SENSITIVITY OF EQUILIBRIUM BEHAVIOR TO HIGHER-ORDER BELIEFS IN NICE GAMES

JONATHAN WEINSTEIN AND MUHAMET YILDIZ

Abstract. We consider "nice" games (where action spaces are compact intervals, utilities continuous and strictly concave in own action), which are used frequently in classical economic models. Without making any "richness" assumption, we characterize the sensitivity of any given Bayesian Nash equilibrium to higher-order beliefs. That is, for each type, we characterize the set of actions that can be played in equilibrium by some type whose lower-order beliefs are all as in the original type. We show that this set is given by a local version of interim correlated rationalizability. This allows us to characterize the robust predictions of a given model under arbitrary common knowledge restrictions. As an application, we show in Cournot competition with many players that even if one assumes that it is common knowledge that the payoffs are within arbitrarily small neighborhood of a given value and the given value is mutually known at arbitrarily high order, he cannot rule out any production level that is below the monopoly production as the equilibrium production of a given firm.

JEL Numbers: C72, C73.

This paper is based on an earlier working paper, Weinstein and Yildiz (2004). We thank Daron Acemoglu, Pierpaolo Battigalli, Tilman Borgers, Glenn Ellison, Drew Fudenberg, Stephen Morris, and the seminar participants at various universities and conferences. Yildiz also thanks Princeton University and Cowles Foundation of Yale University for their generous support and hospitality.

1. Introduction

Most economic applications fix a specific type space. They thus make common knowledge assumptions that are difficult to verify in the modeling stage. Unfortunately, as in the well-known e-mail game example of Rubinstein (1989), the equilibrium behavior may be highly sensitive to these assumptions, and the researcher may not be able to verify whether his predictions are valid. Indeed, under a richness assumption, for finite action games, in Weinstein and Yildiz (2007) [hereafter WY], we showed that whenever there are multiple rationalizable actions, rationalizable strategies are highly sensitive to these assumptions: by relaxing these assumptions suitably one can make any rationalizable action the unique solution. In this paper, focusing on the set of Bayesian Nash equilibria, we extend this result in two important directions.

First, in WY, we assume that the set of underlying payoff parameters is rich enough, so that each action could be dominant at some parameter value. In other words, there are no common knowledge restrictions on payoffs. The application, however, may impose a natural structure on payoffs, and the researcher may be willing to assume that this structure is common knowledge. For example, it may be natural to assume that the bidders in an auction care only about whether they win the object and how much they pay in case they win. Such a bidder would be indifferent towards how he loses the object, and in that case submitting a low bid cannot be a strictly dominant action for him. If the researcher is willing to assume that this is indeed common knowledge, then the richness assumption of WY would fail. In this paper, we do not make any richness assumptions, any yet we characterize the sensitivity of equilibrium strategies by a local version of interim correlated rationalizability, rather than the usual interim correlated rationalizability (Dekel, Fudenberg, Morris (2007)).

Second, in WY, we focus on finite-action games. Here, instead we focus on nice games (Moulin, 1984), which are commonly used in classical economic models. In these games, the action spaces are compact intervals, and the utility functions are continuous and strictly concave in own action, as in Cournot competition and differentiated Bertrand competition. It is crucial for the construction of WY that a particular mapping is measurable, which is vacuously true for finite-action games with finite types. There is no general way to circumvent this problem, but here we are able to circumvent the measurability problem because of a special structure of rationalizability in nice games.

Since we do not make any richness assumption on the set of payoff functions, our result allows us to analyze the robustness of equilibrium predictions in complete information games under weaker robustness concepts, as in uniform topology. Instead of assuming the payoffs are common knowledge, assume that the payoffs are mutually known up to an arbitrarily finite order and that it is common knowledge that the payoffs are in an arbitrarily small neighborhood of the actual payoffs. Then, our result states that an equilibrium prediction remains valid under these slightly weaker assumptions if and only if it is true for all locally interim correlated rationalizable strategies.

In some important games this leads to disturbing conclusions. As an example we consider Cournot oligopoly with linear cost function and sufficiently many firms. We show that in such a game any production level that is less than or equal to the monopoly production is locally interim correlated rationalizable. Therefore, if we weaken the complete information assumption slightly by assuming instead that the payoffs are mutually known up to an arbitrarily high finite order and it is common knowledge that the payoffs are within an arbitrarily small neighborhood of the original payoffs, then all we can conclude is that individual firms' productions do not exceed the monopoly production, a trivial conclusion that follows from profit maximization.

In the next section, we lay out our model. In Section 3, we introduce our notion of sensitivity of equilibrium strategies and present our general result. In section 4, we present our application on Cournot oligopoly. Section 5 concludes. Some of the proofs are relegated to the appendix.

2. Basic Definitions

We consider n-player *nice games* with some possibly unknown payoff-relevant parameter $\theta \in \Theta^*$ where Θ^* is a compact metric space and with a finite set $N = \{1, 2, \ldots, n\}$ of players. In a nice game, the action space of each player i is $A_i = [0, 1]$; the space of action profiles is $A = [0, 1]^n$,[1] and the utility function u_i :

[1]**Notation:** Given any list $X_1, \ldots, X_n$ of sets, write $X = X_1 \times \cdots \times X_n$ with typical element x, $X_{-i} = \prod_{j \neq i} X_j$ with typical element x_{-i}, and $(x_i, x_{-i}) = (x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$.

$\Theta^* \times A \to \mathbb{R}$ of player i is continuous in action profile $a = (a_i, a_{-i}) \in A$ and strictly concave[2] in own action $a_i \in A_i$. We fix the players, action space and utility function and consider the set of games that differ in their specifications of the belief structure on θ, i.e. their type spaces, which we also call models. Formally, by a (*countable*) *model*, we mean a countable set $\Theta \times T_1 \times \cdots \times T_n$ associated with beliefs $\kappa_{t_i} \in \Delta(\Theta \times T_{-i})$ for each $t_i \in T_i$, where $\Theta \subseteq \Theta^*$.

Remark: In our formulation, in any model it is common knowledge that the payoff functions are in $\{u(\theta, \cdot) | \theta \in \Theta^*\}$. Since Θ^* is arbitrary, this allows arbitrary common knowledge restrictions on payoff functions.

Given any type t_i in a type space T, we can compute the first-order belief $h_i^1(t_i) \in \Delta(\Theta^*)$ of t_i (about θ), second-order belief $h_i^2(t_i) \in \Delta(\Theta^* \times \Delta(\Theta^*)^n)$ *of* t_i (about θ and the first-order beliefs), etc., using the joint distribution of the types and θ. Using the mapping $h_i : t_i \mapsto (h_i^1(t_i), h_i^2(t_i), \ldots)$, we can embed all such models in the universal type space (Mertens and Zamir (1985) and Brandenburger and Dekel (1993)). We will consider the subset of universal type space that is generated by countable models. That is, we consider $T^u = T_1^u \times \cdots \times T_n^u$ where

$$T_i^u = \{h_i(t_i) | t_i \in T_i \text{ for some countable model } \Theta \times T\}.$$

A *strategy* of a player i with respect to T_i is any function $s_i : T_i \to A_i$. Given any type t_i and any profile s_{-i} of strategies, we write $\pi(\cdot|t_i, s_{-i}) \in \Delta(\Theta \times A_{-i})$ for the joint distribution of the underlying uncertainty and the other players' actions induced by t_i and s_{-i}. We define $\pi(\cdot|t_i, s_{-i})$ similarly for functions $s_{-i} : \Theta \times T_{-i} \to A_{-i}$. For each $i \in N$ and for each belief $\pi \in \Delta(\Theta \times A_{-i})$, we write $BR_i(\pi)$ for the *unique* action $a_i \in A_i$ that maximizes the expected value of $u_i(\theta, a_i, a_{-i})$ under the probability distribution π. A strategy profile $s^* = (s_1^*, s_2^*, \ldots)$ is a *Bayesian Nash equilibrium* if and only if at each t_i,

$$s_i^*(t_i) = BR_i\left(\pi\left(\cdot|t_i, s_{-i}^*\right)\right).$$

Likewise, for any family of functions $f_j : X_j \to Y_j$, we define $f_{-i} : X_{-i} \to X_{-i}$ by $f_{-i}(x_{-i}) = (f_j(x_j))_{j \neq i}$. Given any metric space (X, d), we write $\Delta(X)$ for the space of probability distributions on X, endowed with Borel σ-algebra and the weak topology.

[2]We use the strict concavity assumption to make sure that a player's utility function for any fixed strategy profile of the others is always single-peaked in his own action. (Single-peakedness is not preserved in presence of uncertainty.)

We will consider singleton selections from Bayesian Nash equilibria of models, picking a Bayesian Nash equilibrium for each model, such that when we put all these models together, the resulting strategy profile remains to be a Bayesian Nash equilibrium of the larger game. This can be thought of a consistency requirement of a theory of selection for various games. More precisely, we will fix a Bayesian Nash equilibrium $s^* : T^u \to A$ in T^u and pick the Bayesian Nash equilibrium $s^*_{|T}$ with

$$s^*_{|T}(t) = s^*(h(t)) \qquad (\forall t \in T)$$

as the solution in type space T. (Notice that $s^*_{|T}$ is a Bayesian Nash equilibrium of T.) Multiple equilibria are introduced to our analysis trivially, by considering sets of equilibria s^* on T^u, which does not affect our analysis.

Our formulation also restricts equilibrium action to depend only on the hierarchy of beliefs. That is, if there are two types t_i and t_i' in possibly two different models with identical belief hierarchies (i.e. $h_i(t_i) = h_i(t_i')$), then the equilibrium actions are the same for t_i and t_i'. In particular, in a model with redundant types, all types with identical belief hierarchies play the same action, ruling out the extra equilibria introduced by the redundant types.

Local Interim Rationalizability. We will show that the sensitivity of equilibrium strategies is characterized by local version of interim (correlated) rationalizability (Dekel, Fudenberg, and Morris (2007), Battigalli (2003), Battigalli and Siniscalchi (2003)). Interim rationalizability allows correlations not only within players' strategies but also between their strategies and θ. For any set $B = B_1 \times \cdots \times B_n \subset A$ and any i and t_i, we set

$$S_i^0[B, t_i] = B_i$$

and define sets $S_i^k[B, t_i]$ for $k > 0$ iteratively by

$$S_i^k[B, t_i] = \{BR_i\left(\mathrm{marg}_{\Theta \times A_{-i}}\pi\right) | \pi \in \Delta(\Theta \times T_{-i} \times A_{-i}), \\ \mathrm{marg}_{\Theta \times T_{-i}}\pi = \kappa_{t_i}, \pi\left(a_{-i} \in S_{-i}^{k-1}[B, t_{-i}]\right) = 1\}.$$

The set $S_i^k[B, t_i]$ consists of best replies to beliefs that assign positive probability only to the actions that are in $S_{-i}^{k-1}[B, \cdot]$. As in Dekel, Fudenberg, and Morris (2007), $S_i^k[B, t_i]$ only depends on $h_i^k(t_i)$, not the particular type space or any higher-order beliefs. That is,

$$S_i^k[B, t_i] = S_i^k\left[B, \tilde{t}_i\right] \text{ whenever } h_i^m(t_i) = h_i^m\left(\tilde{t}_i\right) \text{ for all } m \le k. \tag{2.1}$$

Unlike in the usual rationalizability, the sets can become larger as k increases. Hence, we define the limit set, which consists of the locally interim correlated rationalizable actions with respect to B, by

$$S_i^{\infty}[B,t_i] = \bigcap_{k=0}^{\infty} \bigcup_{m=k}^{\infty} S_i^m[B,t_i].$$

When $B = A$, we simply write $S_i^k[t_i]$ for $S_i^k[B,t_i]$, and $S_i^{\infty}[t_i]$ is the set of interim correlated rationalizable actions for t_i.

The following result is an extension of earlier results by Moulin (1984) and Battigalli (2003) to local interim correlated rationalizability.

Lemma 1. *For any i, t_i, k, and any $a_i \in S_i^k[B,t_i]$,*

$$a_i = BR_i\left(\pi\left(\cdot|t_i,\hat{s}_{-i}\right)\right)$$

for some measurable mapping $\hat{s}_{-i} : \Theta \times T_{-i} \to A_{-i}$ with $\hat{s}_{-i}(\theta,t_{-i}) \in S_{-i}^{k-1}[B,t_{-i}]$.

That is, in a nice game, every rationalizable action is a best reply to a deterministic theory about how the other players' actions are related to their types *and* the underlying parameter. Here, the action of another player j may vary with θ or a third player's type because we allow all possible correlations.

3. Sensitivity to higher-order beliefs

In this section, we will introduce a straightforward measure of sensitivity of a strategy to higher-order beliefs and present our general result, which gives a characterization of sensitivity in terms of local interim correlated rationalizability.

Fix any strategy $s_i^* : T_i^u \to A_i$ on T_i^u and any type t_i of a player i. According to strategy s_i^*, type t_i will play $s_i^*(h_i(t_i))$. Now imagine a researcher who only knows the first k orders of beliefs of player i and knows that he plays s_i^*. All the researcher can conclude from this information is that i will play one of the actions in

$$A_i^k[s_i^*,t_i] \equiv \left\{ s_i^*\left(h_i\left(\tilde{t}_i\right)\right) | h_i\left(\tilde{t}_i\right) \in T_i^u, \quad h_i^m\left(\tilde{t}_i\right) = h_i^m(t_i) \quad \forall m \leq k \right\}.$$

That is, an action is in $A_i^k[s^*,t_i]$ if and only if it is played according to s_i^* by a type $\tilde{t}_i$ that comes from a countable model and whose first k order beliefs are as in t_i. Therefore, $A_i^k[s_i^*,t_i]$ measures precisely how sensitive the strategy s_i^* is to the specification of beliefs at orders higher than k when the first k orders of beliefs are as specified by t_i. Assuming, plausibly, that a researcher can verify only finitely many orders of a player's beliefs, all a researcher can ever know is that player i will play one of the actions in

$$A_i^{\infty}[s_i^*, t_i] = \bigcap_{k=0}^{\infty} A_i^k[s_i^*, t_i].$$

If the researcher knew only that strategy of i is in a given set S_i, rather than knowing what his strategy is, then he could conclude from his information only that i will play an action in

$$A_i^k[S_i, t_i] = \bigcup_{s_i^* \in S_i} A_i^k[s_i^*, t_i].$$

The next result, which is the main general result of this paper, characterizes the sets $A_i^k[s_i^*, t_i]$ for equilibrium strategies s_i^* by local rationalizability.

Proposition 1. *For any equilibrium s^* and any (i, k, t_i),*

$$A_i^k[s_i^*, t_i] = S_i^k[s^*(T^u), t_i].$$

In particular, when $s^*(T^u) = A$, $A_i^k[s_i^*, t_i] = S_i^k[t_i]$. Also, for any $B \subseteq s^*(T)$, $\cup_{m \geq k} S_i^m\left[B, \hat{t}_i\right] \subseteq A_i^k[s_i^*, t_i]$.

Proposition 1 tells us a way of determining how sensitive an arbitrary equilibrium s^* is to the specifications of beliefs at orders higher than k: Consider the set of *all* actions that are played by some type according to s^*, without requiring any connection to the beliefs at hand. Apply the best response operator to this set k times, allowing all possible correlations. The resulting set is precisely the set of actions that could be played by types whose first k orders of beliefs are as specified at the beginning. When the parameter space is rich enough so that all actions are played by some types, this set is simply the set of actions that survive kth-order elimination of strictly dominated actions in the interim stage. When we allow k to be arbitrarily high, this set is simply the set of all (locally) interim correlated rationalizable actions. Notice that the sets $A_i^k[s_i^*, t_i]$ are decreasing in k, i.e. $A_i^k[s_i^*, t_i] \supseteq A_i^{k+1}[s_i^*, t_i]$. The equality in the proposition then implies that the sets $S_i^k[s^*(T^u), t_i]$ are also decreasing in k. When we start with the range $s^*(T^u)$ of an equilibrium, the above process is indeed an elimination process:

Corollary 1. *For any equilibrium* s^* *and any* (i, k, t_i)*,* $S_i^k[s^*(T^u), t_i] \supseteq S_i^{k+1}[s^*(T^u), t_i]$.

Sometimes, it may be difficult to know the set of actions played by arbitrary types according to s^*, but we may still know the behavior of certain types, e.g. the common knowledge types. In that case, we can still use Proposition 1 to find a lower bound: consider the set of actions that are known to be played by some type and apply the best response correspondence k times. In that case, $S_i^k[B, t_i]$ may not be decreasing in k, and one can find a better lower bounds by iterating the procedure further. Since $S_i^m[B, t_i] \subseteq A_i^m[s_i^*, t_i] \subseteq A_i^k[s_i^*, t_i]$ for each $m \geq k$, we have $\cup_{m \geq k} S_i^m\left[B, \hat{t}_i\right] \subseteq A_i^k[s_i^*, t_i]$.

A comparison of this result with that of WY is useful. In WY, we consider a finite action game and assume that the parameter space is so rich that every action becomes dominant at some parameter value. Then, we show that for each $a_i \in S_i^k[t_i]$ and each rationalizable strategy s_i, we can perturb first k order beliefs arbitrarily slightly and change the higher-order beliefs to obtain a type $\tilde{t}_i$ such that $s_i\left(\tilde{t}_i\right) = a_i$. Here, we consider nice games instead of finite-action games. At the expense of focusing on Bayesian Nash equilibria, rather than arbitrary rationalizable strategies, we strengthen the result in two ways. First, we do not make *any* richness assumption, allowing arbitrary common knowledge restriction on payoffs. Instead, we give a general characterization, $A_i^k[s_i^*, t_i] = S_i^k[s^*(T^u), t_i]$, that depends on the range of equilibrium on T^u. Second, since the best reply is always unique, we do not perturb the lower-order beliefs at all, and hence our result does not directly refer to any topology on beliefs.

We will now give the proof for $k = 1$. Our general proof, which is in the appendix, uses the same arguments inductively. The inclusion $A_i^1[s_i^*, t_i] \subseteq S_i^1[s^*(T^u), t_i]$ immediately follows from the definitions and (2.1). Indeed, for any $a_i \in A_i^1[s_i^*, t_i]$, we have $a_i = s_i^*\left(\tilde{t}_i\right)$ for some $\tilde{t}_i$ with $h_i^1\left(\tilde{t}_i\right) = h_i^1(t_i)$, implying also $a_i = BR_i\left(\pi\left(\cdot|\tilde{t}_i, s_{-i}^*\right)\right)$. Then, $a_i \in S_i^1\left[s^*(T^u), \tilde{t}_i\right] = S_i^1[s^*(T^u), t_i]$, where the last equality is by (2.1).

To show the inclusion $S_i^1[s^*(T^u), t_i] \subseteq A_i^1[s_i^*, t_i]$, take any type t_i, from a countable type space $\Theta \times T$, and any $a_i \in S_i^1[s^*(T^u), t_i]$. We need to construct a new type $\tilde{t}_i$, from a countable type space $\tilde{\Theta} \times \tilde{T}$, such that

(1) $h_i^1\left(\tilde{t}_i\right) = h_i^1(t_i)$, i.e., $\sum_{\tilde{t}_{-i} \in \tilde{T}_{-i}} \kappa_{\tilde{t}_i}\left(\theta, \tilde{t}_{-i}\right) = \sum_{t_{-i} \in T_{-i}} \kappa_{t_i}(\theta, t_{-i})$ for each $\theta \in \Theta$, and
(2) $s_i^*\left(\tilde{t}_i\right) = a_i$, i.e., $a_i = BR_i\left(\pi\left(\cdot|\tilde{t}_i, s_{-i}^*\right)\right)$.

By Lemma 1, $a_i = BR_i\left(\pi\left(\cdot|t_i, s_{-i}\right)\right)$ for some function $s_{-i} : \Theta \times T_{-i} \to A_{-i}$ with $s_{-i}\left(\theta, t_{-i}\right) \in s^*_{-i}\left(T^u_{-i}\right)$. By definition, for each $a_{-i} \in s_{-i}\left(\Theta \times T_{-i}\right)$, which is contained in $s^*_{-i}\left(T^u_{-i}\right)$, there exists $t_{-i}\left[a_{-i}\right] \in T^u_{-i}$ such that $s^*_{-i}\left(t_{-i}\left[a_{-i}\right]\right) = a_{-i}$. We will define our type $\tilde{t}_i$ by beliefs

$$\kappa_{\tilde{t}_i}\left(\theta, t_{-i}\left[a_{-i}\right]\right) = \pi\left(\theta, a_{-i}|t_i, s_{-i}\right) \qquad \left(\forall\theta \in \Theta, a_{-i} \in s_{-i}\left(\Theta \times T_{-i}\right)\right).$$

That is, we assign the probability of an action under $\pi\left(\cdot|t_i, s_{-i}\right)$ to a type who plays that action in equilibrium, while we keep the probabilities of θ intact. It is then straightforward to check that the two conditions above are satisfied. First,

$$\sum_{t_{-i}\left[a_{-i}\right]} \kappa_{\tilde{t}_i}\left(\theta, t_{-i}\left[a_{-i}\right]\right) = \sum_{a_{-i}} \pi\left(\theta, a_{-i}|t_i, s_{-i}\right) = \sum_{t_{-i}\in T_{-i}} \kappa_{t_i}\left(\theta, t_{-i}\right)$$

because $\pi\left(\cdot|t_i, s_{-i}\right)$ is induced by type t_i (and s_{-i}). The second condition is satisfied (i.e. $a_i = BR_i\left(\pi\left(\cdot|\tilde{t}_i, s^*_{-i}\right)\right)$) because, for each $\left(\theta, a_{-i}\right)$,

$$\begin{aligned}\pi\left(\theta, a_{-i}|\tilde{t}_i, s^*_{-i}\right) &= \sum_{s^*_{-i}\left(t_{-i}\right)=a_{-i}} \kappa_{\tilde{t}_i}\left(\theta, t_{-i}\right) = \kappa_{\tilde{t}_i}\left(\theta, t_{-i}\left[a_{-i}\right]\right) \\ &= \pi\left(\theta, a_{-i}|t_i, s_{-i}\right)\end{aligned}$$

and $a_i = BR_i\left(\pi\left(\cdot|t_i, s_{-i}\right)\right)$. (It is also clear that $\tilde{t}_i$ comes from a countable space.[3])

It is crucial for the proof and the result that s^* is an equilibrium. Since s^* is an equilibrium, each type plays a best response to the same strategy profile s^*_{-i} of the other players. This puts a strong restriction on the actions played by different types. Indeed, we conclude that $\tilde{t}_i$ plays a_i by assuming that t_i plays a best reply to s^*_{-i}. On the other hand, in a rationalizable strategy, each type's action may be a best response to different rationalizable strategies of the other players. In that case, we could not make such assumptions, and the result need not be true.

Lemma 1 also plays a crucial role in the proof. In order for the belief $\kappa_{\tilde{t}_i}$ to be well-defined, the mapping $a_{-i} \mapsto t_{-i}\left[a_{-i}\right]$ needs to be measurable on the set of actions a_{-i} type t_i assigns positive probability when he plays a_i as a best reply. Lemma 1 allows us to focus on the "degenerate beliefs" for which that set, which is contained in $s_{-i}\left(\Theta \times T_{-i}\right)$, is countable, and therefore the mapping $a_{-i} \mapsto t_{-i}\left[a_{-i}\right]$

[3]For each a_{-i}, we have $t_{-i}\left[a_{-i}\right] \in T^{a_{-i}}_{-i}$ for some countable model $\Theta^{a_{-i}} \times T^{a_{-i}}$. We define the countable model $\tilde{\Theta} \times \tilde{T}$ by $\tilde{\Theta} = \left(\cup_{a_{-i}}\Theta^{a_{-i}}\right) \cup \Theta$, $\tilde{T}_i = \left(\cup_{a_{-i}} T^{a_{-i}}_i\right) \cup \left\{\tilde{t}_i\right\}$, and $\tilde{T}_j = \cup_{a_{-i}} T^{a_{-i}}_j$ for all other players j. The belief of $\tilde{t}_i$ is defined as above, and the beliefs of all the other types are kept as in the original model each comes from.

is trivially measurable on that domain. Since we have uncountably many actions, without Lemma 1, we would need to consider beliefs that put positive probability on an uncountable set of actions. On such a domain, the mapping $a_{-i} \mapsto t_{-i}[a_{-i}]$ need not be measurable.

Proposition 1 is especially useful in determining the robust predictions of complete information models. Instead of assuming that the payoffs are common knowledge, assume that the payoffs are mutually known up to a finite order k, for arbitrarily high k, and that it is common knowledge that the payoffs are within an arbitrarily small neighborhood of the true parameter. Such a relaxation of the complete-information assumption is very weak; it is weaker than many robustness criteria, such as perturbations in the uniform topology. Proposition 1 establishes that in order to check robustness to such mild perturbations, we need to consider all strategies that survive local interim correlated rationalizability for an arbitrarily small neighborhood of the equilibrium. In some important games, this method can allow us to prove that all non-trivial predictions of complete information games will fail to be robust to such mild perturbations. In the next section, we will establish this for Cournot oligopoly.

4. Application: Cournot Oligopoly

In this section, we will show that in Cournot oligopoly with sufficiently firms, even a slight relaxation of the common-knowledge assumption will preclude us from making any prediction beyond the elementary fact that no firm will produce more than the monopoly outcome.

In a Cournot oligopoly with sufficiently many firms, any production level that is less than or equal to the monopoly production is rationalizable (Bernheim (1984), Basu (1992)). We will show that this result extends to local interim correlated rationalizability when the equilibrium of the complete information game is in the interior of B. Then, our Proposition 1 implies that a researcher cannot rule out any such output level as the *equilibrium* output for a firm no matter how many orders of beliefs he specifies, even if he assumes that it is common knowledge that the payoffs are in an arbitrarily small neighborhood of the true value and subscribes to strong refinements of equilibrium that yield unique solutions. Even a slight doubt about the payoffs in very high orders will lead a researcher to fail to rule out any outcome that is less than the monopoly outcome as a firm's equilibrium output.

On the other hand, Borgers and Janssen (1995) show that if we replicate both consumers and the firms in such a way that the cobweb dynamics is stable for the resulting demand and supply curves, then the Cournot oligopoly will be dominance-solvable. In that case, by Proposition 1, equilibrium outcomes will not be sensitive to higher-order beliefs.

Consider n firms with identical constant marginal cost $c > 0$. Simultaneously, each firm i produces q_i at cost $q_i c$ and sell its output at price $P(Q;\theta)$ where $Q = \sum_i q_i$ is the total supply. For some fixed $\bar{\theta}$, we assume that Θ is a closed interval with $\bar{\theta} \in \Theta \neq \{\bar{\theta}\}$. We also assume that $P(0;\bar{\theta}) > 0$, $P(\cdot;\bar{\theta})$ is strictly decreasing when it is positive, and $\lim_{Q\to\infty} P(Q;\bar{\theta}) = 0$. Therefore, there exists a unique $\hat{Q}$ such that $P\left(\hat{Q};\bar{\theta}\right) = c$. We assume that, on $[0,\hat{Q}]$, $P(\cdot;\bar{\theta})$ is continuously twice-differentiable and $P' + QP'' < 0$.

It is well known that, under the assumptions of the model, (i) the profit function, $u(q,Q;\bar{\theta}) = q(P(q+Q) - c)$, is strictly concave in own output q; (ii) the unique best response $q^*(Q_{-i})$ to others' aggregate production Q_{-i} is strictly decreasing on $[0,\hat{Q}]$ with slope bounded away from 0 (i.e., $\partial q^*/\partial Q_{-i} \leq \lambda$ for some $\lambda < 0$); (iii) equilibrium outcome at $t^{CK}(\bar{\theta})$, $s^*(t^{CK}(\bar{\theta}))$, is unique and symmetric (Okuguchi and Suzumura (1971)). We will also assume that θ is a payoff-relevant parameter in the following sense: $q^*(Q_{-i};\theta)$ is a continuous and strictly increasing function of θ at $(Q_{-i},\bar{\theta})$ where $Q_{-i} = (n-1)\, s_j^*(t^{CK}(\bar{\theta}))$.

Lemma 2. *In the Cournot oligopoly above, there exists $\bar{n} < \infty$ such that for any $n > \bar{n}$ and any $B = \left[s_1^*\left(t_1^{CK}(\bar{\theta})\right) - \epsilon, s_1^*\left(t_1^{CK}(\bar{\theta})\right) + \epsilon\right]^n \subset A$ with $\epsilon > 0$, we have*

$$S_i^\infty\left[B; t^{CK}(\bar{\theta})\right] = \left[0, q^M\right] \quad (\forall i \in N),$$

where q^M is the monopoly output under $P(\cdot;\bar{\theta})$ and $s^*(t^{CK}(\bar{\theta}))$ is the unique equilibrium of the complete information game $\{t^{CK}(\bar{\theta})\}$.

This is a straightforward extension of a result by Basu (1992) for rationalizability to local rationalizability. The proof is in the appendix. Together with Proposition 1, this lemma yields the following.

Proposition 2. *In the Cournot oligopoly above, let* $\Theta^* = \left[\bar{\theta} - \varepsilon, \bar{\theta} + \varepsilon\right]$ *for arbitrarily small* $\varepsilon > 0$*. Then, for any equilibrium* s^* *on* T^u*,*

$$A_i^\infty \left[s^*, t_i^{CK}\left(\bar{\theta}\right)\right] = \left[0, q^M\right] \quad (\forall i \in N),$$

where q^M *is the monopoly output under* $P\left(\cdot; \bar{\theta}\right)$*.*

Proof. Since we can put a large upper bound on q, by (i) above, we have a nice game. By the hypothesis, there exists $B \subset s^*(T)$ as in Lemma 2. Hence, Lemma 2 and Proposition 1 imply

$$\left[0, q^M\right] = S_i^\infty \left[B; t_i^{CK}\left(\bar{\theta}\right)\right] \subseteq A_i^\infty \left[s^*, t_i^{CK}\left(\bar{\theta}\right)\right] \subseteq \left[0, q^M\right],$$

yielding the desired equality. □

Our proposition suggests that, with sufficiently many firms, any equilibrium prediction that is not implied by strict dominance will be invalid whenever we slightly deviate from the idealized complete information model. To see this, consider two researchers. One is confident that it is common knowledge that $\theta = \bar{\theta}$. The other is slightly skeptical: he is only willing to concede that it is common knowledge that $\left|\theta - \bar{\theta}\right| \leq \varepsilon$ and agrees with the kth-order mutual knowledge of $\theta = \bar{\theta}$. He is an arbitrarily generous skeptic; he is willing to concede the above for arbitrarily small $\varepsilon > 0$ and arbitrarily large finite k. Our proposition states that the skeptic nonetheless cannot rule out any output level that is not strictly dominated.

Appendix A. Proofs

A.1. **Proof of Lemma 1.** One can easily show that every $\pi \in \Delta(\Theta \times T_{-i} \times A_{-i})$ with $\text{marg}_{\Theta \times T_{-i}} \pi = \kappa_{t_i}$ and $\pi\left(a_{-i} \in S_{-i}^{k-1}[B, t_{-i}]\right) = 1$ is induced by type t_i and a mixed belief $\sigma_{-i} \in \Delta\left(\hat{S}_{-i}^{k-1}[B]\right)$ where $\hat{S}_{-i}^{k-1}[B]$ is the set of all measurable functions $\hat{s}_{-i} : \Theta \times T_{-i} \to A_{-i}$ with $\hat{s}_{-i}(\theta, t_{-i}) \in S_{-i}^{k-1}[B, t_{-i}]$ for each (θ, t_{-i}). We are ready to prove the following result, which immediately implies Lemma 1.

Lemma 3. *For any nice game and for any* i, t_i, k, *the following are true.*

(1) $S_i^k[B, t_i] = \left[\underline{a}^k, \bar{a}^k\right]$ *for some* $\underline{a}_i^k, \bar{a}_i^k \in A_i$*, which depend on* t_i*.*
(2) *For each* $a_i^k \in S_i^k[B, t_i]$*, there exists* $\hat{s}_{-i} \in \hat{S}_{-i}^{k-1}[B]$ *such that*

$$BR_i\left(\pi\left(\cdot | t_i, \hat{s}_{-i}\right)\right) = a_i^k.$$

Proof. We will use induction on k. For $k = 0$, part 1 is true by definition. Assume that part 1 is true for some $k-1$, i.e., $S_j^{k-1}[t_j]$ is a closed interval in $A_j = [0,1]$ for each j. This implies that $\hat{S}_{-i}^{k-1}$ is a closed, convex metric space (with product topology).[4] Moreover, by the Maximum Theorem, $\beta_i(\cdot;t_i)$ that maps each $s_{-i} \in \hat{S}_{-i}^{k-1}$ to $BR_i(\pi(\cdot|t_i, s_{-i}))$ has a closed-graph and hence is continuous. (It is a function.) Since $\hat{S}_{-i}^{k-1}$ is compact and convex, this implies that $\beta_i\left(\hat{S}_{-i}^{k-1};t_i\right)$ is compact and connected, and hence it is convex as it is unidimensional. That is, $\beta_i\left(\hat{S}_{-i}^{k-1};t_i\right) = \left[\underline{a}^k, \bar{a}^k\right]$ for some $\underline{a}_i^k, \bar{a}_i^k \in A_i$. We **claim** that $\beta_i\left(\hat{S}_{-i}^{k-1};t_i\right) = S_i^k[t_i]$. This readily proves part 1. Part 2 follows from the definition of $\beta_i\left(\hat{S}_{-i}^{k-1};t_i\right)$.

Towards proving our **claim**, for each $(\theta, t_{-i}) \in \text{supp}\kappa_{t_i}$ and for each $s_{-i} \in \hat{S}_{-i}^{k-1}$, define function $U_i(\cdot|\theta, t_{-i}, s_{-i})$ by setting $U_i(a_i|\theta, t_{-i}, s_{-i}) = u_i(\theta, a_i, s_{-i}(\theta, t_{-i}))$ at each a_i. Clearly, U_i is strictly concave, and for each $\sigma_{-i} \in \Delta\left(\hat{S}_{-i}^{k-1}\right)$, the expected payoff of type t_i is

$$\int U_i(a_i|\theta, t_{-i}, s_{-i})\, d\kappa_{t_i}(\theta, t_{-i})\, d\sigma_{-i}(s_{-i}). \tag{A.1}$$

Now, take any $a_i > \bar{a}_i^k$. Then, for each (θ, t_{-i}, s_{-i}), by definition of $\bar{a}_i^k$ and strict concavity of $U_i(\cdot|\theta, t_{-i}, s_{-i})$, we have $U_i(a_i|\theta, t_{-i}, s_{-i}) < U_i\left(\bar{a}_i^k|\theta, t_{-i}, s_{-i}\right)$. It then follows from (A.1) that $\bar{a}_i^k$ yields higher expected payoff than a_i for each $\sigma_{-i} \in \Delta\left(\hat{S}_{-i}^{k-1}\right)$, and thus $a_i \notin S_i[t_i]$. Similarly, $a_i \notin S_i[t_i]$ for each $a_i < \underline{a}_i^k$. □

A.2. **Proof of Proposition 1.**

Proof. We proceed inductively on k, showing first $S_i^k[s^*(T^u), t_i] \subseteq A_i^k[s^*, t_i]$. For $k = 0$, both sides are equal to $s^*(T^u)$. For any given k and any player i, write each $h_{-i}(t_{-i})$ as $h_{-i}(t_{-i}) = (\lambda, \eta)$ where $\lambda = \left(h_{-i}^1(t_{-i}), h_{-i}^2(t_{-i}), \ldots, h_{-i}^{k-1}(t_{-i})\right)$ and $\eta = \left(h_{-i}^k(t_{-i}), h_{-i}^{k+1}(t_{-i}), \ldots\right)$ are the lower and higher-order beliefs, respectively. Let $L = \left\{\lambda | \exists \eta : (\lambda, \eta) \in T_{-i}^u\right\}$. The induction hypothesis is that

$$S_{-i}^{k-1}[s^*(T^u), \lambda] \equiv \bigcup_{\eta'} S_{-i}^{k-1}\left[s^*(T^u), (\lambda, \eta')\right] \subseteq A_{-i}^{k-1}[s^*, (\lambda, \eta)] \quad (\forall (\lambda, \eta) \in T_{-i}^u).$$

[4]Proof: Firstly, $\prod_{(\theta, t_j, j)} S_j^{k-1}[B, t_j]$ is a compact space by Tychonoff's theorem. But the space of all measurable functions $f : \Theta \times T_{-i} \to \mathbb{R}^{N\setminus\{i\}}$ is closed. Hence, the intersection of these two spaces, namely $\hat{S}_{-i}^{k-1}$, is compact. Convexity of $\hat{S}_{-i}^{k-1}$ follows from the facts that measurability is preserved under point-wise multiplication and addition and that the range is convex.

Fix any type t_i and any $a_i \in S_i^k[s^*(T^u), t_i]$. We will construct a type $\tilde{t}_i$ such that $s_i^*(\tilde{t}_i) = a_i$ and the first k orders of beliefs are same under t_i and $\tilde{t}_i$, showing that $a_i \in A_i^k[s^*, t_i]$. Now, by Lemma 1 , $a_i = BR_i(\pi(\cdot|t_i, \hat{s}_{-i}))$ for some $\hat{s}_{-i} : \Theta \times T_{-i} \to A_{-i}$ with $\hat{s}_{-i}(\theta, t_{-i}) \in S_{-i}^{k-1}[s^*(T^u), t_{-i}]$. By the induction hypothesis, for each a_{-i} in the image of $\hat{s}_{-i}$, $a_{-i} \in S_{-i}^{k-1}[s^*(T^u), \lambda] \subseteq A_{-i}^{k-1}[s^*, (\lambda, \eta)]$ for some η. Hence, there exists a mapping $\mu : \hat{s}_{-i}(\Theta \times T_{-i}) \to \Theta \times T_{-i}^u$,

$$\mu : (\theta, \lambda, a_{-i}) \mapsto (\theta, \lambda, \tilde{\eta}(a_{-i}, \theta, \lambda)), \tag{A.2}$$

such that

$$s_{-i}^*(\lambda, \tilde{\eta}(a_{-i}, \theta, \lambda)) = a_{-i}. \tag{A.3}$$

We define $\tilde{t}_i$ by

$$\kappa_{\tilde{t}_i} \equiv \left(\text{marg}_{\Theta \times L \times A_{-i}} \pi\right) \circ \mu^{-1},$$

the probability distribution induced on $\Theta \times T_{-i}^u$ by the mapping μ and the probability distribution π. Notice that, since $h_i^k(t_i)$ has countable support and the action spaces are countable, the set $\text{supp}\left(\text{marg}_{\Theta \times L \times A_{-i}} \pi\right)$ is countable, in which case μ is trivially measurable. Hence $\kappa_{\tilde{t}_i}$ is well-defined. By a well-known isomorphism by Mertens and Zamir (1985), $\kappa_{\tilde{t}_i}$ is the belief of a type $\tilde{t}_i$, such that

$$h_i^m(\tilde{t}_i) = \delta_{h_i^{m-1}(\tilde{t}_i)} \times \text{marg}_{\Theta \times [\Delta(X_{m-2})]^{N \setminus \{i\}}} \kappa_{\tilde{t}_i}, \qquad (\forall m > 1) \tag{A.4}$$

and $h_i^1(\tilde{t}_i) = \text{marg}_\Theta \kappa_{\tilde{t}_i}$. Since $\text{supp}(\kappa_{\tilde{t}_i})$ is countable, $h_i(\tilde{t}_i) \in T_i^u$ (as in Footnote 3). By construction of μ, the first k orders of beliefs (about (θ, λ)) are identical under t_i and $\tilde{t}_i$:

$$\begin{aligned} \text{marg}_{\Theta \times L} \kappa_{\tilde{t}_i} &= \text{marg}_{\Theta \times L}\left[\left(\text{marg}_{\Theta \times L \times A_{-i}} \pi\right) \circ \mu^{-1}\right] = \text{marg}_{\Theta \times L}\left(\text{marg}_{\Theta \times L \times A_{-i}} \pi\right) \\ &= \text{marg}_{\Theta \times L} \pi = \text{marg}_{\Theta \times L}\left(\text{marg}_{\Theta \times T_{-i}^u} \pi\right) = \text{marg}_{\Theta \times L} \kappa_{t_i}, \end{aligned}$$

where the second equality is by (A.2). Together with (A.4) and identical equality for t_i, this shows that $h_i^m(\tilde{t}_i) = h_i^m(t_i)$ for each $m \leq k$. Towards showing that $s_i^*(\tilde{t}_i) = a_i$, let $\tilde{\pi} = \kappa_{\tilde{t}_i} \circ \gamma^{-1} \in \Delta\left(\Theta \times T_{-i}^u \times A_{-i}\right)$ be the equilibrium belief of type $\tilde{t}_i$, where $\gamma : (\theta, \lambda, \eta) \mapsto \left(\theta, \lambda, s_{-i}^*(\lambda, \eta)\right)$. By construction,

$$\begin{aligned} \text{marg}_{\Theta \times L \times A_{-i}} \tilde{\pi} &= \kappa_{\tilde{t}_i} \circ \gamma^{-1} \circ \text{proj}_{\Theta \times L \times A_{-i}}^{-1} \\ &= \left(\text{marg}_{\Theta \times L \times A_{-i}} \pi\right) \circ \mu^{-1} \circ \gamma^{-1} \circ \text{proj}_{\Theta \times L \times A_{-i}}^{-1} = \text{marg}_{\Theta \times L \times A_{-i}} \pi. \end{aligned}$$

[By (A.3) and the definition of γ, $\text{proj}_{\Theta \times L \times A_{-i}} \circ \gamma \circ \mu$ is the identity mapping, yielding the last equality.] Therefore,

$$\pi\left(\cdot|\tilde{t}_i, s_{-i}^*\right) = \text{marg}_{\Theta \times A_{-i}} \tilde{\pi} = \text{marg}_{\Theta \times A_{-i}} \pi.$$

Since a_i is the only best reply to these beliefs, $\tilde{t}_i$ must play a_i in equilibrium s^*:

$$s_i^*\left(\tilde{t}_i\right) \in BR_i\left(\pi\left(\cdot|\tilde{t}_i, s_{-i}^*\right)\right) = BR_i\left(\mathrm{marg}_{\Theta\times A_{-i}}\pi\right) = \{a_i\}\,. \tag{A.5}$$

To see the inclusion $A_i^k[s^*, t_i] \subseteq S_i^k[s^*(T^u), t_i]$, observe that for any $\tilde{t}_i$ with $h_i^m\left(\tilde{t}_i\right) = h_i^m(t_i)$ for each $m \leq k$, we have

$$s_i^*\left(\tilde{t}_i\right) \in S_i^k\left[s^*(T^u), \tilde{t}_i\right] = S_i^k\left[s^*(T^u), t_i\right],$$

where the last equality is by (2.1). □

Proof of Lemma 2. Let $\bar{n}$ be any integer greater than $1 + 1/|\lambda|$, where λ is as in (ii). Take any $n > \bar{n}$. By (iii), $B = [\underline{q}^0, \bar{q}^0]^n$ for some $\underline{q}^0, \bar{q}^0$ with $\underline{q}^0 < \bar{q}^0$. By (ii), for any $k > 0$, $S^k\left[B; t^{CK}\left(\bar{\theta}\right)\right] = [\underline{q}^k, \bar{q}^k]^n$, where

$$\bar{q}^k = q^*\left((n-1)\,\underline{q}^{k-1}\right) \text{ and } \underline{q}^k = q^*\left((n-1)\,\bar{q}^{k-1}\right).$$

Define $\underline{Q}^k \equiv (n-1)\,\underline{q}^k$, $\bar{Q}^k \equiv (n-1)\,\bar{q}^k$, and $Q^* = (n-1)\,q^*$, so that

$$\bar{Q}^k = Q^*\left(\underline{Q}^{k-1}\right) \text{ and } \underline{Q}^k = Q^*\left(\bar{Q}^{k-1}\right).$$

Since $(n-1)\,\lambda < 1$, the slope of Q^* is strictly less than -1. Hence $\underline{Q}^k$ decreases with k and becomes 0 at some finite $\bar{k}$, and $\bar{Q}^k$ increases with k and takes value $Q^*(0) = (n-1)\,q^M$ at $\bar{k}+1$. That is, $S^k\left[B; t^{CK}\left(\bar{\theta}\right)\right] = \left[0, q^M\right]^n$ for each $k > \bar{k}$. Therefore, $S^\infty\left[B; t^{CK}\left(\bar{\theta}\right)\right] = \left[0, q^M\right]^n$. □

References

[1] Basu, K. (1992): "A characterization of the class of rationalizable equilibria of oligopoly games," *Economics Letters* 40, 187-191.

[2] Battigalli, P. (2003): "Rationalizability in infinite, dynamic games with complete information," *Research in Economics*, 57, 1-38.

[3] Battigalli, P. and M. Siniscalchi (2002): "Strong Belief and Forward-Induction Reasoning," *Journal of Economic Theory* 106, 356-391.

[4] Battigalli, P. and M. Siniscalchi (2003) "Rationalization and Incomplete Information," *Advances in Theoretical Economics* 3-1, Article 3.

[5] Börgers T. (1994): "Weak Dominance and Approximate Common Knowledge," *Journal of Economic Theory*, 64, 265-276.

[6] Börgers T. and M. Janssen (1995): "On the dominance solvability of large Cournot games," *Games and Economic Behavior* 8, 297-321.

[7] Brandenburger, A. and E. Dekel (1993): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 59, 189-198.

[8] Carlsson, H. and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989-1018.

[9] Dekel, E. and D. Fudenberg (1990): "Rational Behavior with Payoff Uncertainty," *Journal of Economic Theory*, 52, 243-267.

[10] Dekel, E. D. Fudenberg, S. Morris (2006): "Topologies on Types," *Theoretical Economics*, 1, 275-309.

[11] Dekel, E. D. Fudenberg, S. Morris (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15-40.

[12] Harsanyi, J. (1967): "Games with Incomplete Information played by Bayesian Players. Part I: the Basic Model," *Management Science* 14, 159-182.

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

[14] Mertens and Zamir (1985): "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 10, 619-632.

[15] Monderer, D. and D. Samet (1989): "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior*, 1, 170-190.

[16] Morris, S. (2002): "Typical Types," mimeo.

[17] Moulin, H. (1984): "Dominance solvability and Cournot stability," *Mathematical Social Sciences* 7, 83-102.

[18] Okuguchi, K. and Suzumura, K. (1971): "Uniqueness of the Cournot Oligopoly Equilibrium," *Economic Studies Quarterly*, 22, 81-83.

[19] Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior Under 'Almost Common Knowledge'," *The American Economic Review*, Vol. 79, No. 3, 385-391.

[20] Weinstein, J. and M. Yildiz (2004): "Finite-order Implications of Any Equilibrium," MIT Working Paper.

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

Northwestern University and MIT