of Bayes correlated equilibria of $\mathcal{G}$. Since the BCE set is defined by linear inequalities, and $\bar{w}(p)$ is linear in p, one can compute $\bar{w}$ via linear programming.

What about a planner who postulates that agents are rationally inattentive? Theorem 1 provides the answer: such a planner evaluates each base game according to the lowest sum of *uninformed* values that is attainable in some *separated* BCE. More precisely, for an outcome p, let $\underline{w}(p)$ be the utilitarian welfare implied by p in the base game $\mathcal{G}$ if players' payoffs are given by their uninformed value,

$$\underline{w}(p) = \sum_i \underline{v}_i(p).$$

Then Theorem 1 suggests that a planner who assumes information is endogenous would evaluate each game according to the minimum of $\underline{w}(p)$ across all separated BCEs p,

$$\inf_{p \in sBCE} \underline{w}(p).$$

Recall, however, that Theorem 2 says that the separation constraint does not bind for generic games. As such, imposing the separation constraint is only appropriate if the planner is absolutely certain of the structure of the base game. Whereas such certainty might be justifiable in certain cases, here we take the perspective of a cautious planner who, in an economy where agents are rationally inattentive, evaluates $\mathcal{G}$ according to the worst-case value of $\underline{w}(p)$ across all BCE,

$$\underline{w} = \min_{p \in BCE} \underline{w}(p).$$

Since the BCE set is defined by linear inequalities, and $\underline{w}(p)$ is convex in p, one can compute $\underline{w}$ using convex programming.

In the rest of the section we study when the optimal policy under rational inattention differs from the best policy under exogenous information. That is, we are interested in situations where there are two policies $\mathcal{G}$ and $\mathcal{G}'$ such that $\mathcal{G}$ is preferred to $\mathcal{G}'$ under the hypothesis that information is exogenous, $\bar{w} > \bar{w}'$, while $\mathcal{G}'$ is preferred to $\mathcal{G}$ if the planner believes that rational inattention is a relevant feature of the economy, $\underline{w}' > \underline{w}$.

Clearly, the two knowledge regimes can generate different policy prescriptions only if one of the policies leads to a different worst-case value under rational inattention than it does under exogenous information. Whereas Theorem 3 guarantees the existence of a generic $\mathcal{G}$ for which $\bar{w} > \underline{w}$, it does not tell us when such an inequality occurs for a fixed base game. Next, we zoom in on symmetric binary-action games to provide a characterization of the full class of policies $\mathcal{G}$ for which $\bar{w} > \underline{w}$.

A base game $\mathcal{G} = (I, \Theta, \pi, (A_i, u_i)_{i \in I})$ has **binary actions** if for every player i, A_i contains two elements. It is **symmetric** if $A_i = A_j$ for all $i, j \in I$, and if for every permutation $\phi : I \to I$, player i, action profile a, and payoff state θ,

$$u_i(a_\phi, \theta) = u_{\phi(i)}(a, \theta),$$

where $a_\phi = (a_{\phi(j)})_{j \in I}$ is the action profile such that each player j takes action $a_{\phi(j)}$. An outcome $p \in \Delta(A \times \Theta)$ of a symmetric base game is **symmetric** if $p(a, \theta) = p(a_\phi, \theta)$ for every permutation $\phi : I \to I$, action profile a, and payoff state θ. We denote the set of symmetric outcomes by $\Delta_\pi^{sy}(A \times \Theta)$, and the set of symmetric BCEs by BCE^{sy}.

The next result characterizes the binary-action symmetric games for which rational inattention yields strictly lower worst-case welfare than exogenous information.

Proposition 1. *Let $\mathcal{G}$ be a symmetric, binary-action base game. Then, $\underline{w} < \bar{w}$ if and only if all $p^* \in \operatorname{argmin}_{p \in \Delta_\pi^{sy}(A \times \Theta)} \underline{w}(p)$ satisfy the following condition:*

$$a_i \in \operatorname{supp}_i(p^*) \text{ and } BR(p^*_{a_i}) = \{a_i\} \qquad \text{for all } i \in I \text{ and } a_i \in A_i. \tag{5}$$

Thus, one can check whether $\underline{w} < \bar{w}$ by examining the minimizers of $\underline{w}(p)$ among all symmetric outcomes p, ignoring the players' obedience constraints. In particular, one needs to check whether all these minimizers recommend both actions to every player, and only send recommendations that induce unique best responses.[12]

An immediate corollary of Proposition 1 is that in a symmetric binary-action game, if $\underline{w} < \bar{w}$ then

$$\underline{w} = \min_{p \in \Delta_\pi^{sy}(A \times \Theta)} \underline{w}(p).$$

Consistently with Theorem 3, Proposition 1 also implies that $\underline{w} = \bar{w}$ if the base game has one player only, and such player has a binary action.[13]

[12]We note that, while symmetric games are non-generic, the strict inequality identified by Proposition 1 is robust. Formally, whenever the inequality $\underline{w} < \bar{w}$ holds for some binary-action symmetric game $\mathcal{G}$, the inequality also holds for all base games $\mathcal{G}'$ in a neighborhood of $\mathcal{G}$. The result follows from the same logic we use in the proof of Lemma 7. As in Section 5, the distance between $\mathcal{G}$ and $\mathcal{G}'$ is the distance between the profiles of utility functions, keeping fixed I, Θ, π, and $(A_i)_{i \in I}$.

[13]For an explanation, suppose $I = \{i\}$ and $A_i = \{a_i, b_i\}$. Take $p^* \in BCE$ such that agent i have

We now briefly explain the proposition's proof. The key observation is that in a binary action game, a BCE p has the property that $\bar{v}_i(p) > \underline{v}_i(p)$ for every player i if and only if p satisfies (5). To get intuition for the "if" direction, note that having two recommendations that lead to strict best responses means players get a strictly positive benefit from following them. This benefit creates a wedge between the gross value $\bar{v}_i(p)$, which accounts for the value of information, and the uninformed value $\underline{v}_i(p)$, which does not. For the converse direction, note that a violation of (5) means some player i has an action that is optimal across all of the mediator's recommendations. As such, player i loses nothing by ignoring the mediator's recommendations and taking that action. In other words, player i's gross value equals her uninformed value.

Armed with the above observation, we prove the proposition in two steps. The first step shows $\underline{w} < \bar{w}$ holds if and only if all optimal solutions of $\min_{p \in BCE^{sy}} \underline{w}(p)$ satisfy (5). This step follows from applying the above-mentioned observation to symmetric outcomes, and showing such outcomes are sufficient for minimizing utilitarian welfare. The second step shows that all optimal solutions of $\min_{p \in BCE^{sy}} \underline{w}(p)$ satisfy (5) if and only if all optimal solutions of $\min_{p \in \Delta^{sy}_{\pi}(A \times \Theta)} \underline{w}(p)$ satisfy (5). Loosely speaking, the reason is as follows: if the obedience constraint does not bind at the optimum—as (5) dictates—then it can be relaxed, and therefore minimizing over $p \in BCE^{sy}$ is the same as minimizing over $p \in \Delta^{sy}_{\pi}(A \times \Theta)$.

Proposition 1 enables us to find circumstances where the planner's optimal policy depends on whether she believes information is exogenously given, or generated by rational inattention. We demonstrate this fact below in a regime change game.

Example 2. We consider a regime change game whereby a status quo is abandoned if a sufficiently large number of players take an action against it. Such games are well-studied and have been used to model a variety of social phenomena, including currency crises, bank runs, debt crises, and political revolts.[14]

In our application, there are n identical investors, $i \in I = \{1, \ldots, n\}$, each of which decides whether to speculate against (i.e., attack) a distressed financial institution ($a_i = 1$) or not ($a_i = 0$). Speculating costs $k \in (0, 1)$. If enough investors speculate, the institution fails (i.e., the attack succeeds), generating a profit of 1 to

no information and does not randomize between a_i and b_i. Then, $p^* \in \operatorname{argmin}_{p \in \Delta_{\pi}(A \times \Theta)} \underline{v}_i(p)$ and $\operatorname{supp}_i(p^*)$ is a singleton, violating (5).

[14]See Morris and Shin (2003) for a review.

the speculators, and an externality of $-x$ to all passive investors, where $x \in (0, \infty)$. The payoff state $\theta \in \Theta \subseteq \{1, \ldots, n\}$, determines the number of speculators required for the attack to succeed. We assume that $\min \Theta > 1$, and $\max \Theta < n - 1$, meaning no single investor can go against the will of all the others.[15] We summarize these payoffs below:

	$\sum_j a_j \geq \theta$	$\sum_j a_j < \theta$
$u_i(1, a_{-i}, \theta)$	$1 - k$	$-k$
$u_i(0, a_{-i}, \theta)$	$-x$	0

It is easy to verify that, in this example, rational inattention and exogenous information are outcome equivalent, that is, the sBCE set is dense in the BCE set.[16]

Next we demonstrate that a planner who views information as exogenous may adopt different policies than a planner who thinks investors are rationally inattentive. In particular, we show that only a planner who takes information as given has an incentive to change the institution's fundamentals, that is, Θ and π.

We begin by finding conditions under which there is a difference between the worst-case welfare under rational inattention and exogenous information. By Proposition 1, answering this question requires us to minimize the sum of the uninformed values across all symmetric outcomes, ignoring obedience constraints. The following result characterizes the solutions to this optimization problem.

Claim 1. In the regime change game, an outcome p^* minimizes $\underline{w}(p)$ over all $p \in \Delta_\pi^{sy}(A \times \Theta)$ if and only if the following conditions hold: for all payoff states θ,

$$p^* \left(\left\{ (a, \theta) : \theta - 1 \leq \sum_i a_i \leq \theta \right\} \right) = 0, \tag{6}$$

$$p^* \left(\left\{ (a, \theta) : \sum_i a_i > \theta \right\} \right) = \frac{k}{1 + x}. \tag{7}$$

The two optimality conditions have simple interpretations. First, no investor is ever pivotal. And second, the attack succeeds with probability $k/(1 + x)$. To obtain these conditions, we show p^* minimizes $\underline{w}(p)$ across all $p \in \Delta_\pi^{sy}(A \times \Theta)$ only if each

[15]A fortiori, $n > 3$ because $2 \leq \min \Theta \leq \max \Theta < n - 1$.

[16]Indeed, the regime change game admits a strict BCE p where all players take both actions with positive probability (e.g., a convex combinations of the pure Nash equilibria in which everyone speculates or no-one speculates). Thus, any BCE q can be approximated by a separated BCE $(1 - \epsilon)q + \epsilon p$ as $\epsilon \to 0$.

individual investor i is indifferent between never speculating and always speculating when other investors play according to p^*. Minimizing across all symmetric outcomes satisfying this indifference condition then delivers the result.[17]

Combining Proposition 1 and Claim 1, we deduce that $\underline{w} < \bar{w}$ if and only if (5) holds for every symmetric outcome p^* that satisfy (6) and (7). It is easy to find *one* p^* that meets all requirements: have all investors attack together with probability $k/(1+x)$ regardless of the state, and no one speculates otherwise. Calculating the sum of the uninformed values from such p^* immediately give the worst-case welfare under rational inattention,

$$\underline{w} = \underline{w}(p^*) = -\frac{nxk}{1+x}. \tag{8}$$

Hence, under rational inattention, the planner's value decreases in the size of the externality x and the cost of betting on the financial institution's demise k, but does not depend on π; that is, the planner's value does not depend on the institution's fundamentals.

Is there a symmetric outcome p^* that satisfies (6) and (7), but violates (5)? That is, under what conditions on the primitives of game exogenous information and rational inattention generate the same worst-case welfare? The next result provides an answer when the payoff state is binary (see Claim 15 in the online appendix for a many-state generalization).

Claim 2. Suppose $\Theta = \{\underline{\theta}, \bar{\theta}\}$, where $\bar{\theta} \geq \underline{\theta}$. Then $\underline{w} = \bar{w}$ if and only if $\bar{\theta} - \underline{\theta} \geq 3$ and

$$1 - \frac{1}{3}\left(\bar{\theta} - \underline{\theta}\right)\pi(\bar{\theta}) \leq \frac{k}{1+x} \leq \frac{1}{3}\left(\bar{\theta} - \underline{\theta}\right)\left(1 - \pi(\bar{\theta})\right). \tag{9}$$

In other terms, with binary θ, worst-case welfare under rational inattention is the same as under exogenous information if and only if the state is sufficiently uncertain in the sense that $\bar{\theta} - \underline{\theta} \geq 3$, and the probability of $\bar{\theta}$ is not too extreme compared to $k/(1+x)$—e.g., (9) fails if $\pi(\bar{\theta})$ goes to zero or one. In particular, the worst-case welfare under rational inattention is always strictly below the welfare under exogenous information when there's certainty about the institution's fundamentals (i.e., when θ is deterministic).

The proof of Claim 2 is rather detailed; here we provide only a rough intuition. Combining Proposition 1 and Claim 1, one can show the worst-case welfare under ex-

[17] Given the minmax nature of the optimization problem $\min_{p \in \Delta_{\pi}^{sy}(A \times \Theta)} \underline{w}(p)$, one might have conjectured the existence of an indifference condition.

ogenous information coincides with the worst-case welfare under rational inattention only if a symmetric outcome exists that satisfies two conditions. First, an investor who does not see the mediator's recommendation is indifferent between attacking and not attacking. And second, for an investor who does see the mediator's recommendation, either not speculating is a best response to a "speculate" recommendation, or vice versa. Appealing to Bayes rule, one can show that an outcome satisfies these conditions if and only if there is limited overlap between the event where many investors are attacking and the event in which the attack succeeds. Claim 2 follows from showing a sufficient disconnect between these two events is attainable if and only if there is enough uncertainty about θ. Intuitively, disconnecting the two events is easy when θ obtains both high and low values with large probability: one can have the number of speculators just come short of a successful attack when θ is high, and come just above the threshold when θ is low. The same cannot be done when θ is deterministic. In that case, successful attacks necessarily involve more speculating investors than failed ones.

A takeaway is that, unlike under rational inattention, a planner who views information as exogenous may adopt policies that change the institution's fundamentals. For a concrete illustration, consider two policies $\mathcal{G}$ and $\mathcal{G}'$ that differ only in the institution's fundamentals, that is, in the set of states Θ and their distribution $\pi \in \Delta(\Theta)$. Suppose $\mathcal{G}$ satisfies the conditions of Claim 2, but $\mathcal{G}'$ does not. Worst-case welfare under rational inattention is the same for $\mathcal{G}$ and $\mathcal{G}'$: $\underline{w} = \underline{w}'$ by (8). By contrast, $\mathcal{G}$ generates lower welfare under exogenous information: $\bar{w} < \bar{w}'$ by (8) and Claim 2. Consequently, a planner who views information as being exogenous would pay some amount to change the institution's fundamentals, but a planner who believes investors are rationally inattentive should not do so.

7. Further results: Non-Generic Environments

Throughout the paper, we have focused on generic settings. However, many economic institutions of interests, such as auctions, are non-generic. We conclude our work with a discussion of rational inattention in non-generic environments. We characterize the (non-generic) settings in which the separation constraint has substantive impact, and show its force has an all-or-nothing flavor: if the sBCE set is not dense in the BCE set, it is in fact nowhere dense. The proofs are in the online appendix.

We begin by characterizing the environments where the separation constraint has bite. Our characterization is based on two definitions. The first definition represents a situation where, whenever the mediator recommends player i action b_i in a BCE of the game, the player would be equally happy to take action a_i. Myerson (1997) calls this scenario "jeopardization": action a_i **jeopardizes** action b_i if, for every BCE p such that $b_i \in \text{supp}_i(p)$, $a_i \in BR(p_{b_i})$.[18] We denote by $J(b_i)$ the set of actions that jeopardizes b_i.

Every action jeopardizes itself by the obedience constraint; hence, $J(b_i)$ is not empty. A sufficient condition for jeopardization is weak domination: if $u_i(a_i, a_{-i}, \theta) \geq u_i(b_i, a_{-i}, \theta)$ for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, then a_i jeopardizes b_i. But the concept of jeopardization is broader than weak domination. For example, in Matching Pennies, Heads and Tails jeopardize each other, even if neither action is weakly dominant.

Next we introduce the second definition required for our characterization. This definition concerns a class of BCEs where the number of actions that induce different beliefs is minimal. To state this definition, say a BCE p is **maximally supported** if the support of every other BCE is contained by the support of p. A maximally-supported BCE p is **minimally mixed** if $q_{a_i} \neq q_{b_i}$ implies $p_{a_i} \neq p_{b_i}$ for every BCE q, $i \in I$, and $a_i, b_i \in \text{supp}_i(q)$.

For an interpretation of minimal mixing, consider a mediator who wants to implement a BCE p. When $p_{a_i} = p_{b_i}$, the mediator can replace the distinct recommendations of playing a_i and b_i with a single recommendation of mixing between the two actions with probabilities $p(a_i)/(p(a_i) + p(b_i))$ and $p(b_i)/(p(a_i) + p(b_i))$. A BCE p is minimally mixed if a mediator has the least amount of opportunities to implement p recommending mixed actions. Whereas minimally mixed BCEs seem esoteric at first, they are in fact, ubiquitous: the set of minimally mixed BCEs is open and dense in the BCE set (see Lemma 13 in the online appendix).

Our next result uses the concepts of jeopardization and minimally mixed BCEs to characterize when the BCE set and the closure of the sBCE set coincide.

Proposition 2. *The following statements are equivalent:*

(i) The sBCE set is dense in the BCE set.

(ii) A minimally mixed sBCE exists.

[18] Myerson defines jeopardization for games without payoff uncertainty; here we give the obvious extension to games where Θ is not a singleton.

(iii) For every BCE p, $i \in I$, $a_i, b_i \in \operatorname{supp}_i(p)$, if $p_{a_i} \neq p_{b_i}$ then $J(a_i) \cap J(b_i) = \varnothing$.

The result shows how jeopardization and minimal mixing can be used in applications to study sBCE. To verify that the sBCE set is dense in the BCE set, it is enough to produce a minimally mixed sBCE. To verify that the sBCE set is not dense in the BCE set, it is enough to produce a BCE in which two actions induce distinct beliefs and share a common jeopardizing action. As shown by Myerson (1997), the jeopardizing actions can be easily computed from the dual of the system of linear inequalities that defines BCE.

Next, we build on Proposition 2 and obtain that sBCE is an all-or-nothing refinement of BCE.

Proposition 3. *The sBCE set is either dense or nowhere dense in the BCE set.*

For a rough explanation, consider first the case in which a minimally mixed sBCE exists. Then, by Proposition 2, the sBCE set is dense in the BCE set. Consider now the case in which a minimally sBCE does not exist. By Proposition 2, the sBCE set is not dense in the BCE set. To reach the stronger conclusion that the sBCE set is nowhere dense in the BCE set, we use the fact that the set of minimally mixed BCE is open and dense in the BCE set.

Thus, whereas the separation constraint does not bind in most circumstances, whenever it does bind, it significantly restricts the set of attainable outcomes.

Appendix

A. Proof of Theorem 1

The proof of the theorem proceeds as follows. First, we recall a single-agent lemma due Denti (2021). Next, with the help of the lemma, we prove the "if" and "only if" statements of the theorem.

A.1. A single-agent lemma

We take the perspective of an individual i who has to choose an action $a_i \in A_i$ whose utility $w_i(a_i, \omega)$ depends on an uncertain state of nature $\omega \in \Omega$. Both A_i and

Ω are finite. Let $\rho \in \Delta(\Omega)$ be the prior distribution of the state; ρ may not have full support. Before choosing an action, the decision maker can run an experiment $\xi_i : \Omega \to \Delta(X_i)$ at a cost $C_i(\xi_i) \in \mathbb{R}_+$. Let $\mathcal{E}_i$ be the set of feasible experiments. The signal space X_i is finite. Overall, the decision maker faces the following information acquisition problem:

$$\max_{\xi_i \in \mathcal{E}_i, \sigma_i \in \Sigma_i} \left[\sum_{\omega, x_i, a_i} w_i(a_i, \omega)\sigma_i(a_i|x_i)\xi_i(x_i|\omega)\rho(\omega) \right] - C_i(\xi_i) \tag{10}$$

where Σ_i is the set of all action plans $\sigma_i : X_i \to \Delta(A_i)$.

In accordance with the terminology used in the main text, $\xi_i \succsim \xi_i'$ if there is a Markov kernel $g : X_i \to \Delta(X_i)$ such that for every $x_i \in X_i$ and $\omega \in \Omega$ with $\rho(\omega) > 0$,

$$\xi_i'(x_i|\omega) = \sum_{x_i' \in X_i} g(x_i|x_i')\xi(x_i'|\omega).$$

We say that $\mathcal{E}_i$ is **flexible** if, whenever $\xi_i \succsim \xi_i'$ and $\xi_i \in \mathcal{E}_i$, then $\xi_i' \in \mathcal{E}_i$. We also say that $C_i : \mathcal{E}_i \to \mathbb{R}_+$ is **monotone** if, whenever $\xi_i, \xi_i' \in \mathcal{E}_i$ are such that $\xi_i \succsim \xi_i'$ (resp., $\xi_i \succ \xi_i'$), then $C_i(\xi_i) \geq C_i(\xi_i')$ (resp., $C_i(\xi_i) > C_i(\xi_i')$).

Among other results, Denti (2021) characterizes the pairs (ξ_i, σ_i) that are optimal solutions of (10) for some flexible $\mathcal{E}_i$ and monotone C_i. To describe the characterization, let $\mu(x_i) = \sum_\omega \xi_i(x_i|\omega)\rho(\omega)$ be the ex-ante probability that ξ_i generates x_i. For every x_i such that $\mu(x_i) > 0$, we denote by $\mu_{x_i} \in \Delta(\Omega)$ the posterior distribution of the state ω. Finally, let $BR(\mu_{x_i}) = \arg\max_{a_i \in A_i} \sum_\omega w_i(a_i, \omega)\mu_{x_i}(\omega)$ be the set of best responses to μ_{x_i}.

Lemma 1 (Denti, 2021)**.** *A flexible $\mathcal{E}_i$ and a monotone C_i exist such that the pair (ξ_i, σ_i) is an optimal solution of (10) if and only if the following conditions hold:*

(i) For all x_i with $\mu(x_i) > 0$, $\sigma_i(BR(\mu_{x_i})|x_i) = 1$.

(ii) For all x_i and x_i' with $\mu(x_i)\mu(x_i') > 0$, $\mu_{x_i} \neq \mu_{x_i'}$ implies $BR(\mu_{x_i}) \cap BR(\mu_{x_i'}) = \varnothing$.

In addition, one can choose C_i so that

$$C_i(\xi_i) = \sum_{\omega, x_i, a_i} w_i(a_i, \omega)\sigma_i(a_i|x_i)\xi_i(x_i|\omega)\rho(\omega) - \max_{a_i \in A_i} \sum_\omega w_i(a_i, \omega)\rho(\omega).$$

A.2. Proof of the "if" statement of Theorem 1.

Let (p, v) be an outcome-value pair that satisfies Theorem 1-(i) and Theorem 1-(ii). First, we consider the case in which $v = \underline{v}(p)$:

Lemma 2. *There exist a rational-inattention technology* $\mathcal{T} = (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i \in I})$ *and an equilibrium* (ξ, σ) *of* $(\mathcal{G}, \mathcal{T})$ *whose outcome-value pair is* $(p, \underline{v}(p))$.

Proof. For every player i and pair of actions $a_i, b_i \in \text{supp}_i(p)$, let $a_i \sim_i b_i$ be $p_{a_i} = p_{b_i}$. Note that $\sim_i$ is an equivalence relation on $\text{supp}_i(p)$. Let Z_i be the corresponding set of equivalence classes.

We take $Z = \prod_{i \in I} Z_i$. To simply the exposition, we write $a \in z$ if $a_i \in z_i$ for all $i \in I$. Given this convention, we define $\zeta : \Theta \to Z$ by $\zeta(z|\theta) = \sum_{a \in z} p(a, \theta)/\pi(\theta)$. For every player i, we take $X_i = Z_i$, and we define $\xi_i : Z \times \Theta \to \Delta(X_i)$ by $\xi_i(x_i|z, \theta) = 1$ if $x_i = z_i$, and $\xi_i(x_i|z, \theta) = 0$ otherwise. We also define $\sigma_i : X_i \to \Delta(A_i)$ by $\sigma_i(a_i|x_i) = p(a_i)/\sum_{b_i \in x_i} p(b_i)$ if $a_i \in x_i$, and $\sigma_i(a_i|x_i) = 0$ otherwise.

Claim 3. The outcome induced by (ξ, σ) is p, that is, for all $a \in A$ and $\theta \in \Theta$

$$p(a, \theta) = \sum_z \left[\prod_{i \in I} \sigma_i(a_i|z_i)\right] \zeta(z|\theta)\pi(\theta). \tag{11}$$

To prove the claim, we need an intermediate result.

Claim 4. For all $z \in Z$, $a \in z$, and $\theta \in \Theta$,

$$p(a, \theta) = \left[\prod_{i \in I} \sigma_i(a_i|z_i)\right] \sum_{b \in z} p(b, \theta).$$

Proof. It suffices to show that for all $i \in I$,

$$p(a, \theta) = \sigma_i(a_i|z_i) \sum_{b_i \in z_i} p(b_i, a_{-i}, \theta). \tag{12}$$

The desired result then follows from reasoning by induction on the number of players.

So, fix some player i. By Bayes rule, $p(a, \theta) = p(a_i)p_{a_i}(a_{-i}, \theta)$. Recall that $b_i \in z_i$

if and only if $p_{b_i} = p_{a_i}$; in addition, $\sum_{b_i \in z_i} \sigma_i(b_i|z_i) = 1$. Therefore,

$$p(a_i, a_{-i}, \theta) = p(a_i) p_{a_i}(a_{-i}, \theta) = p(a_i) \sum_{b_i \in z_i} \sigma_i(b_i|z_i) p_{a_i}(a_{-i}, \theta)$$
$$= p(a_i) \sum_{b_i \in z_i} \sigma_i(b_i|z_i) p_{b_i}(a_{-i}, \theta).$$

Substituting in the definition of σ_i, one obtains (12). □

Proof of Claim 3. Fix $a \in A$ and $\theta \in \Theta$. If $p(a_i) = 0$ for some player i, then $\sigma(a_i|x_i) = 0$ for all $x_i \in X_i$. Thus, both sides of (11) are equal to zero.

Suppose now that $p(a_i) > 0$ for all players i. Take $z^* \in Z$ such that $a \in z^*$. Using Claim 4 we get that

$$p(a, \theta) = \left[\prod_{i \in I} \sigma_i(a_i|z_i^*)\right] \sum_{b \in z^*} p(b, \theta).$$

The equality (11) follows from noting that $\prod_{i \in I} \sigma_i(a_i|z_i) > 0$ if and only if $z = z^*$. □

Next, for each player i, we use Lemma 1 to construct $\mathcal{E}_i$ and C_i. Given $\Omega = Z \times \Theta$, we define $\rho \in \Delta(\Omega)$ and $w_i : A_i \times \Omega \to \mathbb{R}$ by

$$\rho(z, \theta) = \zeta(z|\theta)\pi(\theta),$$
$$w_i(a_i, z, \theta) = \sum_{a_{-i}} u_i(a_i, a_{-i}, \theta) \left[\prod_{j \neq i} \sigma_j(a_j|z_j)\right].$$

As in Section A.1, let $\mu(x_i)$ be the ex-ante probability that ξ_i generates x_i, and let $\mu_{x_i} \in \Delta(\Omega)$ be i's posterior beliefs about $\omega = (z, \theta)$ when x_i is observed. Note that μ_{x_i} is different from p_{a_i}, which is i's posterior beliefs about (a_{-i}, θ) when the mediator recommends a_i. Recall that $BR(\mu_{x_i}) = \arg\max_{b_i} \sum_\omega w_i(b_i, \omega)\mu_{x_i}(\omega)$, while $BR(p_{a_i}) = \arg\max_{b_i} \sum_{\theta, a_{-i}} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta)$.

To be able to use Lemma 1, we need the following result.

Claim 5. For all $x_i \in X_i$ and $a_i \in x_i$, $BR(p_{a_i}) = BR(\mu_{x_i})$.

Proof. It is enough to show that for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$p_{a_i}(a_{-i}, \theta) = \sum_z \left[\prod_{j \neq i} \sigma_j(a_j|z_j)\right] \mu_{x_i}(z, \theta). \quad (13)$$

Towards this goal, we first observe that for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\begin{aligned}
p(a_i, a_{-i}, \theta) &= \sum_{x_i', x_{-i}} \left[\sigma_i(a_i|x_i') \prod_{j \neq i} \sigma_j(a_j|x_j) \right] \zeta(x_i', x_{-i}|\theta)\pi(\theta) \\
&= \sigma_i(a_i|x_i) \sum_{x_{-i}} \left[\prod_{j \neq i} \sigma_j(a_j|x_j) \right] \zeta(x_i, x_{-i}|\theta)\pi(\theta) \\
&= \frac{p(a_i)}{\sum_{b_i \in x_i} p(b_i)} \sum_{x_{-i}} \left[\prod_{j \neq i} \sigma_j(a_j|x_j) \right] \zeta(x_i, x_{-i}|\theta)\pi(\theta), \qquad (14)
\end{aligned}$$

where the first equality holds by Claim 3, the second equality because $\sigma_i(a_i|x_i') > 0$ if and only if $x_i' = x_i$, and the third equality by definition of $\sigma_i(a_i|x_i)$. Summing both sides of (14) over a_{-i} and θ, we obtain that

$$p(a_i) = \frac{p(a_i)\mu(x_i)}{\sum_{b_i \in x_i} p(b_i)}.$$

Thus, $\mu(x_i) = \sum_{b_i \in x_i} p(b_i)$. We plug this finding in (14) and obtain that for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$p(a_i, a_{-i}, \theta) = \frac{p(a_i)}{\mu(x_i)} \sum_{x_{-i}} \left[\prod_{j \neq i} \sigma_j(a_j|x_j) \right] \zeta(x_i, x_{-i}|\theta)\pi(\theta).$$

The equality (13) follows from observing that $\mu_{x_i}(z, \theta) = \zeta(z|\theta)\pi(\theta)/\mu(x_i)$ if $z_i = x_i$, and $\mu_{x_i}(z, \theta) = 0$ otherwise. $\square$

To verify Lemma 1-(i), take $x_i \in X_i$ and $a_i \in A_i$ such that $\sigma_i(a_i|x_i) > 0$. By the construction of the action plan, $a_i \in x_i$. By the obedience constraint for p, $a_i \in BR(p_{a_i})$. By Claim 5, $BR(p_{a_i}) = BR(\mu_{x_i})$. Therefore, $a_i \in BR(\mu_{x_i})$. We deduce that Lemma 1-(i) holds.

To verify Lemma 1-(ii), take $x_i, x_i' \in X_i$ such that $\mu_{x_i} \neq \mu_{x_i'}$. Let $a_i \in x_i$ and $b_i \in x_i'$. Since $\mu_{x_i} \neq \mu_{x_i'}$, we must have $x_i \neq x_i'$ and, therefore, $p_{a_i} \neq p_{b_i}$. By the separation constraint for p, $BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing$. By Claim 5, $BR(p_{a_i}) = BR(\mu_{x_i})$ and $BR(p_{b_i}) = BR(\mu_{x_i'})$. We deduce $BR(\mu_{x_i}) = BR(\mu_{x_i'}) = \varnothing$: Lemma 1-(ii) holds.

In sum, for every player i, we can invoke Lemma 1 to find a flexible $\mathcal{E}_i$ and a monotone C_i such that (ξ_i, σ_i) is a best reply to (ξ_{-i}, σ_{-i}). This shows that $\mathcal{T} = (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i \in I})$ is a rational-inattention technology such that (ξ, σ) is an equilibrium $(\mathcal{G}, \mathcal{T})$ with outcome p. In addition, as in the last part of Lemma 1, we can choose costs so that, for every player i, $C_i(\xi_i) = \bar{v}_i(p) - \underline{v}_i(p)$, that is, $\bar{v}_i(p) - C_i(\xi_i) = \underline{v}_i(p)$. □

Now we consider the general case in which v may differ from $\underline{v}(p)$. As a starting point, take $\mathcal{T} = (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i\in I})$ and (ξ, σ) as in Lemma 2. Without loss of generality, we assume that $\mathcal{E}_i = \{\xi_i' : \xi_i \succsim \xi_i'\}$ for every player i.

For each $i \in I$, take $\lambda_i \in (0, 1]$ and define $C_{\lambda_i} : \mathcal{E}_i \to \mathbb{R}_+$ by $C_{\lambda_i}(\xi_i') = \lambda_i C_i(\xi_i')$. Since C_i is monotone, C_{λ_i} is monotone. Given $\lambda = (\lambda_i)_{i\in I}$, consider the rational-inattention technology $\mathcal{T}_\lambda = (Z, \zeta, (X_i, \mathcal{E}_i, C_{\lambda_i})_{i\in I})$.

Lemma 3. *The strategy profile (ξ, σ) is an equilibrium of $(\mathcal{G}, \mathcal{T}_\lambda)$.*

Proof. For a player i, consider an alternative strategy (ξ_i', σ_i'). With a slight abuse of notation, we denote by $u_i(\xi_i', \sigma_i', \xi_{-i}, \sigma_{-i})$ the expected utility player i obtains by deviating to (ξ_i', σ_i'). Since (ξ, σ) is an equilibrium of $(\mathcal{G}, \mathcal{T})$,

$$\bar{v}_i(p) - C_i(\xi_i) \geq u_i(\xi_i', \sigma_i', \xi_{-i}, \sigma_{-i}) - C_i(\xi_i').$$

In addition, since $\xi_i \succeq \xi_i'$ (recall: $\mathcal{E}_i = \{\xi_i'' : \xi_i \succsim \xi_i''\}$), $\bar{v}_i(p) \geq u_i(\xi_i', \sigma_i', \xi_{-i}, \sigma_{-i})$. It follows that

$$\bar{v}_i(p) - \lambda_i C_i(\xi_i) \geq u_i(\xi_i', \sigma_i', \xi_{-i}, \sigma_{-i}) - \lambda_i C_i(\xi_i').$$

We deduce the deviation is not profitable. □

Thus, (ξ, σ) is an equilibrium of $(\mathcal{G}, \mathcal{T}_\lambda)$ whose outcome is p. In addition,

$$\bar{v}_i(p) - C_{\lambda_i}(\xi_i) = \bar{v}_i(p) - \lambda_i C_i(\xi_i) = \bar{v}_i(p) - \lambda_i(\bar{v}_i(p) - \underline{v}_i(p)) = (1 - \lambda_i)\bar{v}_i + \lambda_i \underline{v}_i(p).$$

Hence, by appropriately choosing the vector λ, we can be sure that v is the value of (ξ, σ). This concludes the proof of the "if" direction of of Theorem 1.

A.3. Proof of the "only if" statement of Theorem 1.

Let (p, v) be the outcome-value pair induced by an equilibrium (ξ, σ) of an information acquisition game $(\mathcal{G}, \mathcal{T})$ where $\mathcal{T}$ represents rational inattention.

First we show that p is a separated BCE. For every player i, by the definition of equilibrium, the equilibrium strategy (ξ_i, σ_i) maximizes

$$\left[\sum_{a,x,z,\theta} u_i(a,\theta)\sigma_i'(a_i|x_i)\xi_i'(x_i|z,\theta)\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\zeta(z|\theta)\pi(\theta)\right] - C_i(\xi_i)$$

over all $\xi_i' \in \mathcal{E}_i$ and $\sigma_i' \in \Sigma_i$. Since $\mathcal{E}_i$ is flexible and C_i is monotone, we can apply Lemma 1 with appropriate definitions for Ω, ρ, and w_i. Specifically, we take $\Omega = Z \times \Theta$, and we define $\rho \in \Delta(\Omega)$ and $w_i : A_i \times \Omega \to \mathbb{R}$ by

$$\begin{aligned}
\rho(z,\theta) &= \zeta(z|\theta)\pi(\theta), \\
w_i(a_i, z, \theta) &= \sum_{x_{-i}, a_{-i}} u_i(a,\theta)\left[\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right].
\end{aligned}$$

From Lemma 1-(i), we obtain $\sigma_i(BR(\mu_{x_i})|\mu_{x_i}) = 1$ for all x_i with $\mu(x_i) > 0$. From Lemma 1-(ii), we get that for all x_i and x_i' such that $\mu(x_i) > 0$ and $\mu(x_i') > 0$, if $\mu_{x_i} \neq \mu_{x_i'}$ then $BR(\mu_{x_i}) \neq BR(\mu_{x_i'})$.

Next, we relate $BR(\mu_{x_i})$ and $BR(p_{a_i})$. Towards this goal, for each player i and action a_i, let X_{a_i} be the set of positive-probability signals that makes player i take action a_i: $X_{a_i} = \{x_i : \mu(x_i) > 0 \text{ and } \sigma(a_i|x_i) > 0\}$.

Lemma 4. *For all $i \in I$ and $a_i \in \operatorname{supp}_i(p)$, $BR(p_{a_i}) = \bigcap_{x_i \in X_{a_i}} BR(\mu_{x_i})$.*

Proof. First, take $b_i \in BR(p_{a_i})$: we wish to show that $b_i \in BR(\mu_{x_i})$ for all $x_i \in X_{a_i}$. By Lemma 1-(i), $a_i \in BR(\mu_{x_i})$ for all $x_i \in X_{a_i}$. Hence,

$$\sum_{z,\theta} w_i(a_i,z,\theta)\mu_{x_i}(z,\theta) \geq \sum_{z,\theta} w_i(b_i,z,\theta)\mu_{x_i}(z,\theta). \tag{15}$$

For each $x_i \in X_{a_i}$, define $\mu(x_i|a_i) = \sigma_i(a_i|x_i)\nu(x_i)/p(a_i)$. Simple algebra shows that

$$p_{a_i}(a_{-i},\theta) = \sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \sum_{x_{-i},z}\left[\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right]\mu_{x_i}(z,\theta). \tag{16}$$

Therefore,

$$\sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \left(\sum_{z,\theta} w_i(a_i, z, \theta) \mu_{x_i}(z, \theta) \right) = \sum_{a_{-i},\theta} u_i(a_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta),$$
$$\sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \left(\sum_{z,\theta} w_i(b_i, z, \theta) \mu_{x_i}(z, \theta) \right) = \sum_{a_{-i},\theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

As a consequence, since $b_i \in BR(p_{a_i})$, we have that

$$\sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \left(\sum_{z,\theta} w_i(a_i, z, \theta) \mu_{x_i}(z, \theta) \right) \leq \sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \left(\sum_{z,\theta} w_i(b_i, z, \theta) \mu_{x_i}(z, \theta) \right).$$

It follows that (15) holds with equality for all $x_i \in X_{a_i}$. Therefore, since $a_i \in BR(\mu_{x_i})$ for all $x_i \in X_{a_i}$, we deduce that $b_i \in BR(\mu_{x_i})$ for all $x_i \in X_{a_i}$.

Next, take $b_i \in A_i$ such that $b_i \in BR(\nu_{x_i})$ for all $x_i \in X_{a_i}$: we wish to show that $b_i \in BR(p_{a_i})$. By hypothesis, for all $x_i \in X_{a_i}$ and $c_i \in A_i$,

$$\sum_{z,\theta} w_i(c_i, z, \theta) \mu_{x_i}(z, \theta) \leq \sum_{z,\theta} w_i(b_i, z, \theta) \mu_{x_i}(z, \theta).$$

Averaging across inequalities over $x_i \in X_{a_i}$, we obtain that for all $c_i \in A_i$,

$$\sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \left(\sum_{z,\theta} w_i(c_i, z, \theta) \mu_{x_i}(z, \theta) \right) \leq \sum_{x_i \in X_{a_i}} \mu(x_i|a_i) \left(\sum_{z,\theta} w_i(b_i, z, \theta) \mu_{x_i}(z, \theta) \right).$$

Using (16), we deduce that

$$\sum_{a_{-i},\theta} u_i(c_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) \leq \sum_{a_{-i},\theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

We conclude that $b_i \in BR(p_{a_i})$. $\square$

We are ready to show that p is a separated BCE. For the obedience constraint, take $i \in I$ and $a_i \in \text{supp}_i(p)$. By Lemma 1-(i), $a_i \in BR(\mu_{x_i})$ for all $x_i \in X_{a_i}$. By Lemma 4, $a_i \in BR(p_{a_i})$. For the separation constraint, take $i \in I$ and $a_i, b_i \in \text{supp}_i(p)$ such that $p_{a_i} \neq p_{b_i}$. Since $p_{a_i} \neq p_{b_i}$, there are $x_{a_i} \in X_{a_i}$ and $x_{b_i} \in X_{b_i}$ such that $\mu_{x_{a_i}} \neq \mu_{x_{b_i}}$ (this should be obvious, but for confirmation, see (16)). By Lemma 1-(ii), $BR(\mu_{x_{a_i}}) \cap BR(\mu_{x_{b_i}}) = \varnothing$. By Lemma 4, $BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing$.

To complete the proof of the "only if" direction of Theorem 1, we need to show that for every player i, either $v_i = \underline{v}_i(p) = \bar{v}_i(p)$ or $v_i \in [\underline{v}_i(p), \bar{v}_i(p))$.

Fix a player i. Since $C_i(\xi_i) \geq 0$, it must be $v_i \leq \bar{v}_i(p)$. In addition, uninformative experiments have zero cost by hypothesis. Thus, since (ξ_i, σ_i) is a best response to (ξ_{-i}, σ_{-i}), we have that for any uninformative experiment ξ_i',

$$v_i \geq \max_{\sigma_i'} \sum_{a,x,z,\theta} u_i(a,\theta) \left[\sigma_i'(a_i|x_i)\xi_i(x_i|z,\theta) \prod_{j \neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right] \zeta(z|\theta)\pi(\theta) = \underline{v}_i(p).$$

Overall, we conclude that $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$

If $\underline{v}_i(p) = \bar{v}_i(p)$, then $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$ implies $v_i = \underline{v}_i(p) = \bar{v}_i(p)$. Suppose instead that $\underline{v}_i(p) < \bar{v}_i(p)$. If

$$\sum_{a,x,z,\theta} u_i(a,\theta) \left[\prod_{j \in I} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right] \zeta(z|\theta)\pi(\theta) < \bar{v}_i(p),$$

then obviously $v_i < \bar{v}_i(p)$ because $C_i(\xi_i) \geq 0$. If, on the other hand,

$$\sum_{a,x,z,\theta} u_i(a,\theta) \left[\prod_{j \in I} \sigma_j(a_j|x_j)\xi_j(x_j|z,\theta)\right] \zeta(z|\theta)\pi(\theta) = \bar{v}_i(p),$$

then ξ_i cannot be uninformative because $\underline{v}_i(p) < \bar{v}_i(p)$. By monotonicity, $C_i(\xi_i) > 0$, which implies that $v_i < \bar{v}_i(p)$. In sum, if $\underline{v}_i(p) < \bar{v}_i(p)$, then $v_i \in [\underline{v}_i(p), \bar{v}_i(p))$.

B. Proof of Theorem 2

Let $\|u\|$ be the Euclidean norm of $u \in \mathbb{R}^{I \times A \times \Theta}$. We begin with the following lemma.

Lemma 5. *For every $u \in \mathbb{R}^{I \times A \times \Theta}$, $p \in BCE(u)$, and $\epsilon > 0$, there exists $u' \in \mathbb{R}^{I \times A \times \Theta}$ such that $\|u - u'\| \leq \epsilon$ and $p \in sBCE(u')$.*

Proof. For each player $i \in I$, we consider a set $P_i \subseteq \Delta(A_{-i} \times \Theta)$ given by $P_i = \{p_{a_i} : a_i \in \mathrm{supp}_i(p)\}$. Let n_i be the cardinality of P_i (of course, it could be that n_i is smaller than the cardinality of $\mathrm{supp}_i(p)$). Reasoning inductively, we can find an enumeration $p_1, \ldots, p_{n_i}$ of the elements of P_i such that, for every $m_i \in \{1, \ldots, n_i\}$, p_{m_i} is an extreme point of the convex hull of $\{p_1, \ldots, p_{m_i}\}$.

By an hyperplane separation theorem (e.g., Rockafellar, 1970, Corollary 11.4.2) for every $m_i \in \{2, \dots, n_i\}$ we can find a function $f_{m_i} : A_{-i} \times \Theta \to \mathbb{R}$ such that

$$\sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > 0 \geq \max_{l_i \in \{1,\dots,m_i-1\}} \sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{l_i}(a_{-i}, \theta). \tag{17}$$

For $m_i = 1$, we define $f_1(a_{-i}, \theta) = 1$ for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$.

For every $l_i \in \{1, \dots, n_i - 1\}$, we choose $t_{l_i} \in (0, 1]$ such that for every $m_i \in \{l_i + 1, \dots, n_i\}$,

$$\sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > t_{l_i} \sum_{a_{-i}} f_{l_i}(a_{-i}) p_{m_i}(a_{-i}). \tag{18}$$

We can choose such a t_{l_i} because the left-hand side of (18) is positive—see (17). For $l_i = n_i$, we simply define $t_{n_i} = 1$.

For every $l_i \in \{1, \dots, n_i\}$, we define $s_{l_i} = \prod_{m_i=l_i}^{n_i} t_{m_i}$. Using (18), simple algebra shows that for every $l_i \in \{1, \dots, n_i - 1\}$ and $m_i \in \{l_i + 1, \dots, n_i\}$,

$$s_{m_i} \sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > s_{l_i} \sum_{a_{-i},\theta} f_{l_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta). \tag{19}$$

For $a_i \in \text{supp}_i(p)$, we define $g_{a_i} : A_{-i} \times \Theta \to \mathbb{R}$ by $g_{a_i}(a_{-i}, \theta) = s_{m_i} \cdot f_{m_i}(a_{-i}, \theta)$ where m_i is such that $p_{a_i} = p_{m_i}$. For $a_i \notin \text{supp}_i(p)$, we define $g_{a_i} = 0$.

Claim 6. For all $a_i \in \text{supp}_i(p)$ and $b_i \notin \{c_i \in \text{supp}_i(p) : p_{c_i} = p_{a_i}\}$,

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) > \sum_{a_{-i},\theta} g_{b_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta). \tag{20}$$

Proof. Pick $m_i \in \{1, \dots, n_i\}$ such that $p_{a_i} = p_{m_i}$.From the left-hand side of (17) and the fact that $s_{m_i} > 0$, we obtain that

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) = s_{m_i} \sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > 0. \tag{21}$$

Hence, for $b_i \notin \text{supp}_i(p)$, we have

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) > 0 = \sum_{a_{-i},\theta} g_{b_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta),$$

where the equality follows from $g_{b_i} = 0$.

Assume now that $b_i \in \mathrm{supp}_i(p)$. Choose l_i such that $p_{b_i} = p_{l_i}$. Since $p_{a_i} \neq p_{b_i}$, it must be that $m_i \neq l_i$. Suppose that $l_i > m_i$. It follows from the right-hand side of (17)—in (17) the roles of l_i and m_i are inverted—that $0 \geq \sum_{a_{-i},\theta} f_{l_i}(a_{-i},\theta)p_{m_i}(a_{-i},\theta)$. Thus, given that $s_{l_i} > 0$, we deduce that

$$0 \geq \sum_{a_{-i},\theta} g_{b_i}(a_{-i},\theta)p_{a_i}(a_{-i},\theta) = s_{l_i} \sum_{a_{-i},\theta} f_{l_i}(a_{-i},\theta)p_{m_i}(a_{-i},\theta).$$

We obtain that $\sum_{a_{-i},\theta} g_{a_i}(a_{-i},\theta)p_{a_i}(a_{-i},\theta) > 0 \geq \sum_{a_{-i},\theta} g_{b_i}(a_{-i},\theta)p_{a_i}(a_{-i},\theta)$, where we use again (21). For the case $l_i < m_i$, the condition $\sum_{a_{-i},\theta} g_{a_i}(a_{-i},\theta)p_{a_i}(a_{-i},\theta) > \sum_{a_{-i},\theta} g_{b_i}(a_{-i},\theta)p_{a_i}(a_{-i},\theta)$ is equivalent to (19). We conclude that (20) holds. $\square$

To complete the proof of the lemma, for every $\delta > 0$ we define $u' = (u'_i)_{i\in I}$ by $u'_i(a,\theta) = u_i(a,\theta) + \delta g_{a_i}(a_{-i},\theta)$. By choosing δ sufficiently small, we can be make sure that $\|u - u'\| \leq \epsilon$. Since $p \in BCE(u)$, it follows from (20) that for all $i \in I$ and $a_i \in \mathrm{supp}_i(p)$, $a_i \in BR'(p_{a_i}) \subseteq \{b_i \in \mathrm{supp}_i(p) : p_{a_i} = p_{b_i}\}$ where $BR'(p_{a_i})$ is the set of i's best response to a belief p_{a_i} given utility function u'_i. Thus, $p \in sBCE(u')$. $\square$

A subset of a Euclidean space is **semi-algebraic** if it is defined by finite systems of polynomial inequalities. A correspondence between Euclidean spaces is semi-algebraic if its graph is semi-algebraic. The background knowledge on semi-algebraic sets that we use in this proof can be gathered from Blume and Zame (1994, Section 2).

To state the next result, let $\mathrm{cl}(sBCE(u))$ be the closure of the sBCE set.

Lemma 6. *The correspondences $u \mapsto BCE(u)$, $u \mapsto sBCE(u)$, and $u \mapsto \mathrm{cl}(sBCE(u))$ are semi-algebraic.*

Proof. The BCE correspondence is semi-algebraic: for all $u \in \mathbb{R}^{I\times A\times\Theta}$ and $p \in \mathbb{R}^{A\times\Theta}$, $p \in BCE(u)$ if and only if the pair (u,p) is a solution to the following finite system of polynomial inequalities: $p(a,\theta) \geq 0$ for all $a \in A$ and $\theta \in \Theta$; $\sum_a p(a,\theta) = \pi(\theta)$ for all $\theta \in \Theta$; and $\sum_{a_{-i},\theta}(u(a,\theta) - u(b_i,a_{-i},\theta))p(a,\theta) \geq 0$ for all $i \in I$ and $a_i, b_i \in A_i$.

The sBCE correspondence is also semi-algebraic. To prove it, for $u \in \mathbb{R}^{I\times A\times\Theta}$, $p \in \mathbb{R}^{A\times\Theta}$, $i \in I$, and $a_i, b_i, c_i \in A_i$, we denote by $F(u,p,a_i,b_i,c_i)$ the quantity

$$\sum_{a_{-i},\theta} (u(a_i,a_{-i},\theta) - u(c_i,a_{-i},\theta))p(a_i,a_{-i},\theta) + (u(b_i,a_{-i},\theta) - u(c_i,a_{-i},\theta))p(b_i,a_{-i},\theta).$$

We observe that $p \in sBCE(u)$ if and only if $p \in BCE(u)$ and for every $i \in I$ there is $T_i \subseteq A_i \times A_i$ such that the pair (u, p) is a solution of the following finite system of polynomial inequalities:

$$\sum_{a_{-i},\theta} (p(a_i, a_{-i}, \theta)p(b_i) - p(b_i, a_{-i}, \theta)p(a_i))^2 = 0 \quad \text{for all } (a_i, b_i) \in T_i,$$

$$F(u, p, a_i, b_i, c_i) > 0 \quad \text{for all } (a_i, b_i) \notin T_i \text{ and } c_i \in A_i.$$

Thus, $p \in sBCE(u)$ if and only if it the solution of one of finitely many systems of polynomial inequalities; we conclude that the sBCE correspondence is semi-algebraic.

The correspondence $u \mapsto \mathrm{cl}(sBCE(u))$ is also semi-algebraic. Indeed, we have $p \in \mathrm{cl}(sBCE(u))$ if and only if for every $\epsilon > 0$ there exists $q \in \mathbb{R}^{A\times\Theta}$ such that $\|p - q\| \leq \epsilon$ and $q \in sBCE(u)$. Thus, since the sBCE correspondence is semi-algebraic, the graph of the correspondence $u \mapsto \mathrm{cl}(sBCE(u))$ is defined by a first-order formula and therefore semi-algebraic by the Tarski-Seidenberg theorem (Blume and Zame, 1994, page 787). □

We are ready to complete the proof of the theorem. By Lemma 6, the correspondence $u \mapsto \mathrm{cl}(sBCE(u))$ is semi-algebraic. Hence, there is an open subsets U of $\mathbb{R}^{I\times A\times\Theta}$ such that the complement of U has Lebesgue measure zero, and $u \mapsto \mathrm{cl}(sBCE(u))$ is continuous on U (Blume and Zame, 1994, page 786).

We claim that for all $u \in U$, $BCE(u) = \mathrm{cl}(sBCE(u))$. To prove the claim, take any $u \in U$. Since $BCE(u)$ is closed and $sBCE(u) \subseteq BCE(u)$, $\mathrm{cl}(sBCE(u)) \subseteq BCE(u)$. To verify the other inclusion, we use Lemma 5 to find a sequence of games $(u^n)_{n=1}^{\infty}$ such that $u^n \to u$ and, for every n, $p \in sBCE(u^n) \subseteq \mathrm{cl}(sBCE(u^n))$. Since $\mathrm{cl}(sBCE(u))$ is continuous at u, we have $p \in \mathrm{cl}(sBCE(u))$; see Aliprantis and Border (2006, Theorem 17.16). Hence, $BCE(u) \subseteq \mathrm{cl}(sBCE(u))$. We deduce that $BCE(u) = \mathrm{cl}(sBCE(u))$, as desired.

C. Proof of Theorem 3

The first part of the theorem immediately follows from Theorem 2, together with (3) and (4). To prove the theorem's last part, note that, when $|I| = 1$, $\min V_I = \min V_R = \bar{v}(p^0)$, where p^0 is a BCE that provides the player with no information (i.e., a BCE where the player's action does not vary with the state). The result then follows from noting that for a generic game, both V_I and V_R are convex sets that share the same maximal value.

Next, we assume that $|I| \geq 2$, $|\Theta| \geq 2$, and $|A_i| \geq 2$ for at least two distinct player i, and show there is an open set $U \subseteq \mathbb{R}^{I \times A \times \Theta}$ such that $V_I(u) \subset V_R(u)$ for all $u \in U$. To ease the exposition, we actually assume (without loss of generality) that $|A_i| \geq 2$ for all players i.

To prove $V_I(u) \subset V_R(u)$, we will determine that

$$\min_{v \in V_R(u)} \sum_i v_i < \min_{v \in V_I(u)} \sum_i v_i. \tag{22}$$

Consistently with the notation of Section 6, we write $\underline{w}(p,u) = \sum_i \underline{v}_i(p,u)$ and $\bar{w}(p,u) = \sum_i \bar{v}_i(p,u)$. We also denote by $\underline{w}(u)$ the minimum of $\underline{w}(p,u)$ over all $p \in BCE(u)$, and by $\bar{w}(u)$ the minimum of $\bar{w}(p,u)$ over all $p \in BCE(u)$. Note that

$$\min_{v \in V_R(u)} \sum_i v_i = \min_{p \in \mathrm{cl}(sBCE(u))} \underline{w}(p,u) \geq \underline{w}(u), \quad \text{and} \quad \min_{v \in V_I(u)} \sum_i v_i = \bar{w}(u).$$

The next lemma gives a sufficient condition for the existence of an open set $U \subseteq \mathbb{R}^{I \times A \times \Theta}$ such that (22) holds for all $u \in U$.

Lemma 7. *Suppose $u^* \in \mathbb{R}^{I \times A \times \Theta}$ and $p^* \in \Delta_\pi(A \times \Theta)$ satisfy the following properties:*

(i) Each player i takes at least two actions at p^*: $|\mathrm{supp}_i(p^*)| \geq 2$.

(ii) p^* is a strict BCE: $BR(p_{a_i}) = \{a_i\}$ for all $i \in I$ and $a_i \in \mathrm{supp}_i(p)$.

(iii) p^* is the unique minimizer of $\underline{w}(p,u^*)$ over all $p \in BCE(u^*)$.

Then there is a neighborhood U of u^* such that $\underline{w}(p,u) < \bar{w}(u)$ for all $u \in U$ and $p \in \mathrm{cl}(sBCE(u))$.

Proof. Take u^* and p^* that satisfy (i)-(iii). First, we verify that

$$\underline{w}(p^*,u^*) < \bar{w}(u^*). \tag{23}$$

Take $p \in BCE(u^*)$ such that $\bar{w}(u^*) = \bar{w}(p,u^*)$. If $p \neq p^*$, then $\bar{w}(p,u^*) \geq \underline{w}(p,u^*) > \underline{w}(p^*,u^*)$, where the strict inequality holds by (iii); thus, $\bar{w}(u^*) > \underline{w}(p^*,u^*)$. If instead $p = p^*$, then $\bar{w}(p^*,u^*) > \underline{w}(p^*,u^*)$ by (i) and (ii); thus $\bar{w}(u^*) > \underline{w}(p^*,u^*)$. Overall, we conclude that (23) holds, as desired.

The rest of the proof proceed by contradiction. To attain this contradiction, suppose there is a sequence $(u^n)_{n=1}^{\infty}$ converging to u^* such that

$$\min_{p\in \mathrm{cl}(sBCE(u^n))} \underline{w}(p, u^n) = \bar{w}(u^n) \quad \text{for all } n. \tag{24}$$

By (ii), $p^* \in sBCE(u^n)$ for all n sufficiently large. Thus,

$$\underline{w}(p^*, u^n) \geq \min_{p\in \mathrm{cl}(sBCE(u^n))} \underline{w}(p, u^n) \quad \text{for all } n \text{ large enough.} \tag{25}$$

Combining (24) and (25), we obtain that

$$\underline{w}(p^*, u^n) \geq \bar{w}(u^n) \quad \text{for all } n \text{ large enough.} \tag{26}$$

By standard arguments, $\underline{w}(p^*, u)$ is continuous in u. In addition, since the correspondence $u \mapsto BCE(u)$ is upper hemicontinuous, $\bar{w}(u) = \min_{p\in BCE(u)} \bar{w}(p, u)$ is lower semicontinuous in u (e.g., Aliprantis and Border, 2006, Lemma 17.3). It follows from (26) that $\underline{w}(p^*, u^*) = \liminf_{n\to\infty} \underline{w}(p^*, u^n) \geq \liminf_{n\to\infty} \bar{w}(u^n) \geq \bar{w}(u^*)$. Hence, $\underline{w}(p^*, u^*) \geq \bar{w}(u^*)$, which contradicts (23). □

To complete the proof of the theorem, we construct a utility profile u^* and and outcome p^* that satisfy the conditions of Lemma 7. This lemma then delivers a neighborhood U of u^* such that (22) holds for all $u \in U$, which in turn means that the set of u such that $V_I(u) \subset V_R(u)$ has non-empty interior.

We now construct u^* and p^*. Let n be cardinality of I; by hypothesis, $n \geq 2$. For every player i, we order the set of actions from 0 to m_i (where m_i+1 is the cardinality of A_i): $A_i = \{0, \dots, m_i\}$. By hypothesis, A_i contains at least two distinct elements, thus $m_i \geq 1$. We also consider a partition $\Theta = \Theta_l \cup \Theta_h$ of the set of payoff states such that both Θ_l and Θ_h are nonempty; this is feasible because, by hypothesis, Θ contains at least two elements.

For player i, we define u_i^* as follows:

$$u_i^*(a, \theta) = \begin{cases} 0 & \text{if } a_i = 0, \\ \frac{1}{\pi(\Theta_l)} \left(-1 + \frac{1}{n-1} \sum_{j\neq i} a_j\right) & \text{if } a_i = 1 \text{ and } \theta \in \Theta_l, \\ \frac{1}{\pi(\Theta_h)} \left(2 - \frac{1}{n-1} \sum_{j\neq i} a_j\right) & \text{if } a_i = 1 \text{ and } \theta \in \Theta_h, \\ -1 & \text{if } a_i > 1. \end{cases}$$

Thus, action 0 is a safe action. Action 1 has a payoff that depends both on the state and on the average action of the opponents. For states in Θ_l, action 1 generates a *negative* baseline payoff of -1, but there is also a *positive* externality from the actions of others; these payoffs are scaled by $1/\pi(\Theta_l)$. For states in Θ_h, action 1 generates a *positive* baseline payoff of 2, but there is also a *negative* externality from the actions of others; these payoffs are scaled by $1/\pi(\Theta_h)$. Any action outside $\{0,1\}$ is strictly dominated by 0.

Let p^* be the outcome such that all players take action 0 when $\theta \in \Theta_l$, and all players take action 1 when $\theta \in \Theta_h$. Clearly, (i) and (ii) of Lemma 7 hold.

All that remains is to verify (iii). To do so, note first that p^* is the unique minimizer of $\sum_i \sum_{a,\theta} u_i^*(1, a_{-i}, \theta)p(a, \theta)$ over all $p \in \Delta_\pi(\{0,1\}^I \times \Theta)$. Since any action outside $\{0,1\}$ is strictly dominated, we deduce that p^* is the unique minimizer of $\sum_i \sum_{a,\theta} u_i^*(1, a_{-i}, \theta)p(a, \theta)$ over all $p \in BCE(u^*)$. Moreover, simple algebra shows that for all players i,

$$\max_{b_i} \sum_{a,\theta} u_i^*(b_i, a_{-i}, \theta)p^*(a, \theta) = \sum_{a,\theta} u_i^*(1, a_{-i}, \theta)p^*(a, \theta).$$

In turn, this implies that $\underline{w}(p^*, u^*) = \sum_i \sum_{a,\theta} u_i^*(1, a_{-i}, \theta)p^*(a, \theta)$. Therefore, every $p \in BCE(u^*) \setminus \{p^*\}$ has

$$\begin{aligned}\underline{w}(p, u^*) &= \sum_i \max_{b_i} \sum_{a,\theta} u_i^*(b_i, a_{-i}, \theta)p(a, \theta) \geq \sum_i \sum_{a,\theta} u_i^*(1, a_{-i}, \theta)p(a, \theta) \\ &> \sum_i \sum_{a,\theta} u_i^*(1, a_{-i}, \theta)p^*(a, \theta) = \underline{w}(p^*, u^*).\end{aligned}$$

we conclude that $\underline{w}(u^*) = \underline{w}(p^*, u^*)$ if and only if $p = p^*$, that is, (iii) of Lemma 7 holds. The proof is now complete.

References

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis*. Springer, third edition, 2006.

G.-M. Angeletos and K. A. Sastry. Inattentive economies. 2023. URL `https://www.karthiksastry.com/`.

R. J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.

D. Bergemann and S. Morris. Robust predictions in games with incomplete information. *Econometrica*, 81:1251–1308, 2013.

D. Bergemann and S. Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11:487–522, 2016.

D. Bergemann, B. Brooks, and S. Morris. The limits of price discrimination. *American Economic Review*, 105:921–957, 2015.

D. Bergemann, B. Brooks, and S. Morris. First-price auctions with general information structures: Implications for bidding and revenue. *Econometrica*, 85:107–143, 2017.

D. Blackwell. Comparison of experiments. In *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102. University of California Press, 1951.

D. Blackwell. Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, 24:265–272, 1953.

L. E. Blume and W. R. Zame. The algebraic geometry of perfect and sequential equilibrium. *Econometrica*, 62:783–794, 1994.

B. Brooks and S. Du. Optimal auction design with common values: An informationally robust approach. *Econometrica*, 89:1313–1360, 2021.

A. Caplin and M. Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105:2183–2203, 2015.

A. Caplin, M. Dean, and J. Leahy. Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Journal of Political Economy*, 130:1676–1715, 2022.

G. Carroll. Robustness in mechanism design and contracting. *Annual Review of Economics*, 11:139–166, 2019.

L. Colombo, G. Femminis, and A. Pavan. Information acquisition and welfare. *Review of Economic Studies*, 81:1438–1483, 2014.

G. de Clippel and K. Rozen. Communication, perception, and strategic obfuscation. 2022. URL https://sites.google.com/brown.edu/kareenrozen.

M. Dean and N. L. Neligh. Experimental tests of rational inattention. *Journal of Political Economy*, forthcoming.

T. Denti. Costly monitoring in signaling games. 2021. URL https://sites.google.com/site/tommasojdenti.

T. Denti. Posterior separable cost of information. *American Economic Review*, 112:3215–59, 2022.

T. Denti. Unrestricted information acquisition. *Theoretical Economics*, forthcoming.

T. Denti, M. Marinacci, and A. Rustichini. Experimental cost of information. *American Economic Review*, 112:3106–23, 2022.

L. Doval and J. Ely. Sequential information design. *Econometrica*, 88:2575–2608, 2020.

F. Forges. Five legitimate definitions of correlated equilibrium in games with incomplete information. *Theory and Decision*, 35:277–310, 1993.

B. Hebert and J. La'O. Information acquisition, efficiency, and non-fundamental volatility. *Journal of Political Economy*, forthcoming.

C. Hellwig and L. Veldkamp. Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies*, 76:223–251, 2009.

B. L. Lipman and S. Srivastava. Computation as a correlation device. *Games and Economic Behavior*, 2:154–172, 1990.

E. Lipnowski and D. Ravid. Predicting choice from information costs. 2022. URL https://arxiv.org/abs/2205.10434.

B. Maćkowiak and M. Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99:769–803, 2009.

B. Maćkowiak, F. Matějka, and M. Wiederholt. Rational inattention: A review. *Journal of Economic Literature*, 61:226–273, 2023.

F. Matějka and A. McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105:272–298, 2015.

S. Morris and H. S. Shin. Global games: theory and applications. In *Proceedings of the 8th World Congress of the Econometric Society)*, pages 56–114. Cambridge University Press, 2003.

S. Morris and M. Yang. Coordination and continuous stochastic choice. *The Review of Economic Studies*, 89:2687–2722, 2022.

D. P. Myatt and C. Wallace. Endogenous information acquisition in coordination games. *Review of Economic Studies*, 79:340–374, 2012.

R. B. Myerson. Optimal coordination mechanisms in generalized principal–agent problems. *Journal of Mathematical Economics*, 10:67–81, 1982.

R. B. Myerson. Dual reduction and elementary games. *Games and Economic Behavior*, 21:183–202, 1997.

D. Ravid. Ultimatum Bargaining with Rational Inattention. *American Economic Review*, 110:2948–63, 2020.

D. Ravid, A.-K. Roesler, and B. Szentes. Learning before trading: on the inefficiency of ignoring free information. *Journal of Political Economy*, 130:346–387, 2022.

R. T. Rockafellar. *Convex analysis.* Princeton University Press, 1970.

T. Roughgarden. *Twenty lectures on algorithmic game theory.* Cambridge University Press, 2016.

C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50: 665–690, 2003.

M. Yang. Coordination with flexible information acquisition. *Journal of Economic Theory*, 158:721–738, 2015.

W. Zhong. Optimal dynamic information acquisition. *Econometrica*, 90:1537–1582, 2022.

Online Appendix

D. Arbitrary Information Technologies

In this section, we characterize the predictions attainable as one ranges over all information technologies. In particular, we do not require the information technology to be flexible or monotone. We also show it is without loss to require the technology to be flexible, and costs to be weakly monotone.[19] Formally, a cost function C_i is **weakly monotone** if less informative experiment are weakly cheaper to acquire: if $\xi_i, \xi_i' \in \mathcal{E}_i$ are such that $\xi_i \succsim \xi_i'$, then $C_i(\xi_i) \geq C_i(\xi_i')$.

Proposition 4. *Fix a base game $\mathcal{G}$. An information technology $\mathcal{T}$ exists that induces the outcome-value pair (p, v) in an equilibrium of $(\mathcal{G}, \mathcal{T})$ if and only if*

(i) p is a BCE, and

(ii) for every $i \in I$, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$.

In addition, for every player i, one can choose $\mathcal{E}_i$ flexible and C_i weakly monotone.

Proof. "If." Let (p, v) be an outcome-value pair such that p is a BCE and, for every $i \in I$, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$. Since p is a BCE, by Bergemann and Morris (2016) there exist an information structure $\mathcal{S} = (Z, \zeta, (X_i, \xi_i)_{i \in I})$ and a profile of action plans $\sigma = (\sigma_i)_{i \in I}$ such that p is the outcome of (ξ, σ), and for every player i, σ_i maximizes

$$\sum_{a,x,z,\theta} u_i(a, \theta) \left(\sigma_i'(a_i|x_i)\xi_i(x_i|z, \theta) \prod_{j \neq i} \sigma_j(a_j|x_j)\xi_j(x_j|z, \theta) \right) \zeta(z|\theta)\pi(\theta) \tag{27}$$

over all $\sigma_i' \in \Sigma_i$. To ease notation, denote the quantity in (27) by $u_i(\xi_i', \sigma_i', \xi_{-i}, \sigma_{-i})$.

For every player i, let $\mathcal{E}_i = \{\xi_i' : \xi_i \succeq \xi_i'\}$. In addition, take $\lambda_i \in [0, 1]$ such that

$$v_i = \lambda_i \underline{v}_i(p) + (1 - \lambda_i)\bar{v}_i(p).$$

For every $\xi_i' \in \mathcal{E}_i$, define $C_i(\xi_i') = \lambda_i(\max_{\sigma_i'} u_i(\xi_i', \sigma_i', \xi_{-i}, \sigma_{-i}) - \underline{v}_i(p))$. Notice that $\mathcal{E}_i$ is flexible and C_i is weakly monotone.

[19] For an analogous result in single-agent settings, see Caplin and Dean (2015).

It follows from (27) that $C_i(\xi_i) = \lambda_i\left(\bar{v}_i(p) - \underline{v}_i(p)\right)$, which in turn implies that $u_i(\xi,\sigma) - C_i(\xi_i) = v_i$. We also see that for every $\xi_i' \in \mathcal{E}_i$,

$$\begin{aligned}
u_i(\xi,\sigma) - C_i(\xi_i) &= \max_{\sigma_i'} u_i(\xi,\sigma_i',\sigma_{-i}) - C_i(\xi_i) \\
&= \lambda_i \underline{v}_i(p) + (1-\lambda_i)\max_{\sigma_i'} u_i(\xi,\sigma_i',\sigma_{-i}) \\
&\geq \lambda_i \underline{v}_i(p) + (1-\lambda_i)\max_{\sigma_i'} u_i(\xi_i',\xi_{-i},\sigma_i',\sigma_{-i}) \\
&= \max_{\sigma_i'} u_i(\xi_i',\xi_{-i},\sigma_i',\sigma_{-i}) - C_i(\xi_i'),
\end{aligned}$$

where the first equality follows from (27) and the weak inequality from $\xi_i \succeq \xi_i'$. We conclude (ξ,σ) is an equilibrium of $(\mathcal{G},\mathcal{T})$ with $\mathcal{T} = (Z,\zeta,(X_i,\mathcal{E}_i,C_i)_{i\in I})$; in addition, (p,v) is the outcome-value pair corresponding to (ξ,σ).

"Only if." Let (p,v) be the outcome-value pair of an equilibrium (ξ,σ) of an information acquisition game $(\mathcal{G},\mathcal{T})$, with $\mathcal{T} = (Z,\zeta,(X_i,\mathcal{E}_i,C_i)_{i\in I})$. Define the information structure $\mathcal{S} = (Z,\zeta,(X_i,\xi_i)_{i\in I})$. Since (ξ,σ) is an equilibrium of $(\mathcal{G},\mathcal{T})$, σ is an equilibrium of $(\mathcal{G},\mathcal{S})$. By Bergemann and Morris (2016), p is a BCE.

For every player i, $C_i(\xi_i) \geq 0$, which implies that $v_i \leq \bar{v}_i(p)$. In addition, by hypothesis there exists an experiment ξ_i' such that $C_i(\xi_i') = 0$. Thus, since (ξ_i,σ_i) is a best response to (ξ_{-i},σ_{-i}), we have that

$$v_i \geq \max_{\sigma_i'} u_i(\xi_i',\xi_{-i},\sigma_i',\sigma_{-i}) \geq \underline{v}_i(p).$$

We conclude that $v_i \in [\bar{v}_i(p), \underline{v}_i(p)]$. □

E. Examples Where Separation Binds

In this section we present a few simple examples in which the separation constraint has substantial bite, that is, in which the sBCE set is not dense in the BCE set. In all the examples that follow, the sBCE set is nowhere dense is the BCE set. As Theorem 3 predicts, if the sBCE set is not dense in the BCE set, it must be nowhere dense. To ease the exposition, we assume the payoff state is degenerate (i.e., Θ is a singleton), and we omit it.

The simplest example in which the separation constraint has stark effects is the scenario in which the players' utilities are constant: $u_i(a) = u_i(b)$ for all $i \in I$ and $a, b \in A$. In this case, the BCE set is the entire simplex $\Delta(A)$. On the other hand, a BCE is separated if and only if the players' actions are independent. Thus, the sBCE set can be identified with $\prod_{i\in I} \Delta(A_i)$, the set of mixed-action profiles.

The presence of weakly dominated actions is a factor that may put a wedge between BCE and sBCE. For example, consider the following 2×2 game:

	a_2	b_2
a_1	$2, 2$	$2, 2$
b_1	$3, 1$	$0, 0$

This game can be seen as the reduced normal form of a Battle of the Sexes with Outside Option.[20] Note that a_2 weakly dominates b_2.

It is easy to see that the BCEs are all the outcomes p such that $p(b_1, b_2) = 0$ and $p(a_1, a_2) \leq 2p(a_1, b_2)$. However, a BCE p is separated if and only if $p(b_1, a_2) \in \{0, 1\}$. To see why, first consider the case in which $p(b_1, a_2) = \{0, 1\}$. Then, since player 1's action is deterministic (either $p(a_1) = 1$ or $p(b_1) = 1$), the separation constraint is trivially satisfied for both players. Conversely, suppose that $p(b_1, a_2) \in (0, 1)$. In this case, player 2 takes both actions with positive probability, and they induce different beliefs about player 1's action: $p_{a_2}(b_1) > 0 = p_{b_2}(b_1)$. Since a_2 weakly dominates b_2, we have $a_2 \in BR(p_{a_2}) \cap BR(p_{b_2})$. Hence, the separation constraint is not satisfied.

One should not overstate the relationship between weakly dominated actions and separation. As the next example highlights, the separation constraint can have a substantial impact even if no action is weakly dominated:

	a_2	b_2	c_2
a_1	$8, 8$	$3, 7$	$2, 6$
b_1	$7, 3$	$5, 1$	$0, 5$
c_1	$6, 2$	$1, 4$	$4, 0$

The game, which is a variation of Myerson (1997, Figure 6), has no weakly dominated

[20]In the BoS with OO we have in mind, player 1 first chooses between *Out* and *In*. Given *Out*, each player obtains a payoff of 2. Given *In*, the players participate in a coordination game in which they simultaneously choose between a *Bach* concert and a *Stravinsky* concert. If they coordinate on *Bach*, player 1 gets 3 and player 2 gets 1; if they coordinate on *Stravinsky*, player 1 gets 1 and player 2 gets 3; if they mis-coordinate, they both obtain 0.

action. It has one pure Nash equilibrium and one mixed Nash equilibrium:

$$(a_1, a_2) \quad \text{and} \quad \left(\frac{1}{2}b_1 + \frac{1}{2}c_1, \frac{1}{2}b_2 + \frac{1}{2}c_2\right).$$

The BCEs are the convex combinations of the two Nash equilibria: for $t \in [0, 1]$,

$$p^t = t(a_1, a_2) + (1-t)\left(\frac{1}{2}b_1 + \frac{1}{2}c_1, \frac{1}{2}b_2 + \frac{1}{2}c_2\right).$$

The game has only two separated BCE, namely, the two Nash equilibria. Indeed, for every $t \in (0, 1)$ and every player i, the action recommendations a_i and b_i (or c_i) induce distinct posterior beliefs about the action of the opponent: $p^t_{a_i}(a_j) = 1$, while $p^t_{b_i}(b_j) = p^t_{b_i}(c_j) = 1/2$. Yet, a_i is best response to the belief induced by b_i:

$$\frac{1}{2}u_i(a_i, b_j) + \frac{1}{2}u_i(a_i, c_j) = \frac{5}{2} = \frac{1}{2}u_i(b_i, b_j) + \frac{1}{2}u_i(b_i, c_j).$$

F. Strict BCE: Single-Agent Settings

A BCE p is **strict** if all $i \in I$, $a_i \in \operatorname{supp}_i(p)$, and $b_i \in A_i$ with $b_i \neq a_i$,

$$\sum_{a_{-i},\theta} \left(u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta)\right) p(a_i, a_{-i}, \theta) > 0.$$

In the main text, discussing Theorem 2, we mentioned the following result:

Proposition 5. *Let $I = \{i\}$ be a singleton. For generic u_i, the set of strict BCE is dense in the BCE set.*

We expect the result to be known in the literature. However, we could not find a good reference. Thus, next we provide a self-contained proof. The proof relies on two lemmas on dominated actions. A mixed action $\alpha_i \in \Delta(A_i)$ **weakly dominates** a pure action $a_i \in A_i$ if $\sum_{b_i} u_i(b_i, a_{-i}, \theta)\alpha_i(b_i) \geq u_i(a_i, a_{-i}, \theta)$. for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$. The next result provides a characterization of weakly dominated actions:[21]

Lemma 8. *The following statements are equivalent:*

(i) There is no belief $\mu_{a_i} \in \Delta(A_{-i} \times \Theta)$ for which a_i is the unique best response.

[21] See any textbook on statistical decision theory for closely related results on admissibility.

(ii) There is a mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$ that weakly dominates a_i.

Proof. Condition (i) can be rewritten as

$$\max_{\mu_i \in \Delta(A_{-i} \times \Theta)} \min_{b_i \in A_i \setminus \{a_i\}} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\mu_i(a_{-i}, \theta) \leq 0.$$

Equivalently,

$$\max_{\mu_i \in \Delta(A_{-i} \times \Theta)} \min_{\alpha_i \in \Delta(A_i \setminus \{a_i\})} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\mu_i(a_{-i}, \theta)\alpha_i(b_i) \leq 0.$$

By the minimax theorem (e.g., Rockafellar, 1970, Corollary 37.3.2), the above inequality holds if and only if

$$\min_{\alpha_i \in \Delta(A_i \setminus \{a_i\})} \max_{\mu_i \in \Delta(A_{-i} \times \Theta)} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\mu_i(a_{-i}, \theta)\alpha_i(b_i) \leq 0.$$

Equivalently,

$$\min_{\alpha_i \in \Delta(A_i \setminus \{a_i\})} \max_{a_{-i}, \theta} \sum_{a_{-i}, \theta} (u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta))\alpha_i(b_i) \leq 0.$$

which is another way of expressing condition (ii). $\square$

A mixed action $\alpha_i \in \Delta(A_i)$ **strictly dominates** a pure action $a_i \in A_i$ if for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, $\sum_{b_i} u_i(b_i, a_{-i}, \theta)\alpha_i(b_i) > u_i(a_i, a_{-i}, \theta)$. The next result shows that generically, weakly dominated actions are strictly dominated.

Lemma 9. *Let $I = \{i\}$ be a singleton. For generic u_i, if an action a_i is weakly dominated by some mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$, then it is strictly dominated by some mixed action $\beta_i \in \Delta(A_i)$.*

Proof. Let a_i be an action that is weakly dominated by a mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$. Let A'_i be the support of α_i, and let Θ' be set of states θ for which

$$u_i(a_i, \theta) = \sum_{b_i} u_i(b_i, \theta)\alpha_i(b_i). \tag{28}$$

Let m be the cardinality of A'_i, and let n be the cardinality of Θ'. We consider the $m \times n$ matrix $M \in \mathbb{R}^{A'_i \times \Theta'}$ given by $M(b_i, \theta) = u_i(a_i, \theta) - u_i(b_i, \theta)$. For generic u_i, the matrix M has full rank. By (28), the rows of M are linearly dependent. Thus, the rank of M must be n, the number of columns. We obtain that the row space of M has dimension n. Hence, we can find $\beta_i \in \mathbb{R}^{A'_i}$ such that for every $\theta \in \Theta'$ $[\sum_{b_i}(u_i(a_i,\theta) - u(b_i,\theta))\beta_i(b_i) < 0$. For every $t > 0$, we define $\alpha_i^t \in \mathbb{R}^{A'_i}$ by

$$\alpha_i^t(b_i) = \frac{\alpha_i(b_i) + t\beta_i(b_i)}{\sum_{c_i} \alpha_i(c_i) + t\beta_i(c_i)}.$$

For t sufficiently small, α_i^t is a mixed action that strictly dominates a_i. □

We are now ready to prove the proposition on strict BCE.

Proof of Proposition 5. Let A_i^* be the set of actions that are not strictly dominated. Since u_i is generic, it follows from Lemma 9 that each $a_i \in A_i^*$ is not weakly dominated by a mixed action $\alpha_i \in \Delta(A_i \setminus \{a_i\})$. By Lemma 8, there is a belief $\mu_{a_i} \in \Delta(\Theta)$ for which a_i is the unique best response.

Since π has full support, we can find $\nu \in \Delta(\Theta)$ and for every $a_i \in A_i^*$, $t_{a_i} \in (0,1)$—with $\sum_{a_i \in A_i^*} t_{a_i} \leq 1$—such that

$$\pi = \sum_{a_i \in A_i^*} t_{a_i}\mu_{a_i} + \left(1 - \sum_{a_i \in A_i^*} t_{a_i}\right)\nu.$$

Let a_i^* be a best response to ν; necessarily, $a_i^* \in A_i^*$. Define the outcome $p \in \Delta_\pi(A_i \times \Theta)$ as follows:

$$p(a_i,\theta) = \begin{cases} t_{a_i}\mu_{a_i}(\theta) & \text{if } a_i \in A_i^* \setminus \{a_i^*\}, \\ t_{a_i^*}\mu_{a_i^*}(\theta) + \left(1 - \sum_{a_i \in A_i^*} t_{a_i}\right)\nu(\theta) & \text{if } a_i = a_i^*, \\ 0 & \text{otherwise.} \end{cases}$$

The outcome p is a strict BCE. Moreover, if q is a BCE, then $\text{supp}_i(q) \subseteq A_i^* = \text{supp}_i(p)$. Thus, $\{sq + (1-s)p : s \in (0,1) \text{ and } q \in BCE\}$ is a subset of the set of strict BCE, and it is dense in the BCE set. We conclude that (for generic u_i) the set of strict BCE is dense in the BCE set. □

G. Proofs for Section 6

G.1. Proof of Proposition 1

First, we show that focusing on symmetric outcomes is without loss for welfare analysis in symmetric games (the assumption of binary actions has no role in this result).

Claim 7. For every BCE p, there is a symmetric BCE q such that $\bar{w}(q) = \bar{w}(p)$ and $\underline{w}(q) \leq \underline{w}(p)$.

Proof. Fix a BCE p. Let Φ be the set of permutations of I. For every permutation $\phi \in \Phi$, we define the outcome p_ϕ by $p_\phi(a, \theta) = p(a_\phi, \theta)$. Note that player i in p_ϕ behaves as player $j = \phi^{-1}(i)$ in p. One can verify that p_ϕ because p is a BCE and the game is symmetric.

We define the outcome q by $q = \frac{1}{|\Phi|} \sum_{\phi \in \Phi} p_\phi$, where $|\Phi|$ is the cardinality of Φ. As noted above, each p_ϕ is a BCE. Since the BCE set is convex, q is a BCE.

The outcome q is symmetric. Indeed, $\Phi = \{\psi^{-1} \circ \phi : \phi \in \Phi\}$ for every permutation $\psi \in \Phi$. We deduce that

$$\begin{aligned} q(a_\psi, \theta) &= \frac{1}{|\Phi|} \sum_{\phi \in \Phi} p_\phi(a_\psi, \theta) = \frac{1}{|\Phi|} \sum_{\phi \in \Phi} p_{(\psi^{-1} \circ \phi)}(a_\psi, \theta) \\ &= \frac{1}{|\Phi|} \sum_{\phi \in \Phi} p(a_\phi, \theta) = \frac{1}{|\Phi|} \sum_{\phi \in \Phi} p_\phi(a, \theta) = q(a, \theta). \end{aligned}$$

Hence, q is symmetric.

To conclude the proof, we observe $\bar{w}(q) = \frac{1}{|\Phi|} \sum_{\phi \in \Phi} \bar{w}(p_\phi) = \frac{1}{|\Phi|} \sum_{\phi \in \Phi} \bar{w}(p) = \bar{w}(p)$, where the first equality holds because $\bar{w}(p_\phi)$ is affine in p_ϕ, and the second equality because the game is symmetric. Finally, note that $\underline{w}(q) \leq \frac{1}{|\Phi|} \sum_{\phi \in \Phi} \underline{w}(p_\phi) = \frac{1}{|\Phi|} \sum_{\phi \in \Phi} \underline{w}(p) = \underline{w}(p)$, where the first inequality holds because $\underline{w}(p_\phi)$ is convex in p_ϕ, and the second equality because the game is symmetric. $\square$

By Claim 7, $\bar{w}$ is the value of the optimization problem

$$\min_{p \in BCE^{sy}} \bar{w}(p), \tag{29}$$

and $\underline{w}$ is the value of the optimization problem

$$\min_{p \in BCE^{sy}} \underline{w}(p). \tag{30}$$

We consider also the following optimization problem:

$$\min_{p \in \Delta_{\pi}^{sy}(A \times \Theta)} \underline{w}(p), \tag{31}$$

We proceed by successive claims.

Claim 8. The following conditions are equivalent: (i) $\underline{w} < \bar{w}$, and (ii) $\underline{v}_i(p) < \bar{v}_i(p)$ for all players i and optimal solutions p of (30).

Proof. We first prove that (i) implies (ii). Suppose $\underline{w} < \bar{w}$ and let p be an optimal solution of (30). Then, $\underline{w}(p) = \underline{w} < \bar{w} \leq \bar{w}(p)$. The inequality $\underline{w}(p) < \bar{w}(p)$ implies that $\underline{v}_i(p) < \bar{v}_i(p)$ for some player i. Since the game is symmetric and p is symmetric, $\underline{v}_i(p) < \bar{v}_i(p)$ for all players i.

We now prove that (ii) implies (i). Suppose $\underline{v}_i(p) < \bar{v}_i(p)$ for all players i and optimal solutions p of (30). Let $p \in BCE$ be an optimal solution of (29). If p is also an optimal solution of (30), then $\underline{w}(p) = \sum_i \underline{v}_i(p) < \sum_i \bar{v}_i(p) = \bar{w}(p)$ by hypothesis; thus, $\underline{w} < \bar{w}$. If instead p is not an optimal solution of (30), then $\underline{w} < \underline{w}(p) \leq \bar{w}(p) = \bar{w}$; thus, $\underline{w} < \bar{w}$. □

Claim 9. For every $p \in BCE$ and $i \in I$, the following conditions are equivalent: (i) $\underline{v}_i(p) < \bar{v}_i(p)$, and (ii) $a_i \in \text{supp}_i(p)$ and $BR(p_{a_i}) = \{a_i\}$ for all $a_i \in A_i$.

Proof. Condition (i) holds if and only if player i is strictly better by following the action recommendation of the mediator rather then best responding ex ante. In other terms, player i has no action a_i such that for all $b_i \in \text{supp}_i(p)$, $a_i \in BR(p_{b_i})$. Given that A_i has two elements, this is equivalent to condition (ii). □

Claim 10. The following conditions are equivalent: (i) all optimal solutions of (30) satisfy (5), and (ii) all optimal solutions of (31) satisfy (5).

Proof. First we show that (i) implies (ii). Let p be an optimal solution of (30) and let q be an optimal solution of (31). For every $t \in [0, 1]$, define $p^t = (1 - t)p + tq$. Furthermore, set $s = \max\{t : p^t \in BCE^{sy}\}$. Note that s is well defined: the set BCE^{sy} is closed and $p^0 = p \in BCE^{sy}$.

We observe that p^s is an optimal solution of (30): since $\underline{w}(p^t)$ is convex in t, $\underline{w}(p^s) \leq (1-s)\underline{w}(p) + s\underline{w}(q) \leq \underline{w}(p) = \underline{w}$. Thus, p^s must satisfy (5). But this implies that $s = 1$; otherwise, one could find $\epsilon > 0$ sufficiently small so that $p^{s+\epsilon} \in BCE^{sy}$, contradicting the definition of p^s. This implies that $q = p^s$ satisfies (5).

Now we show that (ii) implies (i). Let p be an optimal solution of (30) and let q be an optimal solution of (31). Since q satisfies (5), q is a BCE. Thus, q is an optimal solution of (30). This implies that p is an optimal solution of (31), and therefore satisfies (5). □

By combining the three claims above, we obtain Proposition 1.

G.2. Proof of Claim 1

We begin with a result that establishes a necessary condition for an outcome to solve the relaxed program from Proposition 1. To state the result, let $U_i(a_i, p)$ be player i's payoff if she always takes action a_i while (a_{-i}, θ) is distributed according to p:

$$U_i(a_i, p) = \sum_{b_i, a_{-i}, \theta} u_i(a_i, a_{-i}, \theta) p(b_i, a_{-i}, \theta).$$

Note that for all $p \in \Delta_\pi^{sy}(A \times \Theta)$ and $i \in I$, $\underline{w}(p) = n \max\{U_i(0,p), U_i(1,p)\}$. Thus,

$$\operatorname*{argmin}_{p \in \Delta_\pi^{sy}(A\times\Theta)} \underline{w}(p) = \operatorname*{argmin}_{p \in \Delta_\pi^{sy}(A\times\Theta)} \max\{U_i(0,p), U_i(1,p)\}.$$

Claim 11. Every $p^* \in \operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} \underline{w}(p)$ has $U_i(0,p^*) = U_i(1,p^*)$ for all $i \in I$.

Proof. We prove the contrapositive: if $p^* \in \Delta_\pi^{sy}(A\times\Theta)$ has $U_i(0,p^*) \neq U_i(1,p^*)$, then $p^* \notin \operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} \underline{w}(p)$.

We first consider the case in which $U_i(0,p^*) > U_i(1,p^*)$. Let q be the outcome where all investors always attack. Observe that $q \in \operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} U_i(0,p)$, and that every $r \in \operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} U_i(0,p)$ has the speculative attack succeeding with probability one. Hence, every such r has $U_i(0,r) < U_i(1,r)$, which implies that $p^* \notin \operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} U_i(0,p)$. We deduce that $U_i(0,q) < U_i(0,p^*)$.

For every $\epsilon \in (0,1)$, we define $q^\epsilon = \epsilon q + (1-\epsilon)p^* \in \Delta_\pi^{sy}(A\times\Theta)$. Using the inequality $U_i(0,q) < U_i(0,p^*)$, we obtain that for all $\epsilon > 0$ small enough,

$$\underline{w}(q^\epsilon) = nU_i(0,p^\epsilon) = n\left(\epsilon U_i(0,q) + (1-\epsilon)U_i(0,p^*)\right) < nU_i(0,p^*) = \underline{w}(p^*).$$

We conclude that $p^* \notin \operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} \underline{w}(p)$.

The argument for the case $U_i(0,p^*) < U_i(1,p^*)$ is similar, but with q being replaced by the outcome where no one ever speculates. $\square$

Thanks to Claim 11, to determine $\operatorname{argmin}_{p\in\Delta_\pi^{sy}(A\times\Theta)} \underline{w}(p)$, we can study the following "simpler" optimization problem:

$$\min_{p\in\Delta_\pi^{sy}(A\times\Theta)} U_i(0,p) \quad \text{s.t.} \quad U_i(0,p) = U_i(1,p). \tag{32}$$

Claim 12. An outcome $p \in \Delta_\pi^{sy}(A\times\Theta)$ is an optimal solution of (32) if and only if

$$p\left(\sum_{j\neq i} a_j = \theta - 1\right) = 0 \quad \text{and} \quad p\left(\sum_{j\neq i} a_j \geq \theta\right) = \frac{k}{1+x}.$$

Proof. Simple algebra shows that every p that satisfies $U_i(1,p) = U_i(0,p)$ must yield

$$U_i(0,p) = -xp\left(\sum_{j\neq i} a_j \geq \theta\right) = -x\left[\frac{k - p\left(\sum_{j\neq i} a_j = \theta-1\right)}{1+x}\right].$$

Therefore, we get that (32) is the same as

$$\min_{p\in\Delta_\pi^{sy}(A\times\Theta)} \quad p\left(\sum_{j\neq i} a_j = \theta - 1\right) \quad \text{s.t.} \quad p\left(\sum_{j\neq i} a_j \geq \theta\right) = \frac{k - p\left(\sum_{j\neq i} a_j = \theta-1\right)}{1+x}.$$

Hence, to complete the proof, we only need to be sure that there is $p \in \Delta_\pi^{sy}(A\times\Theta)$ such that

$$p\left(\sum_{j\neq i} a_j = \theta - 1\right) = 0 \quad \text{and} \quad p\left(\sum_{j\neq i} a_j \geq \theta\right) = \frac{k}{1+x}.$$

Such an outcome is easy to construct: with probability $k/1+x$, all players attack; with the remaining probability, no player attacks. $\square$

The following result connects what we have just found with the conditions in the statement of Claim 1.

Claim 13. For an outcome $p \in \Delta_\pi(A\times\Theta)$, the following conditions are equivalent:

(i) For all players i and payoff states θ,

$$p\left(\sum_{j\neq i} a_j = \theta - 1\right) = 0, \tag{33}$$

$$p\left(\sum_{j\neq i} a_j \geq \theta\right) = \frac{k}{1+x}. \tag{34}$$

(ii) For all payoff states θ,

$$p\left(\theta - 1 \leq \sum_{i} a_i \leq \theta\right) = 0, \tag{35}$$

$$p\left(\sum_{i} a_i > \theta\right) = \frac{k}{1+x}. \tag{36}$$

Proof. First we show that (i) implies (ii). Since $\max \Theta < n$,

$$p\left(\sum_i a_i = \theta - 1\right) = p\left(\sum_i a_i = \theta - 1, \text{ and } a_i = 0 \text{ for some } i\right).$$

Thus, $p\left(\sum_i a_i = \theta - 1\right) \leq \sum_i p\left(\sum_{j\neq i} a_j = \theta - 1, \text{ and } a_i = 0\right) = 0$, where the last equality follows from (33). Moreover, since $\min \Theta > 0$,

$$p\left(\sum_i a_i = \theta\right) = p\left(\sum_i a_i = \theta, \text{ and } a_i = 1 \text{ for some } i\right).$$

Thus, $p\left(\sum_i a_i = \theta\right) \leq \sum_i p\left(\sum_{j\neq i} a_j = \theta - 1, \text{ and } a_i = 1\right) = 0$, where the last equality follows from (33). We conclude that (35) holds.

To prove (36), notice that $p\left(\sum_i a_i > \theta\right) = p\left(\sum_i a_i \geq \theta\right)$, because we have just verified that $p\left(\sum_i a_i = \theta\right) = 0$. Then, fixing some player i^*,

$$\begin{aligned}
p\left(\sum_i a_i \geq \theta\right) &= p\left(\sum_{i\neq i^*} a_i \geq \theta, \text{ and } a_{i^*} = 0\right) + p\left(\sum_{i\neq i^*} a_i \geq \theta - 1, \text{ and } a_{i*} = 1\right) \\
&= p\left(\sum_{i\neq i^*} a_i \geq \theta, \text{ and } a_{i^*} = 0\right) + p\left(\sum_{i\neq i^*} a_i \geq \theta, \text{ and } a_{i^*} = 1\right) \\
&= p\left(\sum_{i\neq i^*} a_i \geq \theta\right) = \frac{k}{1+x},
\end{aligned}$$

where the second equality holds by (33), and the last equality by (34). We deduce (36). This completes the proof that (i) implies (ii).

Now we show that (ii) implies (i). Observe that

$$p\left(\sum_{j\neq i} a_j = \theta - 1\right) = p\left(\sum_j a_j = \theta - 1, \text{ and } a_i = 0\right) + p\left(\sum_j a_j = \theta, \text{ and } a_i = 1\right).$$

By (35), the right-hand side is equal to zero: we deduce (33). We obtain (34) from the following chain of equalitites:

$$\begin{aligned} p\left(\sum_{j\neq i} a_j \geq \theta\right) &= p\left(\sum_j a_j \geq \theta, \text{ and } a_i = 0\right) + p\left(\sum_j a_j > \theta, \text{ and } a_i = 1\right) \\ &= p\left(\sum_j a_j > \theta, \text{ and } a_i = 0\right) + p\left(\sum_j a_j > \theta, \text{ and } a_i = 1\right) \\ &= p\left(\sum_j a_j > \theta\right) = \frac{k}{1+x}, \end{aligned}$$

where the second equality follows from (35), and the last equality from (36). This completes the proof that (ii) implies (i). □

Combining the three results above, we obtain Claim 1.

G.3. Proof of Claim 2

First, we obtain necessary and sufficient conditions for $\underline{w} < \bar{w}$ in the regime change game for an arbitrary number of states.

Claim 14. The inequality $\underline{w} < \bar{w}$ holds if and only if all symmetric outcomes p that satisfy (6) and (7), also satisfy

$$p_{a_i=1}\left(\sum_j a_j \geq \theta\right) > \frac{k}{1+x}, \tag{37}$$

where $p_{a_i=1}$ is the conditional probability of (a_{-i}, θ) given $a_i = 1$.

Proof. By Proposition 1, the inequality $\underline{w} < \bar{w}$ holds if and only if all optimal solutions of $\min_{p\in\Delta_\pi^{sy}(A\times\Theta)} \underline{w}(p)$ satisfy (5). By Claim 1, the latter condition is equivalent to the following statement: all symmetric outcomes p that satisfy (6) and (7), also satisfy

(5). Next we verify that, for all symmetric outcomes p that satisfy (6) and (7), the conditions (5) and (37) are equivalent.

Let p be a symmetric outcome that satisfy (6) and (7). First, note (7) implies the attack succeeds with a probability strictly between 0 and 1, and so players must both attack and and not attack with positive probability due to symmetry. Hence $\mathrm{supp}_i(p) = \{0, 1\} = A_i$.

Given $\mathrm{supp}_i(p) = \{0, 1\}$, i's obedience constraints are strict when

$$p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta - 1\right) - k > -x p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta\right), \tag{38}$$

$$p_{a_i=0}\left(\sum_{j\neq i} a_j \geq \theta - 1\right) - k < -x p_{a_i=0}\left(\sum_{j\neq i} a_j \geq \theta\right). \tag{39}$$

By (6)—see also Claim 13—

$$p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta - 1\right) = p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta\right),$$

and

$$p_{a_i=0}\left(\sum_{j\neq i} a_j \geq \theta - 1\right) = p_{a_i=0}\left(\sum_{j\neq i} a_j \geq \theta\right).$$

Thus, (38) and (39) hold if and only if

$$p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta\right) > \frac{k}{1+x} > p_{a_i=0}\left(\sum_{j\neq i} a_j \geq \theta\right).$$

By (6) and (7)—see also Claim 13— $p\left(\sum_{j\neq i} a_j \geq \theta\right) = \frac{k}{1+x}$. Thus, by the law of total probability, (38) and (39) hold if and only if

$$p_{a_i=1}\left(\sum_{j} a_j \geq \theta\right) = p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta - 1\right) = p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta\right) > \frac{k}{1+x}.$$

Overall, we conclude that, for all symmetric outcomes p that satisfy (6) and (7), the conditions (5) and (37) are equivalent. □

Next we refine the characterization $\underline{w} < \bar{w}$ obtained in Claim 14. To state this refinement, denote the CDF of θ by $F(\theta) := \sum_{\theta' \leq \theta} \pi(\theta)$. Define also the cutoff θ^* by

$$\theta^* = \min\left\{\theta \in \Theta : F(\theta) \geq \frac{k}{1+x}\right\}.$$

Claim 15. The inequality $\bar{w} > \underline{w}$ holds if and only if

$$F(\theta^*)\left(\theta^* - \mathbb{E}[\theta|\theta \leq \theta^*]\right) < \frac{k}{1+x}\left(3 - \frac{3k}{1+x} + \theta^* - \mathbb{E}[\theta]\right). \tag{40}$$

Proof. By Claim 14, $\underline{w} < \bar{w}$ is equivalent to

$$\begin{aligned} \frac{k}{1+x} < \min_{p \in \Delta_\pi^{sy}(A\times\Theta)} & p_{a_i=1}\left(\sum_{j\neq i} a_j \geq \theta - 1\right) \\ \text{s.t.} \quad & (6) \text{ and } (7). \end{aligned} \tag{41}$$

Hence, showing (40) and (41) are equivalent is sufficient. To show this equivalence, we first characterize the unique solution to the program on the right hand side of (41). This solution gives the value of the program, which we then compare to $k/(1+x)$.

We begin with an alternative way of representing symmetric outcomes. This representation is based on the observation that an outcome $p \in \Delta_\pi(A\times\Theta)$ is symmetric if and only if, conditional on the state, all action profiles with the same number of attackers have the same probability. Consequently, $p \in \Delta_\pi^{sy}(A\times\Theta)$ if and only if there is $Q : \Theta \to \Delta(\{0,\ldots,n\})$ such that

$$p(a,\theta) = \binom{n}{\sum_j a_j} Q\left(\sum_j a_j \middle| \theta\right)\pi(\theta),$$

where $\binom{n}{\sum_j a_j}$ is the binomial coefficient. Thus, one can write

$$p(a_i = 1) = \sum_\theta \pi(\theta) \sum_{m=1}^{n} \frac{m}{n} Q(m|\theta).$$

Moreover, condition (6) is equivalent to $Q(\theta-1|\theta) = Q(\theta|\theta) = 0$. Therefore,

$$p\left(\sum_{j\neq i} a_j \geq \theta - 1 \text{ and } a_i = 1\right) = \sum_\theta \pi(\theta) \sum_{m \geq \theta+1} \frac{m}{n} Q(m|\theta),$$

and

$$p\left(\sum_j a_j \geq \theta\right) = \sum_\theta \pi(\theta) \sum_{m \geq \theta+1} Q(m|\theta).$$

Hence, letting

$$f(Q) = \frac{\sum_\theta \pi(\theta) \sum_{m\geq\theta+1} mQ(m|\theta)}{\sum_\theta \pi(\theta) \sum_{m=1}^n mQ(m|\theta)},$$

we can write the program on the right hand side of (40) as

$$\begin{aligned}
\min_{Q:\Theta\to\Delta(\{0,\ldots,n\})} & f(Q) \\
\text{s.t.} \quad & \sum_\theta \pi(\theta) \sum_{m\geq\theta+1} Q(m|\theta) = \frac{k}{1+x}, \\
& Q(\theta-1|\theta) = Q(\theta|\theta) = 0 \text{ for all } \theta.
\end{aligned} \tag{42}$$

Since the constraint set is compact and the objective continuous, the above program admits a solution, Q^*. We now use perturbation-based arguments to show Q^* must satisfy a few properties:

1. $Q^*(m|\theta) = 0$ whenever $m \notin \{\theta-2, \theta+1\}$: if $Q^*(m|\theta) > 0$ for $m > \theta+1$ (resp., $m < \theta - 2$), one can reduce the objective without violating the constraints by moving $\epsilon > 0$ mass from $Q^*(m|\theta)$ to $Q^*(\theta+1|\theta)$ (resp., $Q^*(\theta-2|\theta)$).

2. If $Q^*(\theta+1|\theta) > 0$, then $Q^*(\theta'+1|\theta') = 1$ for all $\theta' < \theta$: For a contradiction, suppose $Q^*(\theta+1|\theta) > 0$, but $Q^*(\theta'+1|\theta') < 1$ for some $\theta' < \theta$. For every $\epsilon > 0$, define the following perturbation Q^ϵ of Q:

$$Q^\epsilon(m|\hat{\theta}) = \begin{cases} Q^*(\theta+1|\theta) - \epsilon & \text{if } m = \theta+1, \hat{\theta} = \theta, \\ Q^*(\theta-2|\theta) + \epsilon & \text{if } m = \theta-2, \hat{\theta} = \theta, \\ Q^*(\theta'+1|\theta') + \epsilon\frac{\pi(\theta)}{\pi(\theta')} & \text{if } m = \theta'+1, \hat{\theta} = \theta', \\ Q^*(\theta'-2|\theta') - \epsilon\frac{\pi(\theta)}{\pi(\theta')} & \text{if } m = \theta'-2, \hat{\theta} = \theta', \\ Q^*(m|\hat{\theta}) & \text{otherwise.} \end{cases}$$

The contradiction assumption means Q^ϵ is feasible for all sufficiently small $\epsilon > 0$. Direct computation shows

$$\lim_{\epsilon\searrow 0} \frac{1}{\epsilon}\left(f(Q^\epsilon) - f(Q^*)\right) = \frac{\pi(\theta)(\theta'-\theta)}{\sum_{\hat{\theta}} \pi(\hat{\theta}) \sum_{m=1}^n mQ^*(m|\hat{\theta})} < 0,$$

contradicting the optimality of Q.

3. If $Q^*(\theta-2|\theta) > 0$, then $Q^*(\theta'-2|\theta') = 1$ for all $\theta' > \theta$: For a contradiction, suppose $Q^*(\theta-2|\theta) > 0$, but $Q^*(\theta'-2|\theta') < 1$ for some $\theta' > \theta$. For every $\epsilon > 0$, define the following perturbation Q^ϵ of Q:

$$Q^\epsilon(m|\hat{\theta}) = \begin{cases} Q^*(\theta-2|\theta) - \epsilon & \text{if } m = \theta-2, \hat{\theta} = \theta, \\ Q^*(\theta+1|\theta) + \epsilon & \text{if } m = \theta+1, \hat{\theta} = \theta, \\ Q^*(\theta'-2|\theta') + \epsilon\frac{\pi(\theta)}{\pi(\theta')} & \text{if } m = \theta'-2, \hat{\theta} = \theta', \\ Q^*(\theta'+1|\theta') - \epsilon\frac{\pi(\theta)}{\pi(\theta')} & \text{if } m = \theta'+1, \hat{\theta} = \theta', \\ Q^*(m|\hat{\theta}) & \text{otherwise.} \end{cases}$$

The contradiction assumption means Q^ϵ is feasible for all sufficiently small $\epsilon > 0$. Direct computation shows

$$\lim_{\epsilon \searrow 0} \frac{1}{\epsilon}\left(f(Q^\epsilon) - f(Q^*)\right) = \frac{\pi(\theta)(\theta - \theta')}{\sum_{\hat{\theta}} \pi(\hat{\theta}) \sum_{m=1}^n mQ(m|\hat{\theta})} < 0,$$

contradicting the optimality of Q^*.

The above conditions imply the optimal Q^* admits a cutoff $\tilde{\theta}$ such that $Q^*(\theta+1|\theta) = 1$ for all $\theta < \tilde{\theta}$, $Q^*(\theta-2|\theta) = 1$ for all $\theta > \tilde{\theta}$, and $Q^*(\{\tilde{\theta}+1, \tilde{\theta}-2\}|\tilde{\theta}) = 1$. Then, the constraint

$$\sum_\theta \pi(\theta) \sum_{m \geq \theta+1} Q(m|\theta) = \frac{k}{1+x}$$

pins down the optimum: we must have $\tilde{\theta} = \theta^*$, and

$$Q^*(\theta^*+1|\theta^*) = \frac{1}{\pi(\theta^*)}\left(\frac{k}{1+x} - F(\theta^*-1)\right).$$

Therefore, the inequality (41) becomes

$$\begin{aligned} \frac{k}{1+x} < f(Q^*) &= \frac{\sum_\theta \pi(\theta) \sum_{m \geq \theta+1} mQ^*(m|\theta)}{\sum_\theta \pi(\theta) \sum_{m=1}^n mQ^*(m|\theta)} \\ &= \frac{F(\theta^*)\mathbb{E}[\theta+1|\theta \leq \theta^*] - \left(F(\theta^*) - \frac{k}{1+x}\right)(\theta^*+1)}{\mathbb{E}[\theta] + \frac{k}{1+x} - 2\left(1 - \frac{k}{1+x}\right)}. \end{aligned}$$

Rearranging the above equation gives (40). □

Finally, we prove Claim 2 by specializing Claim 15 to two states.

Proof of Claim 2. We begin the proof by explicitly stating the implication of (40) for the binary case. In particular, we show $\underline{w} < \bar{w}$ if and only if one of the following two conditions hold:

(i) $\frac{k}{1+x} > \pi(\underline{\theta})$ and $\frac{k}{1+x} > \frac{1}{3}\pi(\underline{\theta})(\bar{\theta} - \underline{\theta})$.

(ii) $\frac{k}{1+x} \leq \pi(\underline{\theta})$ and $\frac{k}{1+x} < 1 - \frac{1}{3}\pi(\bar{\theta})(\bar{\theta} - \underline{\theta})$.

To prove the above, we consider two cases, depending on the value of θ^*:

- *Case 1:* $k/(1+x) > \pi(\underline{\theta})$. Then $\theta^* = \bar{\theta}$, and the inequality (40) specializes to

$$\bar{\theta} - \mathbb{E}[\theta] < \frac{k}{1+x}\left(3 - \frac{3k}{1+x} + \bar{\theta} - \mathbb{E}[\theta]\right).$$

  Substituting $\bar{\theta} - \mathbb{E}[\theta] = \pi(\underline{\theta})(\bar{\theta} - \underline{\theta})$ and rearranging gives

$$\left(1 - \frac{k}{1+x}\right)\pi(\underline{\theta})(\bar{\theta} - \underline{\theta}) < 3\frac{k}{1+x}\left(1 - \frac{k}{1+x}\right),$$

  which is equivalent to

$$\frac{1}{3}\pi(\underline{\theta})(\bar{\theta} - \underline{\theta}) < \frac{k}{1+x}.$$

  Thus, we have established (i) is sufficient for $\underline{w} < \bar{w}$, and necessary if $\pi(\underline{\theta}) \geq \frac{k}{1+x}$.

- *Case 2:* Suppose now $k/(1+x) \leq \pi(\underline{\theta})$. Then $\theta^* = \underline{\theta}$. Thus, the inequality (40) is now

$$0 < \frac{k}{1+x}\left(3 - \frac{3k}{1+x} + \underline{\theta} - \mathbb{E}[\theta]\right).$$

  Note $\underline{\theta} - \mathbb{E}[\theta] = -\pi(\bar{\theta})(\bar{\theta} - \underline{\theta})$. Therefore, the above inequality is equivalent to

$$\frac{k}{1+x} < 1 - \frac{1}{3}\pi(\bar{\theta})(\bar{\theta} - \underline{\theta}).$$

  Hence, (ii) is sufficient for $\underline{w} < \bar{w}$, and necessary if $\pi(\underline{\theta}) \leq \frac{k}{1+x}$.

Next, we argue that a violation of one of the claim's conditions implies that either (i) or (ii) above hold. Suppose first $\bar{\theta} - \underline{\theta} < 3$. In this case, $\frac{1}{3}\pi(\underline{\theta})(\bar{\theta} - \underline{\theta}) < \pi(\underline{\theta})$, and so (i) holds whenever $\frac{k}{1+x} > \pi(\underline{\theta})$. If $\frac{k}{1+x} \leq \pi(\underline{\theta})$, then (ii) holds, because

$$1 - \frac{1}{3}\pi(\bar{\theta})(\bar{\theta} - \underline{\theta}) > 1 - \pi(\bar{\theta}) = \pi(\underline{\theta}) \geq \frac{k}{1+x}.$$

Suppose now $\bar{\theta} - \underline{\theta} \geq 3$, but (9) fails. Then one of the following inequality chains must hold: either

$$\frac{k}{1+x} > \frac{1}{3}(\bar{\theta} - \underline{\theta})(1 - \pi(\bar{\theta})) = \frac{1}{3}(\bar{\theta} - \underline{\theta})\pi(\underline{\theta}) \geq \pi(\underline{\theta}),$$

or

$$\frac{k}{1+x} < 1 - \frac{1}{3}(\bar{\theta} - \underline{\theta})\pi(\bar{\theta}) \leq 1 - \pi(\bar{\theta}) = \pi(\underline{\theta}).$$

Either way, $\underline{w} < \bar{w}$ holds: the first inequality chain implies (i), whereas the second inequality chain implies (ii).

To conclude the proof, we show that the claim's condition must hold if neither (i) nor (ii) hold. Suppose first that $\frac{k}{1+x} > \pi(\underline{\theta})$, but (i) fails. Then

$$\frac{1}{3}(\bar{\theta} - \underline{\theta})(1 - \pi(\bar{\theta})) = \frac{1}{3}(\bar{\theta} - \underline{\theta})\pi(\underline{\theta}) \geq \frac{k}{1+x} > \pi(\underline{\theta}),$$

meaning $\bar{\theta} - \underline{\theta} \geq 3$, and the right inequality in (9) holds. For the left inequality, note that

$$1 - \frac{1}{3}(\bar{\theta} - \underline{\theta})\pi(\bar{\theta}) \leq 1 - \pi(\bar{\theta}) = \pi(\underline{\theta}) < \frac{k}{1+x}.$$

Suppose now $\frac{k}{1+x} \leq \pi(\underline{\theta})$, but (ii) fails. Then,

$$1 - \pi(\bar{\theta}) \geq \frac{k}{1+x} \geq 1 - \frac{1}{3}(\bar{\theta} - \underline{\theta})\pi(\bar{\theta}).$$

The right inequality above delivers the left inequality in (9). Moreover, the implied inequality between the left most expression and the right most expression implies

$$\frac{1}{3}(\bar{\theta} - \underline{\theta})\pi(\bar{\theta}) \geq \pi(\bar{\theta}),$$

and so $\bar{\theta} - \underline{\theta} \geq 3$. Finally, to get the right inequality in (9), notice that

$$\frac{k}{1+x} \leq 1 - \pi(\bar{\theta}) \leq \frac{1}{3}(\bar{\theta} - \underline{\theta})(1 - \pi(\bar{\theta})),$$

where the last inequality holds because $\bar{\theta} - \underline{\theta} \geq 3$. $\square$

H. Proofs for Section 7

H.1. Main Results

In this section, we prove Proposition 2 and Proposition 3. As a first step, we prove a basic lemma about best responses. In what follows, for $p \in \Delta(A \times \Theta)$, $a_i \in \text{supp}_i(p)$, and $b_i \in A_i$, take $u_i\left(b_i, p_{a_i}\right) \in \mathbb{R}$ to be

$$u_i\left(b_i, p_{a_i}\right) = \sum_{a_{-i},\theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

Lemma 10. *For every $t \in (0,1)$, $p, q \in BCE$, $i \in I$, and $a_i \in \text{supp}_i(p)$,*

$$BR\left((tp + (1-t)q)_{a_i}\right) \subseteq BR\left(p_{a_i}\right).$$

Proof. Take $b_i \in BR\left((tp + (1-t)q)_{a_i}\right)$. If $a_i \notin \text{supp}_i(q)$, then $(tp + (1-t)q)_{a_i} = p_{a_i}$, which immediately implies the desired result.

Suppose now that $a_i \in \text{supp}_i(q)$. Since $p, q \in BCE$, we have

$$u_i(a_i, p_{a_i}) \geq u_i(b_i, p_{a_i}) \quad \text{and} \quad u_i(a_i, q_{a_i}) \geq u_i(b_i, q_{a_i}).$$

Simple algebra shows that there exists $s \in (0,1)$ such that

$$(tp + (1-t)q)_{a_i} = sp_{a_i} + (1-s)q_{a_i}.$$

Since $b_i \in BR\left((tp + (1-t)q)_{a_i}\right)$, we obtain that

$$\begin{aligned} su_i(b_i, p_{a_i}) + (1-s)u_i(b_i, q_{a_i}) &= u_i(b_i, sp_{a_i} + (1-s)q_{a_i}) \\ &\geq u_i(a_i, sp_{a_i} + (1-s)q_{a_i}) \\ &= su_i(a_i, p_{a_i}) + (1-s)u_i(a_i, q_{a_i}). \end{aligned}$$

We conclude that $u_i(a_i, p_{a_i}) = u_i(b_i, p_{a_i})$ and $u_i(a_i, q_{a_i}) = u_i(b_i, q_{a_i})$. It follows from $p \in BCE$ that $b_i \in BR\left(p_{a_i}\right)$. $\square$

Next, we show that taking convex combinations of BCEs usually preserve the set of action recommendations that lead to different beliefs.

Lemma 11. *For every $p, q \in \Delta(A \times \Theta)$, $i \in I$, and $a_i, b_i \in \mathrm{supp}_i(p)$ with $p_{a_i} \neq p_{b_i}$, there are at most two $t \in (0,1)$ such that*

$$(tp + (1-t)q)_{a_i} = (tp + (1-t)q)_{b_i}. \tag{43}$$

Proof. Note that $t \in (0,1)$ is a solution of (43) if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\begin{aligned} &(tp(a_i, a_{-i}, \theta) + (1-t)q(a_i, a_{-i}, \theta)) \, (tp(b_i) + (1-t)q(b_i)) \\ =&(tp(b_i, a_{-i}, \theta) + (1-t)q(b_i, a_{-i}, \theta)) \, (tp(a_i) + (1-t)q(a_i)) \, . \end{aligned} \tag{44}$$

Each equation (44) is polynomial in t, with degree at most two. Since $p_{a_i} \neq p_{b_i}$, at least one such polynomial equation does not have degree zero and, therefore, has at most two solutions. We deduce that (43) has at most two solutions for $t \in (0,1)$. □

Our next goal is to show that minimally mixed BCEs are the norm rather than the exception. As an intermediate step, we first show the set of minimally mixed BCEs is non-empty.

Lemma 12. *A minimally mixed BCE exists.*

Proof. For every $p \in BCE$, define the set

$$X(p) = \bigcup_i \left\{(a_i, b_i) : a_i, b_i \in \mathrm{supp}_i(p) \text{ and } p_{a_i} \neq p_{b_i}\right\}.$$

Note that $p \in BCE$ is minimally mixed if and only if it has maximal support and for every $q \in BCE$, $X(q) \subseteq X(p)$.

Since the set $A \times \Theta$ is finite and BCE is a convex set, we can find a maximal support $p \in BCE$ such that for every maximal-support $q \in BCE$, the cardinality of $X(p)$ is larger than the cardinality of $X(q)$.

We now show that p is BCE-minimally mixed. Fix an arbitrary $q \in BCE$. For every $t \in (0,1)$, define $p^t = tp + (1-t)q$, which is a BCE because BCE is convex. Since p has maximal support, the same is true for p^t. Thus, the cardinality of $X(p)$ is larger than the cardinality of $X(p^t)$. By Lemma 11, we can find $t \in (0,1)$ such that $X(p) \subseteq X(p^t)$ and $X(q) \subseteq X(p^t)$. This shows that $X(q) \subseteq X(p)$; otherwise, the cardinality of $X(p^t)$ would be strictly larger than the cardinality of $X(p)$. We conclude that p is minimally mixed. $\square$

We now show that the minimally mixed BCEs includes most BCEs in a precise sense.

Lemma 13. *The set of minimally mixed BCEs is open and dense in the set of BCEs.*

Proof. Let P_M denote the set of minimally mixed BCEs. We first argue that P_M is open in BCE. Towards this goal, note the following sets are open in BCE for every $i \in I$ and $a_i, b_i \in A_i$:

$$\{p \in BCE : p(a_i) > 0\}, \quad \text{and} \quad \{p \in BCE : p(a_i)p(b_i) > 0 \text{ and } p_{a_i} \neq p_{b_i}\}.$$

Since A is finite, we obtain that P_M equals the intersection of a finite number of open subsets of P_M. It follows P_M is open in BCE.

To see P_M is dense in BCE, fix some $q \in BCE$. Take p to be a minimally mixed BCE, which exists by Lemma 12. For every $t \in (0,1)$, define $p^t = tp + (1-t)q$. Because p has maximal support, the same is true for p^t for all $t \in (0,1)$. Moreover, by Lemma 11, a finite set $T \subseteq (0,1)$ exists such that for all $t \in (0,1) \setminus T$, $i \in I$, and $a_i, b_i \in \text{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \text{implies} \quad p^t_{a_i} \neq p^t_{b_i}.$$

Thus, p^t is a minimally mixed BCE for all $t \in (0,1) \setminus T$. Thus, q is a limit point of $\{p^t : t \in (0,1) \setminus T\}$, which implies it is a limit point of P_M. $\square$

We are now ready to prove Proposition 2 and Proposition 3.

Proof of Proposition 2. That (i) implies (ii) follows from Lemma 13.

We now show (ii) implies (iii). Let q be a minimally mixed sBCE. Fix any $p \in BCE$, $i \in I$ and $a_i, b_i \in \text{supp}_i(p)$ such that $p_{a_i} \neq p_{b_i}$. Since q is minimally mixed, $a_i, b_i \in \text{supp}_i(q)$ (because q has maximal support) and $q_{a_i} \neq q_{b_i}$. Thus,

$$\varnothing = BR(q_{a_i}) \cap BR(q_{b_i}) \supseteq J(a_i) \cap J(b_i),$$

where we use the separation constraint, and then the fact that $J(c_i) = \cap_{\tilde{p} \in P} BR(\tilde{p}_{c_i})$ for all $c_i \in A_i$. We conclude (ii) implies (iii).

Finally, we argue (iii) implies (i). Fix any $p \in BCE$. Because A is finite and BCE is convex, it follows from Lemma 10 that we can find $q \in BCE$ such that q has maximal support and

$$BR(q_{a_i}) = J(q_{a_i}) \tag{45}$$

for all $i \in I$ and $a_i \in \mathrm{supp}_i(q)$.

For $t \in (0, 1)$, let $p^t = tp + (1 - t)q$. We claim that $p^t \in sBCE$. That $p^t \in BCE$ follows from convexity of the BCE set. To see p^t is a sBCE, take any $i \in I$ and $a_i, b_i \in \mathrm{supp}_i(p^t)$ such that $p^t_{a_i} \neq p^t_{b_i}$. Since q has maximal support, $a_i, b_i \in \mathrm{supp}_i(q)$. Then,

$$BR(p^t_{a_i}) \cap BR(p^t_{b_i}) \subseteq BR(q_{a_i}) \cap BR(q_{b_i}) = J(a_i) \cap J(b_i) = \varnothing,$$

where first we use Lemma 10, then (45), and finally Proposition 2-(iii). We conclude $p^t \in sBCE$ for all $t \in (0, 1)$. Proposition 2-(i) then follows from $p = \lim_{t \to 1} p^t$. □

Proof of Proposition 3. It is enough to prove that if $sBCE$ is not nowhere dense in BCE, then it is dense in BCE. Suppose $sBCE$ is dense in some non-empty set $\tilde{P} \subseteq BCE$ that is open in BCE. Let P_M the set of minimally mixed BCEs. Note $\tilde{P} \cap P_M$ is open (in BCE) and non-empty by Lemma 13. But $sBCE$ is dense in $\tilde{P}$, and so $sBCE \cap (\tilde{P} \cap P_M)$ must also be non-empty. Thus, we have found a minimally mixed sBCE. That $sBCE$ is dense in BCE then follows from Proposition 2. □

H.2. Checking for Equal Beliefs

To check the conditions of Proposition 2, knowing which actions induce different beliefs for some BCE is useful. In this section, we prove a result that shows how to find actions that lead to different beliefs in a closed convex set of outcomes $P \subseteq \Delta(A \times \Theta)$.[22]

For a player i, say an action a_i is **P-coherent** if a $p \in P$ exists with $p(a_i) > 0$.[23] Let $\mathbf{0}$ be the all-zeros vector in $\mathbb{R}^{A_{-i} \times \Theta}$; in what follows, we use the convention that $p_{a_i} = \mathbf{0}$ for every $p \in \Delta(A \times \Theta)$ and $a_i \in A_i$ such that $p(a_i) = 0$. We say an outcome $p \in P$ has **P-maximal support** if the support of every other $q \in P$ is contained by the support of p.

[22] Neither the obedience nor the separation constraint play any role in this section.

[23] Our notion of P-coherent is inspired by the notion of coherence in Nau and McCardle's "Coherent behavior in noncooperative games" (Journal of Economic Theory, vol. 50, pp. 424-444, 1990).

Proposition 6. *Fix a player i and two P-coherent actions $a_i, b_i \in A_i$. Then every $p \in P$ with $a_i, b_i \in \operatorname{supp}_i(p)$ has $p_{a_i} = p_{b_i}$ if and only if one of the following two conditions hold:*

(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $p \in \operatorname{ext}(P)$, $\{p_{a_i}, p_{b_i}\} \subseteq \{\mu, \boldsymbol{0}\}$.

(ii) A constant $\lambda > 0$ exists such that for all $p \in \operatorname{ext}(P)$, $p(a_i)p_{a_i} = \lambda p(b_i)p_{b_i}$.

Thus, to know whether a pair of actions leads to the same beliefs in all outcomes in P, it is enough to check the extreme points of P for one of two properties. The first property states these actions induce the same beliefs in all of the set's extreme points. The second property requires the likelihood ratio for these actions to be constant across all these extreme points.

To prove the proposition, we need the following lemma.

Lemma 14. *Fix a player i and two actions $a_i, b_i \in A_i$. Let $p, q \in \Delta(A \times \Theta)$ such that $\{a_i, b_i\} \subseteq \operatorname{supp}_i(p) \cup \operatorname{supp}_i(q)$. Suppose $r_{a_i} = r_{b_i}$ for all $r \in \{p, q\}$ with $\{a_i, b_i\} \subseteq \operatorname{supp}_i(r)$. If $(tp + (1-t)q)_{a_i} = (tp + (1-t)q)_{b_i}$ for some $t \in (0,1)$, then one of the following two conditions hold:*

(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $r \in \{p, q\}$, $\{r_{a_i}, r_{b_i}\} \subseteq \{\mu, \boldsymbol{0}\}$.

(ii) A constant $\lambda > 0$ exists such that for all $r \in \{p, q\}$, $r(a_i) = \lambda r(b_i)$.

Proof. Let $p^t := tp + (1-t)q$. We proceed by contradiction: we assume that Lemma 14-(i) and Lemma 14-(ii) both fail and show that $p^t_{a_i} \neq p^t_{b_i}$.

We begin by noting that one can rewrite the condition that $r_{a_i} = r_{b_i}$ for all $r \in \{p, q\}$ with $\{a_i, b_i\} \subseteq \operatorname{supp}_i(r)$ as

$$r(a_i)r(b_i)r_{a_i} = r(a_i)r(b_i)r_{b_i} \text{ for all } r \in \{p, q\}. \tag{46}$$

Because $\operatorname{supp}_i(p^t) = \operatorname{supp}_i(p) \cup \operatorname{supp}_i(q)$ and $\{a_i, b_i\} \subseteq \operatorname{supp}_i(p) \cup \operatorname{supp}_i(q)$, we have $\{a_i, b_i\} \subseteq \operatorname{supp}_i(p^t)$. Thus, applying Bayes rule, we obtain that $p^t_{a_i} = p^t_{b_i}$ if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, one has

$$p^t(a_i)p^t(b_i, a_{-i}, \theta) - p^t(b_i)p^t(a_i, a_{-i}, \theta) = 0.$$

Expanding the left hand side of the above equation by substituting in the definition of p^t, rearranging terms as a polynomial in t, and using (46), delivers that the above display equation is equivalent to

$$(t-t^2)\Big[p(a_i)q(b_i,a_{-i},\theta)+q(a_i)p(b_i,a_{-i},\theta)-q(b_i)p(a_i,a_{-i},\theta)-p(b_i)q(a_i,a_{-i},\theta)\Big]=0.$$

Since $t\in(0,1)$, we get that $p^t_{a_i}=p^t_{b_i}$ if and only if for every $a_{-i}\in A_{-i}$ and $\theta\in\Theta$, one has

$$p(a_i)q(b_i,a_{-i},\theta)+q(a_i)p(b_i,a_{-i},\theta)-q(b_i)p(a_i,a_{-i},\theta)-p(b_i)q(a_i,a_{-i},\theta)=0.$$

Writing the above in vector notation delivers that $p^t_{a_i}=p^t_{b_i}$ is equivalent to

$$p(a_i)q(b_i)q_{b_i}+q(a_i)p(b_i)p_{b_i}-q(b_i)p(a_i)p_{a_i}-p(b_i)q(a_i)q_{a_i}=\mathbf{0}. \tag{47}$$

We now divide the proof into cases. Consider first the case in which $\{a_i,b_i\}\subseteq \mathrm{supp}_i(p)\cap\mathrm{supp}_i(q)$. In this case, (46) implies $p_{a_i}=p_{b_i}$ and $q_{a_i}=q_{b_i}$, and so we get that

$$\begin{aligned}p(a_i)q(b_i)q_{b_i}+q(a_i)p(b_i)p_{b_i}-q(b_i)p(a_i)p_{a_i}-p(b_i)q(a_i)q_{a_i}&=\\ &=(p(a_i)q(b_i)-p(b_i)q(a_i))(q_{a_i}-p_{a_i})\neq\mathbf{0},\end{aligned}$$

where the inequality follows from failure of Lemma 14-(i) and Lemma 14-(ii). We conclude (47) fails.

Consider now the case in which $\{a_i,b_i\}\nsubseteq\mathrm{supp}_i(p)\cap\mathrm{supp}_i(q)$. Because Lemma 14-(ii) fails, we can assume $p(a_i)=0<p(b_i)$ without loss of generality. Since the lemma assume $a_i\in\mathrm{supp}_i(p)\cup\mathrm{supp}_i(q)$, it follows $q(a_i)>0$. Therefore, we can use failure of Lemma 14-(ii) to deduce that $p_{b_i}\neq q_{a_i}$. Using these facts, we obtain that

$$p(a_i)q(b_i)q_{b_i}+q(a_i)p(b_i)p_{b_i}-q(b_i)p(a_i)p_{a_i}-p(b_i)q(a_i)q_{a_i}=q(a_i)p(b_i)(p_{b_i}-q_{a_i})\neq\mathbf{0}.$$

It follows that (47) fails. □

We are now ready to prove Proposition 6. The "if" portion is straightforward; the "only if" portion uses Lemma 14.

Proof of Proposition 6. We first prove the "if" portion. Let $p\in P$ and $a_i,b_i\in\mathrm{supp}_i(p)$. Let $t^1,\dots,t^n>0$ and $p^1,\dots,p^n\in\mathrm{ext}(P)$ such that $p=\sum_{m=1}^n t^m p^m$.

Simple algebra shows that for all $c_i \in \mathrm{supp}_i(p)$

$$p_{c_i} = \sum_{m=1}^{n} \frac{t^m p^m(c_i)}{\sum_{l=1}^{n} t^l p^l(c_i)} p^m_{c_i}.$$

If Proposition 6-(i) holds, then

$$\begin{aligned} p_{a_i} = \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} p^m_{a_i} &= \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} \mu \\ &= \mu \\ &= \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} \mu = \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} p^m_{b_i} = p_{b_i}. \end{aligned}$$

Suppose now Proposition 6-(ii) holds. For every m, $p^m(a_i)p^m_{a_i} = \lambda p^m(b_i)p^m_{b_i}$ implies $p^m(a_i) = \lambda p^m(b_i)$ and $p^m_{a_i} = p^m_{b_i}$. Thus,

$$p_{a_i} = \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} p^m_{a_i} = \sum_{m=1}^{n} \frac{t^m \lambda p^m(b_i)}{\sum_{l=1}^{n} t^l \lambda p^l(b_i)} p^m_{b_i} = \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} p^m_{b_i} = p_{b_i}.$$

This concludes the proof of the proposition's "if" portion.

We now show the proposition's "only if" portion. We proceed by contradiction: we assume that Proposition 6-(i) and Proposition 6-(ii) both fail and show that there exists $p \in P$ such that $a_i, b_i \in \mathrm{supp}_i(p)$ and $p_{a_i} \neq p_{b_i}$. As we are done if $p_{a_i} \neq p_{b_i}$ for some $p \in \mathrm{ext}(P)$ with $a_i, b_i \in \mathrm{supp}_i(p)$, assume $p_{a_i} = p_{b_i}$ holds for all such p.

Since Proposition 6-(i) fails, and a_i and b_i are P-coherent, there exist $p, q \in \mathrm{ext}(P)$ such that $p(a_i) > 0$, $q(b_i) > 0$, and $p_{a_i} \neq q_{b_i}$. As we are done if $(0.5p + 0.5q)_{a_i} \neq (0.5p + 0.5q)_{b_i}$, assume $(0.5p + 0.5q)_{a_i} = (0.5p + 0.5q)_{b_i}$. Since $p_{a_i} \neq q_{b_i}$, Lemma 14-(i) fails. Thus, Lemma 14-(ii) must hold: there exist $\lambda > 0$ such that $p(a_i) = \lambda p(b_i)$ and $q(a_i) = \lambda q(b_i)$; in particular, $p(b_i) > 0$ and $q(a_i) > 0$.

Since Proposition 6-(ii) fails, there must exist $r \in \mathrm{ext}(P)$ such that $r(a_i) \neq \lambda r(b_i)$; in particular, $r(a_i) > 0$ or $r(b_i) > 0$. Let $c_i \in \{a_i, b_i\}$ such that $r(c_i) > 0$. Since $p_{a_i} \neq q_{b_i}$, either $r_{c_i} \neq p_{a_i}$, or $r_{c_i} \neq p_{b_i}$, or both. Thus, by Lemma 14, either $(0.5p + 0.5r)_{a_i} \neq (0.5p + 0.5r)_{c_i}$, or $(0.5q + 0.5r)_{b_i} \neq (0.5q + 0.5r)_{c_i}$, or both. In any case, we have found $p \in P$ such that $a_i, b_i \in \mathrm{supp}_i(p)$ and $p_{a_i} \neq p_{b_i}$. $\square$