Incomplete Information Robustness*

Stephen Morris[†] Takashi Ui[‡]

February 2025

Abstract

Consider an analyst who models a strategic situation using an incomplete information game. The true game may involve correlated, duplicated belief hierarchies, but the analyst lacks knowledge of the correlation structure and can only approximate each belief hierarchy. To make predictions in this setting, the analyst uses belief-invariant Bayes correlated equilibria (BIBCE) and seeks to determine which one is justifiable. We address this question by introducing the notion of robustness: a BIBCE is robust if, for every nearby incomplete information game, there exists a BIBCE close to it. Our main result provides a sufficient condition for robustness using a generalized potential function. In a supermodular potential game, a robust BIBCE is a Bayes Nash equilibrium, whereas this need not hold in other classes of games.

JEL classification: C72, D82.

Keywords: Bayes correlated equilibria, belief hierarchies, belief-invariance, generalized potentials, incomplete information games, potential games.

*This work is supported by Grant-in-Aid for Scientific Research Grant Number 18H05217 and National Science Foundation Grant #2001208. We are grateful for the comments of Rafael Veiel, Satoru Takahashi, Takashi Kunimoto, and seminar participants at the 12th Econometric Society World Congress, GAMES 2020, Singapore Management University, the University of Tokyo, and Keio University.

[†]Department of Economics, Massachusetts Institute of Technology.

[‡]Department of Economics, Kanagawa University and Hitotsubashi Institute for Advanced Study.

1 Introduction

When modeling a strategic situation as a game of incomplete information, an analyst must specify the types of players. Types can be decomposed into Mertens-Zamir belief hierarchies and an individually uninformative correlating device (Liu, 2015), as different types may have correlated, duplicated belief hierarchies (Ely and Peski, 2006; Dekel et al., 2007). This decomposition presents two key challenges. First, accurately identifying the true belief hierarchies can be difficult. Second, even if the belief hierarchies are correctly identified, the correlating device may remain unknown.

If these challenges cannot be resolved, what predictions about players' behavior are justifiable? We address this question by considering an analyst who has only approximate knowledge of belief hierarchies and lacks information about their potential duplication and correlation. The analyst assumes that, while the true model may differ from her model, it lies within its neighborhood. If a particular outcome of the analyst's model is qualitatively different from any outcome of some nearby game, as illustrated below, that outcome may not be justifiably adopted as a prediction. Then only outcomes that are (approximately) consistent with some outcomes of all nearby games can be considered justifiable, and we refer to such outcomes as robust.

Motivating example

Let the analyst's model be an incomplete information game with two players 1 and 2; two actions α and β for each player; and two payoff-relevant states θ_1 and θ_2 occurring with equal probability of $1/2$.[1] Each player's set of types is a singleton; that is, it is common knowledge that each player has a uniform prior over $\{\theta_1, \theta_2\}$. The payoffs are summarized in Table 1, where players would be better coordinated by choosing the same actions in state θ_1 and the different actions in state θ_2. This game has an infinite number of Bayes Nash equilibria (BNE) because the expected payoff for each player is $1/2$ under every action profile.

We construct a nearby game with a unique BNE, which is qualitatively different from any BNE in the analyst's model. Let θ_0 be an additional payoff-relevant state associated

[1] This game is discussed by Liu (2015).

θ_1	α	β
α	$1,1$	$0,0$
β	$0,0$	$1,1$

θ_2	α	β
α	$0,0$	$1,1$
β	$1,1$	$0,0$

θ_0	α	β
α	$1,0$	$1,1$
β	$0,1$	$0,0$

Table 1: Payoff matrices under θ_1, θ_2, and θ_0

with the rightmost payoffs in Table 1, where player 1 has a dominant action α, and player 2's best response is not to match player 1's action. Let $n \in \{0, 1, 2, \ldots\}$ be a random variable drawn from the geometric distribution $\Pr(n = k) = \varepsilon(1 - \varepsilon)^k$, which determines a payoff-relevant state as follows: θ_0 occurs if $n = 0$, θ_1 occurs if $n \geq 1$ is odd, and θ_2 occurs if $n \geq 2$ is even.

The types of players are defined by information partitions with respect to n, analogous to the email game (Rubinstein, 1989). Player 1's partition is $\{\{0\}, \{1, 2\}, \{3, 4\}, \{5, 6\}, \ldots\}$; player 2's partition is $\{\{0, 1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}, \ldots\}$. In the limit as $\varepsilon \to 0$, each player believes that either θ_1 or θ_2 occurs with equal probability $1/2$. The limit belief coincides with the belief in the analyst's model, so this game qualifies as a nearby game.

For any $\varepsilon > 0$, we can iteratively eliminate dominated actions as follows.

- For player 1 with $\{0\}$, α is a dominant action because the state is θ_0.
- For player 2 with $\{0, 1\}$, β is the unique best response because the state is θ_0 with probability $1/(2 - \varepsilon)$, and player 1 chooses α with probability at least $1/(2 - \varepsilon)$.
- For player 1 with $\{1, 2\}$, β is the unique best response because the state is θ_1 with probability $1/(2 - \varepsilon)$, and player 2 chooses β with probability at least $1/(2 - \varepsilon)$.

Repeating this, we obtain a unique BNE with the following action profiles that survive iterated elimination of dominated actions: (α, β) when $n \in \{4k\}_{k=0}^{\infty}$, (β, β) when $n \in \{4k + 1\}_{k=0}^{\infty}$, (β, α) when $n \in \{4k + 2\}_{k=0}^{\infty}$, and (α, α) when $n \in \{4k + 3\}_{k=0}^{\infty}$. In the limit as $\varepsilon \to 0$, players choose (α, α) and (β, β) with equal probability when θ_1 occurs, and (α, β) and (β, α) with equal probability when θ_2 occurs. This behavior is summarized in Table 2.

θ_1	α	β
α	$1/4$	0
β	0	$1/4$

θ_2	α	β
α	0	$1/4$
β	$1/4$	0

Table 2: Joint probabilities of actions and a state

The limit outcome in Table 2 is not a BNE of the analyst's model but a Bayes correlated equilibrium (BCE). Players receive recommendations to play one of their actions, randomly generated by the correlating device specified in Table 2. Following these recommendations is optimal because the expected payoff is one, which is the maximum ex-post payoff. Furthermore, the recommended action is individually uninformative, as it does not alter the players' prior beliefs about the state. Such a correlating device is said to be belief-invariant, and a BCE with a belief-invariant correlating device is referred to as a belief-invariant Bayes correlated equilibrium (BIBCE).

Main contribution

We develop a framework to study robust outcomes in incomplete information games and derive sufficient conditions for robustness. The concept of robustness to small perturbations of information structures originated with Kajii and Morris (1997), where the analyst's model is a complete information game. We extend this concept to incomplete information games, addressing issues inherent to the incompleteness of information, such as potential duplication and correlation of belief hierarchies.

To formally define a nearby game, we first consider a game, called an elaboration, that has a collection of belief hierarchies that coincides with that of the analyst's model, but may contain correlated, duplicated belief hierarchies. A nearby game is defined as a game in which the belief hierarchies are approximately those in an elaboration, and is referred to as an ε-elaboration, where ε measures the discrepancies.

We adopt BIBCE as a candidate for the analyst's prediction because, in the motivating example, the BIBCE in Table 2 is the only candidate for a robust outcome. A set of BIBCE is said to be robust if every ε-elaboration has a BIBCE that is close to some BIBCE in that set for sufficiently small ε. If the set consists of a single BIBCE, we simply say that the BIBCE is robust.

To derive a sufficient condition for robust BIBCE, we introduce a generalized potential function (Morris and Ui, 2005) for an incomplete information game. This function is defined on the Cartesian product of the set of states and a covering of each player's action set, i.e., a collection of action subsets whose union equals the action set. It contains information about each player's preferences over these action subsets. A potential function (Monderer and Shapley, 1996) is a special case in which each covering consists of singletons.

A generalized potential function is associated with a belief-invariant correlating device that assigns each player an action subset as a recommendation to choose an action from that subset. An important belief-invariant correlating device is one designed to maximize the expected value of a generalized potential function. For any such correlating device, there exists a BIBCE in which each player chooses an action from the recommended subset. This BIBCE is referred to as a GP-maximizing BIBCE.

Our main result, Theorem 1, establishes that the set of GP-maximizing BIBCE is robust. In particular, if a game is a potential game (i.e., a game that admits a potential function), the set of BIBCE that maximize the expected value of a potential function is robust. This implies that the BIBCE in the motivating example (shown in Table 2) is robust because the game is a potential game in which every player has the same payoff function. Recall that no BNE is robust in this example, even if it maximizes the expected value of a potential function. This contrasts with robustness in complete information games, where a unique potential-maximizing Nash equilibrium is always robust (Ui, 2001; Pram, 2019).

The main result is applied to supermodular games in Section 6. We address the question of when a BNE can be robust and show that if a game is a supermodular pontential game with a unique potential-maximizing BNE, then this BNE is robust. As a numerical example, we consider a binary-action supermodular game with two players, which is a global game (Carlsson and van Damme, 1993) with discretized signals and no dominance regions. This game has a robust BNE, where players always choose ex-post risk-dominant actions. Even in the absence of dominance regions, allowing a richer set of perturbations for the incomplete information game leads to the same outcome as the global game selection.

We also consider a special class of generalized potential functions known as monotone potential functions (Morris and Ui, 2005) and apply it to binary-action supermodular games. In the case of complete information, Oyama and Takahashi (2020) establish a fundamental link between the robustness of equilibria and monotone potentials: an action profile is robust if and only if it is a monotone potential maximizer in generic binary-action supermodular games. We extend this connection to our context of incomplete information.

Related literature

Our study builds on the literature addressing issues arising from the fact that type spaces (Harsanyi, 1967–1968) can embed more correlations than those captured by belief hierarchies (Mertens and Zamir, 1985). Ely and Peski (2006) provide an extended notion of belief hierarchies for two-player games and show that interim independent rationalizability depends on types solely through those hierarchies. Dekel et al. (2007) introduce the concept of interim correlated rationalizability and demonstrate that types with identical belief hierarchies lead to the same set of interim-correlated-rationalizable outcomes.

While these papers treat correlations implicitly, Liu (2015) explicitly introduces correlating devices. He shows that every incomplete information game is equivalent to the conjunction of a non-redundant game, where different types have different hierarchies, and a belief-invariant correlating device. This result implies that a BNE of a game is equivalent to a BIBCE of the non-redundant game in terms of outcomes.

The concept of BIBCE is introduced by Forges (2006) and Lehrer et al. (2010) under the additional restriction that action recommendations cannot depend on the state. A BIBCE is a special case of a BCE, which is introduced by Bergemann and Morris (2013, 2016) as an analyst's prediction when she cannot rule out players having additional information. The work of Bergemann and Morris (2013, 2016) also supports the use of BIBCE in our paper as a candidate for an analyst's prediction under insufficient and imprecise information.[2]

The use of BIBCE is an important distinction between our robustness framework and that of Kajii and Morris (1997), who use BNE as an equilibrium concept for nearby games. We allow correlated decisions in nearby games, which follows Pram (2019). He shows that, when the equilibrium concept in nearby games is an agent-normal form correlated equilibrium, the robustness of Kajii and Morris (1997) and its weaker version of Kajii and Morris (2020a) are equivalent.[3]

Several sufficient conditions for the robustness of Kajii and Morris (1997) have been discussed.[4] Morris and Ui (2005) introduce a generalized potential function and show that an equilibrium maximizing the generalized potential function is robust. A potential

[2]For further motivation to study BIBCE and its relationship to BCE, see Liu (2015) and Bergemann and Morris (2017). For its relation to other special cases of BCE, see Bergemann and Morris (2019).

[3]The weaker version is used by Ui (2001). The difference is shown by Takahashi (2020).

[4]See Kajii and Morris (2020a,b) and Oyama and Takahashi (2020) for recent developments.

function and a monotone potential function are special cases. In the case of binary-action supermodular games, Oyama and Takahashi (2020) show that a monotone potential function provides not only a sufficient but also a necessary condition for the robustness. Our sufficient condition for robust BIBCE extends these insights.

Building on the necessity argument in Oyama and Takahashi (2020), Morris et al. (2024) characterize "smallest equilibrium" implementation in binary-action supermodular games with incomplete information through information design. Their work extends arguments from the robustness literature to incomplete information settings, an approach shared with our paper. However, their focus differs from ours. While Morris et al. (2024) consider arbitrary information structures to induce desirable actions in binary-action supermodular games, our paper examines small perturbations in information structures to assess the robustness of outcomes in general games.

2 Model

2.1 Incomplete information games

Fix a finite set of players I and a finite set of actions A_i for each player $i \in I$. An incomplete information game (T, Θ, π, u) consists of the following elements (we simply refer to it as a game when there is no risk of confusion).

- $T = \prod_{i \in I} T_i$ is an at most countable set of type profiles, where T_i is the set of player i's types.
- $\Theta = \prod_{i \in I} \Theta_i$ is an at most countable set of payoff-relevant states, where player i's payoff function is determined by the i-th component $\theta_i \in \Theta_i$.
- $\pi \in \Delta(T \times \Theta)$ is a common prior.
- $u = (u_i)_{i \in I}$ is a payoff function profile, where $u_i : A \times \Theta \to \mathbb{R}$ is player i's payoff function such that $u_i(\cdot, \theta) = u_i(\cdot, \theta')$ if and only if $\theta_i = \theta'_i$.

Hereafter, we use $C = \prod_{i \in I} C_i$, $C_{-i} = \prod_{j \neq i} C_j$, $C_S = \prod_{i \in S} C_i$, and $C_{-S} = \prod_{i \notin S} C_i$ to denote the Cartesian products of $C_1, C_2, \ldots$ with generic elements $c \in C$, $c_{-i} \in C_{-i}$, $c_S \in C_S$, and $c_{-S} \in C_{-S}$, respectively.

Payoff functions are assumed to be bounded, i.e., $\sup_{i,a,\theta} |u_i(a,\theta)| < \infty$. For each $i \in I$, let $T_i^* \subseteq T_i$ and $\Theta_i^* \subseteq \Theta_i$ denote the sets of player i's types and payoff-relevant states on the support of π, respectively: $T_i^* = \{t_i \in T_i \mid \pi(t_i) > 0\}$ and $\Theta_i^* = \{\theta_i \in \Theta_i \mid \pi(\theta_i)\} > 0$, where $\pi(t_i) \equiv \sum_{t_{-i},\theta} \pi(t,\theta)$ and $\pi(\theta_i) \equiv \sum_{t,\theta_{-i}} \pi(t,\theta)$ are the marginal probabilities. Player i's belief is given by $\pi(t_{-i},\theta|t_i) \equiv \pi(t,\theta)/\pi(t_i)$ when his type is $t_i \in T_i^*$.

Let $\pi^* \in \Delta(T^* \times \Theta^*)$ and $u_i^* : A \times \Theta^* \to \mathbb{R}$ denote the restriction of π to $T^* \times \Theta^*$ and that of u_i to $A \times \Theta^*$, respectively. Note that $(T^*, \Theta^*, \pi^*, u^*)$ is the minimum representation of (T, Θ, π, u) because every player with every type on the support of π in (T, Θ, π, u) has the same belief and payoffs as those in $(T^*, \Theta^*, \pi^*, u^*)$. We say that two games (T, Θ, π, u) and $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$ with the same set of players and the same set of actions are equivalent if they have the same minimum representation.

A decision rule is a mapping $\sigma : T \times \Theta \to \Delta(A)$, under which players choose an action profile $a \in A$ with a probability $\sigma(a|t,\theta)$ when $(t,\theta) \in T \times \Theta$ is realized. Let Σ denote the set of all decision rules. A decision rule σ together with a common prior π determines a joint probability distribution $\sigma \circ \pi \in \Delta(A \times T \times \Theta)$ given by $\sigma \circ \pi(a,t,\theta) \equiv \sigma(a|t,\theta)\pi(t,\theta)$, which is referred to as a distributional decision rule. The set of all distributional decision rules, denoted by $\Sigma \circ \pi \equiv \{\sigma \circ \pi \in \Delta(A \times T \times \Theta) \mid \sigma \in \Sigma\}$, is readily shown to be a compact subset of a linear space $\{f \in \mathbb{R}^{A \times T \times \Theta} \mid \sum_{a,t,\theta} |f(a,t,\theta)| < \infty\}$ with the weak topology, which is metrizable.[5] Each distributional decision rule corresponds to an equivalence class of decision rules, where $\sigma, \sigma' \in \Sigma$ are equivalent if $\sigma \circ \pi(a,t,\theta) = \sigma' \circ \pi(a,t,\theta)$ for all $(a,t,\theta) \in A \times T^* \times \Theta^*$. When we discuss the topology of Σ, we identify Σ with $\Sigma \circ \pi$ by regarding Σ as the set of the equivalence classes and considering the isomorphism from the set of the equivalence classes to $\Sigma \circ \pi$.

2.2 Belief-invariant Bayes correlated equilibrium

A decision rule σ is said to be *obedient* for player i of type t_i if

$$\sum_{a_{-i},t_{-i},\theta} \sigma(a|t,\theta)\pi(t,\theta)u_i(a,\theta) \geq \sum_{a_{-i},t_{-i},\theta} \sigma(a|t,\theta)\pi(t,\theta)u_i((a_i',a_{-i}),\theta)$$

[5] See Lemma A in the appendix for more details.

for all $a_i, a_i' \in A_i$; that is, this player cannot increase the expected payoff by deviating from the action prescribed by the decision rule. In particular, if a decision rule σ is obedient for every player of every type, then σ is simply said to be obedient. An obedient decision rule is referred to as a Bayes correlated equilibrium (BCE) (Bergemann and Morris, 2013, 2016).

A decision rule is *belief-invariant* if $\sigma\left(\{a_i\} \times A_{-i} \middle| \left(t_i, t_{-i}\right), \theta\right)$ is independent of (t_{-i}, θ) for each $a_i \in A_i$, $t_i \in T_i$, and $i \in I$, or equivalently, there exists a player i's strategy $\sigma_i : T_i \to \Delta(A_i)$ such that $\sigma_i(a_i|t_i) = \sigma\left(\{a_i\} \times A_{-i} \middle| \left(t_i, t_{-i}\right), \theta\right)$. Belief-invariance implies that player i's action does not reveal any additional information to the player about the opponents' types and the state. In other words, from the viewpoint of player i who observes (a_i, t_i) as a random variable, t_i is a sufficient statistic for (t_{-i}, θ). This property has played an important role in the literature on incomplete information correlated equilibrium (Forges, 1993, 2006; Lehrer et al., 2010; Liu, 2015). Let Σ^{BI} denote the set of all belief-invariant decision rules. It is readily shown that Σ^{BI} is a compact subset of Σ.

Note that a strategy profile $(\sigma_i)_{i \in I}$, where $\sigma_i : T_i \to \Delta(A_i)$ is player i's strategy, is a special case of a belief-invariant decision rule given by $\sigma(a|t, \theta) = \prod_{i \in I} \sigma_i(a_i|t_i)$. A strategy profile $(\sigma_i)_{i \in I}$ is said to be a Bayes Nash equilibrium (BNE) if it is obedient. Clearly, a BNE is a special case of a belief-invariant BCE (BIBCE). Because a BNE exists in our setting with at most a countable number of states, types, and actions (Milgrom and Weber, 1985), a BIBCE also exists.[6] It is easy to check that the set of all BIBCE has the following property.

Lemma 1. *The set of all BIBCE of (T, Θ, π, u) is a nonempty convex compact subset of Σ containing all BNE, where a convex combination $\alpha\sigma + (1 - \alpha)\sigma'$ for $\sigma, \sigma' \in \Sigma^{BI}$ and $\alpha \in (0, 1)$ is given by $(\alpha\sigma + (1 - \alpha)\sigma')(a|t, \theta) = \alpha\sigma(a|t, \theta) + (1 - \alpha)\sigma'(a|t, \theta)$ for all $a \in A$ and $(t, \theta) \in T \times \Theta$.*

2.3 Elaborations

We introduce a correlating device that sends a signal to player i from a set M_i, which is at most countable. The probability distribution of a signal profile $m = (m_i)_{i \in I} \in M \equiv \prod_{i \in I} M_i$

[6] A BIBCE is also a solution to a linear programming problem with a countable number of variables and constraints.

is given by a mapping $\rho : T \times \Theta \to \Delta(M)$, under which $m \in M$ is drawn with a probability $\rho(m|t, \theta)$ when $(t, \theta) \in T \times \Theta$ is realized. This mapping ρ is referred to as a communication rule. Belief-invariance of a communication rule is defined similarly. That is, ρ is belief-invariant if $\rho(\{m_i\} \times M_{-i} | (t_i, t_{-i}), \theta)$ is independent of (t_{-i}, θ) for each $m_i \in M_i$, $t_i \in T_i$, and $i \in I$, which implies that a signal m_i does not reveal any additional information to player i about the opponents' types and the state.

Combining a game (T, Θ, π, u) and a communication rule ρ, we can construct another game $(\bar{T}, \Theta, \bar{\pi}, u)$ with the same sets of players and actions such that $\bar{T}_i = T_i \times M_i$ for each $i \in I$ and $\bar{\pi}(\bar{t}, \theta) = \pi(t, \theta)\rho(m|t, \theta)$ for each $\bar{t} = ((t_i, m_i))_{i \in I} \in \bar{T}$ and $\theta \in \Theta$. This game is referred to as the conjunction of (T, Θ, π, u) and ρ. Note that player i in $(\bar{T}, \Theta, \bar{\pi}, u)$ receives m_i as well as t_i, where t_i is drawn according to π and m_i is drawn according to ρ, and if ρ is belief-invariant, then this player's knowledge about (t_{-i}, θ) is exactly the same as that in the original game (T, Θ, π, u).

We consider a game that is isomorphic to the conjunction of (T, Θ, π, u) and a belief-invariant communication rule, which is referred to as an elaboration of (T, Θ, π, u).

Definition 1. A game $(\bar{T}, \Theta, \bar{\pi}, u)$ is an elaboration of (T, Θ, π, u) if there exist a belief-invariant communication rule ρ and mappings $\tau_i : \bar{T}_i \to T_i$ and $\mu_i : \bar{T}_i \to M_i$ for each $i \in I$ such that the mapping $\bar{t}_i \mapsto (\tau_i(\bar{t}_i), \mu_i(\bar{t}_i))$ restricted to $\bar{T}^*$ is one-to-one and

$$\bar{\pi}(\bar{t}, \theta) = \pi(\tau(\bar{t}), \theta)\rho(\mu(\bar{t})|\tau(\bar{t}), \theta) \text{ for all } \bar{t} \in \bar{T}, \tag{1}$$

where $\tau(\bar{t}) = (\tau_i(\bar{t}_i))_{i \in I}$ and $\mu(\bar{t}) = (\mu_i(\bar{t}_i))_{i \in I}$.

Note that if $(\bar{T}, \Theta, \bar{\pi}, u)$ is the conjunction of (T, Θ, π, u) and a belief-invariant communication rule ρ, it is an elaboration of (T, Θ, π, u) with τ_i and μ_i given by $\tau_i(t_i, m_i) = t_i$ and $\mu_i(t_i, m_i) = m_i$ for all $\bar{t}_i = (t_i, m_i) \in \bar{T}_i$ and $i \in I$.[7]

The following necessary and sufficient condition for an elaboration does not explicitly refer to a belief-invariant communication rule.

Lemma 2. *A game $(\bar{T}, \Theta, \bar{\pi}, u)$ is an elaboration of (T, Θ, π, u) if and only if, for each*

[7] Equation (1) implicitly requires that $\sum_{\bar{t} \in \tau^{-1}(t)} \rho(\mu(\bar{t})|t, \theta) = 1$ for all $(t, \theta) \in T^* \times \Theta^*$ and $\tau(\bar{T}^*) = T^*$ because $\bar{\pi}(\bar{T} \times \Theta) = 1$.

$i \in I$, there exists a mapping $\tau_i : \bar{T}_i \to T_i^*$ such that

$$\bar{\pi}(\tau^{-1}(t), \theta) = \pi(t, \theta) \text{ for all } (t, \theta) \in T^* \times \Theta^*, \tag{2}$$

$$\bar{\pi}(\tau_{-i}^{-1}(t_{-i}), \theta | \bar{t}_i) = \pi(t_{-i}, \theta | \tau_i(\bar{t}_i)) \text{ for all } \bar{t}_i \in \bar{T}_i^* \text{ and } (t_{-i}, \theta) \in T_{-i}^* \times \Theta^*, \tag{3}$$

where $\tau_i^{-1}(t_i) = \{\bar{t}_i \in \bar{T}_i \mid \tau_i(\bar{t}_i) = t_i\}$, $\tau^{-1}(t) = \prod_{i \in I} \tau_i^{-1}(t_i)$, and $\tau_{-i}^{-1}(t_{-i}) = \prod_{j \neq i} \tau_j^{-1}(t_j)$. The mapping $\tau = (\tau_i(\cdot))_{i \in I}$ is referred to as an elaboration mapping.

Note that a type $\bar{t}_i \in \bar{T}_i$ in an elaboration has the same belief hierarchy over Θ as that of a type $\tau_i(\bar{t}_i) \in T_i$ in the original game by (3). Thus, two types $\bar{t}_i, \bar{t}_i' \in \bar{T}_i$ with $\tau_i(\bar{t}_i) = \tau_i(\bar{t}_i')$ also have the same belief hierarchy over Θ.

2.4 Non-redundant games

We define the notion of a non-redundant game, where different types have different belief hierarchies. In (T, Θ, π, u), we say that $t_i, t_i' \in T_i$ have the same belief hierarchy if there exists another game (T^0, Θ, π^0, u) such that (T, Θ, π, u) is an elaboration of (T^0, Θ, π^0, u) with an elaboration mapping τ^0 satisfying $\tau_i^0(t_i) = \tau_i^0(t_i')$.[8] If all different types in (T^0, Θ, π^0, u) have different belief hierarchies, we say that (T^0, Θ, π^0, u) is a non-redundant representation of the original game (T, Θ, π, u).

Definition 2. A game (T^0, Θ, π^0, u) is a non-redundant representation of (T, Θ, π, u) if it satisfies the following two conditions: (i) (T, Θ, π, u) is an elaboration of (T^0, Θ, π^0, u) with an elaboration mapping τ^0; (ii) for $t_i, t_i' \in T_i^*$, $\tau_i^0(t_i) = \tau_i^0(t_i')$ if and only if t_i and t_i' have the same belief hierarchy. If (T, Θ, π, u) is a non-redundant representation of itself, then we simply say that (T, Θ, π, u) is a non-redundant game.

A non-redundant game[9] will serve as the analyst's model in the next section. Liu (2015)

[8] To state it formally, let $Z_i^1 = \Delta(\Theta)$, $Z_{-i}^{k-1} = \prod_{j \in I \setminus \{i\}} Z_j^{k-1}$, and $Z_i^k = \Delta(\Theta \times Z_{-i}^1 \times \cdots \times Z_{-i}^{k-1})$ for $k \geq 2$. For each $t_i \in T_i$, define $h_i^k(t_i) \in Z_i^k$ by $h_i^1(\theta | t_i) = \pi(\{(t_{-i}, \theta) \mid t_{-i} \in T_{-i}\} | t_i)$ for each $\theta \in \Theta$, and $h_i^k(\theta, \hat{h}_{-i}^1, \ldots, \hat{h}_{-i}^{k-1} | t_i) = \pi(\{(t_{-i}, \theta) \mid h_j^l(t_j) = \hat{h}_j^l \text{ for all } j \neq i \text{ and } l \in \{1, \ldots, k-1\}\} | t_i)$ for each $\theta \in \Theta$ and $\hat{h}_{-i}^l \in Z_{-i}^l$. The belief hierarchy of t_i is $h_i(t_i) \equiv (h_i^k(t_i))_{k=1}^{\infty}$. As shown by Liu (2015), $h_i(t_i) = h_i(t_i')$ if and only if (T, Θ, π, u) is an elaboration of another game (T^0, Θ, π^0, u) with an elaboration mapping τ^0 satisfying $\tau_i^0(t_i) = \tau_i^0(t_i')$.

[9] We can characterize a non-redundant game using belief hierarchies. Let $[t_i] \equiv \{t_i' \in T_i \mid h_i(t_i) = h_i(t_i')\}$ and $[t] = ([t_i])_{i \in I}$. Consider (T^0, Θ, π^0, u) with $T^0 = \{[t] \mid t \in T\}$ and $\pi^0([t], \theta) = \pi(\{(t', \theta) \mid t' \in [t]\})$. Then, it is straightforward to show that (T^0, Θ, π^0, u) is a non-redundant game of (T, Θ, π, u).

shows that if two games have the same set of belief hierarchies and one is non-redundant, then the other is an elaboration of the non-redundant game.[10]

A BIBCE of a non-redundant game can be understood as a BNE of its elaboration. To see this, let $(\bar{T}, \Theta, \bar{\pi}, u)$ be an elaboration of (T, Θ, π, u) with τ. For a decision rule σ of (T, Θ, π, u) and a decision rule $\bar{\sigma}$ of $(\bar{T}, \Theta, \bar{\pi}, u)$, we say that σ and $\bar{\sigma}$ are outcome equivalent if both decision rules induce the same probability distribution over $A \times T \times \Theta$:

$$\sigma \circ \pi(a, t, \theta) = \bar{\sigma} \circ \bar{\pi}(a, \tau^{-1}(t), \theta) \tag{4}$$

for all $(a, t, \theta) \in A \times T \times \Theta$, where $\bar{\sigma} \circ \bar{\pi}(a, \tau^{-1}(t), \theta) = \sum_{\bar{t} \in \tau^{-1}(t)} \bar{\sigma} \circ \bar{\pi}(a, \bar{t}, \theta)$. In particular, we say that σ and $\bar{\sigma}$ are equivalent if

$$\sigma(a|\tau(\bar{t}), \theta) = \bar{\sigma}(a|\bar{t}, \theta)$$

for all $(a, \bar{t}, \theta) \in A \times \bar{T} \times \Theta$. Clearly, if σ and $\bar{\sigma}$ are equivalent, then they are outcome equivalent, but not vice versa.

The following two lemmas are due to Liu (2015).[11] The first lemma shows that the set of BIBCE of (T, Θ, π, u) coincides with the set of BNE of all elaborations of (T, Θ, π, u) in terms of outcome equivalence. This implies that if the analyst uses a non-redundant game as her model but allows for all elaborations as the true model, then each BIBCE serves as a candidate for her prediction.

Lemma 3. *A decision rule σ is a BIBCE of (T, Θ, π, u) if and only if there exists an elaboration $(\bar{T}, \Theta, \bar{\pi}, u)$ with an elaboration mapping τ and a BNE $\bar{\sigma} = (\bar{\sigma}_i)_{i \in I}$ such that σ and $\bar{\sigma}$ are outcome equivalent, i.e., for all $(a, t, \theta) \in A \times T \times \Theta$,*

$$\sigma(a|t, \theta)\pi(t, \theta) = \sum_{\bar{t} \in \tau^{-1}(t)} \prod_{i \in I} \bar{\sigma}_i(a_i|\bar{t}_i)\bar{\pi}(\bar{t}, \theta).$$

Moreover, the set of BIBCE of (T, Θ, π, u) also coincides with the set of BIBCE of an arbitrary elaboration of (T, Θ, π, u) in terms of outcome equivalence, as shown by the second lemma. Thus, once BIBCE is adopted as the equilibrium concept for analysis, a

[10] Liu (2015) assumes a finite type space.

[11] For completeness, we provide the proofs in the supplementary material.

non-redundant game suffices for consideration, since every elaboration has the same set of BIBCE.

Lemma 4. *Let $(\bar{T}, \Theta, \bar{\pi}, u)$ be an elaboration of (T, Θ, π, u). If σ is a BIBCE of (T, Θ, π, u) and a decision rule $\bar{\sigma}$ of $(\bar{T}, \Theta, \bar{\pi}, u)$ is equivalent to σ, then $\bar{\sigma}$ is a BIBCE of $(\bar{T}, \Theta, \bar{\pi}, u)$. If $\bar{\sigma}$ is a BIBCE of $(\bar{T}, \Theta, \bar{\pi}, u)$ and a decision rule σ of (T, Θ, π, u) is outcome equivalent to $\bar{\sigma}$, then σ is a BIBCE σ of (T, Θ, π, u).*

3 Robustness

This section introduces a framework to construct nearby games of the analyst's model using elaborations and provides a formal definition of robust BIBCE.

Let (T, Θ, π, u) be a non-redundant game, which we regard as the analyst's model. The analyst believes that a belief-invariant communication rule may exist but has no information about it. Thus, the analyst predicts players' behavior using the outcomes of BNE of elaborations. By Lemma 3, such outcomes are BIBCE of (T, Θ, π, u).

We further assume that the analyst believes that the true game is in a "neighborhood" of elaborations of the analyst's model. We call such a nearby game an ε-elaboration, which is approximately an elaboration of (T, Θ, π, u). An elaboration is a special case of ε-elaboration with $\varepsilon = 0$.

Definition 3. For $\varepsilon \geq 0$, a game $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$ is an ε-elaboration of (T, Θ, π, u) if the following condition is satisfied.

1. $\Theta^* \subseteq \bar{\Theta}$ and $\bar{u}_i(\cdot, \theta) = u_i(\cdot, \theta)$ for all $\theta \in \Theta^*$.

2. $\bar{\pi}(\bar{T}^\sharp) \geq 1 - \varepsilon$, where $\bar{T}^\sharp = \prod_{i \in I} \bar{T}_i^\sharp$ and $\bar{T}_i^\sharp = \{\bar{t}_i \in \bar{T}_i \mid \bar{\pi}(\Theta_i^* \times \bar{\Theta}_{-i} | \bar{t}_i) = 1\}$.

3. There exist a mapping $\tau_i : \bar{T}_i \to T_i^*$ and $\bar{T}_i^\flat \subseteq \bar{T}_i$ with $\bar{\pi}(\bar{T}_i^\flat) \geq 1 - \varepsilon$ such that

$$\sup_{E \subseteq T \times \Theta} \Big| \sum_{(t,\theta) \in E} \bar{\pi}(\tau^{-1}(t), \theta) - \sum_{(t,\theta) \in E} \pi(t, \theta) \Big| \leq \varepsilon, \tag{5}$$

$$\sup_{E_{-i} \subseteq T_{-i} \times \Theta} \Big| \sum_{(t_{-i},\theta) \in E_{-i}} \bar{\pi}(\tau_{-i}^{-1}(t_{-i}), \theta | \bar{t}_i) - \sum_{(t_{-i},\theta) \in E_{-i}} \pi(t_{-i}, \theta | \tau_i(\bar{t}_i)) \Big| \leq \varepsilon \tag{6}$$

for all $\bar{t}_i \in \bar{T}_i^{\flat}$, where $\tau(\bar{t}) = (\tau_i(\bar{t}_i))_{i \in I}$ and $\tau_{-i}(\bar{t}_{-i}) = (\tau_j(\bar{t}_j))_{n \neq i}$. We call the mapping τ an elaboration mapping (with some abuse of terminology).

By the first condition, the set of payoff-relevant states in an ε-elaboration includes that in the original game. All players in an ε-elaboration have the same payoff functions as those in the original game with probability greater than $1 - \varepsilon$ by the second condition (recall that $\theta_i \in \Theta_i$ determines player i's payoff function). The third condition implies that an ε-elaboration is approximately an elaboration because, when $\varepsilon = 0$, (5) and (6) reduce to (2) and (3) in Lemma 2. It is straightforward to verify that the nearby game in the motivating example in the introduction satisfies the conditions in Definition 3.

Remark 1. Kajii and Morris (1997) define an ε-elaboration of a complete information game, which is (T, Θ, π, u) with T and Θ being singletons. Their ε-elaboration is an incomplete information game that satisfies the first two conditions of Definition 3. Thus, in the case of complete information, any ε-elaboration under our definition is also an ε-elaboration in the sense of Kajii and Morris (1997). See Section 7.1 for further details.

We are now ready to define robustness. A set of BIBCE of a non-redundant game is said to be robust if, for sufficiently small $\varepsilon > 0$, every ε-elaboration has a BIBCE that is close to some BIBCE in this set.[12]

Definition 4. Let (T, Θ, π, u) be a non-redundant game. A set of BIBCE of (T, Θ, π, u), $\mathcal{E} \subseteq \Sigma^{BI}$, is robust if, for every $\delta > 0$, there exists $\bar{\varepsilon} > 0$ such that, for all $\varepsilon \leq \bar{\varepsilon}$, every ε-elaboration $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$ with an elaboration mapping τ has a BIBCE $\bar{\sigma}$ such that

$$\sup_{E \subseteq A \times T \times \Theta} \left| \sum_{(a,t,\theta) \in E} \bar{\sigma} \circ \bar{\pi}(a, \tau^{-1}(t), \theta) - \sum_{(a,t,\theta) \in E} \sigma \circ \pi(a, t, \theta) \right| \leq \delta \tag{7}$$

for some $\sigma \in \mathcal{E}$, where $\bar{\sigma} \circ \bar{\pi}(a, \bar{t}, \theta) = \bar{\sigma}(a|\bar{t}, \theta)\bar{\pi}(\bar{t}, \theta)$ and $\sigma \circ \pi(a, t, \theta) = \sigma(a|t, \theta)\pi(t, \theta)$. If $\mathcal{E} = \{\sigma\}$ is a singleton, then σ is said to be robust.

To interpret (7), recall Lemma 4, which says every 0-elaboration has a BIBCE $\bar{\sigma}$ that is equivalent to a BIBCE σ of (T, Θ, π, u), where (4) holds for all $(a, t, \theta) \in A \times T \times \Theta$. Equation (7) means that (4) holds approximately for sufficiently small $\varepsilon > 0$ in the case of ε-elaborations.

[12] We can also consider robustness in "redundant" games. However, robust BIBCE in such games are outcome equivalent to those in non-redundant games, so it suffices to consider the non-redundant case.

Remark 2. Given the definition of robustness, the motivating example in Section 1 is summarized as follows and illustrates why we focus on BIBCE in our robustness concept. We have shown that there exists an ε-elaboration with a unique BNE. This BNE is close to the BIBCE in Table 2 when ε is small, but it is not close to any BNE of the analyst's model. This implies that the set of BNE for the analyst's model is not robust. This is why we adopt BIBCE as the equilibrium concept for the analyst's model. We also adopt BIBCE as the equilibrium concept for ε-elaborations. This is because there exists an ε-elaboration whose BNE is not the BIBCE in Table 2. For instance, the analyst's model is a 0-elaboration, where every BNE differs from the BIBCE in Table 2.

Remark 3. Robustness is defined in terms of a distributional decision rule $\sigma \circ \pi$. Thus, a set of BIBCE of (T, Θ, π, u) is robust if and only if the corresponding set of BIBCE of the minimum representation is robust. We use this observation to prove our main result.

Remark 4. In Kajii and Morris (1997), the analyst's model is a complete information game, and the solution concept for the model is a correlated equilibrium. A correlated equilibrium is defined to be robust if, for sufficiently small $\varepsilon > 0$, every ε-elaboration has a BNE that is close to the correlated equilibrium. See Section 7.1 for further details.

4 Generalized potentials

This section introduces a generalized potential function, which we use to derive a sufficient condition for robustness. The function is associated with a special class of communication rules, where each signal received by a player is a subset of the player's actions, interpreted as a recommendation to choose an action from this subset. We start with discussing such communication rules.

4.1 $\mathcal{A}$-Decision rule

As a set of signals for communication rules, let $\mathcal{A}_i \subset 2^{A_i} \backslash \emptyset$ be a covering of A_i for each $i \in I$; that is, $\mathcal{A}_i$ is a collection of nonempty subsets of A_i such that $\bigcup_{X_i \in \mathcal{A}_i} X_i = A_i$. Each signal $X_i \in \mathcal{A}_i$ vaguely prescribes an action in X_i; that is, it recommends that player i choose an action from X_i. We write $\mathcal{A} = \{X \mid X = \prod_{i \in I} X_i,\ X_i \in \mathcal{A}_i\}$ and $\mathcal{A}_{-i} = \{X_{-i} \mid X_{-i} = \prod_{j \neq i} X_j,\ X_j \in \mathcal{A}_j\}$.

For example, let $A_i = \{0, 1\}$ and $\mathcal{A}_i = \{\{0, 1\}, \{1\}\}$ for each $i \in I$. If player i receives $X_i = \{1\}$, the recommendation is to choose action 1. If player i receives $X_i = \{0, 1\}$, the recommendation allows the player to freely choose an action. This example will be discussed in Section 6.2.

We can describe both a decision rule and a communication rule using a single mapping $\gamma : T \times \Theta \to \Delta(A \times \mathcal{A})$ such that $\gamma(a, X|t, \theta) = 0$ whenever $a \notin X$, which assigns a joint probability distribution over $(a, X) \in A \times \mathcal{A}$ to each $(t, \theta) \in T \times \Theta$. This mapping is referred to as an $\mathcal{A}$-decision rule, under which player i receives a signal X_i as a vague recommendation and selects an action a_i from X_i.

We say that an $\mathcal{A}$-decision rule γ is obedient if the corresponding decision rule in the conjunction is obedient; that is, for each $i \in I$ and $(t_i, X_i) \in T_{-i} \times \mathcal{A}_{-i}$, it holds that

$$\sum_{a_{-i}, X_{-i}, t_{-i}, \theta} \gamma(a, X|t, \theta)\pi(t, \theta)u_i(a, \theta) \geq \sum_{a_{-i}, X_{-i}, t_{-i}, \theta} \gamma(a, X|t, \theta)\pi(t, \theta)u_i((a_i', a_{-i}), \theta) \tag{8}$$

for all $a_i, a_i' \in A_i$. Belief-invariance of an $\mathcal{A}$-decision rule is defined similarly: γ is belief-invariant if $\gamma(\{a_i\} \times A_{-i} \times \{X_i\} \times \mathcal{A}_{-i}|t, \theta)$ is independent of (t_{-i}, θ) for all $(a_i, X_i) \in A_i \times \mathcal{A}_i$, $t_i \in T_i$, and $i \in I$. An obedient belief-invariant $\mathcal{A}$-decision rule γ is referred to as a BIBCE, with some abuse of terminology, since the "decision rule" component of γ is a BIBCE, as demonstrated by the next lemma.

Lemma 5. *If an $\mathcal{A}$-decision rule γ is a BIBCE, then $\sigma \in \Sigma^{BI}$ with $\sigma(a|t, \theta) = \gamma(\{a\} \times \mathcal{A}|t, \theta)$ for all $(a, t, \theta) \in A \times T \times \Theta$ is a BIBCE of (T, Θ, π, u).*

4.2 Generalized potentials and BIBCE

A generalized potential function (Morris and Ui, 2005) is defined as a function over $\mathcal{A} \times \Theta$, which contains information about players' preferences.

Definition 5. A bounded function $F : \mathcal{A} \times \Theta \to \mathbb{R}$ is a generalized potential function of (T, Θ, π, u) if, for each $i \in I$ and $P_i \in \Delta(A_{-i} \times \mathcal{A}_{-i} \times \Theta)$ such that $P_i(A_{-i} \times \mathcal{A}_{-i} \times \Theta^*) = 1$ and $P_i(a_{-i}, X_{-i}, \theta) = 0$ whenever $a_{-i} \notin X_{-i}$,

$$X_i \in \arg\max_{X_i' \in \mathcal{A}_i} \sum_{X_{-i}, \theta} P_i(X_{-i}, \theta) F((X_i', X_{-i}), \theta) \tag{9}$$

implies

$$X_i \cap \arg\max_{a_i' \in A_i} \sum_{a_{-i},\theta} P_i(a_{-i},\theta) u_i((a_i', a_{-i}),\theta) \neq \emptyset, \tag{10}$$

where $P_i(X_{-i},\theta) = \sum_{a_{-i}\in A_{-i}} P_i(a_{-i}, X_{-i},\theta)$ and $P_i(a_{-i},\theta) = \sum_{X_{-i}\in\mathcal{A}_{-i}} P_i(a_{-i}, X_{-i},\theta)$. The value of $F(X,\theta)$ for $\theta \notin \Theta^*$ can be arbitrary.

A probability distribution P_i is typically derived from an $\mathcal{A}$-decision rule γ as player i's belief over $A_{-i} \times \mathcal{A}_{-i} \times \Theta$ when the opponents' actions and signals follow γ:

$$P_i(a_{-i}, X_{-i}, \theta) = \sum_{t_{-i}} \sum_{a_i, X_i} \gamma(a, X|t,\theta)\pi(t_{-i},\theta|t_i).$$

Suppose that player i receives a recommendation $X_i \in \mathcal{A}_i$ that satisfies (9) under this belief; that is, X_i maximizes the expected value of F. Condition (10) then requires that at least one action in X_i maximizes the expected value of player i's payoff function.

At one extreme, let $\mathcal{A}_i = \{A_i\}$ for each $i \in I$, where a signal contains no information. Clearly, every incomplete information game has a generalized potential with this domain.

At the other extreme, let $\mathcal{A}_i = \{\{a_i\} \mid a_i \in A_i\}$ for each $i \in I$, where each signal recommends a single action. A potential function (Monderer and Shapley, 1996) is a generalized potential function with this domain. A function $v : A \times \Theta \to \mathbb{R}$ is a potential function of (T, Θ, π, u) if there exists $q_i : A_{-i} \times \Theta \to \mathbb{R}$ such that

$$u_i(a,\theta) = v(a,\theta) + q_i(a_{-i},\theta) \tag{11}$$

for all $i \in I$, $a \in A$, and $\theta \in \Theta^*$. The next lemma shows that a potential function is a special case of a generalied potential function.

Lemma 6. *Suppose that (T, Θ, π, u) has a potential function $v : A \times \Theta \to \mathbb{R}$. Then, $F : \mathcal{A} \times \Theta \to \mathbb{R}$ with $\mathcal{A}_i = \{\{a_i\} \mid a_i \in A_i\}$ for each $i \in I$ and $F(\{a\},\theta) = v(a,\theta)$ for each $(a,\theta) \in A \times \Theta$ is a generalized potential function.*

One of the key properties of a potential function is that if a strategy profile maximizes its expected value, then it constitutes a BNE. A similar result also holds for BIBCE, even for generalized potential functions.

To see this, fix a generalized potential function F with a domain $\mathcal{A}$. Let Γ^{BI} denote the set of all belief-invariant $\mathcal{A}$-decision rules. For each $\gamma \in \Gamma^{BI}$, let $\gamma(a|t,\theta)$ and $\gamma(X|t,\theta)$ denote the conditional marginal probabilities of a and X, respectively; that is, $\gamma(a|t,\theta) = \sum_{X\in\mathcal{A}} \gamma(a, X|t,\theta)$ and $\gamma(X|t,\theta) = \sum_{a\in A} \gamma(a, X|t,\theta)$.

Let $\Gamma^F \subset \Gamma^{BI}$ be the set of belief-invariant $\mathcal{A}$-decision rules that maximize the expected value of a generalized potential function F:

$$\Gamma^F \equiv \arg\max_{\gamma\in\Gamma^{BI}} \sum_{X,t,\theta} \gamma(X|t,\theta)\pi(t,\theta)F(X,\theta).$$

The next lemma shows that Γ^F contains a BIBCE.

Lemma 7. *There exists a BIBCE $\gamma \in \Gamma^F$.*

If $\gamma \in \Gamma^F$ is a BIBCE, then $\sigma \in \Sigma^{BI}$ with $\sigma(a|t,\theta) = \gamma(a|t,\theta)$ is a BIBCE by Lemma 5. Such a BIBCE is referred to as a GP-maximizing BIBCE. We denote the set of all GP-maximizing BIBCE by

$$\mathcal{E}^F \equiv \{\sigma \in \Sigma^{BI} \mid \gamma \in \Gamma^F \text{ is a BIBCE and } \sigma(a|t,\theta) = \gamma(a|t,\theta)\}.$$

5 Main results

5.1 Robustness of GP-maximizing BIBCE

The following main result of this paper shows that $\mathcal{E}^F$ is robust.

Theorem 1. *If a non-redundant game (T, Θ, π, u) has a generalized potential function $F : \mathcal{A} \times \Theta \to \mathbb{R}$, then $\mathcal{E}^F$ is nonempty and robust.*

Every game admits a generalized potential function with the domain $\mathcal{A} = \{\{A\}\}$, in which case $\mathcal{E}^F$ is the set of all BIBCE. Thus, if a BIBCE is unique, which is the only element of $\mathcal{E}^F$, it is robust.

Corollary 2. *If a non-redundant game (T, Θ, π, u) has a unique BIBCE, then it is robust.*

This result is analogous to the finding in Kajii and Morris (1997) that a unique correlated equilibrium of a complete information game is robust.

If a game has a potential function $v : A \times \Theta \to \mathbb{R}$, it has a generalized potential function $F : \mathcal{A} \times \Theta \to \mathbb{R}$ with $\mathcal{A} = \{\{a\} \mid a \in A\}$ and $F(\{a\}, \theta) = v(a, \theta)$ by Lemma 6. In this case, it is straightforward to show that the set of all GP-maximizing BIBCE coincides with the set of all potential maximizing (P-maximizing, henceforth) belief-invariant decision rules:

$$\mathcal{E}^F = \mathcal{E}^v = \Sigma^v \equiv \arg\max_{\sigma \in \Sigma^{BI}} \sum_{a,t,\theta} \sigma(a|t,\theta)\pi(t,\theta)v(a,\theta).$$

Thus, we obtain the following corollary of Theorem 1.

Corollary 3. *If a non-redundant game (T, Θ, π, u) has a potential function $v : A \times \Theta \to \mathbb{R}$, then Σ^v is nonempty and robust.*

We apply this corollary to the motivating example in the introduction. The game has a potential function that is identical to the payoff function in Table 1. Thus, the expected value of the potential function with respect to the probability distribution in Table 2 is equal to one, which no other distribution can achieve. This implies that the BIBCE given by Table 2 uniquely maximizes the potential function over all decision rules. Consequently, by Corollary 3, the BIBCE is robust.

5.2 Proof of Theorem 1

It suffices to show that every sequence of ε^k-elaborations $\{(\bar{T}^k, \bar{\Theta}^k, \bar{\pi}^k, \bar{u}^k)\}_{k=1}^{\infty}$ with $\lim_{k\to\infty} \varepsilon^k = 0$ has a sequence of BIBCE $\{\bar{\sigma}^k\}_{k=1}^{\infty}$ such that

$$\lim_{k\to\infty} \inf_{\sigma\in\mathcal{E}^F} \sup_{E\subseteq A\times T\times\Theta} \left| \sum_{(a,t,\theta)\in E} \bar{\sigma}^k \circ \bar{\pi}^k(a, (\tau^k)^{-1}(t), \theta) - \sum_{(a,t,\theta)\in E} \sigma \circ \pi(a,t,\theta) \right| = 0,$$

where τ^k is an an elaboration mapping of $(\bar{T}^k, \bar{\Theta}^k, \bar{\pi}^k, \bar{u}^k)$.

We first show that a weaker version of the above condition holds. The following lemma lemma focuses on a special sequence of ε-elaborations that share the same sets of types, states, and payoff functions, differing only in their priors. Such a sequence of ε-elaborations is shown to admit a corresponding sequence of BIBCE that converges to some GP-maximizing BIBCE as ε approaches zero.

Lemma 8. *Let (T, Θ, π, u) be an incomplete information game with a generalized potential function $F : \mathcal{A} \times \Theta \to \mathbb{R}$. Then, every sequence of ε^k-elaborations $\{(T, \Theta, \pi^k, u)\}_{k=1}^{\infty}$ satisfying $\lim_{k\to\infty} \varepsilon^k = 0$ and sharing a common elaboration mapping $\tau : T \to T^*$ has a sequence of BIBCE $\{\sigma^k\}_{k=1}^{\infty}$ such that*

$$\lim_{k\to\infty} \inf_{\sigma\in\mathcal{E}^F} \sum_{(a,t,\theta)\in A\times T\times\Theta} \left|\sigma^k \circ \pi^k(a, \tau^{-1}(t), \theta) - \sigma \circ \pi(a, t, \theta)\right| = 0. \tag{12}$$

In the following proof of Theorem 1, we show that any sequence of ε-elaborations can be transformed into an equivalent special sequence considered in Lemma 8. We then apply Lemma 8 to the transformed sequence, thereby establishing the robustness of GP-maximizing BIBCE.

Proof of Theorem 1. Let $\{(\bar{T}^k, \bar{\Theta}^k, \bar{\pi}^k, \bar{u}^k)\}_{k=1}^{\infty}$ be a sequence of ε^k-elaborations such that $\lim_{k\to\infty} \varepsilon^k = 0$. Without loss of generality, assume that $\bar{T}^k \cap T = \emptyset$, $\bar{T}^k \cap \bar{T}^l = \emptyset$, and $\bar{\Theta}^k \cap \bar{\Theta}^l = \Theta^*$ for all $k \neq l$. Define $\bar{T} = T \cup \left(\bigcup_{k=1}^{\infty} \bar{T}^k\right)$ and $\bar{\Theta} = \bigcup_{k=1}^{\infty} \bar{\Theta}^k$, which are countable.

We construct an equivalent sequence of ε^k-elaborations $\{(\bar{T}, \bar{\Theta}, \bar{\lambda}^k, \bar{u})\}_{k=1}^{\infty}$. Let $\bar{\lambda}^k \in \Delta(\bar{T} \times \bar{\Theta})$ be an extension of $\bar{\pi}^k$ to $\bar{T} \times \bar{\Theta}$: $\bar{\lambda}^k(\bar{t}, \bar{\theta}) = \bar{\pi}^k(\bar{t}, \bar{\theta})$ if $(\bar{t}, \bar{\theta}) \in \bar{T}^k \times \bar{\Theta}^k$ and $\bar{\lambda}^k(\bar{t}, \bar{\theta}) = 0$ otherwise. Let $\bar{u}_i : A \times \bar{\Theta} \to \mathbb{R}$ be such that $\bar{u}_i(\cdot, \theta) = u_i(\cdot, \theta)$ if $\theta \in \Theta^*$ and $\bar{u}_i(\cdot, \bar{\theta}) = \bar{u}_i^k(\cdot, \bar{\theta})$ if $\bar{\theta} \in \bar{\Theta}^k \setminus \Theta^*$ for each $i \in I$. Given an elaboration mapping τ^k of $(\bar{T}^k, \bar{\Theta}^k, \bar{\pi}^k, \bar{u}^k)$, an elaboration mapping τ of $(\bar{T}, \bar{\Theta}, \bar{\lambda}^k, \bar{u})$ is defined as $\tau_i(\bar{t}_i) = \tau_i^k(\bar{t}_i)$ if $\bar{t}_i \in \bar{T}_i^k$ for each $i \in I$. Clearly, $(\bar{T}, \bar{\Theta}, \bar{\lambda}^k, \bar{u})$ and $(\bar{T}^k, \bar{\Theta}^k, \bar{\pi}^k, \bar{u}^k)$ have the same minimum representation (every player of every type on the common support has the same belief and the same payoffs in both games), and $(\bar{T}, \bar{\Theta}, \bar{\lambda}^k, \bar{u})$ is also an ε^k-elaboration of (T, Θ, π, u).

We introduce another incomplete information game $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$, where $\bar{\pi} \in \Delta(\bar{T} \times \bar{\Theta})$ is an extension of π to $\bar{T} \times \bar{\Theta}$, i.e., $\bar{\pi}(t, \theta) = \pi(t, \theta)$ if $(t, \theta) \in T \times \Theta$ and $\bar{\pi}(t, \theta) = 0$ otherwise. Note that $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$ and (T, Θ, π, u) have the same minimum representation, and $(\bar{T}, \bar{\Theta}, \bar{\lambda}^k, \bar{u})$ is an ε^k-elaboration of $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$. Because an arbitrary extension of F to $\mathcal{A} \times \bar{\Theta}$ is a generalized potential function of $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$, $\bar{\sigma} : \bar{T} \times \bar{\Theta} \to \Delta(A)$ is a GP-maximizing BIBCE of $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$ if and only if the restriction of $\bar{\sigma}$ to $T \times \Theta$ is a GP-maximizing BIBCE of (T, Θ, π, u).

Now, let $\overline{\mathcal{E}}^F$ be the set of all GP-maximizing BIBCE of $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$. Then, by Lemma

8, $\{(\bar{T}, \bar{\Theta}, \bar{\lambda}^k, \bar{u})\}_{k=1}^{\infty}$ has a sequence of BIBCE $\{\bar{\sigma}^k\}_{k=1}^{\infty}$ such that

$$\lim_{k\to\infty} \inf_{\bar{\sigma}\in\overline{\mathcal{E}}^F} \sum_{(a,t,\theta)\in A\times\bar{T}\times\bar{\Theta}} \left|\bar{\sigma}^k \circ \bar{\lambda}^k(a, \tau^{-1}(t), \theta) - \bar{\sigma} \circ \bar{\pi}(a, t, \theta)\right| = 0,$$

which is equivalent to

$$\lim_{k\to\infty} \inf_{\bar{\sigma}\in\overline{\mathcal{E}}^F} \sup_{E\subseteq A\times\bar{T}\times\bar{\Theta}} \left| \sum_{(a,t,\theta)\in E} \bar{\sigma}^k \circ \bar{\lambda}^k(a, \tau^{-1}(t), \theta) - \sum_{(a,t,\theta)\in E} \bar{\sigma} \circ \bar{\pi}(a, t, \theta) \right| = 0.$$

Let $\bar{\xi}^k : \bar{T}^k \times \bar{\Theta}^k \to \Delta(A)$ be the restriction of $\bar{\sigma}^k : \bar{T} \times \bar{\Theta} \to \Delta(A)$ to $\bar{T}^k \times \bar{\Theta}^k$. Then, $\bar{\xi}^k$ is a BIBCE of $(\bar{T}^k, \bar{\Theta}^k, \bar{\pi}^k, \bar{u}^k)$, and $\{\bar{\xi}^k\}_{k=1}^{\infty}$ satisfies

$$\lim_{k\to\infty} \inf_{\sigma\in\mathcal{E}^F} \sup_{E\subseteq A\times T\times\Theta} \left| \sum_{(a,t,\theta)\in E} \bar{\xi}^k \circ \bar{\pi}^k(a, (\tau^k)^{-1}(t), \theta) - \sum_{(a,t,\theta)\in E} \sigma \circ \pi(a, t, \theta) \right| = 0$$

because $\{\bar{\sigma} \circ \bar{\pi} \in \Delta(A \times \bar{T} \times \bar{\Theta}) \mid \bar{\sigma} \in \overline{\mathcal{E}}^F\}$ coincides with $\{\sigma \circ \pi \in \Delta(A \times T \times \Theta) \mid \sigma \in \mathcal{E}^F\}$ on their common support. Therefore, $\mathcal{E}^F$ is robust. □

5.3 Proof outline for Lemma 8

We sketch the proof of Lemma 8, focusing on the special case where (T, Θ, π, u) has a potential function v and a P-maximizing BIBCE is unique, i.e., $\mathcal{E}^v = \Sigma^v = \{\sigma^*\}$. We adopt a potential function satisfying the following condition. Let $\phi : \Theta \to \Theta^*$ be a mapping such that the i-th component of $\phi(\theta) = (\phi_i(\theta))_{i\in I}$ equals θ_i if $\theta_i \in \Theta_i^*$. Note that ϕ is the identity mapping when restricted to Θ^*. Assume that v satisfies

$$v(a, \theta) = v(a, \phi(\theta)) \text{ for all } (a, \theta) \in A \times \Theta, \tag{13}$$

which is possible because the definition of potential functions does not impose any conditions on their values when $\theta \notin \Theta^*$. Then, for $\theta \in \Theta$ and $i \in I$ with $\theta_i \in \Theta_i^*$,

$$u_i(a, \theta) = u_i(a, \phi(\theta)) = v(a, \phi(\theta)) + q_i(a_{-i}, \phi(\theta)) = v(a, \theta) + q_i(a_{-i}, \phi(\theta)),$$

where the first equality follows from the assumption that θ_i determines player i's payoff function, the second equality from (11), and the third equality from (13). This implies that the best response of player i with a payoff function $u_i(a,\theta)$ coincides with that of the same player with a payoff function $v(a,\theta)$ whenever $\theta_i \in \Theta_i^*$.

The proof proceeds in three steps. In the first step, we construct a candidate for a sequence of BIBCE $\{\sigma^k\}_{k=1}^\infty$ satisfying (12). In so doing, let $T_i^k \equiv \{t_i \in T_i \mid \pi^k(\Theta_i^* \times \Theta_{-i}|t_i) = 1\}$ denote the set of player i's types in an ε^k-elaboration (T,Θ,π^k,u) who believe that their payoff functions are the same as those in the analyst model (T,Θ,π,u). Note that $\pi^k(T^k) \geq 1-\varepsilon^k$ by the second condition of Definition 3. We call types in T_i^k standard types, and all other types perturbed types. For a type profile $t \in T$, we denote the set of players with standard types by $S^k(t) \equiv \{i \in I \mid t_i \in T_i^k\}$.

Consider a decision rule σ^k of (T,Θ,π^k,u), where standard types and perturbed types choose actions independently of each other, conditional on (t,θ); that is,

$$\sigma^k(a|t,\theta) = \sigma^k_{S^k(t)}(a_{S^k(t)}|t,\theta)\cdot\sigma^k_{-S^k(t)}(a_{-S^k(t)}|t,\theta), \tag{14}$$

where $\sigma_S(a_S|t,\theta) \equiv \sigma(\{a_S\}\times A_{-S}|t,\theta)$ and $\sigma_{-S}(a_{-S}|t,\theta) \equiv \sigma(\{a_{-S}\}\times A_S|t,\theta)$. Let Σ^k denote the set of all belief-invariant decision rules of (T,Θ,π^k,u) that satisfy (14). For each $\sigma^k \in \Sigma^k$, we refer to $\sigma_{S^k(t)}$ and $\sigma_{-S^k(t)}$ as the standard types' rule and the perturbed types' rule, respectively.

Using the Kakutani-Fan-Glicksberg fixed point theorem, we can show the existence of a BIBCE $\sigma^k \in \Sigma^k$ satisfying the following condition: The standard types' rule $\sigma^k_{S^k(t)}$ maximizes the expected value of a potential function:

$$\sum_{a,t,\theta}\sigma^k\circ\pi^k(a,t,\theta)v(a,\theta) \geq \sum_{a,t,\theta}\sigma'\circ\pi^k(a,t,\theta)v(a,\theta) \tag{15}$$

for any $\sigma' \in \Sigma^k$ such that the perturbed types' rule is the same as that of σ^k, i.e., $\sigma^k_{-S^k(t)} = \sigma'_{-S^k(t)}$. The condition (15) ensures that σ^k is obedient for players of standard types, which follows from the choice of a potential function discussed above.

Claim 1. *For each k, (T,Θ,π^k,u) has a BIBCE $\sigma^k \in \Sigma^k$ that satisfies (15) for any $\sigma' \in \Sigma^k$ with $\sigma^k_{-S^k(t)} = \sigma'_{-S^k(t)}$.*

Consider the sequence of BIBCE $\{\sigma^k\}_{k=1}^{\infty}$. In the second step, we show that the limit of the left-hand side of (15) as $k \to \infty$ is at least as large as that under the potential-maximizing BIBCE σ^* of the analyst's model.

Claim 2. *The sequence of BIBCE $\{\sigma^k\}_{k=1}^{\infty}$ satisfies*

$$\liminf_{k\to\infty} \sum_{a,t,\theta} \sigma^k \circ \pi^k(a,t,\theta) v(a,\theta) \geq \sum_{a,t,\theta} \sigma^* \circ \pi(a,t,\theta) v(a,\theta). \tag{16}$$

To see why, let $\hat{\sigma}^k \in \Sigma^k$ be the decision rule in which the perturbed types follow σ^k, while the standard types follow σ^*:

$$\hat{\sigma}^k_{-S^k(t)}(a_{-S^k(t)}|t,\theta) = \sigma^k_{-S^k(t)}(a_{-S^k(t)}|t,\theta) \text{ and } \hat{\sigma}^k_{S^k(t)}(a_{S^k(t)}|t,\theta) = \sigma^*_{S^k(t)}(a_{S^k(t)}|\tau(t),\theta)$$

for all (a,t,θ). Then, by (15), we obtain

$$\sum_{a,t,\theta} \sigma^k \circ \pi^k(a,t,\theta) v(a,\theta) \geq \sum_{a,t,\theta} \hat{\sigma}^k \circ \pi^k(a,t,\theta) v(a,\theta). \tag{17}$$

Since $\hat{\sigma}^k(a|t,\theta) = \sigma^*(a|\tau(t),\theta)$ holds for $t \in T^k$ and $\pi^k(T^k) \geq 1-\varepsilon^k$, we can show that the right-hand side of (17) converges to that of (16) through algebraic manipulation.

In the final step, let $\eta^k \in \Delta(A \times T \times \Theta)$ be given by $\eta^k(a,t,\theta) \equiv \sigma^k \circ \pi^k(a,\tau^{-1}(t),\theta)$. Then,

$$\sum_{a,t,\theta} \sigma^k \circ \pi^k(a,t,\theta) v(a,\theta) = \sum_{a,t,\theta} \sum_{t' \in \tau^{-1}(t)} \sigma^k(a|t',\theta)\pi^k(t',\theta) v(a,\theta) = \sum_{a,t,\theta} \eta^k(a,t,\theta) v(a,\theta).$$

Thus, (16) is rewritten as

$$\liminf_{k\to\infty} \sum_{a,t,\theta} \eta^k(a,t,\theta) v(a,\theta) \geq \sum_{a,t,\theta} \sigma^* \circ \pi(a,t,\theta) v(a,\theta). \tag{18}$$

Using (18), we establish the following, which implies that $\{\sigma^k\}_{k=1}^{\infty}$ satisfies (12).

Claim 3. *It holds that*

$$\lim_{k\to\infty} \sum_{a,t,\theta} \left|\eta^k(a,t,\theta) - \sigma^* \circ \pi(a,t,\theta)\right| = 0. \tag{19}$$

This claim can be established through the following reasoning. The sequence of probability distributions $\{\eta^k\}_{k=1}^{\infty}$ is tight, and thus it has a convergent subsequence $\{\eta^{k_l}\}_{l=1}^{\infty}$ with $\lim_{l\to\infty} \eta^{k_l} = \eta^\star$. The limit $\eta^\star$ satisfies $\pi(t,\theta) = \sum_{a\in A} \eta^\star(a,t,\theta)$, and the decision rule $\sigma^\star : T^* \times \Theta^* \to \Delta(A)$ defined by $\sigma^\star(a|t,\theta) = \eta^\star(a,t,\theta)/\pi(t,\theta)$ for all $(t,\theta) \in T^* \times \Theta^*$ is belief-invariant. From (18), we have

$$\sum_{a,t,\theta} \eta^\star(a,t,\theta) v(a,\theta) = \sum_{a,t,\theta} \sigma^\star \circ \pi(a,t,\theta) v(a,\theta) \geq \sum_{a,t,\theta} \sigma^* \circ \pi(a,t,\theta) v(a,\theta). \qquad (20)$$

Since $\sigma^\star$ is belief-invariant and σ^* is the unique P-maximizing belief-invariant decision rule, equality must hold in (20). This implies that $\eta^\star = \sigma^\star \circ \pi = \sigma^* \circ \pi$. Since any convergent subsequence of $\{\eta^k\}_{k=1}^{\infty}$ has the same limit $\eta^\star = \sigma^* \circ \pi$, $\{\eta^k\}_{k=1}^{\infty}$ is a convergent sequence with the limit $\eta^\star$, which in turn implies (19).

6 Supermodular games

6.1 Supermodular potential games

The robust BIBCE in the motivating example is not a BNE. This raises a natural question: in what class of games can a BNE be robust? We show that the robust set of BIBCE given by Theorem 1 contains BNE if the game is a supermodular potential game. In particular, if the robust set is a singleton, then the robust BIBCE is a BNE.

Let A_i be linearly ordered with $\geq_i$. We write $\geq_I$ for the product order: for $a, b \in A$, $a \geq_I b$ if and only $a_i \geq_i b_i$ for all $i \in I$. We say that (T, Θ, π, u) is a supermodular game if, for each $\theta \in \Theta$, the ex-post game with a payoff function profile $(u_i(\cdot,\theta))_{i\in I}$ exhibits strategic complementarities; that is, for each $i \in I$ and $a, b \in A$ with $a \geq_I b$,

$$u_i((a_i, a_{-i}), \theta) - u_i((b_i, a_{-i}), \theta) \geq u_i((a_i, b_{-i}), \theta) - u_i((b_i, b_{-i}), \theta).$$

For each $\sigma \in \Sigma^{BI}$, we define a pure strategy profile $\overline{\sigma} = (\overline{\sigma}_i)_{i\in I}$, where for each $i \in I$ and $t_i \in T_i$, player i of type t_i chooses the maximum action a_i such that $\sigma_i(a_i|t_i) = \sigma(\{a_i\} \times A_{-i} | (t_i, t_{-i}), \theta) > 0$. Similarly, we define another pure strategy profile $\underline{\sigma} = (\underline{\sigma}_i)_{i\in I}$, where player i of type t_i chooses the minimum action that satisfies the same condition.

The following proposition establisehs that if σ is a P-maximizing BIBCE, then the strategy profiles $\overline{\sigma}$ and $\underline{\sigma}$ are also P-maximizing BIBCE.

Proposition 1. *Assume that a supermodular game (T, Θ, π, u) has a potential function v. Then, for each P-maximizing BIBCE $\sigma \in \mathcal{E}^v$, it holds that $\overline{\sigma}, \underline{\sigma} \in \mathcal{E}^v$.*

This proposition implies that $\overline{\sigma}$ and $\underline{\sigma}$ are BNE that maximize the expected value of the potential function (P-maxizing BNE). Thus, if a P-maximizing BNE is unique, then we must have $\sigma = \overline{\sigma} = \underline{\sigma}$, which is a robust BNE.

As an example of such a BNE, we consider a binary-action supermodular game with two players, each choosing either action 1 or action 0 (which can be interpreted as "invest" and "not invest," respectively). The set of states is given by $\Theta = \{0, 1, 2, 3, \ldots\}$. A payoff table and a potential function for each state $n \in \Theta$ are given below, where $0 < r < 1$.

A payoff table

	action 1	action 0
action 1	r^{n+1}, r^{n+1}	$r^{n+1} - 1, 0$
action 0	$0, r^{n+1} - 1$	$0, 0$

A potential function

	action 1	action 0
action 1	r^{n+1}	0
action 0	0	$1 - r^{n+1}$

State $n \in \Theta$ occurs with probability $p(1-p)^n > 0$, where $p \in (0, 1)$. Each player has the following information partition of Θ: player 1's partition is $\{\{0\}, \{1, 2\}, \{3, 4\}, \ldots\}$, and player 2's partition is $\{\{0, 1\}, \{2, 3\}, \{4, 5\}, \ldots\}$, which are analogous to those in the email game. The type of a player with a partition $\{k-1, k\}$ (or $\{0\}$) is referred to as type k (or type 0). Thus, the sets of player 1's types and player 2's types are $T_1 = \{0, 2, 4, \ldots\}$ and $T_2 = \{1, 3, 5, \ldots\}$, respectively. This game can be interpreted as a global game (Carlsson and van Damme, 1993) with discretized signals and no dominance regions.

We represent a pure-strategy profile as a sequence of actions $x = (x_k)_{k=0}^{\infty} \in \{0, 1\}^{T_1 \cup T_2}$, where x_k is the action of type k. Let $s^\tau = (s_k^\tau)_{k=0}^{\infty}$ denote a monotone strategy profile:

$$s_k^\tau = \begin{cases} 1 & \text{if } k \leq \tau - 1, \\ 0 & \text{if } k \geq \tau, \end{cases}$$

where τ is a non-negative integer or infinite ($\tau = \infty$). If r is sufficiently close to one, this game has exactly three pure-strategy BNE: s^0, s^∞, and s^{τ^*}, where $\tau^* \geq 2$ is the smallest integer satisfying $r^{\tau^*} < (1-p)/(1+r(1-p))$, provided no integer τ satisfies this condition with equality.[13]

The expected value of the potential function under a strategy profile x is calculated as

$$f(x) \equiv \sum_{k=0}^{\infty} p(1-p)^k \Big(r^{k+1} x_k x_{k+1} + (1 - r^{k+1})(1 - x_k)(1 - x_{k+1}) \Big).$$

Using this formula, we can verify that $f(s^{\tau^*}) > f(s^0)$ and $f(s^{\tau^*}) > f(s^\infty)$. Thus, s^{τ^*} is the unique P-maximizing BNE, which is a robust BNE by Proposition 1.

6.2 Binary-action supermodular games

Let (T, Θ, π, u) be a supermodular game with $A_i = \{0, 1\}$ and $\Theta = \Theta^*$. For $S \subseteq I$, we denote by $\mathbf{1}_S$ the action profile in which all players in S play action 1, while the others play action 0. By convention, we write $\mathbf{1} = \mathbf{1}_I$. We assume that this game has a generalized potential function $F : \mathcal{A} \times \Theta \to \mathbb{R}$ satisfying the following conditions:

G1. $\mathcal{A}_i = \{\{0, 1\}, \{1\}\}$ for each $i \in I$.

G2. $F(X^{\mathbf{1}}, \theta) > F(X, \theta)$ for all $X \neq X^{\mathbf{1}} = \prod_{i \in I} X_i^{\mathbf{1}}$ and $\theta \in \Theta$, where $X_i^{\mathbf{1}} = \{1\}$.

It can be readily shown that the decision rule with "always **1**" is the unique GP-maximizing BIBCE, regardless of the prior $\pi \in \Delta(T \times \Theta)$. We denote this decision rule by $\sigma^{\mathbf{1}}$; that is, $\sigma^{\mathbf{1}}(\mathbf{1}|t, \theta) = 1$ for all $(t, \theta) \in T \times \Theta$.

By Theorem 1, $\sigma^{\mathbf{1}}$ is robust for arbitrary prior $\pi \in \Delta(T \times \Theta)$. Moreover, the converse is also true: if $\sigma^{\mathbf{1}}$ is a robust BIBCE of a binary-action supermodular game for arbitrary prior $\pi \in \Delta(T \times \Theta)$, then it must (generically) be a GP-maximizing BIBCE with a genelized potential function satisfying G1 and G2, as formalized in the next proposition.

Proposition 2. *Let (T, Θ, π, u) be a binary-action supermodular game. Then, the following results hold.*

[13] This condition ensures that type τ^*'s best response is action 0 when type $\tau^* - 1$ chooses action 1 and type $\tau^* + 1$ chooses action 0. The conditional expected value of the potential function for type τ^* is $r^{\tau^*}/(2-p)$ under action 1 and $(1-p)(1-r^{\tau^*+1})/(2-p)$ under action 0. If the latter is greater, action 0 is the best response.

(1) *If* (T, Θ, π, u) *has a generalized potential function satisfying* G1 *and* G2*, then* $\sigma^{\mathbf{1}}$ *is robust in* (T, Θ, π', u) *for arbitrary* $\pi' \in \Delta(T \times \Theta)$.

(2) *For a generic payoff function profile* u*, if* (T, Θ, π, u) *does not have a generalized potential function satisfying* G1 *and* G2*, then there exists* $\pi' \in \Delta(T \times \Theta)$ *such that* $\sigma^{\mathbf{1}}$ *is not robust in* (T, Θ, π', u).

Part (1) follows directly from Theorem 1, while part (2) is a consequence of Theorem 3 of Oyama and Takahashi (2020), who use an alternative representation of a generalized potential function.

Given F, let $v : A \times \Theta \to \mathbb{R}$ be a function defined by

$$v((\min X_i)_{i \in I}, \theta) = F(X, \theta) \tag{21}$$

for all $(X, \theta) \in \mathcal{A} \times \Theta$. Note that G2 is equivalent to

$$v(\mathbf{1}, \theta) > v(a, \theta) \text{ for all } a \neq \mathbf{1} \text{ and } \theta \in \Theta. \tag{22}$$

The following lemma provides a sufficient condition on v to ensure that F is a generalized potential function.

Lemma 9. *Suppose that there exist a function* $v : A \times \Theta \to \mathbb{R}$ *and a constant* $\lambda_i > 0$ *for each* $i \in I$ *such that*

$$\lambda_i(u_i((1, a_{-i}), \theta) - u_i((0, a_{-i}), \theta)) \geq v((1, a_{-i}), \theta) - v((0, a_{-i}), \theta) \tag{23}$$

for all $a_{-i} \in A_{-i}$ *and* $\theta \in \Theta$*. If* (T, Θ, π, u) *or* v *is supermodular, then* $F : \mathcal{A} \times \Theta \to \mathbb{R}$ *given by* (21) *is a generalized potential function.*

The function v in this lemma is referred to as a monotone potential function (Morris and Ui, 2005). If $\lambda_i = 1$ for all $i \in I$ and equality holds in (23), then a monotone potential function is a potential function.

Oyama and Takahashi (2020) focus on complete information games and show that, for a generic binary-action supermodular game, if there does not exist a monotone potential function satisfying (22), then, for every $\varepsilon > 0$, there exists an ε-elaboration in the sense of Kajii and Morris (1997) such that "all 0" survives the iterated elimination of strictly dominated strategies. Since an ε-elaborations in Kajii and Morris (1997) is equivalent to our ε-elaboration (see Subsection 7.1), the action profile **1** is not robust in our sense.

This result leads us to part (2) of Proposition 2. Suppose there is no generalized potential function satisfying G1 and G2. Then, there cannot be a monotone potential function satisfying (22). This, in turn, means that there exists $\theta' \in \Theta$ such that the complete information game in which θ' is common knowledge does not have a monotone potential function satisfying (22) with $\theta = \theta'$. Now, consider a degenerate prior $\pi' \in \Delta(T \times \Theta)$ with $\pi'(t', \theta') = 1$ for some specific type $t' \in T$. Since (T, Θ, π', u) is a complete information game, there exists an ε-elaboration in which "all 0" survives the iterated elimination of strictly dominated strategies. This establishes that $\sigma^{\mathbf{1}}$ is not robust in (T, Θ, π', u).

7 Discussion

We discuss the remaining issues regarding the relationship between the robustness analysis in the complete information setting and our analysis in the incomplete information setting. For proofs of the stated results, see Morris and Ui (2020) or the supplementary material.

7.1 ε-elaborations in Kajii and Morris (1997)

The framework of Kajii and Morris (1997) is summarized as follows. The analyst's model is a complete information game, where $T = T^* = \{t\}$ and $\Theta = \Theta^* = \{\theta\}$ are singletons. An ε-elaboration is defined as an incomplete information game that satisfies the first two conditions of Definition 3. The solution concept for the analyst's model is a correlated equilibrium, and that for ε-elaborations is a BNE.

The set of ε-elaborations defined in Kajii and Morris (1997) includes the set of ε-elaborations defined in our paper in the case of complete information games. However, our definition is equivalent to theirs in the following sense.

Lemma 10. *Suppose that $T = T^* = \{t\}$ and $\Theta = \Theta^* = \{\theta\}$. If $(\bar{T}, \bar{\Theta}, \bar{\pi}, \bar{u})$ satisfies the first two conditions of Definition 3 for $\varepsilon > 0$, then it is a $\sqrt{\varepsilon}$-elaboration.*

Thus, the key distinction between our robustness concept and that of Kajii and Morris (1997) lies in the equilibrium concept used for ε-elaborations: we adopt BIBCE, whereas

Kajii and Morris (1997) adopt BNE. Pram (2019) proposes a weaker definition of robust equilibria in complete information games than that of Kajii and Morris (1997) by adopting agent-normal form correlated equilibria as the solution concept for ε-elaborations. Since both BNE and agent-normal form correlated equilibria are special cases of BIBCE, our definition of robustness, when applied to complete information games, is weaker than those of Pram (2019) and Kajii and Morris (1997).

7.2 The critical path theorem

Kajii and Morris (1997) study common belief events and derive a lower bound on their probability. This result, known as the critical path theorem, is used to establish a sufficient condition for the robustness. Oyama and Takahashi (2020) extend the theorem and prove its generic converse, thus establishing a fundamental link between robustness and monotone potentials as discussed in Section 6.2.

Given the significance of the critical path theorem, it is natural to examine its counterpart in the incomplete information setting. To sketch this, consider a binary-action supermodular game (T, Θ, π, u). For each $i \in I$, fix a subset of types $E_i \subset T_i$. Construct an associated fictitious game by preserving the belief and payoff function for player i of type $t_i \in E_i$, while forcing types $t_i \notin E_i$ to choose action 0 as their dominant action. This fictitious game then constitutes an ε-elaboration with $\varepsilon \equiv 1 - \pi(E)$, where $E = \prod_{i \in I} E_i$.

Consider the largest BNE in the fictitious game, which is the largest pure strategy profile that survives the iterated elimination of strictly dominated strategies. Let $CB_i^u(E)$ denote the set of player i's types who choose action 1 in the largest BNE and define $CB^u(E) = \prod_{i \in I} CB_i^u(E) \subset E \subset T$. This set $CB^u(E)$ can be interpreted as a common belief event in the sense that players have a common belief that action 1 is a best response in the fictitious game when $t \in CB^u(E)$ is realized.

The critical path theorem provides a lower bound for the probability of $CB^u(E)$ in terms of ε. We can establish the following lower bound.

Lemma 11. *If (T, Θ, π, u) admits a monotone potential function v satisfying* (22)*, then*

$$\pi(CB^u(E)) \geq 1 - \kappa(v)\varepsilon, \text{ where } \kappa(v) \equiv 1 + \frac{\sup_{S \subseteq S' \neq I, \theta \in \Theta} v(\mathbf{1}_S, \theta) - v(\mathbf{1}_{S'}, \theta)}{\inf_{S \neq I, \theta \in \Theta} v(\mathbf{1}, \theta) - v(\mathbf{1}_S, \theta)} > 0. \quad (24)$$

By (24), $\pi(CB^u(E))$ converges to one as $\varepsilon \to 0$. This implies that if ε is sufficiently small, the ε-elaboration has a BNE in which every player chooses action 1 with probability close to one. Oyama and Takahashi (2020) establish (24) in the special case where u and v are independent of θ.

Appendix

This appendix contains omitted proofs. Throughout these proofs, for any countable set S, we regard the set of probability distributions $\Delta(S)$ as a subset of the linear space $\{f : S \to \mathbb{R} \mid \sum_{s\in S} |f(s)| < \infty\}$ endowed with the l_1-norm $\|f\|_1 = \sum_{s\in S} |f(s)|$. Because S is countable, it is straightforward to show that the topology of weak convergence in $\Delta(S)$ coincides with the topology induced by the l_1-norm in $\Delta(S)$. Thus, the following result holds by Prohorov's theorem (see Billingsley, 1999), which will be used throughout.

Lemma A. *Let $\Delta(S)$ be endowed with the topology induced by the l_1-norm. If $P \subset \Delta(S)$ is tight, i.e., for any $\varepsilon > 0$, there exists a finite set $K^\varepsilon \subset S$ such that $p(K^\varepsilon) > 1 - \varepsilon$ for all $p \in P$, then the closure of P is compact. Conversely, if the closure of $P \subset \Delta(S)$ is compact, then P is tight.*

A Proofs for Section 2

A.1 Proof of Lemma 2

To show the "if" part, suppose that, for each $i \in I$, there exists a mapping $\tau_i : \bar{T}_i \to T_i$ satisfying (2) and (3). Let $M_i = \bar{T}_i$ and $\mu_i : \bar{T}_i \to M_i$ be such that $\mu_i(\bar{t}_i) = \bar{t}_i$ for all $\bar{t}_i \in \bar{T}_i$, by which the mapping $\bar{t}_i \mapsto (\tau_i(\bar{t}_i), \mu_i(\bar{t}_i))$ is one-to-one. Consider a communication rule $\rho : T \times \Theta \to \Delta(M)$ satisfying (1); that is, $\rho(\bar{t}|t,\theta)\pi(t,\theta) = \bar{\pi}(\bar{t},\theta)$ if $\tau(\bar{t}) = t$ and $\rho(\bar{t}|t,\theta) = 0$ otherwise, which is well-defined by (2). Then, for $(t,\theta) \in T^* \times \Theta^*$ and $\bar{t}_i \in \bar{T}_i^*$ with $t_i = \tau_i(\bar{t}_i)$,

$$\begin{aligned}
&\rho(\{\bar{t}_i\} \times \bar{T}_{-i}|t,\theta) \\
&\quad = \rho(\{\bar{t}_i\} \times \tau_{-i}^{-1}(t_{-i})|t,\theta) = \frac{\bar{\pi}(\{\bar{t}_i\} \times \tau_{-i}^{-1}(t_{-i}),\theta)}{\pi(t,\theta)} = \frac{\bar{\pi}(\tau_{-i}^{-1}(t_{-i}),\theta|\bar{t}_i) \times \bar{\pi}(\bar{t}_i)}{\pi(t_{-i},\theta|t_i) \times \pi(t_i)} = \frac{\bar{\pi}(\bar{t}_i)}{\pi(t_i)}
\end{aligned}$$

by (3). Because $\rho(\{\bar{t}_i\} \times \bar{T}_{-i}|t, \theta) = \bar{\pi}(\bar{t}_i)/\pi(t_i)$ is independent of t_{-i} and θ, ρ is a belief-invariant communication rule, and thus $(\bar{T}, \Theta, \bar{\pi}, u)$ is an elaboration of (T, Θ, π, u).

To show the "only if" part, suppose that $(\bar{T}, \Theta, \bar{\pi}, u)$ is an elaboration of (T, Θ, π, u); that is, there exists a belief-invariant communication rule ρ and mappings $\tau_i : \bar{T}_i \to T_i$ and $\mu_i : \bar{T}_i \to M_i$ for each $i \in I$ satisfying the condition in Definition 1. It is enough to show that τ and $\bar{\pi}$ satisfy (2) and (3). Let $(t, \theta) \in T^* \times \Theta^*$ and $\bar{t}_i \in \bar{T}_i$ be such that $\tau_i(\bar{t}_i) = t_i$. By (1),

$$\bar{\pi}(\tau^{-1}(t), \theta) = \pi(t, \theta) \sum_{\bar{t}:\tau(\bar{t})=t} \rho(\mu(\bar{t})|t, \theta) = \pi(t, \theta).$$

Thus, (2) holds. Moreover, by (1) again,

$$\begin{aligned}
\bar{\pi}(\tau_{-i}^{-1}(t_{-i}), \theta|\bar{t}_i) &= \frac{\bar{\pi}(\{\bar{t}_i\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\sum_{t'_{-i},\theta'} \bar{\pi}(\{\bar{t}_i\} \times \tau_{-i}^{-1}(t'_{-i}), \theta')} \\
&= \frac{\pi(t, \theta) \left(\sum_{\bar{t}_{-i}:\tau_{-i}(\bar{t}_{-i})=t_{-i}} \rho(\mu(\bar{t})|t, \theta)\right)}{\sum_{t'_{-i},\theta'} \pi((t_i, t'_{-i}), \theta') \left(\sum_{\bar{t}_{-i}:\tau_{-i}(\bar{t}_{-i})=t'_{-i}} \rho(\mu(\bar{t})|(t_i, t'_{-i}), \theta')\right)} \\
&= \frac{\pi(t, \theta)}{\sum_{t'_{-i},\theta'} \pi((t_i, t'_{-i}), \theta')} = \pi(t, \theta|t_i),
\end{aligned}$$

where the third equality holds because

$$\sum_{\bar{t}_{-i}:\tau_{-i}(\bar{t}_{-i})=t_{-i}} \rho(\mu(\bar{t})|t, \theta) = \sum_{m_{-i}\in M_{-i}} \rho((\mu_i(\bar{t}_i), m_{-i})|t, \theta) = \rho(\{\mu_i(\bar{t}_i)\} \times M_{-i}|t, \theta)$$

is independent of t_{-i} and θ by the belief-invariance of ρ. Thus, (3) holds. □

B Proofs for Section 4

For each $\gamma \in \Gamma^{BI}$, the conditional distribution of (a_i, X_i) given t_i is denoted by $\gamma_i(a_i, X_i|t_i)$. Belief-invariance implies that $\gamma_i(a_i, X_i|t_i) = \gamma(\{a_i\} \times A_{-i} \times \{X_i\} \times \mathcal{A}_{-i}|t, \theta)$, which will be used in the proofs. We write $\gamma_i(a_i|t_i) = \sum_{X_i\in\mathcal{A}_i} \gamma_i(a_i, X_i|t_i)$ and $\gamma_i(X_i|t_i) = \sum_{a_i\in A_i} \gamma_i(a_i, X_i|t_i)$.

B.1 Proof of Lemma 5

Let an $\mathcal{A}$-decision rule γ be a BIBCE. Because γ is obedient, (8) holds. By taking the summation of each side of (8) over $X_i \in \mathcal{A}_i$, we obtain

$$\sum_{a_{-i}, t_{-i}, \theta} \gamma(\{a\} \times \mathcal{A} | t, \theta) \pi(t, \theta) u_i(a, \theta) \geq \sum_{a_{-i}, t_{-i}, \theta} \gamma(\{a\} \times \mathcal{A} | t, \theta) \pi(t, \theta) u_i((a_i', a_{-i}), \theta),$$

and thus σ is obedient. In addition, because γ is belief-invariant, $\gamma(\{a_i\} \times A_{-i} \times \{X_i\} \times \mathcal{A}_{-i} | t, \theta) = \gamma_i(a_i, X_i | t_i)$, which implies $\sigma(\{a_i\} \times A_{-i} | t, \theta) = \sum_{X_i \in \mathcal{A}_i} \gamma_i(a_i, X_i | t_i)$. Thus, σ is belief-invariant as well. □

B.2 Proof of Lemma 6

When $\mathcal{A}_i = \{\{a_i\} \mid a_i \in A_i\}$ for each $i \in I$, we can rewrite (9) and (10) as

$$a_i \in \arg\max_{a_i' \in A_i} \sum_{a_{-i}, \theta} P_i(a_{-i}, \theta) v((a_i', a_{-i}), \theta), \tag{B.1}$$

$$a_i \in \arg\max_{a_i' \in A_i} \sum_{a_{-i}, \theta} P_i(a_{-i}, \theta) u_i((a_i', a_{-i}), \theta), \tag{B.2}$$

respectively, since $P_i(a_{-i}, \theta) = P_i(\{a_{-i}\}, \theta)$. Then, (B.1) is equivalent to (B.2) by (11). □

B.3 Proof of Lemma 7

Each $\gamma \in \Gamma^{BI}$, together with a prior π, defines a joint probability distribution $\gamma \circ \pi$ over $A \times \mathcal{A} \times T \times \Theta$ given by $\gamma \circ \pi(a, X, t, \theta) = \gamma(a, X | t, \theta) \pi(t, \theta)$, which is referred to as a distributional (belief-invariant) $\mathcal{A}$-decision rule. We denote the set of all such rules by

$$\Gamma^{BI} \circ \pi = \{\gamma \circ \pi \in \Delta(A \times \mathcal{A} \times T \times \Theta) \mid \gamma \in \Gamma^{BI}\}$$

and identify it with Γ^{BI}; that is, we regard Γ^{BI} as the set of equivalence classes induced by each $\gamma \circ \pi \in \Gamma^{BI} \circ \pi$, where γ and γ' are equivalent if $\gamma \circ \pi = \gamma' \circ \pi$. It can be verified that $\Gamma^{BI} \circ \pi$ is a tight closed subset of $\Delta(A \times \mathcal{A} \times T \times \Theta)$. Thus, by Lemma A, it is compact.

Therefore,

$$\Gamma^F \circ \pi \equiv \arg \max_{\gamma\circ\pi \in \Gamma^{BI}\circ\pi} \sum_{a,X,t,\theta} \gamma \circ \pi(a, X, t, \theta) F(X, \theta)$$

is nonempty, implying that Γ^F is also nonempty.

Then, for each $\gamma \in \Gamma^F$ and $(X_i, t_i) \in \mathcal{A}_i \times T_i$ with $\gamma_i(X_i|t_i) > 0$, it holds that

$$X_i \in \arg \max_{X_i' \in \mathcal{A}_i} \sum_{a,\, X_{-i},\, t_{-i},\, \theta} \gamma(a, X|t, \theta)\pi(t, \theta)F((X_i', X_{-i}), \theta), \tag{B.3}$$

which implies that

$$X_i \cap \arg \max_{a_i' \in A_i} \sum_{a,\, X_{-i},\, t_{-i},\, \theta} \gamma(a, X|t, \theta)\pi(t, \theta)u_i((a_i', a_{-i}), \theta) \neq \emptyset \tag{B.4}$$

by Definition 5.

Fix $\hat{\gamma} \in \Gamma^F$ and let $\hat{\rho} : T \times \Theta \to \Delta(\mathcal{A})$ be the communication rule with $\hat{\rho}(X|t, \theta) = \hat{\gamma}(X|t, \theta)$. Consider the conjunction of (T, Θ, π, u) and $\hat{\rho}$ and let $\Sigma_i^{\mathcal{A}}$ be the set of player i's strategies in the conjunction that always assign some $a_i \in X_i$ whenever player i receives $X_i \in \mathcal{A}_i$:

$$\Sigma_i^{\mathcal{A}} = \{\sigma_i : T_i \times \mathcal{A}_i \to \Delta(A_i) \mid \sigma_i(a_i|t_i, X_i) = 0 \text{ for } a_i \notin X_i\}.$$

For each $\sigma \in \Sigma^{\mathcal{A}}$, a $\mathcal{A}$-decision rule γ given by $\gamma(a, X|t, \theta) = \prod_{i\in I} \sigma_i(a_i|t_i, X_i)\hat{\rho}(X|t, \theta)$ is an element of Γ^F since $\gamma(X|t, \theta) = \hat{\rho}(X|t, \theta) = \hat{\gamma}(X|t, \theta)$.

We show that the conjunction has a BNE in which, for each i, player i follows a strategy in $\Sigma_i^{\mathcal{A}}$. Observe that, for any $\sigma_{-i} \in \Sigma_{-i}^{\mathcal{A}}$ and $(X_i, t_i) \in \mathcal{A}_i \times T_i$ with $\hat{\rho}_i(X_i|t_i) > 0$, it holds that

$$X_i \cap \arg \max_{a_i' \in A_i} \sum_{a_{-i},\, X_{-i},\, t_{-i},\, \theta} \prod_{j\neq i} \sigma_j(a_j|t_j, X_j)\hat{\rho}(X|t, \theta)\pi(t, \theta)u_i(a_i', a_{-i}) \neq \emptyset$$

by (B.4). That is, for any $\sigma_{-i} \in \Sigma_{-i}^{\mathcal{A}}$, there is $\sigma_i' \in \Sigma_i^{\mathcal{A}}$ that is a best response to σ_{-i}. Thus, the restricted best response correspondence $B^{\mathcal{A}} : \Sigma^{\mathcal{A}} \rightrightarrows \Sigma^{\mathcal{A}}$ given by

$$B^{\mathcal{A}}(\sigma) = \{\sigma' \in \Sigma^{\mathcal{A}} \mid \sigma_i' \text{ is a best response strategy to } \sigma_{-i} \text{ for each } i \in I\}$$

has nonempty convex values. Moreover, it can be readily shown that it has a closed graph in the distributional strategy space $\Sigma^{\mathcal{A}} \circ \pi$. Applying the Kakutani-Fan-Glicksberg fixed point theorem, we obtain a BNE $\sigma^* \in \Sigma^{\mathcal{A}}$ of the conjunction. Then, the $\mathcal{A}$-decision rule $\gamma^* \in \Gamma^F$ given by $\gamma^*(a, X|t, \theta) = \prod_{i\in I} \sigma_i^*(a_i|t_i, X_i)\hat{\rho}(X|t, \theta)$ is a desired BIBCE. □

C Proofs for Section 5

For a belief-invariant $\mathcal{A}$-decision rule $\gamma \in \Gamma^{BI}$, we write $\gamma_S(a_s, X_S|t, \theta) \equiv \gamma(\{a_S\} \times A_{-S} \times \{X_S\} \times \mathcal{A}_{-S}|t, \theta)$. We also write $\gamma_i(a_i, X_i|t_i) \equiv \gamma_{\{i\}}(a_i, X_i|t, \theta)$, which is possible since γ is belief-invariant, $\gamma_S(a_S|t, \theta) = \gamma_S(\{a_S\} \times \mathcal{A}_{-S}|t, \theta)$, and $\gamma_S(X_S|t, \theta) = \gamma_S(A_{-S} \times \{X_S\}|t, \theta)$.

C.1 Proof of Lemma 8

We adopt a generalized potential function F satisfying the following condition. Let ϕ : $\Theta \to \Theta^*$ be a mapping satisfying $\phi_i(\theta) = \theta_i$ if $\theta_i \in \Theta_i^*$ for all $i \in I$, where $\phi_i(\theta)$ is the i-th component of $\phi(\theta) \in \prod_{i\in I} \Theta_i^*$. We assume that F satisfies

$$F(X, \theta) = F(X, \phi(\theta)) \text{ for all } (X, \theta) \in \mathcal{A} \times \Theta^*, \tag{C.1}$$

which is possible because the definition of generalized potential functions does not impose any conditions on their values when $\theta \notin \Theta^*$.

Consider $\{(T, \Theta, \pi^k, u)\}_{k=1}^{\infty}$. For each k, let $T_i^k \equiv \{t_i \in T_i \mid \pi^k(\Theta_i^* \times \Theta_{-i}|t_i) = 1\}$ and $S^k(t) \equiv \{i \in I \mid t_i \in T_i^k\}$. We call types in T_i^k standard types, and all other types perturbed types, which are the same as those in Section 5.2.

Let

$$\Gamma^k \equiv \{\gamma \in \Gamma^{BI} \mid \gamma(a, X|t, \theta) = \gamma_{S^k(t)}(a_{S^k(t)}, X_{S^k(t)}|t, \theta) \prod_{i\notin S^k(t)} \gamma_i(a_i, X_i|t_i)\}$$

denote the collection of belief-invariant $\mathcal{A}$-decision rules in which perturbed types receive independent signals and choose independent actions. Note that Γ^k is convex with respect to the following convex combination: for $\gamma, \gamma' \in \Gamma^k$, $\gamma'' = \lambda\gamma + (1 - \lambda)\gamma' \in \Gamma^k$ is given by

$$\gamma''_{S^k(t)}(\cdot|t, \theta) = \lambda\gamma_{S^k(t)}(\cdot|t, \theta) + (1 - \lambda)\gamma'_{S^k(t)}(\cdot|t, \theta),$$

$$\prod_{i \notin S^k(t)} \gamma_i''(\cdot|t_i) = \prod_{i \notin S^k(t)} \Big(\lambda \gamma_i(\cdot|t_i) + (1-\lambda)\gamma_i'(\cdot|t_i)\Big).$$

For each $\gamma \in \Gamma^k$, let

$$\Gamma^k[\gamma] = \{\gamma' \in \Gamma^k \mid \gamma_i'(a_i|t_i) = \gamma_i(a_i|t_i) \text{ for all } a_i \in A_i,\, t_i \notin T_i^k, \text{ and } i \in I\}$$

denote the subset of Γ^k in which perturbed types follow γ, which is also convex.

In the next lemma, we construct a candidate for $\{\sigma^k\}_{k=1}^{\infty}$ satisfying (12).

Lemma B. *An ε^k-elaboration (T, Θ, π^k, u) has is a BIBCE $\gamma^k \in \Gamma^k$ satisfying*

$$\gamma^k \in \arg\max_{\gamma \in \Gamma^k[\gamma^k]} \sum_{(X,t,\theta) \in \mathcal{A} \times T \times \Theta} \gamma(X|t,\theta)\pi^k(t,\theta)F(X,\theta). \tag{C.2}$$

Thus, there exists a BIBCE $\sigma^k \in \Sigma^{BI}$ given by $\sigma^k(a|t,\theta) = \gamma^k(a|t,\theta)$.

Note that in a BIBCE γ^k, standard types follow vague recommendations $X_i \in \mathcal{A}_i$ designed to maximize the expected value of the generalized potential function, whereas perturbed types make decisions independently.

Proof. We characterize γ^k as a fixed point of a correspondence on Γ^k, which is defined by the composition of two correspondences. The existence of a fixed point is then established using a standard argument. The proof consists of the following three steps.

Step 1: The first correspondence $\Psi^1 : \Gamma^k \rightrightarrows \Gamma^k$ is give by

$$\begin{aligned}\Psi^1(\gamma) = \{\gamma' \in \Gamma^k \mid & \text{ for each } t_i \notin T_i^k,\, \gamma_i'(a_i|t_i) > 0 \text{ implies} \\ & a_i \in \arg\max_{a_i' \in A_i} \sum_{a_{-i}, t_{-i}, \theta} \gamma_{-i}(a_{-i}|t,\theta)\pi^k(t,\theta)u_i((a_i', a_{-i}), \theta), \\ & \gamma'(a, X|t,\theta) = \gamma_{S^k(t)}(a_{S^k(t)}, X_{S^k(t)}|t,\theta) \prod_{i \notin S^k(t)} \gamma_i'(a_i, X_i|t_i) \text{ for all } (a,t,\theta)\}.\end{aligned}$$

Under $\gamma' \in \Psi^1(\gamma)$, perturbed types choose best responses to γ, whereas standard types follow γ. Note that $\Psi^1(\gamma)$ is a nonempty convex subset of Γ^k.

Step 2: The second correspondence $\Psi^2 : \Gamma^k \rightrightarrows \Gamma^k$ is given by

$$\Psi^2(\gamma) = \{\gamma' \in \Gamma^{k,F}[\gamma] \mid \gamma' \text{ is obedient for each } i \in I \text{ with } t_i \in T_i^k\},$$

where

$$\Gamma^{k,F}[\gamma] \equiv \arg\max_{\gamma' \in \Gamma^k[\gamma]} \sum_{X,t,\theta} \gamma'(X|t,\theta)\pi^k(t,\theta)F(X,\theta) = \arg\max_{\gamma' \in \Gamma^k[\gamma]} \sum_{X,t,\theta} \gamma' \circ \pi^k(X,t,\theta)F(X,\theta).$$

Note that (C.2) is written as $\gamma^k \in \Gamma^{k,F}[\gamma^k]$, and that $\Gamma^{k,F}[\gamma]$ is nonempty because $\{\gamma' \circ \pi^k \in \Delta(A \times \mathcal{A} \times T \times \Theta) \mid \gamma' \in \Gamma^k[\gamma]\}$ is tight and closed, and compact by Lemma A. Under $\gamma' \in \Psi^2(\gamma)$, perturbed types follow γ, whereas standard types simultaneously choose best responses; that is, given the mixed actions of perturbed types, standard types behave as if they follow a BIBCE. Note that $\Psi^2(\gamma)$ is a convex subset of Γ^k. We show that $\Psi^2(\gamma)$ is nonempty.

The proof is essentially the same as that of Lemma 7. It begins with the observation that, for any $\gamma' \in \Gamma^{k,F}[\gamma]$ and $(X_i, t_i) \in \mathcal{A}_i \times T_i^k$ with $\gamma_i'(X_i|t_i) > 0$, (B.3) and (B.4) hold, where γ and π are replaced by γ' and π^k, respectively.

Fix $\hat{\gamma} \in \Gamma^{k,F}[\gamma]$ and let $\hat{\rho} : T \times \Theta \to \Delta(\mathcal{A})$ be the communication rule with $\hat{\rho}(X|t,\theta) = \hat{\gamma}(X|t,\theta)$. Consider the conjunction of (T, Θ, π^k, u) and $\hat{\rho}$. Define $\Sigma_i^{k,\mathcal{A}}$ as the set of player i's strategies that always assign some $a_i \in X_i$ whenever player i receives $X_i \in \mathcal{A}_i$, with the mixed actions of perturbed types remaining the same as those in γ:

$$\Sigma_i^{k,\mathcal{A}} = \{\sigma_i : T_i \times \mathcal{A}_i \to \Delta(A_i) \mid \sigma_i(a_i|t_i, X_i) = 0 \text{ for } a_i \notin X_i,\ \sigma_i(a_i|t_i, X_i) = \gamma_i(a_i|t_i) \text{ for } t_i \notin T_i^k\}.$$

Then, $\gamma' \in \Gamma^{BI}$ given by $\gamma'(a, X|t,\theta) = \prod_{i \in I} \sigma_i(a_i|t_i, X_i)\hat{\rho}(X|t,\theta)$ is an element of $\Gamma^{k,F}[\gamma]$ for each $\sigma \in \Sigma^{k,\mathcal{A}}$.

We show that the conjunction admits $\sigma \in \Sigma^{k,\mathcal{A}}$ where all standard types simultaneously choose best responses. Observe that, for any $\sigma_{-i} \in \Sigma_{-i}^{k,\mathcal{A}}$ and $(X_i, t_i) \in \mathcal{A}_i \times T_i^k$ with $\hat{\rho}_i(X_i|t_i) > 0$, we have

$$X_i \cap \arg\max_{a_i' \in A_i} \sum_{a_{-i}, X_{-i}, t_{-i}, \theta} \prod_{j \neq i} \sigma_j(a_j|t_j, X_j)\hat{\rho}(X|t,\theta)\pi^k(t,\theta)u_i(a_i', a_{-i}) \neq \emptyset$$

by (B.4). This implies that, for any $\sigma_{-i} \in \Sigma_{-i}^{k,\mathcal{A}}$, there exists $\sigma_i' \in \Sigma_i^{k,\mathcal{A}}$ that is a best response to σ_{-i} for every standard type. Thus, the restricted best response correspondence

$B^{k,\mathcal{A}} : \Sigma^{k,\mathcal{A}} \rightrightarrows \Sigma^{k,\mathcal{A}}$ given by

$$B^{k,\mathcal{A}}(\sigma) = \{\sigma' \in \Sigma^{k,\mathcal{A}} \mid \sigma'_i \text{ is a best response strategy to } \sigma_{-i} \text{ for each } i \in I \text{ of } t_i \in T^k\}$$

has nonempty convex values.

As in the proof of Lemma 7, the fixed point theorem guarantees the existence of $\sigma^* \in \Sigma^{k,\mathcal{A}}$ where all standard types simultaneously choose best responses. Then, $\mathcal{A}$-decision rule $\gamma^* \in \Gamma^F$ given by $\gamma^*(a, X|t, \theta) = \prod \sigma_i^*(a_i|t_i, X_i)\hat{\rho}(X|t, \theta)$ belongs to $\Psi^2(\gamma)$, confirming that $\Psi^2(\gamma)$ is nonempty.

Step 3: We define a new correspondence $\Psi : \Gamma^k \rightrightarrows \Gamma^k$ using Ψ^1 and Ψ^2:

$$\Psi(\gamma) = \{\gamma' \in \Gamma^k \mid \gamma'(a, X|t, \theta) = \gamma^2_{S^k(t)}(a_{S^k(t)}, X_{S^k(t)}|t, \theta) \prod_{i \notin S^k(t)} \gamma^1_i(a_i, X_i|t_i) \text{ for all } (a, X, t, \theta),$$
$$\text{where } \gamma^1 \in \Psi^1(\gamma) \text{ and } \gamma^2 \in \Psi^2(\gamma)\}.$$

If γ is a fixed point of Ψ, then $\gamma \in \Psi^1(\gamma)$ and $\gamma \in \Psi^2(\gamma)$. Thus, γ satisfies (C.2) and is obedient for standard types since $\gamma \in \Psi^2(\gamma)$. Moreover, it is also obedient for perturbed types because $\gamma \in \Psi^1(\gamma)$. Therefore, γ is the desired BIBCE.

To show the existence of a fixed point, consider the set

$$\Gamma^k \circ \pi^k \equiv \{\gamma \circ \pi^k \in \Delta(A \times \mathcal{A} \times T \times \Theta) \mid \gamma \circ \pi^k(a, X, t, \theta) = \gamma(a, X|t, \theta)\pi^k(t, \theta),\ \gamma \in \Gamma^k\}.$$

Since $\Gamma^k \circ \pi^k$ is a tight closed subset of $\Delta(A \times \mathcal{A} \times T \times \Theta)$, it follows from Lemma A that it is compact. Now, we regard Ψ as a correspondence defined on $\Gamma^k \circ \pi^k$ by identifying Γ^k with $\Gamma^k \circ \pi^k$. It can then be readily shown that Ψ has a closed graph and nonempty convex values, since Ψ^1 and Ψ^2 possess these properties. Thus, by the Kakutani-Fan-Glicksberg fixed point theorem, Ψ has a fixed point.

- $\Psi(\gamma)$ is nonempty because $\Psi^1(\gamma)$ and $\Psi^2(\gamma)$ are nonempty.
- $\Psi(\gamma)$ is convex because $\Psi^1(\gamma)$ and $\Psi^2(\gamma)$ are convex.
- $\Psi(\gamma)$ has a closed graph because $\Psi^1(\gamma)$ and $\Psi^2(\gamma)$ have closed graphs.

Then, by the Kakutani-Fan-Glicksberg fixed point theorem and the above argument, Ψ has a fixed point. □

Let γ^k and σ^k be the BIBCE of (T, Θ, π^k, u) constructed in Lemma B. We now prove Lemma 8 using them. Define $\eta^k \in \Delta(A \times \mathcal{A} \times T \times \Theta)$ by

$$\eta^k(a, X, t, \theta) \equiv \gamma^k \circ \pi^k(a, X, \tau^{-1}(t), \theta) = \sum_{t' \in \tau^{-1}(t)} \gamma^k(a, X|t', \theta)\pi^k(t', \theta).$$

Note that $\eta^k(t, \theta) \equiv \sum_{(a,X) \in A \times \mathcal{A}} \eta^k(a, X, t, \theta) = \pi^k(\tau^{-1}(t), \theta)$, which is zero if $t \notin \tau(T)$. We write $\eta^k(a, X|t, \theta) \equiv \eta^k(a, X, t, \theta)/\eta^k(t, \theta)$ for $(t, \theta) \in T \times \Theta$ with $\eta^k(t, \theta) > 0$. It can be readily shown that $\{\eta^k\}_{k=1}^\infty$ is tight, so it has a convergent subsequence by Lemma A, which is denoted by $\{\eta^{k_l}\}_{l=1}^\infty$ with $\lim_{l \to \infty} \eta^{k_l} = \eta^*$. Note that $\eta^*(t, \theta) = \pi(t, \theta)$ for each $(t, \theta) \in T^* \times \Theta^*$. Thus, we have $\eta^*(a, X|t, \theta)\pi(t, \theta) = \eta^*(a, X, t, \theta)$. We will regard $\eta^*(a, X|t, \theta)$ as an $\mathcal{A}$-decision rule in $(T^*, \Theta^*, \pi^*, u^*)$.

To prove Lemma 8, we show that $\{\sigma^k\}_{k=1}^\infty$ satisfies (12), which can be rewritten as

$$\lim_{k \to \infty} \inf_{\sigma \in \mathcal{E}^F} \sum_{(a,t,\theta) \in A \times T \times \Theta} \left| \eta^k(a, t, \theta) - \sigma \circ \pi(a, t, \theta) \right| = 0. \tag{C.3}$$

To establish (C.3), it suiffices to show that, for every convergent subsequence $\{\eta^{k_l}\}_{l=1}^\infty$, there exists $\sigma \in \mathcal{E}^F$ such that $\eta^*(a, t, \theta) = \sigma \circ \pi(a, t, \theta)$. This is because if (C.3) does not hold, then there exists a convergent subsequence $\{\eta^{k_l}\}_{l=1}^\infty$ such that

$$\lim_{l \to \infty} \inf_{\sigma \in \mathcal{E}^F} \sum_{(a,t,\theta) \in A \times T \times \Theta} \left| \eta^{k_l}(a, t, \theta) - \sigma \circ \pi(a, t, \theta) \right| > 0,$$

which implies that no $\sigma \in \mathcal{E}^F$ satisfies $\eta^*(a, t, \theta) = \sigma \circ \pi(a, t, \theta)$.

We denote an arbitrary convergent subsequence by $\{\eta^k\}_{k=1}^\infty$ with $\lim_{k \to \infty} \eta^k = \eta^*$, instead of $\{\eta^{k_l}\}_{l=1}^\infty$, to simplify notation. The existence of $\sigma \in \mathcal{E}^F$ satisfying $\eta^*(a, t, \theta) = \eta^*(a|t, \theta)\pi(t, \theta) = \sigma \circ \pi(t, \theta)$ follows from the next three lemmas.

Lemma C. *The $\mathcal{A}$-decision rule $\eta^*(a, X|t, \theta)$ in $(T^*, \Theta^*, \pi^*, u^*)$ is belief-invariant.*

Lemma D. *The $\mathcal{A}$-decision rule $\eta^*(a, X|t, \theta)$ in $(T^*, \Theta^*, \pi^*, u^*)$ is obedient.*

Lemma E. *The $\mathcal{A}$-decision rule $\eta^*(a, X|t, \theta)$ in $(T^*, \Theta^*, \pi^*, u^*)$ is an element of Γ^F.*

Proof of Lemma C. Fix $i \in I$ and $(t, \theta) \in T^* \times \Theta^*$. Consider sufficiently large k satisfying $\varepsilon^k < \pi^k(\tau^{-1}(t)) \leq \pi^k(\{\tau_i^{-1}(t_i)\} \times T_{-i})$, which exists since $\lim_{k\to\infty} \pi^k(\tau^{-1}(t))) = \pi(t) > 0$. Define

$$\zeta_i^k(a_i, X_i|t_i) \equiv \sum_{t_i' \in \tau_i^{-1}(t_i)} \gamma_i^k(a_i, X_i|t_i')\pi^k(t_i')/\pi^k(\tau_i^{-1}(t_i)),$$

where $\gamma_i^k(a_i, X_i|t_i') \equiv \sum_{a_{-i}, X_{-i}} \gamma^k(a, X|t', \theta)$, which is well-defined due to the belief-invariance of γ^k. Recall that

$$\eta^k(a, X|t, \theta) = \sum_{t' \in \tau^{-1}(t)} \gamma^k(a, X|t', \theta)\pi^k(t', \theta)/\pi^k(\tau^{-1}(t), \theta),$$

and thus

$$\eta^k(a_i, X_i|t, \theta) = \sum_{t_i' \in \tau_i^{-1}(t_i)} \gamma_i^k(a_i, X_i|t_i')\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)/\pi^k(\tau^{-1}(t), \theta).$$

To establish the belief-invariance of $\eta^*(\cdot|t, \theta)$, it is enough to show that

$$\begin{aligned} &\lim_{k\to\infty} |\eta^k(a_i, X_i|t, \theta) - \zeta_i^k(a_i, X_i|t_i)| \\ &= \lim_{k\to\infty} \left| \sum_{t_i' \in \tau_i^{-1}(t_i)} \gamma_i^k(a_i, X_i|t_i') \underbrace{\left(\frac{\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi^k(\tau_i^{-1}(t_i))} \right)}_{\text{(C.4)*}} \right| = 0 \qquad \text{(C.4)} \end{aligned}$$

because this implies that

$$\begin{aligned} &|\eta^*(a_i, X_i|t, \theta) - \eta^*(a_i, X_i|(t_i, t_{-i}'), \theta')| \\ &\leq \lim_{k\to\infty} |\eta^k(a_i, X_i|t, \theta) - \eta^k(a_i, X_i|(t_i, t_{-i}'), \theta')| \\ &\leq \lim_{k\to\infty} |\eta^k(a_i, X_i|t, \theta) - \zeta_i^k(a_i, X_i|t_i)| + \lim_{k\to\infty} |\eta^k(a_i, X_i|(t_i, t_{-i}'), \theta') - \zeta_i^k(a_i, X_i|t_i)| = 0. \end{aligned}$$

To show (C.4), observe that the absolute value of (C.4)* satisfies the following inequality:

$$\begin{aligned} \left| \frac{\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi^k(\tau_i^{-1}(t_i))} \right| &\leq \left| \frac{\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi(t_i)} \frac{\pi(t, \theta)}{\pi^k(\tau^{-1}(t), \theta)} \right| \\ &+ \left| \frac{\pi^k(t_i')}{\pi(t_i)} \frac{\pi(t, \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi(t_i)} \right| + \left| \frac{\pi^k(t_i')}{\pi(t_i)} - \frac{\pi^k(t_i')}{\pi^k(\tau_i^{-1}(t_i))} \right|. \end{aligned}$$

Using the definition of an ε-elaboration, we can evaluate each term in the right-hand side as follows: there exists $T_i^{\flat,k} \subset T_i$ with $\pi^k(T_i^{\flat,k}) \geq 1 - \varepsilon^k$ such that, for all $t_i' \in \tau_i^{-1}(t_i) \cap T_i^{\flat,k}$ (which is nonempty because $\pi^k(\tau_i^{-1}(t_i)) > \varepsilon^k$),

$$\left| \frac{\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi(t_i)} \frac{\pi(t, \theta)}{\pi^k(\tau^{-1}(t), \theta)} \right| = \frac{\pi^k(t_i')}{\pi^k(\tau^{-1}(t), \theta)} \left| \pi^k(\tau_{-i}^{-1}(t_{-i}), \theta | t_i') - \pi(t_{-i}, \theta | t_i) \right|$$
$$\leq \frac{\pi^k(t_i')}{\pi^k(\tau^{-1}(t), \theta)} \varepsilon^k,$$

$$\left| \frac{\pi^k(t_i')}{\pi(t_i)} \frac{\pi(t, \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi(t_i)} \right| = \frac{\pi^k(t_i')}{\pi(t_i)\pi^k(\tau^{-1}(t), \theta)} \left| \pi(t, \theta) - \pi^k(\tau^{-1}(t), \theta) \right| \leq \frac{\pi^k(t_i')}{\pi(t_i)\pi^k(\tau^{-1}(t), \theta)} \varepsilon^k,$$

$$\left| \frac{\pi^k(t_i')}{\pi(t_i)} - \frac{\pi^k(t_i')}{\pi^k(\tau_i^{-1}(t_i))} \right| = \frac{\pi^k(t_i')}{\pi(t_i)\pi^k(\tau_i^{-1}(t_i))} \left| \pi^k(\tau_i^{-1}(t_i)) - \pi(t_i) \right| \leq \frac{\pi^k(t_i')}{\pi(t_i)\pi^k(\tau_i^{-1}(t_i))} \varepsilon^k.$$

Thus,

$$\left| \frac{\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi^k(\tau_i^{-1}(t_i))} \right|$$
$$\leq \varepsilon^k \pi^k(t_i') \left(\frac{1}{\pi^k(\tau^{-1}(t), \theta)} + \frac{1}{\pi(t_i)\pi^k(\tau^{-1}(t), \theta)} + \frac{1}{\pi(t_i)\pi^k(\tau_i^{-1}(t_i))} \right).$$

Then,

$$|\eta^k(a_i, X_i | t, \theta) - \zeta_i^k(a_i, X_i | t_i)|$$
$$\leq \left| \sum_{t_i' \in \tau_i^{-1}(t_i) \cap T_i^{\flat,k}} \gamma_i^k(a_i, X_i | t_i') \left(\frac{\pi^k(\{t_i'\} \times \tau_{-i}^{-1}(t_{-i}), \theta)}{\pi^k(\tau^{-1}(t), \theta)} - \frac{\pi^k(t_i')}{\pi^k(\tau_i^{-1}(t_i))} \right) \right| + \pi^k(\tau_i^{-1}(t_i) \setminus T_i^{\flat,k})$$
$$\leq \varepsilon^k \pi^k(\tau_i^{-1}(t_i) \cap T_i^{\flat,k}) \left(\frac{1}{\pi^k(\tau^{-1}(t), \theta)} + \frac{1}{\pi(t_i)\pi^k(\tau^{-1}(t), \theta)} + \frac{1}{\pi(t_i)\pi^k(\tau_i^{-1}(t_i))} \right) + \varepsilon^k \to 0$$

as $k \to \infty$, which implies (C.4). □

Proof of Lemma D. Because γ^k is obedient in (T, Θ, π^k, u), it holds that

$$\sum_{a_{-i}, X_{-i}} \left(\sum_{t_{-i}, \theta} \gamma^k(a, X|t, \theta)\pi^k(t, \theta) \right) u_i(a, \theta) \geq \sum_{a_{-i}, X_{-i}} \left(\sum_{t_{-i}, \theta} \gamma^k(a, X|t, \theta)\pi^k(t, \theta) \right) u_i((a_i', a_{-i}), \theta)$$

for all $t_i \in T_i$, $X_i \in \mathcal{A}_i$, $a_i \in X_i$, $a_i' \in A_i$, and $i \in I$. By summing over all $t_i \in \tau_i^{-1}(t_i')$ for each $t_i' \in \tau_i(T_i)$, we obtain

$$\sum_{a_{-i}, X_{-i}, t_{-i}', \theta} \eta^k(a, X, t', \theta) u_i(a, \theta) \geq \sum_{a_{-i}, X_{-i}, t_{-i}', \theta} \eta^k(a, X, t', \theta) u_i((a_i', a_{-i}), \theta), \tag{C.5}$$

which follows from

$$\begin{aligned}
\sum_{t_i \in \tau_i^{-1}(t_i')} \left(\sum_{t_{-i}, \theta} \gamma^k(a, X|t, \theta)\pi^k(t, \theta) \right) &= \sum_{t_i \in \tau_i^{-1}(t_i')} \left(\sum_{(t_{-i}', \theta) \in \tau_{-i}(T_{-i}) \times \Theta} \; \sum_{t_{-i} \in \tau_{-i}^{-1}(t_{-i}')} \gamma^k(a, X|t, \theta)\pi^k(t, \theta) \right) \\
&= \sum_{(t_{-i}', \theta) \in \tau_{-i}(T_{-i}) \times \Theta} \; \sum_{t \in \tau^{-1}(t')} \gamma^k(a, X|t, \theta)\pi^k(t, \theta) \\
&= \sum_{(t_{-i}', \theta) \in \tau_{-i}(T_{-i}) \times \Theta} \eta^k(a, X, t', \theta).
\end{aligned}$$

Taking the limit of (C.5) as $k \to \infty$, we obtain

$$\sum_{a_{-i}, X_{-i}, t_{-i}, \theta} \eta^*(a, X|t, \theta)\pi(t, \theta) u_i(a, \theta) \geq \sum_{a_{-i}, X_{-i}, t_{-i}, \theta} \eta^*(a, X|t, \theta)\pi(t, \theta) u_i((a_i', a_{-i}), \theta)$$

since $\eta^*(t, \theta) = \pi(t, \theta)$. Therefore, $\eta^*(a, X|t, \theta)$ is obedient. □

Proof of Lemma E. We show that

$$\sum_{X, t, \theta} \eta^*(X|t, \theta)\pi(t, \theta) F(X, \theta) \geq \sum_{X, t, \theta} \hat{\gamma}(X|t, \theta)\pi(t, \theta) F(X, \theta) \tag{C.6}$$

for arbitrary $\hat{\gamma} \in \Gamma^F$. Fix $\hat{\gamma} \in \Gamma^F$ and let $\hat{\gamma}^k \in \Gamma^k[\gamma^k]$ be such that

$$\hat{\gamma}^k_{S^k(t)}(a_{S^k(t)}, X_{S^k(t)}|t, \theta) = \hat{\gamma}_{S^k(t)}(a_{S^k(t)}, X_{S^k(t)}|\tau(t), \theta),$$

which is well-defined because $\hat{\gamma}$ is belief-invariant. Note that $\hat{\gamma}^k(a, X|t, \theta) = \hat{\gamma}(a, X|\tau(t), \theta)$

if $t \in T^k$. Then, $\gamma^k \in \Gamma^{k,F}[\gamma^k]$ implies that

$$\sum_{X,t,\theta} \gamma^k(X|t,\theta)\pi^k(t,\theta)F(X,\theta) \geq \sum_{X,t,\theta} \hat{\gamma}^k(X|t,\theta)\pi^k(t,\theta)F(X,\theta). \tag{C.7}$$

We show that taking the limit of (C.7) yields (C.6).

Consider the left-hand side of (C.7). Then,

$$\begin{aligned}
&\sum_{X,t,\theta} \gamma^k(X|t,\theta)\pi^k(t,\theta)F(X,\theta) \\
&\qquad = \sum_{X,t',\theta}\left(\sum_{t\in\tau^{-1}(t')} \gamma^k(X|t,\theta)\pi^k(t,\theta)\right)F(X,\theta) = \sum_{X,t',\theta} \eta^k(X,t',\theta)F(X,\theta) \\
&\qquad \to \sum_{X,t',\theta} \eta^*(X,t',\theta)F(X,\theta) = \sum_{X,t',\theta} \eta^*(X|t',\theta)\pi(t',\theta)F(X,\theta)
\end{aligned} \tag{C.8}$$

as $k \to \infty$.

Consider the right-hand side of (C.7). Then,

$$\sum_{X,t,\theta} \hat{\gamma}^k(X|t,\theta)\pi^k(t,\theta)F(X,\theta) \geq \sum_{X,\theta}\sum_{t\in T^k} \hat{\gamma}(X|\tau(t),\theta)\pi^k(t,\theta)F(X,\theta) + \pi^k(T\setminus T^k)\inf_{X,\theta} F(X,\theta)$$

since $\hat{\gamma}^k(X|t,\theta) = \hat{\gamma}(X|\tau(t),\theta)$ for $t \in T^k$. The first term in the right-hand side satisfies

$$\begin{aligned}
\sum_{X,\theta}\sum_{t\in T^k} \hat{\gamma}(X|\tau(t),\theta)\pi^k(t,\theta)F(X,\theta) &\geq \sum_{X,t,\theta} \hat{\gamma}(X|\tau(t),\theta)\pi^k(t,\theta)F(X,\theta) - \pi^k(T\setminus T^k)\sup_{X,\theta} F(X,\theta) \\
&= \sum_{X,t,\theta} \hat{\gamma}(X|t,\theta)\pi^k(\tau^{-1}(t),\theta)F(X,\theta) - \pi^k(T\setminus T^k)\sup_{X,\theta} F(X,\theta) \\
&= \sum_{X,t,\theta} \hat{\gamma}(X|t,\theta)\eta^k(t,\theta)F(X,\theta) - \pi^k(T\setminus T^k)\sup_{X,\theta} F(X,\theta).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\lim_{k\to\infty} \sum_{X,t,\theta} \hat{\gamma}^k(X|t,\theta)\pi^k(t,\theta)F(X,\theta) \\
&\qquad \geq \lim_{k\to\infty} \sum_{X,t,\theta} \hat{\gamma}(X|t,\theta)\eta^k(t,\theta)F(X,\theta) + \pi^k(T\setminus T^k)\left(\inf_{X,\theta} F(X,\theta) - \sup_{X,\theta} F(X,\theta)\right) \\
&\qquad = \sum_{X,t,\theta} \hat{\gamma}(X|t,\theta)\eta^*(t,\theta)F(X,\theta) = \sum_{X,t,\theta} \hat{\gamma}(X|t,\theta)\pi(t,\theta)F(X,\theta).
\end{aligned} \tag{C.9}$$

It follows from (C.7), (C.8), and (C.9) that (C.6) holds. □

D Proofs for Section 6

D.1 Proof of Proposition 1

A function $f : A \to \mathbb{R}$ is supermodular if, for any $a, b \in A$,

$$f(a \vee b) + f(a \wedge b) \geq f(a) + f(b),$$

where $a \vee b = (\max\{a_i, b_i\})_{i \in I}$ is the join of a and b, and $a \wedge b = (\min\{a_i, b_i\})_{i \in I}$ is the meet of a and b. It is known that a potential function $v(\cdot, \theta) : A \to \mathbb{R}$ of a supermodular game is a supermodular function for each $\theta \in \Theta$.

A supermodular function has the following property.

Lemma F. *Let $f : A \to \mathbb{R}$ be a supermodular function. For arbitrary $\mu \in \Delta(A)$, there exists $\mu^* \in \Delta(A)$ satisfying the following conditions.*

- *The marginal probability distributions of μ and μ^* on A_i are the same for each $i \in I$.*
- $\sum_a \mu^*(a) f(a) \geq \sum_a \mu(a) f(a)$.
- *The support of μ^* is linearly ordered with respect to the product order $\geq_I$.*

Proof. Let $\geq^L$ denote the lexicographic order on A; that is, $a >^L b$ if there exists $i \in I$ such that $a_i > b_i$ and $a_j = b_j$ for all $j < i$. For $A' = \{a^1, \dots, a^K\} \subseteq A$, which is lexicographically ordered as $a^1 >^L a^2 >^L \cdots >^L a^K$, define $\kappa_1(A') \in \{1, \dots, K\}$ as the smallest index such that there exists $k \geq \kappa_1(A') + 1$ satisfying $a^{\kappa_1(A')} \not\geq_I a^k$. This index exists unless A' is linearly ordered with respect to $\geq_I$. If A' is linearly ordered, we set $\kappa_1(A') = K$. When $\kappa_1(A') < K$, define $\kappa_2(A') \geq \kappa_1(A') + 1$ as the smallest index satisfying $a^{\kappa_1(A')} \not\geq_I a^{\kappa_2(A')}$. Note that, for each $k < \kappa_1(A')$ and $l > k$, we have $a^k >_I a^l$ as well as $a^k >^L a^l$. We write $\alpha(A') = a^{\kappa_1(A')}$ and $\beta(A') = a^{\kappa_2(A')}$.

For each $\mu \in \Delta(A)$ and $a, b \in A$ with $a \not\geq_I b$ and $b \not\geq_I a$, let $\phi[\mu | a, b] \in \Delta(A)$ be such that

$$\phi[\mu|a,b](x) = \begin{cases} \mu(x) - \min_{x'\in\{a,b\}} \mu(x') & \text{if } x \in \{a,b\}, \\ \mu(x) + \min_{x'\in\{a,b\}} \mu(x') & \text{if } x \in \{a \vee b, a \wedge b\}, \\ \mu(x) & \text{if } x \notin \{a,b,a \vee b, a \wedge b\}. \end{cases}$$

Then,

$$\sum_x \phi[\mu|a,b](x) f(x) - \sum_x \mu(x) f(x) = \min_{x'\in\{a,b\}} \mu(x')(f(a\vee b) + f(a\wedge b) - f(a) - f(b)) \geq 0$$

because f is a supermodular function. In addition, the marginal probability distributions of $\phi[\mu|a,b]$ and μ on A_i are the same for each $i \in I$.

Fix $\mu \in \Delta(A)$. Let $\mu^0 = \mu$ and $A^0 = \text{supp}(\mu^0)$. For each $n \geq 1$, we construct μ^n and A^n from μ^{n-1} and A^{n-1} as follows. If $\kappa_1(A^{n-1}) < |A^{n-1}|$, define $\mu^n \in \Delta(A)$ and $A^n = \text{supp}(\mu^n)$ by

$$\mu^n = \phi[\mu^{n-1}|\alpha(A^{n-1}), \beta(A^{n-1})].$$

The support satisfies

$$A^n = (A^{n-1} \cup \{\alpha(A^{n-1}) \vee \beta(A^{n-1}), \alpha(A^{n-1}) \wedge \beta(A^{n-1})\}) \setminus \arg \min_{x \in \{\alpha(A^{n-1}), \beta(A^{n-1})\}} \mu(x).$$

Repeat the construction as long as $\kappa_1(A^n) < |A^n|$.

If $\kappa_1(A^n) = |A^n|$, we terminate the construction and set $\mu^* = \mu^n$, which is the desired probability distribution. To see why, note that $\text{supp}(\mu^*) = A^n$ is linearly ordered with respect to $\geq_I$ since $\kappa_1(A^n) = |A^n|$, and that μ and μ^* have the same marginal distributions because, for each k, μ^k and μ^{k+1} have the same marginal distributions. Moreover, $\sum_x \mu^*(x) f(x) = \sum_x \mu^n(x) f(x) \geq \sum_x \mu^{n-1}(x) f(x) \geq \cdots \geq \sum_x \mu(x) f(x)$.

We show that there exists n satisfying $\kappa_1(A^n) = |A^n|$. Define $\kappa_1^n = \kappa_1(A^n)$ and let $A^n = \{a^{1,n}, \ldots, a^{K^n,n}\}$ be lexicographically ordered as $a^{1,n} \geq^L \cdots \geq^L a^{K^n,n}$.

Assume $\kappa_1^n < |A^n|$. By construction, $a^{1,n} \vee \cdots \vee a^{K^n,n}$ and $a^{1,n} \wedge \cdots \wedge a^{K^n,n}$ are independent of n. Moreover, we have $a^{1,n+1} \geq_I a^{1,n}$ because if $\kappa_1^n = 1$ then $a^{1,n+1} = a^{1,n} \vee \beta(A^n)$ and if $\kappa_1^n \geq 2$ then $a^{1,n+1} = a^{1,n}$. In the latter case, $a^{1,n} \geq_I a$ for all $a \in A^n$, which means $a^{1,n} = a^{1,n} \vee \cdots \vee a^{K^n,n}$. As a result, there exists a minimum integer $n_1 \geq 0$ such that $\kappa_1^n \geq 2$ and $a^{1,n} = a^{1,n_1}$ for all $n \geq n_1$.

Next, assume $\kappa_1^{n_1} < |A^{n_1}|$. For $n \geq n_1$, we must have $a^{2,n+1} \geq_I a^{2,n}$ since if $\kappa_1^n = 2$ then $a^{2,n+1} = a^{2,n} \vee \beta(A^n)$ and if $\kappa_1^n \geq 3$ then $a^{2,n+1} = a^{2,n}$. In the latter case, $a^{2,n} \geq_I a$ for all $a \in A^n \setminus \{a^{1,n}\}$, which means $a^{2,n} = a^{2,n} \vee \cdots \vee a^{K^n,n}$. As a result, there exists a minimum integer $n_2 \geq n_1$ such that $\kappa_1^n \geq 3$ and $a^{2,n} = a^{2,n_2}$ for all $n \geq n_2$.

By repeating this argument, we can construct a sequence of integers $\{n_k\}_{k=1}$ such that, for each $k \geq 1$ with $\kappa_1^{n_k} < |A^{n_k}|$, the next integer $n_{k+1} \geq n_k$ is the smallest integer satisfying $\kappa_1^n \geq k+2$ and $a^{k+1,n} = a^{k+1,n_{k+1}}$ for all $n \geq n_{k+1}$. In particular, we have $\kappa_1^{n_{k+1}} \geq k+2$. Since $\kappa_1^n \leq |A|$, $\{n_k\}_{k=1}$ cannot be an infinite sequence. Therefore, there must exsit n with $\kappa_1^n = |A^n|$. □

We are ready to prove Proposition 1. Let σ be a P-maximizing BIBCE. By Lemma F, there exists $\sigma^* \in \Sigma^{BI}$ that satisfies the following conditions for each $(t,\theta) \in T \times \Theta$.

- $\sigma^*(\{a_i\} \times A_{-i}|t,\theta) = \sigma(\{a_i\} \times A_{-i}|t,\theta) = \sigma_i(a_i|t_i)$ for all $a_i \in A_i$.
- $\sum_a \sigma^*(a|t,\theta)v(a,\theta) \geq \sum_a \sigma(a|t,\theta)v(a,\theta)$.
- The support of $\sigma^*(\cdot|t,\theta)$ is linearly ordered with respect to $\geq_I$.

For each $i \in I$, let $\overline{a_i}(t_i)$ and $\underline{a_i}(t_i)$ denote the maximum and minimum actions in the support of $\sigma_i^*(\cdot|t_i) = \sigma_i(\cdot|t_i) \in \Delta(A_i)$, respectively. By the third condition, the maximum and minimum action profile in the support of $\sigma^*(\cdot|t,\theta)$ must be $\overline{a}(t) \equiv (\overline{a_i}(t_i))_{i\in I}$ and $\underline{a}(t) \equiv (\underline{a_i}(t_i))_{i\in I}$, respectively.

By the second condition, it holds that

$$\sum_{a,t,\theta} \sigma^*(a|t,\theta)\pi(t,\theta)v(a,\theta) \geq \sum_{a,t,\theta} \sigma(a|t,\theta)\pi(t,\theta)v(a,\theta),$$

which implies that $\sigma^* \in \mathcal{E}^v$. This further implies that, for each $(t,\theta) \in T^* \times \Theta^*$, the function $v(\cdot,\theta)$ is constant over the support of $\sigma^*(\cdot|t,\theta)$; that is, $v(\overline{a}(t),\theta) = v(\underline{a}(t),\theta) = v(a,\theta)$ for any a in the support. Otherwise, there would exist an action profile a in the support that attains the maximum value of $v(\cdot,\theta)$ which is strictly greater than that of some other action profile in the support. In that case, we could construct a belief-invariant decision rule that achieves a strictly higher expected value of the potential function than σ^* by assigning probability one to a when (t,θ) is realized, which contradicts $\sigma^* \in \mathcal{E}^v$.

Since $\overline{\sigma}(\overline{a}(t)|t,\theta) = 1$ and $\underline{\sigma}(\underline{a}(t)|t,\theta) = 1$ for all $(t,\theta) \in T^* \times \Theta^*$, we have

$$\sum_{a,t,\theta} \sigma^*(a|t,\theta)\pi(t,\theta)v(a,\theta) = \sum_{a,t,\theta} \underline{\sigma}(a|t,\theta)\pi(t,\theta)v(a,\theta) = \sum_{a,t,\theta} \overline{\sigma}(a|t,\theta)\pi(t,\theta)v(a,\theta),$$

which implies that $\underline{\sigma}, \overline{\sigma} \in \mathcal{E}^v$. □

D.2 Proof of Lemma 9

In the proof, we use a one-to-one mapping $\Lambda_i : A_i \to \mathcal{A}_i$ defined by $\Lambda_i(1) = \{1\}$ and $\Lambda_i(0) = \{0,1\}$ for each $i \in I$. Note that $v(a,\theta) = F(\Lambda(a),\theta)$ for all $(a,\theta) \in A \times \Theta$.

Let $P_i \in \Delta(A_{-i} \times \mathcal{A}_{-i} \times \Theta)$ satisfy $P_i(a_{-i}, X_{-i}, \theta) = 0$ if $a_{-i} \notin X_{-i}$. For each increasing subset $B_{-i} \subseteq A_{-i}$ (i.e. $a_{-i} \in B_{-i}$ and $a'_{-i} \geq a_{-i}$ imply $a'_{-i} \in B_{-i}$), we have

$$\begin{aligned}\sum_{a_{-i} \in B_{-i}} P_i(\Lambda_{-i}(a_{-i}),\theta) &= \sum_{a_{-i} \in B_{-i}} \sum_{a'_{-i} \in B_{-i}} P_i(a'_{-i}, \Lambda_{-i}(a_{-i}),\theta) \\ &\leq \sum_{X_{-i} \in \mathcal{A}_{-i}} \sum_{a'_{-i} \in B_{-i}} P_i(a'_{-i}, X_{-i},\theta) = \sum_{a_{-i} \in B_{-i}} P_i(a_{-i},\theta).\end{aligned}$$

Thus, $P_i(a_{-i}|\theta) \in \Delta(A_{-i})$ first-order stochastically dominates $P_i(\Lambda_{-i}(a_{-i})|\theta) \in \Delta(A_{-i})$.

Using the first-order stochastic dominance and supermodularity, we show that F is a generalized potential function; that is, if X_i satisfies (9), it also satisfies (10). Since the case of $X_i = \{0,1\}$ is trivially true, we consider the case of $X_i = \{1\}$. If (9) is true when $X_i = \{0,1\}$, we have

$$\sum_{a_{-i},\theta} P_i(\Lambda_{-i}(a_{-i}),\theta)(v((1,a_{-i}),\theta) - v((0,a_{-i}),\theta) \geq 0.$$

Thus, if the game is supermodular, (10) holds because

$$\begin{aligned}&\sum_{a_{-i},\theta} P_i(a_{-i},\theta)(u_i((1,a_{-i}),\theta) - u_i((0,a_{-i}),\theta)) \\ &\qquad \geq \sum_{a_{-i},\theta} P_i(\Lambda_{-i}(a_{-i}),\theta)(u_i((1,a_{-i}),\theta) - u_i((0,a_{-i}),\theta)) \\ &\qquad \geq \lambda_i^{-1} \sum_{a_{-i},\theta} P_i(\Lambda_{-i}(a_{-i}),\theta)(v((1,a_{-i}),\theta) - v((0,a_{-i}),\theta)) \geq 0,\end{aligned}$$

where the first inequality follows from the first-order stochastic dominance of $P_i(a_{-i}|\theta)$ over $P_i(\Lambda_{-i}(a_{-i})|\theta)$, and the second from (23). Similarly, if the monotone potential function is supermodular, (10) also holds because

$$\begin{aligned}
&\sum_{a_{-i},\theta} P_i(a_{-i},\theta)(u_i((1,a_{-i}),\theta) - u_i((0,a_{-i}),\theta)) \\
&\qquad \geq \lambda_i^{-1} \sum_{a_{-i},\theta} P_i(a_{-i},\theta)(v((1,a_{-i}),\theta) - v((0,a_{-i}),\theta)) \\
&\qquad \geq \lambda_i^{-1} \sum_{a_{-i},\theta} P_i(\Lambda_{-i}(a_{-i}),\theta)(v((1,a_{-i}),\theta) - v((0,a_{-i}),\theta)) \geq 0,
\end{aligned}$$

where the first inequality follows from (23), and the second from the first-order stochastic dominance. Therefore, F is a generalized potential function. □

References

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

Bergemann, D., Morris, S., 2016. Bayes correlated equilibrium and the comparison of information structures in games. Theor. Econ. 11, 487–522.

Bergemann, D., Morris, S., 2017. Belief-free rationalizability and informational robustness. Games Econ. Behav. 104, 744–759.

Bergemann, D., Morris, S., 2019. Information design: A unified perspective. J. Econ. Lit. 57, 44–95.

Billingsley, P., 1999. Convergence of Probability Measures, 2nd Edition. John Wiley & Sons.

Carlsson, H., van Damme, E., 1993. Global games and equilibrium selection. Econometrica 61, 989–1018.

Dekel, E., Fudenberg, D., Morris, S., 2007. Interim correlated rationalizability. Theor. Econ. 2, 15–40.

Ely, J., Peski, M., 2006. Hierarchies of belief and interim rationalizability. Theor. Econ. 1, 19–65.

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

Forges, F., 2006. Correlated equilibrium in games with incomplete information revisited. Theory Dec. 61, 329–344.

Harsanyi, J. C., 1967–1968. Games with incomplete information played by Bayesian players. Manage. Sci. 14, 159–182, 320–334, 486–502.

Kajii, A., Morris, S. 1997. The Robustness of equilibria to incomplete information. Econometrica 65, 1283–1309.

Kajii, A., Morris, S. 2020a. Refinements and higher order beliefs: A unified survey. Jpn. Econ. Rev. 71, 7–34.

Kajii, A., Morris, S. 2020b. Notes on "Refinements and higher order beliefs." Jpn. Econ. Rev. 71, 35–41.

Lehrer, E., Rosenberg, D., Shmaya, E., 2010. Signaling and mediation in games with common interest. Games Econ. Behav. 68, 670–682.

Liu, Q., 2015. Correlation and common priors in games with incomplete information. J. Econ. Theory 157, 49–75.

Mertens, J.-F., Zamir, S., 1985. Formulation of Bayesian analysis for games with incomplete information. Int. J. Game Theory 14, 1–29.

Milgrom, P. R., Weber, R. J., 1985. Distributional strategies for games with incomplete information. Math. Oper. Res. 10, 619–632.

Morris, S., Oyama, D., Takahashi, S., 2024. Implementation via information design in binary-action supermodular games. Econometrica 92, 775–813.

Morris, S., Ui, T., 2005. Generalized potentials and robust sets of equilibria. J. Econ. Theory 124, 45–78.

Morris, S., Ui, T., 2020. Incomplete Information Robustness. Working Papers on Central Bank Communication 019, University of Tokyo.

Monderer, D., Shapley, L. S., 1996. Potential games. Games Econ. Behav. 14, 124–143.

Oyama, D., Takahashi, S., 2020. Generalized belief operator and robustness in binary-action supermodular games. Econometrica 88, 693–726.

Pram, K., 2019. On the Equivalence of robustness to canonical and general elaborations. J. Econ. Theory 180, 1–10.

Rubinstein, A., 1989. The electronic mail game: Strategic behavior under "Almost Common Knowledge." Amer. Econ. Rev. 385–391.

Takahashi, S., 2020. Non-equivalence between all and canonical elaborations. Jpn. Econ. Rev. 71, 43–57.

Ui, T., 2001. Robust equilibria of potential games. Econometrica 69, 1373–1380.