Robust prediction in games with uncertain parameters*

Shintaro Miura[†] Takuro Yamashita[‡]

April 27, 2022

Abstract

We consider games where an analyst is not confident about players' true information structure for payoff-relevant parameters. We define a robust prediction by a set of action profiles such that, given any information structure among the players, there is a Bayesian Nash equilibrium given that information structure whose equilibrium action profiles are in this set. We then show that, in order to identify a robust prediction, it is sufficient to focus on the *canonical type space*, a Harsanyi type space constructed as an analogy of level-k theory. As applications of our approach, we derive robust predictions in Cournot competition, Diamond search games, and auctions.

*We are grateful to Pierpaolo Battigalli, Takashi Kunimoto, Jiangtao Li, Satoru Takahashi, and seminar participants at Singapore Management University, Bocconi University, 26th Decentralization Conference in Japan, and SAET 2021. Seyed Kourosh Khounsari and David Morales Ruiz provide excellent research assistance. Shintaro Miura appreciates the financial support from JSPS Grant-in-Aid (18H03640, 19H01471, 19K13655). Takuro Yamashita acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 714693), and ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). All remaining errors are our own.

[†]Department of Economics, Kanagawa University, smiura@kanagawa-u.ac.jp

[‡]Toulouse School of Economics, takuro.yamashita@tse-fr.eu

1 Introduction

It is often the case that the actual environment analyzed using a game model is much more complicated than the model itself. For example, to model an auction, we typically assume some simple information structure for bidders, even if we are not completely confident about such an assumption. Many researchers have argued that we should keep in mind that our prediction based on a simple model/information structure may be fragile because the actual players may enjoy a different information structure than the one we assume.

The goal of this paper is to provide a "robust" prediction in such a situation, in the sense that it provides a reasonable prediction even if the true information structure differs from the assumed one.

More specifically, we consider a game with uncertain payoff parameters. A prediction assigns a subset of action profiles for each possible realization of the payoff parameters, with the interpretation that any action profile in that subset could occur given that parameter realization. We say that a prediction is *robust* if for any possible information structure, there exists some Bayesian Nash equilibrium given each possible information structure that the equilibrium action profile is in the prediction. Intuitively, if a prediction is *not robust*, then it fails to provide a reasonable prediction of the equilibrium play for some information structure. Hence, we assert that a robust prediction provides a reasonable prediction regardless of the actual information structure.

To encourage the importance of robust predictions, consider the following example.

Example 1. Consider a first-price auction with two risk-neutral bidders and private values: bidder $i \in \{1, 2\}$ has value v_i which is independent and identically determined by a cdf F.

For simplicity, assume that the analyst is confident that F is common knowledge (e.g., imagine a lab experiment where the bidders are informed about F). A standard "textbook" assumption is that they play a Bayesian Nash equilibrium, where each i bids:

$$a_i^{\mathrm{BNE}}(v_i) = v_i - \int_0^{v_i} \frac{F(x)}{F(v_i)} dx. \tag{1}$$

However, as in recent behavioral and experimental studies, some bidders may be behaving "irrationally", for example, by systematically bidding their true valuations.[1] If some bidders are such truthful types, or some bidders believe that others may be truthful, then the bidders no longer behave the same way as the textbook formula. Thus, if it is concerned that the above textbook assumption may be wrong, a more robust prediction may be required.

Even if the space of the payoff parameter values is relatively simple, the space of all possible information structures (formally represented by Harsanyi type spaces) could be very large, as the players may enjoy nontrivial higher-order beliefs. It is therefore not straightforward to determine whether a given prediction is robust or not. The main contribution of this paper is to provide a condition for a robust prediction (Theorem 1). The condition is simple in the sense that it comprises two properties: a product structure and a best-reply property. As demonstrated in the applications in Section 6, these properties are relatively easy to check. It is also worth noting that our robust prediction could have some reasonable predictive power in applications, as it could be strictly smaller than *belief-free rationalizability* (Bergemann and Morris, 2017). We discuss the underlying equilibrium selection idea behind in this paper.

[1] Truthful bidding is commonly observed in lab experiments. Kagel and Levin (1993), for example, report that about 11% of subjects bid their own value truthfully in the first-price auction. Besides, this assumption is often adopted for explaining overbidding behaviors: see, for example, Crawford and Iriberri (2007).

The proof of Theorem 1 proceeds roughly as follows. We begin with any exogenously given type space and a Bayesian Nash equilibrium given that type space as the "baseline" prediction. As in the above auction example, one might typically consider a simple type space and a "standard" equilibrium given that type space (in case there exist multiple Bayesian Nash equilibria). Based on this baseline type space, we construct a sequence of type spaces, made increasingly richer, as an analogy to *level-k theory* (referred to as *level-k type spaces*), and construct a Bayesian Nash equilibrium given those type spaces. Thus, by construction, a collection of the equilibrium action profiles given those type spaces provides a reasonable prediction if any of those level-k type spaces is the actual one.

Of course, the true type space could be different from any of them. Nevertheless, we show that this prediction is robust; there is a Bayesian Nash equilibrium given any possible (not necessarily level-k) type space whose equilibrium action profile is in the above constructed prediction. In this sense, we regard the level-k type spaces as *canonical* ones in identifying the robust prediction.

Although Theorem 1 only concerns finite games for simplicity, we also show that our argument can be extended to infinite games with additional assumptions. This extension is important, as the applications considered in this paper are all non-finite games.

The rest of the paper is structured as follows. Section 1.1 discusses some related papers in the literature, after which Section 2 introduces a model and our notion of a robust prediction. Section 3 provides a necessary condition for a robust prediction based on level-k type spaces, and we show in Section 4 that the constructed prediction based on the level-k type spaces is robust. Section 5 extends the result to non-finite games with additional assumptions, and Section 6 analyzes three applications: Cournot competition, Diamond search games, and auctions. Finally, Section 7 concludes.

1.1 Related literature

The literature of robust prediction in games can be roughly divided into the following two strands depending on the meaning of robustness. In the first strand, a robust prediction represents a "collection" of possible equilibrium outcomes under possible information structures, based on the idea that the analyst has limited knowledge about the true information structure. Bergemann and Morris (2016) consider an environment in which there is a common prior over the payoff parameters and show that their robust prediction is characterized as a set of *Bayes correlated equilibria*.[2] Likewise, Bergemann and Morris (2017) show that the robust prediction without the common prior restriction is given by a set of *belief-free rationalizable* strategies.[3] Our paper belongs to this first strand of inquiry. In the sense that we allow for non-common-prior type spaces, our approach is closer to Bergemann and Morris (2017) than to Bergemann and Morris (2016). However, we obtain sharper predictions than belief-free rationalizability of Bergemann and Morris (2017), as our robust prediction collects *some* Bayesian Nash equilibrium outcome given any possible information structure, while Bergemann and Morris (2017) collect *all* Bayesian Nash equilibrium outcomes. In other words, Bergemann and Morris (2017) require robustness for both the players' information structure and equilibrium selection, while we only require the former. The difference is stark in some applications such as first-price auction, where belief-free rationalizability predicts a large set of bidding

[2]More specifically, Bergemann and Morris (2016) consider the scenario in which there is a baseline information structure but where each player could face a different information structure by observing additional private signals. If the private signals are not restricted to payoff irrelevant ones, the robust prediction is given by the set of Bayes correlated equilibrium. Otherwise, it ends up with the set of *belief-invariant Bayes correlated equilibria* à la Liu (2015). Bergemann and Morris (2013) characterize a Bayes correlated equilibrium in a game with quadratic payoff functions and normally distributed uncertainties.

[3]If the private signals are restricted to payoff irrelevant ones, then the robust prediction is given by a set of *interim correlated rationalizable strategies* (Dekel, Fudenberg, and Morris, 2007).

functions (e.g., Battigalli and Siniscalchi, 2003), while our prediction can be much sharper. We discuss this point later in more detail.[4]

Though less related, the second strand of robust predictions concerns "selections" of equilibrium outcomes that are immune to small misspecifications based on perturbations of the baseline information structure. Weinstein and Yildiz (2007b) consider perturbations of belief hierarchies up to finite orders and then show that any rationalizable strategy of the baseline case can be their robust prediction.[5] Chen, Takahashi, and Xiong (2014) extend their notion of robust prediction by incorporating the perturbation of payoff structures, requiring that a reasonable strategy should be uniquely rationalizable not only in nearby types but also in nearby games with slightly perturbed payoff structures. They show that there generically exists no robust prediction due to types exhibiting a payoff tie. Weinstein and Yildiz (2011) select Bayesian Nash equilibrium outcomes based on the robustness to slight perturbations of belief hierarchies.[6] They call the set of outcomes that are supported by an equilibrium in all of the nearby games the *minimally-robust prediction*. They show that a lower bound of the minimally robust prediction is given by a local version of the *interim correlated rationalizability* (ICR, hereafter) of the baseline game with complete information.[7]

[4]Miura and Yamashita (2020) characterizes a robust prediction in our sense in a cheap-talk game à la Crawford and Sobel (1982) where the common knowledge on the players' behavioral types is relaxed.

[5]The validity of the conclusion is investigated from several perspectives. See, for example, Chen (2012) and Penta (2012) are for dynamic games, and Penta (2013) and Chen, Takahashi, and Xiong (2021) relax the *richness assumption*, which is a crucial property for Weinstein and Yildiz (2007b).

[6]More specifically, they consider nearby games in which (i) the payoff parameter is slightly misspecified and (ii) the true parameter is mutually known up to some finite order.

[7]While the papers on robust selection adopt perturbations of interim beliefs (or belief hierarchies), the literature also adopts perturbations of ex ante beliefs (or priors). Kajii and Morris (1997) fix a complete-information game as a baseline scenario and consider its nearby incomplete-information games, where the players share a common prior belief such that the baseline game is played with sufficiently high probability and games with different

The construction of the canonical type space is based on *level-k theory* à la Stahl and Wilson (1994, 1995) and Nagel (1995). The literature adopts the level-k reasoning for explaining "non-equilibrium behaviors" by not requiring the correct beliefs about the opponents' behaviors, as opposed to the notion of Nash equilibrium. For instance, Crawford and Iriberri (2007) exploit level-k reasoning to explain overbidding in auctions, Crawford, Kugler, Neeman, and Pauzner (2009) discuss the optimal auction when bidders have the level-k thinking and de Clippel, Saran, and Serrano (2019) study mechanism design with level-k agents.[8] It is worthwhile emphasizing that we adopt the level-k reasoning, not to justify non-equilibrium behaviors but to predict robust "equilibrium behaviors", whereas our analysis in Section 6.3 is reminiscent of Crawford and Iriberri (2007). We revisit this point later.

2 Model

Consider a strategic-form game with I players. Each player $i \in \{1, \dots, I\}$ is endowed with a *payoff parameter* $\theta_i \in \Theta_i$ and takes an action $a_i \in A_i$. His payoff is $u_i(a, \theta)$ given action profile $a = (a_i)_{i=1}^I \in A = \prod_{i=1}^I A_i$ and parameter profile $\theta = (\theta_i)_{i=1}^I \in \Theta = \prod_{i=1}^I \Theta_i$. For technical simplicity, assume that each A_i and Θ_i is finite, which is relaxed in Section 5. The maintained assumption is that each i knows (at least) his true payoff parameter θ_i. However, he may not know θ_{-i}, and may have an arbitrary belief (and higher-order belief) about it. In some applications (See Sections 6.1 and 6.2, for example), one may want to include an additional parameter, say $\theta_0 \in \Theta_0$, which is unknown to any player. That is straightforwardly possible by adding

payoff structures are played with the remaining probability. They regard an equilibrium of the baseline game as robust if, for any associated incomplete-information game, there exists a Bayesian Nash equilibrium that is sufficiently close to the original equilibrium. See also Ui (2001), Morris and Ui (2005, 2020), and Oyama and Tercieux (2010).

[8] See Crawford, Costa-Gomes, and Iriberri (2013) for a comprehensive survey.

a dummy player (say player 0, with a trivial action space A_0) and set $\theta_0 \in \Theta_0$ as his payoff parameter.

Our goal is to provide a "robust prediction", in the sense that it is a reasonable prediction given any (admissible) information structure of the players. For this purpose, we first introduce the notion of a *prediction*.

Definition 1. A *prediction* is a correspondence $\Gamma : \Theta \to 2^A$.

Our prediction Γ is interpreted as predicting that, if the true parameter profile is θ, then the players would play some action profile in $\Gamma(\theta)$ ($\subseteq A$). Even if one prediction is reasonable given some information structure, it may not be reasonable given another information structure, as in the following example.

Example 2. We revisit a first-price auction game mentioned in Example 1. Each bidder is either rational or truthful. While he knows his own behavioral type, he may be uncertain about the opponent's behavioral type. Let $\Theta_i = \{\mathrm{rat}, \mathrm{tru}\}$ be the parameter set and bidding function $a_i : v_i \mapsto a_i(v_i) \in \mathbb{R}_+$ represent an action.

First, suppose that the rational bidder i certainly believes that the opponent is also the rational type. A reasonable prediction given this information structure may be $\Gamma^*(\theta) = \{\gamma_1^*(\theta_1), \gamma_2^*(\theta_2)\}$, where $\gamma_i^*(\mathrm{tru})$ is the identity mapping and $\gamma_i^*(\mathrm{rat}) = a_i^{\mathrm{BNE}}$ specified by (1).

Alternatively, if rational bidder i certainly believes that the opponent is truthful, then a reasonable prediction given this alternative information structure may be $\Gamma^{**}(\theta) = \{\gamma_1^{**}(\theta_1), \gamma_2^{**}(\theta_2)\}$, where $\gamma_i^{**}(\mathrm{tru})$ is the identity mapping and $\gamma_i^{**}(\mathrm{rat}) = a_i^1(\cdot)$ with $a_i^1(v_i) \in \arg\max_b (v_i - b)F(b)$.

Note that prediction Γ^* (resp. Γ^{**}) is reasonable under the first (resp. second) information structure, but it is unreasonable under the other. What about their "union" Γ^{***} (i.e., $\Gamma^{***}(\theta) = \Gamma^*(\theta) \cup \Gamma^{**}(\theta)$ for each θ)? It provides a reasonable prediction if one of the above two information structures is the true one, but not necessarily with other (possibly more complicated) information structures. In situations where we have little information about the players' information structure, it is important to consider *robust* prediction, which provides a reasonable prediction given any possible information structure.

Formally, an information structure is represented by a type space à la Harsanyi (1967-68).

Definition 2. A *(Harsanyi) type space* is $\mathcal{T} = (T_i, \hat{\theta}_i, \hat{\beta}_i)_{i=1}^I$ where

(i) each T_i is the (finite) set of i's types;

(ii) $\hat{\theta}_i(t_i) \in \Theta_i$ denotes t_i's "payoff-parameter";

(iii) $\hat{\beta}_i(t_i) \in \Delta(T_{-i})$ denotes t_i's belief for the others' types.

A type-space representation is quite general in that many popular information structures, e.g., complete information and incomplete information with common/different priors, can be described as special cases. Hereafter, a type space is simply represented by $\mathcal{T} = (T, \hat{\theta}, \hat{\beta})$ unless there is a fear of confusion, where $T = \prod_{i=1}^I T_i$, $\hat{\theta} : T \to \Theta$, and $\hat{\beta} : T \to \prod_{i=1}^I \Delta(T_{-i})$. Once a type space is fixed, then we can define player i's strategy as mapping $\sigma_i : T_i \to \Delta(A_i)$, and a Bayesian Nash equilibrium given that type space as follows.

Definition 3. $\sigma = (\sigma_i)_{i=1}^I$ is a *Bayesian Nash equilibrium* (hereafter, BNE) given $\mathcal{T}$ if for all i, t_i, a_i, and a_i', $\sigma_i(a_i|t_i) > 0$ implies that:

$$\mathbb{E}_{t_{-i}\sim\hat{\beta}_i(t_i)} \left[\sum_{a_{-i}} u_i \left((a_i, a_{-i}), \left(\hat{\theta}_i(t_i), \hat{\theta}_{-i}(t_{-i}) \right) \right) \sigma_{-i} \left(a_{-i} | t_{-i} \right) \right]$$
$$\geq \mathbb{E}_{t_{-i}\sim\hat{\beta}_i(t_i)} \left[\sum_{a_{-i}} u_i \left((a_i', a_{-i}), \left(\hat{\theta}_i(t_i), \hat{\theta}_{-i}(t_{-i}) \right) \right) \sigma_{-i}(a_{-i}|t_{-i}) \right].$$

Now we are ready to formally introduce the concept of *robust prediction*.

Definition 4. Prediction Γ is *robust* if for any finite type space $\mathcal{T} = (T, \hat{\theta}, \hat{\beta})$, there exists a (mixed) BNE σ given $\mathcal{T}$ such that $\operatorname{supp}(\sigma(t)) \subseteq \Gamma\left(\hat{\theta}(t)\right)$ for any $t \in T$.

Put differently, a prediction Γ is *not* robust if there is some type space $\mathcal{T}$ such that given whatever BNE given $\mathcal{T}$, its equilibrium action profile for some θ is outside $\Gamma(\theta)$. In this sense, a non-robust prediction has a problem of "overlooking" possible information structures and the corresponding equilibrium plays. A robust prediction is so constructed that it contains some equilibrium action profile for each possible type space and, in this sense, does not suffer from this overlooking problem.

Remark 1. Notably, our robustness is not necessarily for equilibrium selection because we do not require that robust prediction contains *any* equilibrium outcomes for each possible type space. It contrasts with Bergemann and Morris (2017), which require robustness for both the information structure and equilibrium selection, obtaining *belief-free rationalizability*.[9] An interpretation of our (information-structure-only) robustness is that it provides the *necessary* implications of whatever assumptions we impose on the baseline model.

Our approach would be reasonable if the analyst is not confident about the players' information structure but more confident about certain equilibrium selection. For example, in many applications such as voting and trading, the game often has a trivial weakly dominated equilibrium, which cannot be eliminated by (belief-free) rationalizability. In first-price auctions, many bidding functions are rationalizable (e.g., Battigalli and Siniscalchi, 2003), which means little predictive power. Our notion of robust prediction is "less

[9]See Brandenburger and Dekel (1987) and Battigalli, Di Tillio, Grillo, and Penta (2011).

robust" in equilibrium selection but instead gains more predictive power (see Section 6.3).

Our approach could also be useful in robust mechanism design, where the robustness is sometimes for the information structure while the mechanism designer is allowed to select his most favored equilibrium (e.g., Chung and Ely, 2007; Bögers and Smith, 2014; Chen and Li, 2018; Yamashita and Zhu, 2021).[10]

Remark 2. Formally, $\Gamma(\cdot) = A$ is a trivial robust prediction, but clearly this would be a useless prediction unless it is the *unique* robust prediction. What we are really interested in is the *sharpest* robust prediction in the sense that Γ is robust and any sub-correspondence is not robust. The robust prediction we construct below, Γ^*, satisfies this sharpness requirement as it is originally constructed by considering the *necessary* conditions of a robust prediction.

3 Necessary conditions for robust prediction

Typically, we assume a (simple) type space in applications even if we are not fully confident of its use. Besides, we often make some equilibrium selection. In this sense, a typical prediction in applications is based on a joint assumption on the information structure and equilibrium play. The concern is that the actual information structure of the players may differ from what is assumed. This section aims to obtain implications about the assumption of the information structure under the standard equilibrium selection.

Let $\mathcal{T}^0 = (T^0, \hat{\theta}^0, \hat{\beta}^0)$ represent such a (simple) baseline type space, called a *level*-0 type space.[11] Let σ^0 denote a BNE given $\mathcal{T}^0$, and let Γ^0 denote the

[10] Contrarily, Brooks and Du (2021) consider a pessimistic mechanism designer concerning both information and equilibrium selection. Brooks and Du (2021) show that, in certain auction contexts, the same revenue is guaranteed regardless of his attitude toward equilibrium selection.

[11] As demonstrated below, this name is based on the analogy of our construction to

"baseline prediction" based on σ^0.[12] More specifically, for each θ,

$$\Gamma^0(\theta) = \left\{ a \in A \;\middle|\; \exists t \in T^0 \text{ such that } \hat{\theta}^0(t) = \theta \text{ and } a \in \operatorname{supp}\left(\sigma^0(t)\right) \right\}.$$

The specification of the baseline scenario is highly flexible in the sense that we can arbitrarily choose a type space and an equilibrium. For example, as in Section 6, we may regard the level-0 type space as a complete-information type space such that $\theta = \theta^*$ for some θ^* is common knowledge (i.e., $T^0 = \{t^*\}$ and $\hat{\theta}^0(t^*) = \theta^*$).

We now construct a robust prediction based on this baseline prediction. In general, the baseline prediction we begin with affects the constructed robust prediction. In this sense, our robust prediction should be interpreted as the implication of the assumptions that we (often implicitly) impose on the baseline information structure and equilibrium play.

With the concern of a wrong assumption in level-0 type space, a natural candidate for an alternative type space may be one that "includes" the original type space $\mathcal{T}^0$. We then consider *level*-1 type space $\mathcal{T}^1 = (T^1, \hat{\theta}^1, \hat{\beta}^1)$, which should satisfy the following conditions: for any i,

(i) $T_i^1 \supseteq T_i^0$;

(ii) If $t_i \in T_i^0 (\subseteq T_i^1)$, then $\left(\hat{\theta}_i^1(t_i), \hat{\beta}_i^1(t_i)\right) = \left(\hat{\theta}_i^0(t_i), \hat{\beta}_i^0(t_i)\right)$;

(iii) For all $t_i \in T_i^1$, $\hat{\beta}_i^1(t_i) \in \Delta(T_{-i}^0)$;

(iv) For each $(\theta_i, \beta_i) \in \Theta_i \times \Delta(T_{-i}^0)$, there exists $t_i \in T_i^1$ with $\left(\hat{\theta}_i^1(t_i), \hat{\beta}_i^1(t_i)\right) = (\theta_i, \beta_i)$.

level-k theory (e.g., Stahl and Wilson, 1994).

[12] Although we consider the case where only one equilibrium is selected given $\mathcal{T}^0$, the analysis does not change even if we consider multiple equilibria as the baseline prediction, in which case Γ^0 is defined as the union of all the considered equilibria.

(v) If there exists type $t_i \in T_i^0$ such that $\text{supp}\,(\sigma_i^0(t_i))$ is not a singleton set, then for any $a_i \in \text{supp}\,(\sigma_i^0(t_i))$, there exists type $t_i(a_i) \in T_i^1$ such that $\hat{\theta}_i^1\,(t_i(a_i)) = \hat{\theta}_i^0(t_i)$ and $\hat{\beta}_i^1\,(t_i(a_i)) = \hat{\beta}_i^0(t_i)$.

Condition (i) requires that $\mathcal{T}^0$ is a *belief-closed subspace* of $\mathcal{T}^1$. Intuitively, T_i^1 is constructed by adding "new" types. Condition (ii) imposes consistency in the sense that type $t_i \in T_i^0$ should have the same parameter and belief regardless of the "level". Condition (iii) is analogous to the standard level-k idea, requiring that any level-1 type certainly believes that his opponents are of level-0 types. Condition (iv) requires a sense of richness property, guaranteeing the existence of mapping from any pair of parameters and beliefs to an associated type.[13] Condition (v) requires the existence of "copies" of types who are to play (pure) actions that are supported in σ^0. Such copy types can avoid technical difficulties when the baseline equilibrium is a mixed-strategy equilibrium. Notice that the copies are added as different types from the original.

Given type space $\mathcal{T}^1$, define σ^1 as follows: for any i,

$$\sigma_i^1 = \begin{cases} \sigma_i^0(t_i) & \text{if } t_i \in T_i^0, \\ a_i & \text{if } t_i = t_i(a_i) \\ a_i^1 & \text{otherwise,} \end{cases}$$

where

$$a_i^1 \in \arg\max_{a_i \in A_i} \mathbb{E}_{t_{-i} \sim \hat{\beta}_i^1(t_i)} \left[\sum_{a_{-i}} u_i \left((a_i, a_{-i}), \left(\hat{\theta}_i^1(t_i), \hat{\theta}_{-i}^0(t_{-i}) \right) \right) \sigma_{-i}^0\,(a_{-i} | t_{-i}) \right].$$

[13]Because of this property, formally, the level-1 type space is not a *finite* type space. However, as in Remark 3, we can find a finite type space that yields the same prediction as to the (infinite) level-1 type space. Because of this, we treat the level-1 type space as if it is a finite type space. The same remark applies to any level-k type space constructed subsequently.

Conditions (ii) and (iii) imply that σ^1 is a BNE given $\mathcal{T}^1$. Because the consistency condition requires that any level-0 type $t_i \in T_i^0$ still holds the same "perspective" in the level-1 type space, it seems reasonable to assume that he plays according to σ_i^0 as in the baseline scenario. In contrast, because any type $t_i \in T_i^1 \backslash T_i^0$ believes that the other players are of level-0 types, he simply chooses a best response to the opponents' level-0 strategy σ_{-i}^0, which is essentially determined (up to indifference).[14] It is worth noting that any type in $T_i^1 \backslash T_i^0$ never randomizes under strategy σ_i^1. Let Γ^1 denote the prediction based on the new equilibrium σ^1: for each θ,

$$\begin{aligned}\Gamma^1(\theta) &= \left\{ a \in A \,\middle|\, \exists t \in T^1 \text{ such that } \hat{\theta}^1(t) = \theta \text{ and } a \in \operatorname{supp}\left(\sigma^1(t)\right) \right\} \\ &\supseteq \Gamma^0(\theta).\end{aligned}$$

Now, we have a (weakly) larger prediction, but there remains the concern that the actual type space is different both from $\mathcal{T}^0$ and $\mathcal{T}^1$. As a natural alternative space, we can consider type spaces with higher levels, constructed as an analogy of level-1 type space. Formally, given level-$(k-1)$ type space $\mathcal{T}^{k-1} = (T^{k-1}, \hat{\theta}^{k-1}, \hat{\beta}^{k-1})$ with $k \geq 2$, level-k type space $\mathcal{T}^k = (T^k, \hat{\theta}^k, \hat{\beta}^k)$ is inductively constructed, which satisfies the following properties: for any i,

(i) $T_i^k \supseteq T_i^{k-1}$;

(ii) If $t_i \in T_i^{k-1} (\subseteq T_i^k)$, then $\left(\hat{\theta}_i^k(t_i), \hat{\beta}_i^k(t_i)\right) = \left(\hat{\theta}_i^{k-1}(t_i), \hat{\beta}_i^{k-1}(t_i)\right)$;

(iii) For all $t_i \in T_i^k$, $\hat{\beta}_i^k(t_i) \in \Delta\left(T_{-i}^{k-1}\right)$;

(iv) For each $(\theta_i, \beta_i) \in \Theta_i \times \Delta\left(T_{-i}^{k-1}\right)$, there exists $t_i \in T_i^k$ with $\left(\hat{\theta}_i^k(t_i), \hat{\beta}_i^k(t_i)\right) = (\theta_i, \beta_i)$.

[14] If there exist multiple best responses, we arbitrarily choose one of them. While the selection of a_i^1 could affect the characterization, its qualitative properties do not change.

Likewise, we can inductively construct a BNE σ^k given $\mathcal{T}^k$ and σ^{k-1} as follows: for any i,

$$\sigma_i^k = \begin{cases} \sigma_i^{k-1}(t_i) & \text{if } t_i \in T_i^{k-1}, \\ a_i^k & \text{otherwise,} \end{cases}$$

where

$$a_i^k \in \arg\max_{a_i \in A_i} \mathbb{E}_{t_{-i} \sim \hat{\beta}_i^k(t_i)} \left[\sum_{a_{-i}} u_i \left((a_i, a_{-i}), \left(\hat{\theta}_i^k(t_i), \hat{\theta}_{-i}^{k-1}(t_{-i}) \right) \right) \sigma_{-i}^{k-1} \left(a_{-i} | t_{-i} \right) \right].$$

Finally, prediction Γ^k based on BNE σ^k is defined as follows: for any θ,

$$\begin{aligned} \Gamma^k(\theta) &= \left\{ a \in A \,\middle|\, \exists t \in T^k \text{ such that } \hat{\theta}^k(t) = \theta \text{ and } a \in \operatorname{supp}\left(\sigma^k(t)\right) \right\} \\ &\supseteq \Gamma^{k-1}(\theta). \end{aligned}$$

We continue such an "expanding" procedure until no action profile is added. More specifically, owing to the finiteness of A, there exists some $K \in \mathbb{N}$ such that $\Gamma^{K+1} = \Gamma^K$, where the procedure terminates. Denote $\mathcal{T}^* = \mathcal{T}^{K+1}$, $\sigma^* = \sigma^{K+1}$, and $\Gamma^* = \Gamma^{K+1}$. We refer to $\mathcal{T}^*$ and Γ^* as the *canonical type space* and the *canonical prediction*, respectively.

Remark 3. Formally, the canonical type space $\mathcal{T}^*$ constructed above is not a finite type space. However, we can always construct a finite type space $\hat{\mathcal{T}}^*$ by selecting finitely many types from T^*, which are "sufficient" in the sense that any action profile in the canonical prediction Γ^* is played by some type profile in this set of finitely many types.[15]

[15] More specifically, we start from a finite level-0 type space, and construct type space $\mathcal{T}^1$ and the associated prediction Γ^1, as mentioned above. Now, for any parameter profile θ and action profile $a \in \Gamma^1(\theta)$, we appropriately select a type profile t such that $\hat{\theta}^1(t) = \theta$ and $a \in \operatorname{supp}\left(\sigma^1(t)\right)$. Let $\hat{T}^1$ be the set of such type profiles, and define type space $\hat{\mathcal{T}}^1 = (\hat{T}^1, \hat{\theta}^1, \hat{\beta}^1)$ where $\hat{\theta}^1$ and $\hat{\beta}^1$ are restricted on $\hat{T}^1$. We then replace $\mathcal{T}^1$ with $\hat{\mathcal{T}}^1$. By replacing with $\hat{\mathcal{T}}^k$ for any k, we can construct a desired type space.

Clearly, a robust prediction Γ that contains Γ^0 must contain Γ^*, i.e., $\Gamma^*(\theta) \subseteq \Gamma(\theta)$ for any θ.

4 Γ^* as a robust prediction

So far, we only consider a particular class of alternative type spaces. The constructed canonical prediction Γ^* provides a reasonable prediction *if* any of those level-k type spaces is the actual type space. However, the actual type space may differ from all of them, and then the concern is that Γ^* may not be able to provide a reasonable prediction for those other type spaces.

Fortunately, Γ^* is a robust prediction, as we show below.[16] In other words, the canonical type space $\mathcal{T}^*$ is sufficient for identifying a robust prediction even if it looks artificial when it first appears.

Theorem 1. Γ^* is a robust prediction.

To prove Theorem 1, we need the following lemma that characterizes desired properties of canonical prediction Γ^*.

Lemma 1. Γ^* has the following properties.

(i) *Product structure:* For any i and $\theta \in \Theta$, there exists $\Gamma_i^*(\theta_i) \subseteq A_i$ such that $\Gamma^*(\theta) = \prod_{i=1}^{I} \Gamma_i^*(\theta_i)$.

(ii) *Best-reply property:* For any i, θ_i, and $q \in \Delta(\Theta_{-i} \times A_{-i})$ such that $q(\theta_{-i}, a_{-i}) > 0$ implies $a_{-i} \in \Gamma_{-i}^*(\theta_{-i})$, there exists $a_i \in \Gamma_i^*(\theta_i)$ such that for any $a_i' \in A_i$,

$$\mathbb{E}_{(\theta_{-i},a_{-i})\sim q}\left[u_i\left((a_i, a_{-i}), (\theta_i, \theta_{-i})\right)\right] \geq \mathbb{E}_{(\theta_{-i},a_{-i})\sim q}\left[u_i\left((a_i', a_{-i}), (\theta_i, \theta_{-i})\right)\right]$$

[16] All omitted proofs are in the Appendix.

holds, i.e., action a_i is a best response to conjecture q.[17]

With Lemma 1, the proof of Theorem 1 is straightforward. First, fix an arbitrary finite type space $\mathcal{T} = (T, \hat{\theta}, \hat{\beta})$, and consider a restricted game, where player i with type t_i can play only within $\Gamma_i^*\left(\hat{\theta}_i(t_i)\right)$ for any i and $t_i \in T_i$. Because the restricted game is a finite game, the standard argument guarantees the existence of BNE σ given type space $\mathcal{T}$. To complete the proof, it is sufficient to show that BNE σ is still a BNE in the original game. The best-reply property assures that, as long as $\sigma(t) \subseteq \Gamma^*\left(\hat{\theta}(t)\right)$ holds, no one has an incentive to deviate from $\sigma_i(t_i)$ even if the action set expands from $\Gamma_i^*\left(\hat{\theta}_i(t_i)\right)$ to A_i, implying that σ is an equilibrium of the original game.

Remark 4. Because of the best-reply property, the canonical prediction Γ^* is a variant of a *curb set*, proposed by Basu and Weibull (1991) in complete-information games. We say that a set of action profiles $X \subseteq A$ is a curb set if $BR(X) \subseteq X$, where $BR_i(X_{-i})$ is the set of player i's best responses given that the opponents' behaviors are restricted to X_{-i} and $BR(X) = \prod_{i=1}^{I} BR_i(X_{-i})$.[18] Furthermore, we say that curb set X is *tight* if $X = BR(X)$. Intuitively, a curb set is generalization of *rationalizability* (Bernheim, 1984) in the sense that it is a maximal tight curb set.

5 Robust prediction for infinite games

While the characterization of robust prediction so far is for finite games, many economic applications are described as infinite games (e.g., Θ and A

[17] In other words, it means that as long as the opponents' play stays in $\Gamma_{-i}^*(\cdot)$, player i has a best response in $\Gamma_i^*(\cdot)$, which should be distinct from the *best response property* by Pearce (1984).

[18] The original definition by Basu and Weibull (1991) also requires some technical conditions.

are infinite), which prevents the direct application of the results to the specific contexts. As a preliminary of the following section discussing applications, this section extends the results so far to infinite games with additional structures.

An extension of the above argument to infinite games is nontrivial. Specifically, the construction process of the canonical type space $\mathcal{T}^*$ may not stop within finite steps due to the infiniteness of the action space. Of course, as a direct analogy of the finite-environment counterpart, we can construct an *infinite* sequence of level-k type spaces and define canonical prediction Γ^* as the union of possible action profiles given those type spaces. However, as opposed to the finite case, the constructed infinite type space may not admit a *finite* subspace yielding the same Γ^* as a prediction, which does not achieve our purpose.

To overcome the above concern, we justify prediction Γ^* by using a collection of *finite* type spaces instead of the (infinite version of) canonical type space $\mathcal{T}^*$. Specifically, we discretize the sets of parameters and beliefs and construct a level-k type space in this discretized environment by applying the previous argument. Importantly, for each level, the associated type space is finite. We then construct a collection of those finite type spaces for the level and the degree of discretization. Note that a robust prediction should include the prediction associated with each such type space in the collection. Furthermore, we show that the associated prediction becomes larger as the level goes up or the discretization becomes finer, and it converges to the closure of Γ^* in the limit. This eliminates the concern by guaranteeing the necessary property as in the finite environment: a robust prediction must include Γ^*.[19]

[19] Conversely, the sufficiency part continues to be similar, as long as only finite type spaces are considered.

5.1 Preliminaries

Throughout this section, we impose the following two assumptions.

Assumption 1. (i) For each i, Θ_i and A_i are compact metric spaces, endowed with their Borel-σ algebra, and u_i is continuous in this metric.[20]

(ii) For each i, θ_i and $q_i \in \Delta(\Theta_{-i} \times A_{-i})$, his best response, denoted by $a_i^*(\theta_i, q_i)$, is unique; that is, $\arg\max_{a_i \in A_i} \int_{\Theta_{-i} \times A_{-i}} u_i\left((a_i, a_{-i}), (\theta_i, \theta_{-i})\right) dq_i$ is a singleton set.

We also treat $\Delta(\Theta_{-i} \times A_{-i})$ as a metric space by adopting a Prokhorov metric, which makes $\Delta(\Theta_{-i} \times A_{-i})$ a compact metric space.[21] Unless confusing, we always use the same symbol $d(\cdot, \cdot)$ to represent the distance of two elements in the same metric space, and $\mathbb{B}_\varepsilon(\cdot)$ to represent the open ε-ball around an element of any of those metric spaces with the representation that $\mathbb{B}_\varepsilon(X) = \bigcup_{x \in X} \mathbb{B}_\varepsilon(x)$ for set X. A product space is associated with a product metric.

By the maximum theorem, $a_i^*(\theta_i, q_i)$ is continuous both in θ_i and q_i. That is, for each $\varepsilon > 0$ and θ_i, there exists $\delta(\varepsilon, \theta_i) > 0$ such that for any q_i and $q_i' \in \Delta(\Theta_{-i} \times A_{-i})$ with $d(q_i, q_i') < \delta(\varepsilon, \theta_i)$, we have $d\left(a_i^*(\theta_i, q_i), a_i^*(\theta_i, q_i')\right) < \varepsilon$. In what follows, by normalizing the metric for q_i, we set $\delta(\varepsilon, \theta_i) = \varepsilon$ for each θ_i without loss of generality.

Finally, it is worthwhile noting that we adopt the same definition for a robust prediction. That is, we restrict our attention to finite type spaces

[20] Weinstein and Yildiz (2011) assume a *nice structure*, which is stronger than our requirement.

[21] More specifically, the distance between $q_i, q_i' \in \Delta(\Theta_{-i} \times A_{-i})$ is defined as follows:

$$d(q_i, q_i') = \inf \left\{ d > 0 \;\middle|\; \begin{array}{l} \text{for any measurable } X \subseteq \Theta_{-i} \times A_{-i}, \\ \text{(i) } q_i\left(\mathbb{B}_d(X)\right) + d \geq q_i'(X) \text{ and} \\ \text{(ii) } q_i'\left(\mathbb{B}_d(X)\right) + d \geq q_i(X) \end{array} \right\},$$

where $\mathbb{B}_d(X)$ denotes the open d-ball around X. It is worthwhile to remark that this metric implies a weak-* topology. As noted later, we may normalize this metric in an equivalent way to simplify the notation.

even in infinite games.[22]

5.2 Infinite level-k construction

The infinite-game version of the canonical type space and the canonical prediction are defined as follows. We begin with a *finite* T_i^0 for each i. This restriction seems reasonable and is maintained in all the applications in Section 6. Then we inductively construct level-k type space $\mathcal{T}^k$ and the associated prediction Γ^k analogously to those in Section 3 (i.e., those satisfying the Conditions (i)-(iv)). Note that they may be infinite type spaces at this point.

As mentioned above, the construction process may not stop within finite steps. Hence, the canonical type space and the associated prediction are defined as the limits of sequences: $\left\{\mathcal{T}^k\right\}_{k=0}^{\infty}$ and $\left\{\Gamma^k\right\}_{k=0}^{\infty}$, respectively. More precisely, given sequence $\left\{\mathcal{T}^k\right\}_{k=0}^{\infty}$, the *infinite canonical type space* $\mathcal{T}^{**} = (T^{**}, \hat{\theta}^{**}, \hat{\beta}^{**})$ is defined as follows: for each i,

(i) $T_i^{**} = \bigcup_{k=0}^{\infty} T_i^k = \lim_{k\to\infty} T_i^k$;

(ii)-(a) $\hat{\theta}^{**} : T_i^{**} \to \Theta_i$ is a function such that if $t_i \in T_i^k$, then $\hat{\theta}_i^{**}(t_i) = \hat{\theta}_i^k(t_i)$;

(ii)-(b) $\hat{\beta}_i^{**} : T_i^{**} \to \Delta\left(T_{-i}^{**}\right)$ is a function such that if $t_i \in T_i^k$, then $\hat{\beta}_i^{**}(t_i) = \hat{\beta}_i^k(t_i)$.

Note that, by the properties imposed during the construction process of level-k type space $\mathcal{T}^k$, the above is well defined. Furthermore, as an implication of the modified definition, type space $\mathcal{T}^{**}$ also satisfies the following properties: for each i,

(iii) For all $t_i \in T_i^{**}$, $\hat{\beta}_i^{**}(t_i) \in \Delta\left(T_{-i}^{**}\right)$;

[22] Weinstein and Yildiz (2011) also adopts the similar restriction.

(iv) For each $(\theta_i, \beta_i) \in \Theta_i \times \Delta\left(T_{-i}^{**}\right)$, there exists $t_i \in T_i^{**}$ with $\left(\hat{\theta}_i^{**}(t_i), \hat{\beta}_i^{**}(t_i)\right) = (\theta_i, \beta_i)$.

Hence, infinite canonical type space $\mathcal{T}^{**}$ has the same properties as its finite counterpart.

Likewise, given sequence $\left\{\Gamma^k\right\}_{k=0}^{\infty}$, the *infinite canonical prediction* Γ^{**} is defined as follows: for each θ,

$$\Gamma^{**}(\theta) = \mathrm{cl}\left(\bigcup_{k=0}^{\infty} \Gamma^k(\theta)\right) = \mathrm{cl}\left(\lim_{k\to\infty} \Gamma^k(\theta)\right),$$

where $\mathrm{cl}(\cdot)$ denotes the closure operator. Taking the closure makes $\Gamma^{**}(\theta)$ compact, and it is another modification from the finite counterpart. This modification is essential for the following theorem, which is a counterpart of Theorem 1.

Theorem 2. Γ^{**} is a robust prediction.

The proof is analogous to the finite case: we first guarantee the existence of BNE σ^* in the restricted game, where the available action set is restricted to $\Gamma^{**}\left(\hat{\theta}(t_i)\right)$ for each i and t_i, and then show that σ^* continues to be a BNE in the unrestricted game. Owing to the compactness of Γ^{**}, the existence of a BNE is assured by the Kakutani-Glicksberg-Fan fixed-point theorem. Furthermore, Γ^{**} still has the product structure and satisfies the best-reply property. Note that taking the closure does not affect these nice properties of Γ^{**} thanks to the continuity. As a result, the same argument used above is still valid here.

Contrary to the previous section, showing that $\mathcal{T}^{**}$ yields Γ^{**} as a prediction is not satisfactory because the corresponding type space $\mathcal{T}^{**}$ may not admit an equivalent finite type space that yields Γ^{**} as a prediction. Potentially, Γ^{**} might admit a proper sub-correspondence that itself is a robust prediction. In order to show that such a concern is not real, we identify a collection of finite type spaces such that (the closure of) the union of the predictions given those finite type spaces yields Γ^{**}. A detailed discussion is found in Appendix B.

6 Applications

This section derives a robust prediction in the following well-known environments and discusses its implications. Cournot competition, a Diamond search game, and a first-price auction are studied in Sections 6.1, 6.2, and 6.3, respectively. More specifically, we characterize infinite canonical prediction Γ^{**} in each scenario based on the discussion in Section 5.

6.1 Cournot competition

We consider the Cournot competition with I firms. Each firm simultaneously chooses nonnegative quantity $a_i \in A_i = \mathbb{R}_+$ with zero marginal costs of production. The inverse demand function of the good is given by $P(a, \theta) = \max\left\{\theta - \sum_{i=1}^{I} a_i, 0\right\}$. It is common knowledge that θ is included in interval $[\underline{\theta}, \bar{\theta}]$, but there is no common prior governing the distribution. That is, $\theta = \theta_0 \in \Theta_0 = [\underline{\theta}, \bar{\theta}]$ and $\theta_i \in \Theta_i = \emptyset$ for any $i \in \{1, 2, \ldots, I\}$.[23] Hence, with abuse of some notation, the set of parameter profiles is simply represented by $\Theta = [\underline{\theta}, \bar{\theta}]$. The firm i's payoff function is given by $u_i((a_i, a_{-i}), \theta) = \max\{(\theta - a_i - a_{-i})a_i, 0\}$, where, with some abuse of notation, $a_{-i} = \sum_{j \neq i} a_j$ throughout this subsection.

To guarantee interior solutions, we impose the following. Intuitively, it requires that the uncertainty over the parameter is sufficiently small.

[23] Recall that θ_0 refers to the parameter that no player knows. Formally, imagine a "dummy player" 0 whose action space is trivial.

Assumption 2. $2\underline{\theta} - \bar{\theta} \geq 0$.

Let us take a complete-information environment of some $\tilde{\theta}$ and its Nash equilibrium as the baseline scenario. Specifically, the type set and the belief function are defined as follows: for each i,

- $T_i^0 = \{t^0\}$;
- $\hat{\beta}_i^0 \left(\{\tilde{\theta}\} \times T_{-i}^0 \;\middle|\; t^0 \right) = 1$, where $\tilde{\theta} \in (\underline{\theta}, \bar{\theta})$.

That is, under type space $\mathcal{T}^0 = \left(T^0, \hat{\beta}^0 \right)$, each firm certainly believes that (i) the demand parameter is $\tilde{\theta}$, and (ii) the opponents also have the same beliefs.[24] Note that the specification of parameter $\tilde{\theta}$ is highly flexible.

Now, we construct the robust prediction Γ^{**} based on infinite canonical type space $\mathcal{T}^{**}$, exactly as in Section 5 (and hence it is omitted).[25]

First, we consider a duopoly model (i.e., $I = 2$).

Proposition 1. Consider the Cournot duopoly with Assumption 2. Then, $\Gamma^{**}(\theta) = \left[(2\underline{\theta} - \bar{\theta})/3, (2\bar{\theta} - \underline{\theta})/3 \right]^2$ holds for each $\theta \in [\underline{\theta}, \bar{\theta}]$.

We have the following remarks. First, the infinite canonical prediction does not depend on $\tilde{\theta}$; that is, it is irrelevant to the specification of the level-0 scenario. Intuitively, the irrelevance comes from the fact that the specification of the level-0 scenario only influences the firm's best response through the conjecture of the opponents' behaviors, which becomes less influential as the level goes up and vanishes in the limit. Suppose, for example, that $t_i \in T_i^1$, and consider his best response. Type t_i's best response is given by $a_i(t_i) = \mathbb{E}_i[\theta]/2 - \mathbb{E}_i[a_j]/2$, where $\mathbb{E}_i[\cdot] = \mathbb{E}_{(\theta, t_j) \sim \hat{\beta}_i(t_i)}[\cdot]$. As type t_i certainly believes that the opponent's type is $t_j = t^0$, his conjecture about a_j is $\mathbb{E}_i[a_j] = \sigma^0(t^0) = \tilde{\theta}/3$, which is highly sensitive to the specification of $\tilde{\theta}$. The

[24] For simplification, parameter function $\hat{\theta}(\cdot)$ is omitted from the definition of the type space.

[25] See Appendix A.4.1 for the detail.

impact of the specification of the level-0 scenario only influences the level-1 behavior through this channel. Furthermore, as the absolute value of the slope of the best-response function with respect to the opponent's behavior is $1/2$, the marginal effect of the opponent's behavior to t_i's best response is less than 1.[26] The same arguments hold for any subsequent levels. More specifically, k times iteration implies that the best response of type $t_i \in T_i^k$ is as follows:

$$a_i(t_i) = \frac{1}{2}\mathbb{E}_i[\theta] - \frac{1}{4}\mathbb{E}_i\mathbb{E}_j[\theta] + \cdots + \frac{1}{2^k}\mathbb{E}_i\mathbb{E}_j\mathbb{E}_i \cdots \mathbb{E}_i[\theta] - \frac{1}{2^k}\mathbb{E}_i\mathbb{E}_j\mathbb{E}_i \cdots \mathbb{E}_i[a_j]. \tag{2}$$

While the impact of the specification of $\tilde{\theta}$ remains in the last term of (2), it is less influential than the previous levels and vanishes in the limit. As a result, the canonical prediction becomes independent of the specification of the level-0 behavior.

Second, as $\bar{\theta} - \underline{\theta} \to 0$, the canonical prediction converges to a singleton set consisting of the associated Nash equilibrium (hereafter, NE). The difference $\bar{\theta} - \underline{\theta}$ measures the degree of misspecification of the true demand parameter. Hence, our characterization suggests that the prediction also becomes more precise as the precision of the model specification improves, which would be an attractive property of predictions.

Next, we consider an oligopoly model with $I = 3$ and derive infinite canonical prediction Γ^{**}, likewise. Although we adopt the same type space constructed above, the characterization is quite different from that of the duopoly case, as demonstrated in the following.

Proposition 2. Consider the Cournot oligopoly with $I = 3$. Then, $\Gamma^{**}(\theta) = [0, \bar{\theta}/2]^3$ holds for each $\theta \in [\underline{\theta}, \bar{\theta}]$.

[26] This property of the best-response function implies the *global stability* by Weinstein and Yildiz (2007a).

While the canonical prediction is irrelevant to the specification of the level-0 scenario as in the duopoly, the underlying mechanism is different. As in the duopoly model, the specification of $\tilde{\theta}$ only affects firm i's best response through the conjecture of the opponents' behaviors with the coefficient being less than 1. However, the impact never disappears even at higher levels because there is more than one opponent. Specifically, type t_i's best response is given by $a_i(t_i) = \mathbb{E}_i[\theta]/2 - \mathbb{E}_i[a_j]/2 - \mathbb{E}_i[a_h]/2$. Because the specification of $\tilde{\theta}$ affects both actions a_j and a_h, its marginal effect to type t_i's best response is not less than 1 even though the marginal effect from each opponent is $1/2$. As a result, the "expansion rate" of level-k prediction Γ^k is not discounted at higher levels, and then its lower bound reaches 0 (i.e., the lower bound of A_i) at some level. Because the lower bound of Γ^k is given by 0 independent of the specification of $\tilde{\theta}$ after it reaches 0, the upper bound of Γ^k also becomes independent from the specification of $\tilde{\theta}$. Furthermore, prediction Γ^{**} no longer converges to a singleton set as $\bar{\theta} - \underline{\theta} \to 0$, which is another contrast to the duopoly.

Remark 5. Propositions 1 and 2 clarify the relationship with the existing works. First, Weinstein and Yildiz (2007a) show that the Cournot duopoly is dominance solvable as in the complete-information games (Bernheim, 1984) if the possible belief hierarchies are restricted to those in the class of the "finite-order perturbation" à la Weinstein and Yildiz (2007b). That is, the robust prediction (in the sense of Weinstein and Yildiz (2007b)) of the Cournot duopoly specifies a unique outcome, which is a sharp contrast to our result. The difference lies in the definition of the robust prediction: we do not restrict our attention to the nearby types in the sense of the finite-order perturbation. As a result, there remain disagreements over the first-order and second-order beliefs (and higher-order beliefs), which makes prediction Γ^{**} non-singleton.

Second, contrary to the duopoly, the difference in the definition of robust prediction does not appear in the oligopoly model. Proposition 2 is reminiscent of Basu (1992), characterizing rationalizability in complete-information Cournot oligopoly. Weinstein and Yildiz (2011) and Chen, Takahashi, and Xiong (2021) obtain analogous characterizations of incomplete-information Cournot oligopoly based on their versions of robust predictions.

6.2 Diamond search

Next, we consider a Diamond search model with two players. Each player simultaneously chooses his search intensity $a_i \in A_i = [0, 2]$ with search cost $c_i(a_i) = a_i^3/3$. The gain from finding a trading partner depends on the players' search intensities and parameter θ characterizing the search environment. Specifically, player i's payoff function is given by $u_i((a_i, a_{-i}), \theta) = \theta a_i a_{-i} - a_i^3/3$, where $-i$ represents player i's opponent. We assume that nobody knows the true parameter θ, but it is common knowledge that the true parameter is included in interval $[1, 1 + \varepsilon]$ with $\varepsilon > 0$. Hence, as in the Cournot competition, $\theta = \theta_0 \in \Theta_0 = [1, 1 + \varepsilon]$ and $\theta_i \in \Theta_i = \emptyset$ for each i. For the simplification, the set of parameter profiles is represented by $\Theta = [1, 1 + \varepsilon]$.[27]

Let us take the complete-information situation with $\theta = 1$ as the level-0 type space. Contrary to the previous application, there exist multiple NEs even though there is no uncertainty over θ: both action profiles $(0, 0)$ and $(1, 1)$ are NEs. In what follows, we assume that action profile $(1, 1)$ is played under the level-0 type space and construct the robust prediction given that level-0 equilibrium selection.[28] Applying the level-k construction in Section 5, we obtain the robust prediction Γ^{**} as follows.

Proposition 3. $\Gamma^{**}(\theta) = [1, 1 + \varepsilon]^2$ holds for each $\theta \in [1, 1 + \varepsilon]$.

[27] This setup is associated with Example 1 of Weinstein and Yildiz (2011).

[28] Selecting payoff-dominating equilibrium (if it exists) is common in the applied literature.

We have the following two implications. First, contrary to the Cournot competition, the infinite canonical prediction of the Diamond search depends on the baseline scenario. For example, if we select the other equilibrium $(0, 0)$ as the level-0 equilibrium, then the corresponding robust prediction satisfies $\Gamma^{**}(\theta) = \{ (0, 0) \}$ for each θ.

Second, our result may be interpreted as a version of monotone comparative statics with respect to the change in the payoff parameter. In the baseline, level-0 information structure, it is common knowledge that $\theta = 1$, and we assume that the players play $(1, 1)$. In case $\Theta = [1, 1 + \varepsilon]$, which is "higher" than $\Theta = \{1\}$ in the strong set-order sense, we predict that the players must play an action profile in $[1, 1+\varepsilon]^2$, which is "higher" than $(1, 1)$. Notably, this comparative statics does not depend on the players' information structure, and in this sense, it is a "robust" monotone comparative statics with respect to the information structure.

This property is based on each level-k type's comparative statics in his best response. Note that the equilibrium strategy of type $t_i \in T_i^k$ is given by

$$\sigma_i^*(t_i) = \sigma_i^k(t_i) = \sqrt{\mathbb{E}_i \left[\theta \sigma_{-i}^{k-1}(t_{-i}) \right]}.$$

Hence, monotone comparative statics holds in the sense that if the expectation of the parameter θ monotonically changes, then the equilibrium behavior also changes monotonically.[29]

Interestingly, this "robust" monotone comparative statics holds not only for stable equilibria but also for unstable equilibria, in contrast to the standard "complete-information" monotone comparative statics. For instance, consider a slightly different payoff function given by $u_i((a_i, a_{-i}), \theta) = \theta a_i a_{-i} - (a_i/2 + a_i^3/3)$, and let $\theta = 3/2$ be common knowledge in the level-0 type space. There are three NEs: stable equilibria $(0, 0)$ and $(1, 1)$ and an unstable equi-

[29] This observation is related to the *Van Zandt-Vives order* on type spaces. See Van Zandt and Vives (2007) and Kunimoto and Yamashita (2020) for more details.

librium $(1/2, 1/2)$. Note that, independent of the specification of σ^0, the equilibrium strategy of type $t_i \in T_i^k$ in the modified setup is given by

$$\sigma_i^k(t_i) = \begin{cases} \sqrt{\mathbb{E}_i[\theta\sigma_{-i}^{k-1}(t_{-i})] - \frac{1}{2}} & \text{if } \mathbb{E}_i[\theta\sigma_{-i}^{k-1}(t_{-i})] \geq \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, even if we construct the level-k type space in which the unstable equilibrium is played in the baseline scenario, our "robust" monotone comparative statics still holds, which is a contrast to the standard monotone comparative statics that may not be valid for unstable equilibria.[30]

Remark 6. Weinstein and Yildiz (2011) study their version of robust prediction in Diamond search games. Their robust prediction concerns a selection of equilibrium outcomes robust to small (finite-order) perturbations of belief hierarchies. In this sense, their notion is quite different from our robust prediction concept, which concerns a collection of possible outcomes across all possible information structures. More specifically, they fix a complete-information game as a baseline scenario, a sufficiently large type space $\mathcal{T}^u$ containing the baseline scenario, and BNE σ^{WY} given $\mathcal{T}^u$. They focus on subsets of action profiles under σ^{WY} induced by types whose belief hierarchies are slightly different from those of the baseline scenario, which is referred to as *minimally-robust prediction* (hereafter, MRP) of σ^{WY}. Weinstein and Yildiz (2011) show that a *locally rationalizable set* (hereafter, LRS), a local version of ICR, is a lower bound of the MRP. In the above Diamond search context, LRS and MRP of BNE σ^{WY} such that $(1, 1)$ is played in the baseline scenario and $(0, 0)$ is played in all nearby types are given by $\{\ (1, 1)\ \}$ and $\{\ (0, 0), (1, 1)\ \}$, respectively, which is a contrast to Proposition 3.

[30] The characterization of this alternative setup is available from the authors upon request.

6.3 First-price auctions

As the third application, we consider a first-price auction with two bidders who might be irrational in the sense of Examples 1 and 2. Each bidder i has private value $v_i \in [0,1]$ for the good. As in the standard formulation, we assume that value v_i is bidder i's private information, and it is common knowledge that each v_i is independently and identically distributed according to a uniform distribution $F(v) = v$. To keep consistency in terminology, bidding function $a_i : v_i \mapsto a_i(v_i) \in \mathbb{R}_+$ is regarded as "action a_i". There are two types of bidders: *truthful* and *rational bidders*. Let $\theta_i \in \Theta_i = \{0,1\}$ represent bidder i's behavioral type: $\theta_i = 0$ (resp. $\theta_i = 1$) means that bidder i is a truthful bidder (resp. rational bidder). We assume that there is no common prior over parameter set $\Theta = \prod_{i=1}^2 \Theta_i$ although parameter θ_i is also bidder i's private information.[31] If bidder i is truthful, then he always bids his value v_i; otherwise, he behaves as in the standard manner. Specifically, the payoff function is given by the following:

$$u_i((a_i, a_{-i}), \theta_i) = \begin{cases} \mathbb{E}_{v \sim F^2}\left[(v_i - a_i(v_i))\,\mathrm{Prob}\,(a_{-i}(v_{-i}) \leq a_i(v_i))\right] & \text{if } \theta_i = 1, \\ 0 & \text{if } \theta_i = 0 \text{ and } a_i(v_i) = v_i, \\ -\infty & \text{otherwise.} \end{cases}$$

Let us consider the complete-information situation of $\theta = (0,0)$ and the truthful bidding equilibrium as the baseline, level-0 prediction. The result does not change at all even if we begin with $\theta = (1,1)$ and the associated symmetric, monotone equilibrium as the baseline prediction.

It is worthwhile to remark on the following. First, at any level, truthful

[31] Another possible formulation is regarding private value v_i as a part of parameter θ_i in our original framework. Although such an alternative can be considered, it appears too different from the standard formulation of first-price auctions, where the value distribution is common knowledge. To clarify the impact of the minimal departure from the standard environment, we adopt the current formulation.

bidding is always the best action for the truthful bidder, and hence, we can essentially focus on the rational bidder's behavior. Second, in terms of notation, let $\Gamma^{**}(\theta, v) = \prod_i \Gamma_i^{**}(\theta_i, v_i)$ denote the predicted set of bid profiles in case θ is the profile of the bidders' behavioral types and v is the profile of their private values.

Proposition 4. Consider the first-price auction with uniform distribution F. Then, for each i, $\Gamma_i^{**}(\theta_i, v_i) = \{v_i\}$ if $\theta_i = 0$ and $\{v_i/2\}$ otherwise.

It is worthwhile to remark that the rational bidder's equilibrium strategy is also uniquely determined irrelevant to his belief about the opponent; that is, any type of the rational bidder adopts bidding strategy $a_i(v_i) = v_i/2$. It comes from the uniform distribution of private values. Specifically, for any type of the rational bidder in the level-k construction, his expected payoff by bidding b is given by

$$(1 + \mu)\left((v_i - b)b\right),$$

where μ is the probability that the opponent is also the rational bidder derived from his belief. Therefore, its maximizer is identical whatever the level is. As long as the private value is governed by a power distribution $F(x) = x^c$ with $c > 0$ (the uniform distribution is a special case with $c = 1$), we have qualitatively the same conclusion.[32]

This observation provides an implication for revenue comparisons. Consider the second-price auction with the same environment (SPA, hereafter) and focus on a BNE where the rational bidder also bids truthfully. The revenue of SPA is equivalent to that under the second-price auction without the truthful bidder (SPA*, hereafter). By the revenue equivalence theorem, the revenue of the SPA* is equivalent to that under the first-price auction without the truthful bidder (FPA*, hereafter). Furthermore, the standard

[32]The detail is available from the authors upon request.

argument implies that the symmetric and monotone BNE of FPA* is given by $a_i^{BNE}(v_i) = v_i/2$, which is identical to the equilibrium strategy of the rational bidder in the original first-price auction (FPA, hereafter) as specified in Proposition 4. Because the truthful bidder in FPA bids more aggressively than the rational bidder, the revenue of FPA is weakly higher than that of SPA. Therefore, we conclude that, if some bidders might be truthful, even if there is no consensus as to how likely they are truthful types, the first-price auction is weakly better for the seller than the second-price auction.

Remark 7. Proposition 4 is reminiscent of Crawford and Iriberri (2007). They show that, as long as bidders' valuations are drawn from a uniform distribution, the level-k and the equilibrium predictions coincide in independent private-value first-price auctions. While they consider the bidders' level-k reasoning as a potentially useful way to understand their non-equilibrium behaviors, we show that this level-k idea is relevant not only when the bidders are truly level-k minded but also as a canonical way to obtain a robust prediction.[33]

7 Conclusion

This paper introduces a concept of robust prediction when the analyst does not know the true information structure about the uncertain payoff-relevant parameters. In particular, we focus on a set of action profiles that contains *some* Bayesian Nash equilibrium for any information structures. We show that the canonical type space, constructed as an analogy of level-k theory, is sufficient for finding the robust prediction in the sense that, as long as we focus on the prediction under the canonical type space, there should be no concern about the ignorance of possible outcomes caused by the misspec-

[33]Because of this motivation, they focus on investigating the finite-level types instead of the limit/infinite level as in this paper based on the equilibrium reasoning.

ification of the true information structure. Our approach is applicable to both finite and infinite games. We also characterize the robust predictions in several economic applications, clarifying the relationship with the related notions. Our robust prediction is often easy to find thanks to its iterative definition but provides a sharp prediction compared with the existing notions related to rationalizability. Therefore, our approach offers new insights in several contexts (e.g., robust mechanism design, strategic communication, etc.), which we leave for future research.

A Appendix: proofs

A.1 Proof of Lemma 1

Let Γ^* be the canonical prediction based on BNE σ^* given canonical type space $\mathcal{T}^* = (T^*, \hat{\theta}^*, \hat{\beta}^*)$.

(i) Fix i and $\theta_i \in \Theta_i$, arbitrarily, and let $\Gamma_i^*(\theta_i)$ be the set of all actions that player i with θ_i may play in σ^*, defined by

$$\Gamma_i^*(\theta_i) = \left\{ a_i \in A_i \;\middle|\; \exists t_i \in T_i^* \text{ such that } \hat{\theta}_i^*(t_i) = \theta_i \text{ and } a_i \in \operatorname{supp}(\sigma_i^*) \right\}.$$

By construction of Γ^*, for each $\theta \in \Theta$, $a \in \Gamma^*(\theta)$ is equivalent to that there exists type profile $t \in T^*$ such that $\hat{\theta}^*(t) = \theta$ and $a \in \operatorname{supp}(\sigma^*(t))$. Furthermore, by definition of $\Gamma_i^*(\theta_i)$, it is obviously equivalent to that $a_i \in \Gamma_i^*(\theta_i)$ for any i, which implies the first statement.

(ii) Fix i, $\theta_i \in \Theta_i$, and $q_i \in \Delta(\Theta_{-i} \times A_{-i})$ such that $q_i(\theta_{-i}, a_{-i}) > 0$ implies that $a_{-i} \in \Gamma_{-i}^*(\theta_{-i})$, arbitrarily. It is worthwhile to notice that in the level-k construction, we add actions at each level, and the process terminates at level $K \geq 1$. Thus, for each j, $\theta_j \in \Theta_j$, and $a_j \in \Gamma_j^*(\theta_j)$, there exists type $t_j \in T_j^K$ such that $\hat{\theta}^K(t_j) = \hat{\theta}^*(t_j) = \theta_j$

and $\sigma_j^K(t_j) = \sigma_j^*(t_j) = a_j$.[34] Denote that type by $t_j^{(\theta_j, a_j)} \in T_j^K$. Now, we observe that there exists type $t_i \in T_i^{K+1}$ who "essentially has belief q" and $\hat{\theta}_i^{K+1}(t_i) = \theta_i$. More specifically, consider belief $\beta_i \in \Delta\left(T_{-i}^K\right)$ defined by:

$$\beta_i(t_{-i}) = \begin{cases} q(\theta_{-i}, a_{-i}) & \text{if } t_{-i} = t_{-i}^{(\theta_{-i}, a_{-i})} \text{ for any } \theta_{-i} \in \Theta_{-i} \text{ and } a_{-i} \in \Gamma_{-i}^*(\theta_{-i}), \\ 0 & \text{otherwise.} \end{cases}$$

By the richness property that level-k type spaces should satisfy, there exists type $t_i \in T_i^{K+1}$ such that $\hat{\theta}_i^{K+1}(t_i) = \hat{\theta}_i^*(t_i) = \theta_i$ and $\hat{\beta}_i^{K+1}(t_i) = \hat{\beta}_i^*(t_i) = \beta_i$. Note that action $\sigma_i^{K+1}(t_i)$ is a best response to belief β_i. Because $\sigma_i^{k+1}(t_i) = \sigma_i^*(t_i) \in \Gamma_i^*(\theta_i)$, we conclude that Γ^* has a best-reply property. $\square$

A.2 Proof of Theorem 1

Fix a finite type space $\mathcal{T} = (T, \hat{\theta}.\hat{\beta})$, arbitrarily. First, we consider a restricted game, where for any i and $t_i \in T_i$, available actions to player i with type t_i is restricted to $\Gamma_i^*\left(\hat{\theta}_i(t_i)\right) \subseteq A_i$. Because the restricted game is a finite game, there exists a (mixed) BNE σ given $\mathcal{T}$. By construction, it is obvious that $\text{supp}(\sigma(t)) \subseteq \Gamma^*\left(\hat{\theta}(t)\right)$ holds for any $t \in T$. We then consider the original game, where the players' action sets are no longer restricted. By Lemma 1-(ii), canonical prediction Γ^* has a best-reply property, implying that BNE σ constructed above is still a BNE given $\mathcal{T}$ in the original game. Therefore, because type space $\mathcal{T}$ is arbitrarily chosen, we conclude that canonical prediction Γ^* is robust. $\square$

[34] If $a_j \in \Gamma_j^*(\theta_j) \backslash \Gamma_j^0(\theta_j)$, then the existence of type t_j is obvious from the definition of strategy σ_j^k. Otherwise, the existence of such a type is also guaranteed with Condition (v) of level-1 type space $\mathcal{T}^1$should satisfy.

A.3 Proof of Theorem 2

First, it is worthwhile to note that the infinite-counterpart of Lemma 1 holds. That is, infinite canonical prediction Γ^{**} has the product structure and satisfies the best-reply property. Now, arbitrarily fix a *finite* type space $\mathcal{T} = (T, \hat{\theta}, \hat{\beta})$, and consider a restricted game, where the available action is restricted to $\Gamma_i^{**}\left(\hat{\theta}_i(t_i)\right)$ for each i and $t_i \in T_i$. By the Kakutani-Glicksberg-Fan fixed-point theorem, there exists a BNE σ^* in the restricted game.[35] Owing to the best-reply property, we say that σ^* is still a BNE in the original game. Therefore, we conclude that Γ^{**} is a robust prediction. $\square$

A.4 Proof of Proposition 1

A.4.1 Construction of the infinite canonical type space

Now, we construct infinite canonical type space $\mathcal{T}^{**} = \left(T^{**}, \hat{\beta}^{**}\right)$ as follows.[36] As mentioned in the body of the paper, the level-0 type space $\mathcal{T}^0 = \left(T^0, \hat{\beta}^0\right)$ is defined as follows: for each i,

- $T_i^0 = \{t^0\}$;
- $\hat{\beta}_i^0\left(\{\tilde{\theta}\} \times T_{-i}^0 \,\middle|\, t^0\right) = 1$, where $\tilde{\theta} \in (\underline{\theta}, \bar{\theta})$.

Given level-0 type space $\mathcal{T}^0$, type space $\mathcal{T}^1 = \left(T^1, \hat{\beta}^1\right)$ is defined as follows: for each i,

- $T_i^1 = \{\, t_i^1(\theta') \mid \theta' \in \Theta \,\}$;
- $\hat{\beta}_i^1\left(\{\theta'\} \times T_{-i}^0 \,\middle|\, t_i^1(\theta')\right) = 1$ for any $\theta' \in \Theta$.

[35] See Fudenberg and Tirole (1991) for the detail.

[36] For the sake of exposition, the domain of belief function $\hat{\beta}_i$ is represented by $\Theta \times T_{-i}$ for each i given type space $\mathcal{T} = (T, \hat{\beta})$, where $T_{-i} = T \backslash T_i$.

That is, type $t_i^1(\theta')$ certainly believes that (i) the demand parameter is $\theta = \theta'$, and (ii) opponent j's type is $t_j = t^0$ for any $j \neq i$. Note that type space $\mathcal{T}^1$ satisfies Conditions (i) to (v) in Section 3.

For any $k \geq 2$, given type space $\mathcal{T}^{k-1}$, type space $\mathcal{T}^k = \left(T^k, \hat{\beta}^k\right)$ is recursively defined as follows: for any i,

- $T_i^k = \left\{ t_i^k(\theta', \beta_i) \;\middle|\; \theta' \in \Theta \text{ and } \beta_i \in \Delta\left(T_{-i}^{k-1}\right) \right\}$;
- For any $\theta' \in \Theta$ and $\beta_i \in \Delta\left(T_{-i}^{k-1}\right)$, $\text{marg}_\Theta \hat{\beta}_i^k \left(\{\theta'\} \;\middle|\; t_i^k(\theta', \beta_i)\right) = 1$ and $\text{marg}_{T_{-i}^{k-1}} \hat{\beta}_i^k \left(t_i^k(\theta', \beta_i)\right) = \beta_i$.

As in the previous levels, type $t_i(\theta', \beta_i)$ (i) certainly believes that the demand parameter is $\theta = \theta'$, and (ii) his belief about the opponents is β_i. Note that type space $\mathcal{T}^k$ also satisfies Condition (i) to (iv) in Section 3. Given sequence $\left\{\mathcal{T}^k\right\}_{k=0}^{\infty}$, type space $\mathcal{T}^{**}$ is defined as the limit of the sequence.

A.4.2 Proof of Proposition 1

Once type space $\mathcal{T}$ is fixed, each firm's best response is given by the standard arguments. As the expected payoff of type $t_i \in T_i$ is

$$\mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}\left[u_i(a_i, a_{-i}, \theta)\right] = \begin{cases} \left(\mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[\theta - a_{-i}] - a_i\right) a_i & \text{if } 0 \leq a_i \leq \mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[\theta - a_{-i}], \\ 0 & \text{otherwise,} \end{cases}$$

the first-order condition implies that the best response of type t_i is

$$a_i(t_i) = \begin{cases} \mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[\theta - a_{-i}]/2 & \text{if } \mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[\theta - a_{-i}] \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

As type space $\mathcal{T}^0$ represents a complete-information game with demand parameter $\theta = \tilde{\theta}$, (3) implies that BNE σ^0 given $\mathcal{T}^0$ is characterized by

$\sigma_i^0(t^0) = \tilde{\theta}/3$ for each i. Hence, level-0 prediction Γ^0 is given by

$$\Gamma^0(\theta) = \begin{cases} \left\{\left(\frac{1}{3}\tilde{\theta}, \frac{1}{3}\tilde{\theta}\right)\right\} & \text{if } \theta = \tilde{\theta}, \\ \emptyset & \text{otherwise.} \end{cases}$$

For type space $\mathcal{T}^1$, BNE σ^1 given $\mathcal{T}^1$ is defined as in Section 3. By (3) and the construction of $\mathcal{T}^1$, the type $t_i^1(\theta')$'s behavior is

$$\sigma_i^1\left(t_i^1(\theta')\right) = \frac{1}{2}\left(\mathbb{E}_{\theta \sim \mathrm{marg}_\Theta \hat{\beta}_i^1\left(t_i^1(\theta')\right)}[\theta] - \sigma_j^0\left(t^0\right)\right) = \frac{1}{6}\left(3\theta' - \tilde{\theta}\right).$$

Hence, level-1 prediction Γ^1 is characterized as follows: for each θ,

$$\Gamma^1(\theta) = \left[\frac{1}{6}\left(3\underline{\theta} - \tilde{\theta}\right), \frac{1}{6}\left(3\bar{\theta} - \tilde{\theta}\right)\right]^2,$$

which is well defined under Assumption 2. Note that Γ^1 is constant in θ.

Consider level-k prediction Γ^k for $k \geq 2$. As the induction hypothesis, we assume that level-$(k-1)$ prediction $\Gamma^{k-1}(\theta) = \left[\underline{a}^{k-1}, \bar{a}^{k-1}\right]^2$ holds for any θ. Fix $\theta \in \Theta$, arbitrarily. By (3) and the construction of type space $\mathcal{T}^k$, type $t_i(\theta', \beta_i)$ $(\in T_i^k)$'s behavior is

$$\begin{aligned} \sigma_i^k\left(t_i^k(\theta', \beta_i)\right) &= \frac{1}{2}\left(\mathbb{E}_{\theta \sim \mathrm{marg}_\Theta \hat{\beta}_i^k\left(t_i^k(\theta', \beta_i)\right)}[\theta] - \mathbb{E}_{t_j \sim \mathrm{marg}_{T_j^{k-1}} \hat{\beta}_i^k\left(t_i^k(\theta', \beta_i)\right)}\left[\sigma_j^{k-1}(t_j)\right]\right) \\ &= \frac{1}{2}\left(\theta' - \mathbb{E}_{t_j \sim \mathrm{marg}_{T_j^{k-1}} \hat{\beta}_i^k\left(t_i^k(\theta', \beta_i)\right)}\left[\sigma_j^{k-1}(t_j)\right]\right). \end{aligned}$$

Let $\underline{a}^k \leq \sigma_i^k\left(t_i^k(\theta', \beta_i)\right) \leq \bar{a}^k$. Specifically,

$$\begin{aligned} \underline{a}^k &= \frac{1}{2}\left(\underline{\theta} - \bar{a}^{k-1}\right), \\ \bar{a}^k &= \frac{1}{2}\left(\bar{\theta} - \underline{a}^{k-1}\right). \end{aligned}$$

Letting $\alpha^k = \underline{a}^k + \bar{a}^k$ implies that

$$\alpha^k = \frac{1}{2}\left(\underline{\theta} + \bar{\theta} - \alpha^{k-1}\right) \Longleftrightarrow \alpha^k - \frac{1}{3}\left(\underline{\theta} + \bar{\theta}\right) = -\frac{1}{2}\left(\alpha^{k-1} - \frac{1}{3}\left(\underline{\theta} + \bar{\theta}\right)\right).$$

Because $\alpha^1 - \left(\underline{\theta} + \bar{\theta}\right)/3 = \left(\underline{\theta} + \bar{\theta} - 2\tilde{\theta}\right)/6$,

$$\alpha^k = \frac{1}{3}\left(\underline{\theta} + \bar{\theta}\right) + \frac{1}{6}\left(\underline{\theta} + \bar{\theta} - 2\tilde{\theta}\right)\left(-\frac{1}{2}\right)^{k-1}. \tag{4}$$

Likewise, letting $\gamma^k = \bar{a}^k - \underline{a}^k$ implies that

$$\gamma^k = \frac{1}{2}\left(\bar{\theta} - \underline{\theta} + \gamma^{k-1}\right) \Longleftrightarrow \gamma^k - \left(\bar{\theta} - \underline{\theta}\right) = \frac{1}{2}\left(\gamma^{k-1} - \left(\bar{\theta} - \underline{\theta}\right)\right).$$

Because $\gamma^1 - \left(\bar{\theta} - \underline{\theta}\right) = -\left(\bar{\theta} - \underline{\theta}\right)/2$,

$$\gamma^k = \left(\bar{\theta} - \underline{\theta}\right) - \frac{1}{2}\left(\bar{\theta} - \underline{\theta}\right)\left(\frac{1}{2}\right)^{k-1}. \tag{5}$$

By (4) and (5), we have

$$\underline{a}^k = \frac{1}{3}\left(2\underline{\theta} - \bar{\theta}\right) + \frac{1}{6}\left(\underline{\theta} + \bar{\theta} - 2\tilde{\theta}\right)\left(\frac{1}{2}\right)^k (-1)^{k-1} + \frac{1}{2}\left(\bar{\theta} - \underline{\theta}\right)\left(\frac{1}{2}\right)^k, \tag{6}$$

$$\bar{a}^k = \frac{1}{3}\left(2\bar{\theta} - \underline{\theta}\right) + \frac{1}{6}\left(\underline{\theta} + \bar{\theta} - 2\tilde{\theta}\right)\left(\frac{1}{2}\right)^k (-1)^{k-1} - \frac{1}{2}\left(\bar{\theta} - \underline{\theta}\right)\left(\frac{1}{2}\right)^k. \tag{7}$$

Because θ is arbitrarily chosen, we conclude that $\Gamma^k(\theta) = \left[\underline{a}^k, \bar{a}^k\right]^2$ holds for any θ, as specified above.

As the bounds of prediction Γ^k is characterized by (6) and (7), we have $\Gamma^k(\theta) \subseteq \Gamma^{k+1}(\theta)$ for any θ and $k \geq 0$. Therefore,

$$\Gamma^{**}(\theta) = \mathrm{cl}\left(\lim_{k\to\infty} \Gamma^k(\theta)\right) = \left[\frac{1}{3}\left(2\underline{\theta} - \bar{\theta}\right), \frac{1}{3}\left(2\bar{\theta} - \underline{\theta}\right)\right]^2$$

holds for any θ. $\square$

A.5 Proof of Proposition 2

Note that, as in the duopoly model, we adopt the infinite canonical type space $\mathcal{T}^{**}$ constructed in Appendix A.4.1, and each type's best response is given by (3) once type space $\mathcal{T}$ is fixed. Because level-0 type space $\mathcal{T}^0$ represents a complete-information game with demand parameter $\tilde{\theta}$, BNE σ^0 given $\mathcal{T}^0$ is uniquely determined by $\sigma_i\left(t^0\right) = \tilde{\theta}/4$. Hence, level-0 prediction Γ^0 is given by

$$\Gamma^0(\theta) = \begin{cases} \left\{\left(\frac{1}{4}\tilde{\theta}, \frac{1}{4}\tilde{\theta}, \frac{1}{4}\tilde{\theta}\right)\right\} & \text{if } \theta = \tilde{\theta}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Given level-1 type space $\mathcal{T}^1$, (3) implies that the best response of type $t_i^1(\theta') \in T_i^1$ is

$$\sigma_i^1\left(t_i^1(\theta')\right) = \frac{1}{2}\left(\mathbb{E}_{\theta \sim \operatorname{marg}_\Theta \hat{\beta}_i^1\left(t_i^1(\theta')\right)}[\theta] - \sigma_{-i}^0(t^0)\right) = \frac{1}{4}\left(2\theta' - \tilde{\theta}\right).$$

Hence, level-1 prediction Γ^1 is characterized as follows: for each θ,

$$\Gamma^1(\theta) = \left[\frac{1}{4}\left(2\underline{\theta} - \tilde{\theta}\right), \frac{1}{4}\left(2\bar{\theta} - \tilde{\theta}\right)\right]^3,$$

where we assume that $2\underline{\theta} - \tilde{\theta} \geq 0$ without loss of generality.[37]

We then consider level-k prediction Γ^k for $k \geq 2$. As an induction hypothesis, we assume that $\Gamma^{k-1}(\theta) = \left[\underline{a}^{k-1}, \bar{a}^{k-1}\right]$ for each θ, and arbitrarily

[37] If $2\underline{\theta} - \tilde{\theta} < 0$, then we directly jump to the case where the lower bound is characterize as a corner solution, which is discussed below.

fix $\theta \in \Theta$. (3) implies that the best response of type $t_i(\theta', \beta_i) \in T_i^k$ is

$$\begin{aligned}
\sigma_i^k\left(t_i^k(\theta', \beta_i)\right) &= \frac{1}{2}\left(\mathbb{E}_{\theta \sim \mathrm{marg}_\Theta \hat{\beta}_i^k\left(t_i^k(\theta',\beta_i)\right)}[\theta] - \mathbb{E}_{t_{-i} \sim \mathrm{marg}_{T_{-i}^{k-1}} \hat{\beta}_i^k\left(t_i^k(\theta',\beta_i)\right)}\left[\sigma_{-i}^{k-1}(t_{-i})\right]\right) \\
&= \frac{1}{2}\left(\theta' - \mathbb{E}_{t_{-i} \sim \mathrm{marg}_{T_{-i}^{k-1}} \hat{\beta}_i^k\left(t_i^k(\theta',\beta_i)\right)}\left[\sigma_{-i}^{k-1}(t_{-i})\right]\right).
\end{aligned}$$

Let $\underline{a}^k \leq \sigma_i^k\left(t_i^k(\theta', \beta_i)\right) \leq \bar{a}^k$. Specifically,

$$\begin{aligned}
\underline{a}^k &= \frac{1}{2}\underline{\theta} - \bar{a}^{k-1}, \\
\bar{a}^k &= \frac{1}{2}\bar{\theta} - \underline{a}^{k-1}.
\end{aligned}$$

Letting $\alpha^k = \underline{a}^k + \bar{a}^k$ implies that

$$\alpha^k = \frac{1}{2}\left(\underline{\theta} + \bar{\theta}\right) - \alpha^{k-1} \Longleftrightarrow \alpha^k - \frac{1}{4}\left(\underline{\theta} + \bar{\theta}\right) = -\left(\alpha^{k-1} - \frac{1}{4}\left(\underline{\theta} + \bar{\theta}\right)\right).$$

Because $\alpha^1 - \left(\underline{\theta} + \bar{\theta}\right)/4 = \left(\underline{\theta} + \bar{\theta} - 2\tilde{\theta}\right)/4$,

$$\alpha^k = \frac{1}{4}\left(\underline{\theta} + \bar{\theta}\right) + \frac{1}{4}\left(\underline{\theta} + \bar{\theta} - 2\tilde{\theta}\right)(-1)^{k-1}. \tag{8}$$

Likewise, letting $\gamma^k = \bar{a}^k - \underline{a}^k$ implies that

$$\gamma^k = \frac{1}{2}\left(\bar{\theta} - \underline{\theta}\right) + \gamma^{k-1} \Longleftrightarrow \gamma^k - \gamma^{k-1} = \frac{1}{2}\left(\bar{\theta} - \underline{\theta}\right).$$

Because $\gamma^1 = \left(\bar{\theta} - \underline{\theta}\right)/2$, we have

$$\gamma^k = \frac{1}{2}k\left(\bar{\theta} - \underline{\theta}\right). \tag{9}$$

By (8) and (9), we have

$$\underline{a}^k = \frac{1}{8}\left(\underline{\theta}+\bar{\theta}\right) - \frac{1}{4}k\left(\bar{\theta}-\underline{\theta}\right) + \frac{1}{8}\left(\underline{\theta}+\bar{\theta}-2\tilde{\theta}\right)(-1)^{k-1}, \tag{10}$$

$$\bar{a}^k = \frac{1}{8}\left(\underline{\theta}+\bar{\theta}\right) + \frac{1}{4}k\left(\bar{\theta}-\underline{\theta}\right) + \frac{1}{8}\left(\underline{\theta}+\bar{\theta}-2\tilde{\theta}\right)(-1)^{k-1}. \tag{11}$$

Notice that for any $k \geq 2$, $\underline{a}^k < \underline{a}^{k-1}$ and $\bar{a}^k > \bar{a}^{k-1}$ hold because of $\tilde{\theta} \in (\underline{\theta}, \bar{\theta})$.

Contrary to the duopoly, (10) implies that there exists $k^* \in \mathbb{N}$ such that $\underline{a}^{k^*+1} < 0 \leq \underline{a}^{k^*}$. Hence, for any $2 \leq k \leq k^*$, prediction Γ^k is given by $\Gamma^k(\theta) = \left[\underline{a}^k, \bar{a}^k\right]^3$ as an analogy of the duopoly. For the characterization of Γ^k with $k > k^*$, consider level-(k^*+1). Because $\underline{a}^{k^*+1} < 0 \leq \underline{a}^{k^*}$ and $A_i = \mathbb{R}_+$, the lower bound of $\sigma_i^{k^*+1}\left(t_i^{k^*+1}(\theta', \beta_j)\right)$ should be 0, whereas its upper bound is given as in the previous levels. Hence, we have

$$\Gamma^{k^*+1}(\theta) = \left[0, \frac{1}{8}\left(\underline{\theta}+\bar{\theta}\right) + \frac{1}{4}(k^*+1)\left(\bar{\theta}-\underline{\theta}\right) + \frac{1}{8}\left(\underline{\theta}+\bar{\theta}-2\tilde{\theta}\right)(-1)^{k^*}\right]^3.$$

For level-(k^*+2), as $\bar{a}^{k^*+1} > \bar{a}^{k^*}$, it also holds that $\underline{a}^{k^*+2} < 0$, implying that the lower bound of $\sigma_i^{k^*+2}\left(t_i^{k^*+2}(\theta', \beta_j)\right)$ should be 0. Because the minimum quantity that level-(k^*+1) types choose is 0, the upper bound of $\sigma_i^{k^*+2}\left(t_i^{k^*+2}(\theta', \beta_j)\right)$ is given by $\bar{\theta}/2$. The similar arguments hold for any level $k > k^*+2$. Hence, for $k \geq k^*+2$, level-k prediction is given by $\Gamma^k(\theta) = [0, \bar{\theta}/2]^3$. Because θ is arbitrarily chosen, we conclude that Γ^k is independent of θ.

The arguments so far imply that $\Gamma^k(\theta) \subseteq \Gamma^{k+1}(\theta)$ holds for each θ and $k \geq 0$. Therefore,

$$\Gamma^{**}(\theta) = \mathrm{cl}\left(\lim_{k\to\infty} \Gamma^k(\theta)\right) = \left[0, \frac{1}{2}\bar{\theta}\right]^3$$

holds for each θ. $\square$

A.6 Proof of Proposition 3

Note that player i's expected payoff under type space $\mathcal{T}$ is given by $\mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[u_i(a_i,a_{-i},\theta)] = \mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[\theta a_j]a_i - a_i^3/3$. Hence, the first-order condition implies that type t_i's best response is

$$a_i(t_i) = \sqrt{\mathbb{E}_{(\theta,t_{-i})\sim\hat{\beta}_i(t_i)}[\theta a_{-i}]}. \tag{12}$$

We adopt infinite canonical type space $\mathcal{T}^{**}$ that is identical to that constructed in Appendix A.4.1 except that $\theta = 1$ is common knowledge among level-0 types. Furthermore, as mentioned in the body of the paper, we assume that the level-0 types play the payoff-dominating equilibrium: $a_i = 1$ for each i. Specifically, because type t^0 certainly believes that (i) $\theta = 1$ and (ii) the opponent chooses action $a = 1$. Hence, level-0 prediction Γ^0 is given by

$$\Gamma^0(\theta) = \begin{cases} \{\,(1,1)\,\} & \text{if } \theta = 1, \\ \emptyset & \text{otherwise.} \end{cases}$$

Likewise, (12) implies that BNE σ^1 given $\mathcal{T}^1$ is characterized as follows: for each i and type $t_i\,(\theta',\beta_i)$,

$$\sigma_i^1\left(t_i^1(\theta',\beta_i)\right) = \sqrt{\theta' \mathbb{E}_{t_{-i}\sim \mathrm{marg}_{\mathcal{T}^0_{-i}}\hat{\beta}_i^1\left(t_i^1(\theta',\beta_i)\right)}\left[\sigma^0_{-i}(t_{-i})\right]}.$$

Hence, we have $\Gamma^1(\theta) = [1,(1+\varepsilon)^{1/2}]^2$ for each θ.

Now, consider level-k prediction Γ^k with $k \geq 2$. As an induction hypothesis, we assume that $\Gamma^{k-1}(\theta) = [1,\bar{a}^{k-1}]^2$ for each θ. Under arbitrarily fixed θ, (12) implies that BNE σ^k given $\mathcal{T}^k$ is characterized as follows: for each i

and $t_i^k(\theta', \beta_i)$,

$$\sigma_i^k\left(t_i^k(\theta', \beta_i)\right) = \sqrt{\theta' \mathbb{E}_{t_{-i} \sim \mathrm{marg}_{T_{-i}^{k-1}} \hat{\beta}_i^k\left(t_i^k(\theta', \beta_i)\right)}\left[\sigma_{-i}^{k-1}(t_{-i})\right]}.$$

Let $\underline{a}^k$ and $\bar{a}^k$ be the upper and lower bounds of σ_i^k. Specifically, $\underline{a}^k = 1$ and $\bar{a}^k = \sqrt{(1+\varepsilon)\bar{a}^{k-1}}$. Note that

$$\bar{a}^k = \left((1+\varepsilon)\bar{a}^{k-1}\right)^{1/2} \Longleftrightarrow \alpha^k - 1 = \frac{1}{2}\left(\alpha^{k-1} - 1\right),$$

where $\alpha^k = \log_{1+\varepsilon} \bar{a}^k$. Because $\alpha^1 - 1 = -1/2$, we have

$$\alpha^k = 1 - \frac{1}{2^k} \Longleftrightarrow \bar{a}^k = (1+\varepsilon)^{1-1/2^k}.$$

Hence, as θ is arbitrarily chosen, we conclude that $\Gamma^k(\theta) = \left[1, (1+\varepsilon)^{1-1/2^k}\right]$ holds for each θ. Furthermore, because $\Gamma^k(\theta) \subseteq \Gamma^{k+1}(\theta)$ holds for any θ and $k \geq 0$, we have

$$\Gamma^{**}(\theta) = \mathrm{cl}\left(\lim_{k\to\infty} \Gamma^k(\theta)\right) = [1, 1+\varepsilon]^2$$

holds for any θ. $\square$

A.7 Proof of Proposition 4

First, we construct infinite canonical type space $\mathcal{T}^{**}$ as follows. We assume that the bidders are certainly the truthful type under level-0 type space $\mathcal{T}^0 = \left(T^0, \hat{\theta}^0, \hat{\beta}^0\right)$. Specifically, for each i, (i) $T_i^0 = \{t^0\}$, and (ii) $\hat{\theta}_i^0(t^0) = 0$ and $\hat{\beta}_i^0\left(T_{-i}^0 \mid t^0\right) = 1$. In the subsequent levels, the bidders would be the rational type, and they have subjective beliefs over the opponent's type. That is, level-1 type space $\mathcal{T}^1 = \left(T^1, \hat{\theta}^1, \hat{\beta}^1\right)$ is defined as follows: for each i, (i) $T_i^1 = \{\, t_i^1(\theta_i) \mid \theta_i \in \Theta_i \,\}$, and (ii) $\hat{\theta}_i^1\left(t_i^1(\theta_i)\right) = \theta_i$

and $\hat{\beta}_i^1\left(T^0_{-i} \mid t_i^1(\theta_i)\right) = 1$ for each θ_i. Likewise, level-k type space $\mathcal{T}^k = \left(T^k, \hat{\theta}^k, \hat{\beta}^k\right)$ for $k \geq 2$ is recursively defined as follows: for each i, (i) $T_i^k = \left\{ t_i^k(\theta_i, \beta_i) \;\middle|\; \theta_i \in \Theta_i \text{ and } \beta_i \in \Delta\left(T_{-i}^{k-1}\right) \right\}$, and (ii) $\hat{\theta}_i^k\left(t_i^k(\theta_i, \beta_i)\right) = \theta_i$ and $\hat{\beta}_i^k\left(t_i^k(\theta_i, \beta_i)\right) = \beta_i$ for each θ_i and β_i. Infinite canonical type space $\mathcal{T}^{**} = \left(T^{**}, \hat{\theta}^{**}, \hat{\beta}^{**}\right)$ is then defined as the limit of sequence $\left\{\mathcal{T}^k\right\}_{k=0}^{\infty}$.

Next, we derive infinite canonical prediction Γ^{**} given infinite canonical type space $\mathcal{T}^{**}$. As mentioned in the body of the paper, the best response of the truthful bidder is uniquely determined independent of his subjective belief: for each i, k, and β_i, $\sigma_i^k\left(t_i^k(0, \beta_i)\right) = a^I$, where $a^I(v_i) = v_i$ for each v_i. Hence, it is straightforward that the unique BNE σ^0 given $\mathcal{T}^0$ is characterized by $\sigma_i^0(t^0) = a^I$ for each i, implying that

$$\Gamma_i^0(\theta_i, v_i) = \begin{cases} \{\, v_i \,\} & \text{if } \theta_i = 0, \\ \emptyset & \text{otherwise} \end{cases}$$

holds for each v_i.

Now, we consider the optimal behavior of type $t_i^1(1)$. Because he certainly believes that his opponent is the truthful type, his expected payoff from bid b is given by $(v_i - b)b$. By the first-order condition, the best response is $a_i^1(v_i) = v_i/2$. Hence, level-1 prediction Γ_i^1 is characterized as follows:

$$\Gamma_i^1(\theta_i, v_i) = \begin{cases} \{\, v_i \,\} & \text{if } \theta_i = 0, \\ \{\, \frac{1}{2} v_i \,\} & \text{otherwise.} \end{cases}$$

Given the behavior characterized so far, type $t_i^2(1, \beta_i)$'s expected payoff from bid b is $(v_i - b)\left(b(1 + \mu)\right)$, where $\mu = \beta_i\left(\{t_{-i}^1 \mid \hat{\theta}_{-i}^1(t_{-i}^1) = 1\}\right)$ represents type $t_i^2(1, \beta_i)$'s belief that his opponent is the rational bidder. By the first-order condition, his best response is $a_i^2(v_i) = v_i/2$. Hence, level-2 prediction

Γ_i^2 is characterized as follows:

$$\Gamma_i^2(\theta_i, v_i) = \begin{cases} \{\, v_i \,\} & \text{if } \theta_i = 0, \\ \{\, \frac{1}{2} v_i \,\} & \text{otherwise.} \end{cases}$$

Because $\Gamma^1 = \Gamma^2$, we have the same characterization for the higher levels; that is, $\Gamma^k = \Gamma^1$ for any $k \geq 2$. Thus, we conclude that $\Gamma^{**} = \Gamma^1$. $\square$

B Appendix: Supplementary materials

B.1 Finite discretization

This section challenges the concern by considering a finitely discretized environment. Our stance is to define a robust prediction as the one providing a reasonable prediction in any given *finite* type space. In the finite environment, focusing on a single type space, i.e., (finite) canonical type space $\mathcal{T}^*$, is sufficient in the sense that Γ^* is attained as a prediction given that type space. Here, we show an analogous necessity result but based on a *collection* of finite type spaces constructed by the level-k argument in finitely discretized environments. Hereafter, such a finite type space is referred to as an *m-discretized level-k type space* with $m \in \mathbb{N}$. Intuitively, the m-discretized level-k type space is the level-k type space in the environment where the possible payoff parameters and beliefs about the opponents' types are finite, and the distance between two distinct elements in each set is at least $1/m$. In the following, we will show that the prediction generated from the collection of m-discretized level-k type spaces approximate infinite canonical prediction Γ^{**}.

We inductively construct m-discretized level-k type spaces analogously to the infinite level-k type spaces with the modification that the sets of possible payoff parameters and beliefs are finite. More specifically, as a starting point, m-discretized level-0 type space $\mathcal{T}^{m,0} = (T^{m,0}, \hat{\theta}^{m,0}, \hat{\beta}^{m,0})$ is defined by $\mathcal{T}^{m,0} = \mathcal{T}^0$ for each m. That is, we adopt the baseline scenario that is identical to that of the infinite level-k construction. Note that type space $\mathcal{T}^{m,0}$ is finite by assumption.

Given type space $\mathcal{T}^{m,0}$, m-discretized level-1 type space $\mathcal{T}^{m,1} = (T^{m,1}, \hat{\theta}^{m,1}, \hat{\beta}^{m,1})$ is constructed as follows. Let $\hat{\Theta}_i = \left\{ \theta_i \in \Theta_i \;\middle|\; \exists t_i \in T_i^{m,0} \text{ such that } \hat{\theta}_i^{m,0}(t_i) = \theta_i \right\}$ denote the set of player i's payoff parameters when his type is in m-discretized level-0. Because $T_i^{m,0}$ is finite for each i, so is $\hat{\Theta}_i$. Now, we fix $\theta^* \in \Theta$, arbitrarily. For each i and m, we take a finite subset $\Theta_i^m \subseteq \Theta_i$ such that (i) $\theta^* \in \Theta_i^m$, (ii) $\hat{\Theta}_i \subseteq \Theta_i^m \subseteq \Theta_i^{m+1}$, and (iii) $\Theta_i \subseteq \mathbb{B}_{1/m}(\Theta_i^m)$. Likewise, let $B_i^{m,0} \subseteq \Delta(T_{-i}^{m,0})$ be a *finite* subset of beliefs that player i may have about the opponents' level-0 types, such that (i) $B_i^{m,0} \subseteq B_i^{m+1,0}$ and (ii) $\Delta\left(T_{-i}^{m,0}\right) \subseteq \mathbb{B}_{1/m}\left(B_i^{m,0}\right)$. Note that Θ_i^m and $B_i^{m,0}$ are the index sets of finite subcovers of Θ_i and $\Delta\left(T_{-i}^{m,0}\right)$, respectively.[38] In the m-discretized level-1 type space, the possible payoff parameters and beliefs are restricted to finite sets Θ_i^m and $B_i^{m,0}$ for each i, respectively. Given the finite restriction, type space $\mathcal{T}^{m,1}$ is constructed to satisfy the properties that are analogous to the finite counterpart. That is, for each i,

(i) $T_i^{m,1}$ is finite and $T_i^{m,1} \supseteq T_i^{m,0} = T_i^0$;

(ii) $\hat{\theta}_i^{m,1}(t_i) \in \Theta_i^m$ and $\hat{\beta}_i^{m,1}(t_i) \in B_i^{m,0}$ for each $t_i \in T_i^{m,1}$;

(iii) If $t_i \in T_i^{m,0}$, then $\left(\hat{\theta}_i^{m,1}(t_i), \hat{\beta}_i^{m,1}(t_i)\right) = \left(\hat{\theta}_i^{m,0}(t_i), \hat{\beta}_i^{m,0}(t_i)\right)$;

(iv) For each $(\theta_i, \beta_i) \in \Theta_i^m \times B_i^{m,0}$, there exists $t_i \in T_i^{m,1}$ with $\left(\hat{\theta}_i^{m,1}(t_i), \hat{\beta}_i^{m,1}(t_i)\right) = (\theta_i, \beta_i)$.[39]

[38] The existence of Θ_i^m and $B_i^{m,0}$ is guaranteed because of the compactness of Θ_i and $\Delta\left(T_{-i}^{m,0}\right)$.

[39] The counterpart of Condition (v) imposed in finite games is unnecessary here. The condition is used for showing the best-reply property, which is an essential property for the sufficiency part. However, because we show the sufficiency part by directly focusing

Given type space $\mathcal{T}^{m,k-1}$ with $k \geq 2$, we can analogously define type space $\mathcal{T}^{m,k}$. As in the m-discretized level-1 type space, player i's possible payoff parameters are restricted to Θ_i^m at any level. The possible beliefs in type space $\mathcal{T}^{m,k}$ are restricted to *finite* set $B_i^{m,k-1} \subseteq \Delta\left(T_{-i}^{m,k-1}\right)$ such that (i) $B_i^{m,k-2} \subseteq B_i^{m,k-1} \subseteq B_i^{m+1,k-1}$ and (ii) $\Delta\left(T_{-i}^{m,k-1}\right) \subseteq \mathbb{B}_{1/m}\left(B_i^{m,k-1}\right)$. Given finite sets Θ_i^m and $B_i^{m,k-1}$, type space $\mathcal{T}^{m,k}$ is constructed with satisfying the analogous properties required in the level-1.

As an analogy of the finite counterpart, we construct a BNE $\sigma^{m,k}$ given type space $\mathcal{T}^{m,k}$ for each k, which implies prediction $\Gamma^{m,k}$. Especially, BNE $\sigma^{m,0}$ is fixed to that focused in the infinite level-k construction for each m. Because we share the same baseline scenario, we have that $\Gamma^{m,0} = \Gamma^0$ for each m. We then define $\Gamma^m = \bigcup_{k\in\mathbb{N}} \Gamma^{m,k}$. It is worthwhile to remark the following. Because type space $\mathcal{T}^{m,k}$ is finite for each m and k, prediction $\Gamma^{m,k}$ should be included in a robust prediction. However, as mentioned above, this process may not stop at any finite level, as opposed to the finite environment. In this case, the union of all these m-discretized level-k type spaces may not be a finite space as itself. Therefore, in what follows, instead of considering a single type space that has countably many levels, we interpret this construct as a collection of countably many finite type spaces. That is, mathematically, we continue to use Γ^m as the union given as above, but with the interpretation that it is literally the union of the predictions in the above constructed m-discretized level-k type spaces, instead of regarding it as a prediction of a single infinite type space. Now, we have the following theorem.

Theorem 3. For each infinite canonical prediction Γ^{**} and $\theta \in \Theta$, there exists a collection of m-discretized level-k type spaces $\{\mathcal{T}^{m,k}\}_{m,k\in\mathbb{N}}$ such that $\mathrm{cl}\left(\bigcup_{m\in\mathbb{N}} \Gamma^m(\theta)\right) = \Gamma^{**}(\theta)$.

This relationship comes from the fact that $\Gamma^m(\theta) \subseteq \Gamma^{m+1}(\theta) \subseteq \Gamma^{**}(\theta) \subseteq$

on the infinite canonical type space, it is unnecessary to guarantee the best-reply property in m-discretized type spaces.

$\mathbb{B}_{1/m}(\Gamma^m(\theta))$ for each m and $\theta \in \Theta^m$. Intuitively, as the grid of the discretized environment becomes finer, the prediction in the discretized environment becomes expands whereas its open ball shrinks more. As a result, both sets coincide in the limit. In this sense, we say that Γ^m approximates infinite canonical prediction Γ^{**} by considering a discretized environment with a sufficiently fine but finite grid. As a final remark, the adopted collection of m-discretized level-k type spaces depends on θ. That is, to show the above relationship, we might use different collections for distinct parameters. Thus, by collecting all of these discretized level-k type spaces, we conclude that: for any $\theta \in \Theta$ and $a \in \Gamma^{**}(\theta)$, there exists a finite type space in which θ plays a, and vice versa. This is precisely the desired necessity property of Γ^{**}.

B.2 Proof of Theorem 3

B.2.1 Preliminaries

First, we show the following key lemma.

Lemma 2. For each $m \in \mathbb{N}$ and $\theta \in \Theta^m$, prediction $\Gamma^m(\theta)$ satisfies the following: $\Gamma^m(\theta) \subseteq \Gamma^{m+1}(\theta) \subseteq \Gamma^{**}(\theta) \subseteq \mathbb{B}_{1/m}(\Gamma^m(\theta))$.

Proof of Lemma 2. The first inclusion is obvious from the fact that (i) $\sigma^{m,0} = \sigma^0$, (ii) $\Theta_i^m \subseteq \Theta_i^{m+1}$, and (iii) $B_i^{m,k} \subseteq B_i^{m+1,k}$ for each i, m, and k. The second inclusion is because the m-discretized level-k type spaces (that induce Γ^m) and the infinite canonical type space (that induce Γ^{**}) (i) have the same level-0 type space (and the same equilibrium behavior for the level-0 types), but (ii) at each $k \geq 1$, the latter essentially allows for more types than the former. Therefore, $\Gamma^{**}(\theta)$ contains more elements.

To show the third inclusion, take m, $\theta \in \Theta^m$ and $a \in \Gamma^{**}(\theta)$. The proof goes by induction. First, suppose that level-0 types have θ as their payoff parameters. That is, we say that $a \in \Gamma^0(\theta)$. Because of the construction of type space $\mathcal{T}^{m,0}$, we have that $\Gamma^{m,0}(\theta) = \Gamma^0(\theta)$, implying that

$a \in \Gamma^{m,0}(\theta) \subseteq \Gamma^m(\theta) \subseteq \mathbb{B}_{1/m}\left(\Gamma^m(\theta)\right)$. We suppose, as an induction hypothesis, that if level-$(k-1)$ types have θ as their payoff parameters, then we have $a \in \Gamma^{k-1}(\theta) \subseteq \mathbb{B}_{1/m}\left(\Gamma^{m,(k-1)}(\theta)\right)$. We now consider the level k, and suppose that level-k types in T^k have θ as their payoff parameters. Because $a \in \Gamma^{**}(\theta)$, for each i, there exists $t_i \in T_i^k \subseteq T_i^{**}$ such that (i) $\hat{\theta}_i^k(t_i) = \theta_i$, (ii) q_i is a conjecture about (θ_{-i}, a_{-i}) induced by $\hat{\beta}_i^k(t_i)$, and (iii) $a_i = a_i^*(\theta_i, q_i) \in \Gamma_i^k(\theta_i)$. Note that, by the induction hypothesis, we have $\operatorname{supp}\left(q_i(\theta_{-i}, \cdot)\right) \subseteq \Gamma_{-i}^{k-1}(\theta_{-i}) \subseteq \mathbb{B}_{1/m}\left(\Gamma_{-i}^{m,(k-1)}(\theta_{-i})\right)$ for each $\theta_{-i} \in \Theta_{-i}$. Hence, there exists conjecture $q_i' \in \Delta\left(\Theta_{-i}^m \times A_{-i}\right)$ such that (i) $d(q_i, q_i') < 1/m$ and (ii) $\operatorname{supp}\left(q_i'(\theta_{-i}, \cdot)\right) \subseteq \Gamma_{-i}^{m,(k-1)}(\theta_{-i})$ for each $\theta_{-i} \in \Theta_{-i}^m$. Because of Property (ii) of conjecture q' and the construction of m-discretized level-k type spaces, there exists type $t_i'' \in T_i^k$ such that belief $\hat{\beta}_i^{m,k}(t_i'')$ implies conjecture $q_i'' \in \Delta\left(\Theta_{-i}^m \times A_{-i}\right)$ that is sufficiently close to conjecture q_i'. Hence, without loss of generality, we assume that $q_i'' = q_i$ for simplicity. Therefore, Property (i) of conjecture q_i' implies that

$$a_i = a_i^*(\theta_i, q_i) \in \mathbb{B}_{1/m}\left(a_i^*(\theta_i, q_i')\right) \subseteq \mathbb{B}_{1/m}\left(\Gamma_i^{m,k}(\theta_i)\right) \subseteq \mathbb{B}_{1/m}\left(\Gamma_i^m(\theta_i)\right).$$

That is, $\Gamma_i^k(\theta_i) \subseteq \mathbb{B}_{1/m}\left(\Gamma_i^{m,k}(\theta_i)\right)$ holds, implying the third inclusion. □

B.2.2 Proof of Theorem 3.

First, we show that $\operatorname{cl}\left(\bigcup_{m\in\mathbb{N}} \Gamma^m(\theta)\right) \subseteq \Gamma^{**}(\theta)$. For each $\theta \in \Theta$, there exists a collection of m-discretized level-k type spaces $\{\mathcal{T}^{m,k}\}_{m,k\in\mathbb{N}}$ such that $\theta \in \Theta^m$ for some m. Then, the first and second inclusions of Lemma 2 implies that

$$\operatorname{cl}\left(\bigcup_{m\in\mathbb{N}} \Gamma^m(\theta)\right) \subseteq \operatorname{cl}\left(\Gamma^{**}(\theta)\right) = \Gamma^{**}(\theta),$$

where the last equality comes from the definition of Γ^{**}.

Next, we show that $\Gamma^{**}(\theta) \subseteq \operatorname{cl}\left(\bigcup_{m\in\mathbb{N}} \Gamma^m(\theta)\right)$. Fix $\theta \in \Theta$ and $a \in$

$\Gamma^{**}(\theta)$, arbitrarily. Again, there exists a collection of m-discretized level-k type spaces $\{\mathcal{T}^{m,k}\}_{m,k\in\mathbb{N}}$ such that $\theta \in \Theta^m$ for some m. Let m represent the first integer such that $\theta \in \Theta^m$. Then, by the third inclusion of Lemma 2,

$$a \in \Gamma^{**}(\theta) \subseteq \mathbb{B}_{\frac{1}{m}}\left(\Gamma^m(\theta)\right) \subseteq \mathbb{B}_{\frac{1}{m}}\left(\bigcup_{\tilde{m}\in\mathbb{N}} \Gamma^{\tilde{m}}(\theta)\right).$$

Because of the monotonicity of Θ^m, $\theta \in \Theta^{m'}$ holds for any $m' \geq m$. Hence:

$$a \in \Gamma^{**}(\theta) \subseteq \mathbb{B}_{\frac{1}{m'}}\left(\Gamma^{m'}(\theta)\right) \subseteq \mathbb{B}_{\frac{1}{m'}}\left(\bigcup_{\tilde{m}\in\mathbb{N}} \Gamma^{\tilde{m}}(\theta)\right).$$

Therefore, we have that

$$a \in \bigcap_{m'\geq m} \mathbb{B}_{\frac{1}{m'}}\left(\bigcup_{\tilde{m}\in\mathbb{N}} \Gamma^{\tilde{m}}(\theta)\right) = \mathrm{cl}\left(\bigcup_{\tilde{m}\in\mathbb{N}} \Gamma^{\tilde{m}}(\theta)\right). \qquad \square$$

References

Basu, K. (1992): "A Characterization of the Class of Rationalizable Equilibria of Oligopoly Games," *Economics Letters*, 40(2), 187–191.

Basu, K., and J. W. Weibull (1991): "Strategy Subsets Closed under Rational Behavior," *Economics Letters*, 36(2), 141–146.

Battigalli, P., A. Di Tillio, E. Grillo, and A. Penta (2011): "Interactive Epistemology and Solution Concepts for Games with Asymmetric Information," *B.E. Journal of Theoretical Economics*, 11(1), Article 6.

Battigalli, P., and M. Siniscalchi (2003): "Rationalizable Bidding in First-Price Auctions," *Games and Economic Behavior*, 45(1), 38–72.

Bergemann, D., and S. Morris (2013): "Robust Prediction in Games with Incomplete Information," *Econometrica*, 81(4), 1251–1308.

——— (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structure in Games," *Theoretical Economics*, 11(2), 487–522.

——— (2017): "Belief-Free Rationalizability and Informational Robustness," *Games and Economic Behavior*, 104, 744–759.

Bernheim, B. D. (1984): "Rationalizable Strategic Behavior," *Econometrica*, 52(4), 1007–1028.

Bögers, T., and D. Smith (2014): "Robust Mechanism Design and Dominant Strategy Voting Rules," *Theoretical Economics*, 9(2), 339–360.

Brandenburger, A., and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55(6), 1391–1402.

Brooks, B., and S. Du (2021): "Optimal Auction Design with Common Values: An Informationally Robust Approach," *Econometrica*, 89(3), 1313–1360.

Chen, Y.-C. (2012): "The Structure Theorem for Rationalizability in the Normal Form of Dynamic Games," *Games and Economic Behavior*, 75(2), 587–597.

Chen, Y.-C., and J. Li (2018): "Revisiting the Foundations of Dominant-Strategy Mechanisms," *Journal of Economic Theory*, 178, 294–317.

Chen, Y.-C., S. Takahashi, and S. Xiong (2014): "The Robust Selection of Rationalizability," *Journal of Economic Theory*, 151, 448–475.

——— (2021): "Robust Refinements of Rationalizability with Arbitrarily Payoff Uncertainty," mimeo, National University of Singapore and University of California, Riverside.

CHUNG, K.-S., AND J. ELY (2007): "Foundations of Dominant-Strategy Mechanisms," *Review of Economic Studies*, 74(2), 447–476.

CRAWFORD, V. P., M. A. COSTA-GOMES, AND N. IRIBERRI (2013): "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, 51(1), 5–62.

CRAWFORD, V. P., AND N. IRIBERRI (2007): "Level-k Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-value Auctions?," *Econometrica*, 75(6), 1721–1770.

CRAWFORD, V. P., T. KUGLER, Z. NEEMAN, AND A. PAUZNER (2009): "Behaviorally Optimal Auction Design: Examples and Observations," *Journal of the European Economic Association*, 7(2-3), 377–387.

CRAWFORD, V. P., AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451.

DE CLIPPEL, G., R. SARAN, AND R. SERRANO (2019): "Level-k Mechanism Design," *The Review of Economic Studies*, 86(3), 1207–1227.

DEKEL, E., D. FUDENBERG, AND S. MORRIS (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2(1), 15–40.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory.* MIT Press, Cambridge, MA.

HARSANYI, J. C. (1967-68): "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, 159–182,320–334,486–502.

KAGEL, J. H., AND D. LEVIN (1993): "Independent Private Value Auctions: Bidder Behaviour in First-, Second- and Third-Price Auctions with Varying Numbers of Bidders," *Economic Journal*, 103(419), 868–879.

Kajii, A., and S. Morris (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65(6), 1283–1309.

Kunimoto, T., and T. Yamashita (2020): "Order on Types Based on Monotone Comparative Statics," *Journal of Economic Theory*, 189, 105082.

Liu, Q. (2015): "Correlation and Common Priors in Games with Incomplete Information," *Journal of Economic Theory*, 157, 49–75.

Miura, S., and T. Yamashita (2020): "Maximal Miscommunication," *Economics Letters*, 188, 108962.

Morris, S., and T. Ui (2005): "Generalized Potentials and Robust Sets of Equilibria," *Journal of Economic Theory*, 124(1), 45–78.

——— (2020): "Incomplete Information Robustness," mimeo, MIT and Hitotsubashi University.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85(5), 1313–1326.

Oyama, D., and O. Tercieux (2010): "Robust Equilibria under Non-Common Priors," *Journal of Economic Theory*, 145(2), 752–784.

Pearce, D. G. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52(4), 029–1050.

Penta, A. (2012): "Higher-Order Uncertainty and Information: Static and Dynamic Games," *Econometrica*, 80(2), 631–660.

——— (2013): "On the Structure of Rationalizability for Arbitrary Spaces of Uncertainty," *Theoretical Economics*, 8(2), 405–430.

Stahl, D. O., and P. W. Wilson (1994): "Experimental Evidence on Players' Models of Other Players," *Journal of Economic Behavior & Organization*, 25(3), 309–327.

——— (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10(1), 218–254.

Ui, T. (2001): "Robust Equilibria of Potential Games," *Econometrica*, 69(5), 1373–1380.

Van Zandt, T., and X. Vives (2007): "Monotone Equilibria in Bayesian Games of Strategic Complementarities," *Journal of Economic Theory*, 134(1), 339–360.

Weinstein, J., and M. Yildiz (2007a): "Impact of Higher-Order Uncertainty," *Games and Economic Behavior*, 60(1), 200–212.

——— (2007b): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75(2), 365–400.

——— (2011): "Sensitivity of Equilibrium Behavior to Higher-Order Beliefs in Nice Games," *Games and Economic Behavior*, 72(1), 288–300.

Yamashita, T., and S. Zhu (2021): "On the Foundation of Ex Post Incentive Compatible Mechanisms," *American Economic Journal: Microeconomics*, forthcoming.