

# Rationalizable conjectural equilibrium: A framework for robust predictions

IGNACIO ESPONDA

Stern School of Business, New York University

I introduce a new framework to study environments with both structural and strategic uncertainty, different from Harsanyi's (1967–1968) “Bayesian games”, that allows a researcher to test the robustness of Nash predictions while maintaining certain desirable restrictions on players' beliefs. The solution concept applied to this environment is rationalizable conjectural equilibrium (RCE), which integrates both learning from feedback (in the spirit of self-confirming equilibrium) and from introspection (in the spirit of rationalizability). I provide an epistemic definition of RCE and obtain a characterization in terms of a procedure that generalizes iterated deletion of strategies that are not a best response.

**KEYWORDS.** Rationalizability, self-confirming equilibrium, epistemic framework, robust equilibrium predictions.

**JEL CLASSIFICATION.** C70, C72.

## 1. INTRODUCTION

The standard notion of Nash (1951) equilibrium implicitly requires each player to play a strategy that is optimal given a correct belief about both the strategies of other players and any structural (e.g., payoff) uncertainty that may exist. In this paper, I propose a novel approach to studying environments with both structural and strategic uncertainty, different from Harsanyi's (1967–1968) “Bayesian game” framework, that allows a researcher to test the robustness of Nash predictions while maintaining certain desirable restrictions on players' beliefs.

To motivate the need to relax the Nash assumption, consider a Cournot game where firms repeatedly compete by simultaneously choosing quantities in the face of uncertain demand. Suppose that firms never observe each other's quantities, but instead observe the market price that results from their interaction. Then it may be unwarranted to assume that, in a steady state, firms have correct beliefs about one another's quantity. A weaker restriction is that firms have correct beliefs about the resulting market price,

---

Ignacio Esponda: [iesponda@stern.nyu.edu](mailto:iesponda@stern.nyu.edu)

I thank Pierpaolo Battigalli, Adam Brandenburger, Amanda Friedenberg, Bart Lipman, and two anonymous referees for detailed feedback, and Susan Athey, Heski Bar-Isaac, Liran Einav, Yossi Feinberg, Drew Fudenberg, Jon Levin, Stephen Morris, Ilya Segal, Andy Skrzypacz, Federico Weinschelbaum, and Muhamet Yildiz for helpful comments. This paper is a revised version of a chapter of my Stanford Ph.D. dissertation; I thank SIEPR for financial support through the Bradley Graduate and Post-Graduate Fellowship Program.

Copyright © 2013 Ignacio Esponda. Licensed under the [Creative Commons Attribution-NonCommercial License 3.0](https://creativecommons.org/licenses/by-nc/3.0/). Available at <http://econtheory.org>.

DOI: 10.3982/TE928

and then use that information to make inferences about both other firms' quantities and the state of demand.<sup>1</sup>

In addition to allowing for more realistic assumptions, relaxing the Nash assumption provides additional benefits. First, insofar as multiple beliefs may be consistent with the available evidence, players may want to supplement their data with introspection. For example, to pin down her beliefs, a player in the Cournot game may not limit herself to inferences from information on prices, but may make additional inferences if she believes that other players optimally choose their quantities and also make inferences from prices. Thus, weakening the Nash assumption provides an explicit role for strategic thinking that would otherwise be absent. Second, it may be of interest to relax not only the restriction that beliefs over certain equilibrium outcomes must be correct, but also other standard restrictions, such as the restrictions that players have correct beliefs over the number of other players in the game or over the private information possessed by these players.<sup>2</sup>

The proposed framework can be summarized as follows. A simultaneous-move game specifies the objective strategic environment that players are facing, including the set of players, the strategy sets, a set of fundamentals or payoff parameters, and preferences. The description of a game is completed by adding two exogenous elements. The first element is a state space that describes the elements of the game over which players are allowed to be uncertain and over which their (higher-order) beliefs are defined. In this paper, the state space is given by the product of the strategy sets and the set of fundamentals, so that players have both strategic and structural uncertainty.

The second element is a collection of feedback partitions over the state space, one for each player, that capture both *ex ante* (exogenous) and *ex post* (endogenous) restrictions on beliefs. An example of an *ex ante* restriction is the requirement that players know that a particular player knows the objective game (i.e., fundamental) being played. An example of an *ex post* restriction is the requirement that each player has a correct belief about the payoff she obtains given everyone's play. Since everyone's play is determined endogenously (in equilibrium), then the specific restriction on beliefs also is endogenous. *Ex post* restrictions are intended to capture a learning environment where players repeatedly play the game and obtain feedback about the outcome (such as own payoff feedback in the example). The idea of using a feedback correspondence (or a partition) to relax the Nash assumption is borrowed from the conjectural or self-confirming equilibrium literature (Battigalli 1987, Fudenberg and Levine 1993a, Rubinstein and Wolinsky 1994, Dekel et al. 2004).

For a fixed game as described above, an epistemic model (as in Aumann 1987, Brandenburger and Dekel 1987, Tan and da Costa Werlang 1988, Aumann and Brandenburger 1995) is used to formalize player's hierarchies of beliefs over both structural

<sup>1</sup>The feature of this game that is common to many other settings is that players need to disentangle the extent to which their observations are explained by others' actions or by the fundamentals of the economy.

<sup>2</sup>In addition, modeling how information about past outcomes affects equilibrium behavior can provide new research directions on optimal information policy (Esponda 2008b, Jehiel 2011) and bounded rationality (Esponda 2008a).

and strategic uncertainty.<sup>3</sup> The epistemic model formalizes the notion of beliefs about an event, and  $k$ th order and common belief about an event. Two events play a crucial role: the event that players are rational (i.e., strategies maximize utility given beliefs) and the event that beliefs are consistent (i.e., not falsifiable by evidence that results from feedback about actual equilibrium outcomes).

The solution concept applied to this game is rationalizable conjectural equilibrium (RCE), which intends to capture the steady state of a learning process that combines learning and introspection. A strategy profile is a  $k$ -rationalizable conjectural equilibrium ( $k$ -RCE) if there exists an epistemic model where players' strategies maximize their expected utility given their beliefs and their beliefs satisfy two restrictions: (1) consistency with respect to the feedback partition and the equilibrium being played, and (2)  $k$ th order belief that players are rational and have consistent beliefs.<sup>4</sup>

The challenge with the implementation of the above definition is that it requires searching over all epistemic states and verifying which states satisfy the equilibrium conditions. I show, however, that there is a straightforward characterization of equilibrium that dispenses with the notion of an epistemic model. The main characterization is in terms of a generalization of the familiar procedure that iteratively removes strategies that are not a best response. The iterative procedure has a long history as a solution concept or characterization of a solution concept. It dates back at least to Gale (1953) and Luce and Raiffa (1957), with subsequent developments by Moulin (1979), Bernheim (1984), Pearce (1984), Brandenburger and Dekel (1987), and Battigalli and Siniscalchi (2003), among others.<sup>5</sup>

The approach to modeling games with structural and strategic uncertainty developed in this paper differs from the standard Harsanyi (1967–1968) approach of defining a Bayesian game and focusing on a Bayesian Nash equilibrium (BNE) of this game. In the Bayesian Nash framework, the researcher first exogenously specifies players' (higher-order) beliefs over the fundamentals (i.e., a type space) and then applies the notion of a Nash equilibrium to strategies that are defined as type-contingent plans of action. Hence, a BNE endogenously restricts beliefs about others' strategies (to be correct) but places no endogenous restrictions on players' beliefs about fundamentals. In contrast, in this paper, the description of a game specifies a *set* of feasible higher-order beliefs over *both* strategies and fundamentals as well as *joint* restrictions on these beliefs that depend on the endogenous play of the game. In particular, an (epistemic) type space is not part of the description of a game, but rather a convenient tool to express a particular hierarchy of beliefs that support equilibrium play. One implication is that the standard

<sup>3</sup>The distinguishing feature of an epistemic model is that players' beliefs about every element of the game of which they are uncertain, including the strategies of other players, is included in the representation of higher-order beliefs. The idea that hierarchies of beliefs can be formalized via the notion of a type space appears in the seminal work of Harsanyi (1967–1968) and is formalized and refined by Mertens and Zamir (1985), Brandenburger and Dekel (1993), and Heifetz and Samet (1998), among others.

<sup>4</sup>The approach is to provide an epistemic definition of equilibrium and then obtain a characterization, rather than the more standard route of first providing an algorithmic definition of a solution concept and then providing an epistemic characterization.

<sup>5</sup>Molinari (1991) and Gilli (1999) propose a similar iterative procedure to characterize RCE in the case of complete information (see also Battigalli 1999).

equivalence between games of incomplete information and games of asymmetric information no longer holds in this paper.<sup>6</sup> The relationship to BNE is illustrated by an example in [Section 2](#) and further discussed in [Section 6.2](#).

The development of an alternative to BNE is partly motivated by two important results.<sup>7</sup> First, extending a result by [Brandenburger and Dekel \(1987\)](#), [Battigalli and Siniscalchi \(2003\)](#) show that if no restrictions are placed on the hierarchies of beliefs over fundamentals (i.e., the Harsanyi type space), then any rationalizable outcome can be played in a BNE. In particular, any restrictions on beliefs imposed by BNE, beyond common belief of rationality, depend on the choice of particular type spaces. This result suggests the need to better understand how different type spaces ultimately restrict equilibrium beliefs. Second, [Dekel et al. \(2004\)](#) show that quite restrictive assumptions are necessary to justify BNE without a common prior as a steady state of a learning process. The reason is that whatever data players use to learn the strategies of other players, it is unlikely that they will not use that data to update their beliefs about the fundamentals of the game. One implication is that if one interprets equilibrium as a steady state of a learning process, then beliefs about strategies and fundamentals should be determined *jointly*. Given these two results, the door remains open for a framework that provides a transparent and tractable link between exogenous restrictions on players' higher-order beliefs about certain outcomes (e.g., it is common belief that each player has correct beliefs about her own equilibrium payoff) and the consequent *joint* restrictions on equilibrium beliefs about strategies and fundamentals. This paper provides one such framework.

The notion of RCE was introduced by [Rubinstein and Wolinsky \(1994\)](#) for simultaneous-move games *without* structural (payoff) uncertainty. As they emphasized, RCE captures outcomes that are between Nash equilibrium and rationalizability, and it can be viewed as a refinement of either rationalizability or self-confirming equilibrium.<sup>8</sup> The notion of RCE in this paper reduces to [Rubinstein and Wolinsky's \(1994\)](#) notion in games without structural uncertainty. The main contributions of this paper are to propose a novel approach to studying games with *structural* uncertainty, to extend the definition of RCE for such games and to formalize it via an epistemic model, and to provide a general characterization result. As mentioned above and highlighted with examples throughout the paper, the framework provides a tractable way to verify the robustness of Nash predictions to relaxations of standard assumptions.

Two related literatures investigate robustness of Nash predictions within the context of Harsanyi's framework. The first of these literatures investigates robustness to higher-order beliefs (over fundamentals) by studying how BNE predictions change with the choice of different Harsanyi type spaces; see, for example, [Rubinstein \(1989\)](#),

<sup>6</sup>As is well known, a Bayesian game equivalently represents a situation where players have certain higher-order beliefs about the fundamentals (incomplete information) and a situation where players receive private information about a randomly drawn fundamental (asymmetric information). [Battigalli et al. \(2011\)](#) argue that this distinction is also relevant for the notion of rationalizability.

<sup>7</sup>See also the discussion by [Battigalli \(2003, Section 1.2\)](#).

<sup>8</sup>The RCE concept is explored in the context of extensive-form games by [Battigalli and Guaitoli \(1997\)](#), [Dekel et al. \(1999\)](#), [Battigalli \(1999\)](#), and [Fudenberg and Kamada \(2011\)](#).

Carlsson and van Damme (1993), Morris et al. (1995), Morris and Shin (1998), Neeman (2004), Heifetz and Neeman (2006), Feinberg and Skrzypacz (2005), Bergemann and Morris (2005), and Weinstein and Yildiz (2007a, 2007b). A second literature models certain aspects of bounded rationality by maintaining the typical common prior assumption over fundamentals but relaxing the assumption that players must have correct beliefs about other players' strategies; see Eyster and Rabin (2005) and Jehiel and Koessler (2008).<sup>9</sup> The difference in the approach to robustness between these literatures and the current paper stems from the differences between the Bayesian Nash framework and the framework in this paper. As pointed out by Dekel et al. (2004) for the first literature and Fudenberg (2006) and Esponda (2008a) for the second literature, the implications of these one-sided restrictions are sometimes hard to interpret as steady states of reasonable learning processes. In contrast, the approach to robustness in this paper explicitly specifies the aspects of the game, including the equilibrium outcomes, over which players must have correct beliefs.<sup>10</sup>

Finally, a growing literature is dispensing with the specification of a Harsanyi type space and studying solution concepts that do not depend on specifying a particular hierarchy of beliefs over fundamentals (Battigalli 2003, Battigalli and Siniscalchi 2003, 2007, Bergemann and Morris 2005, 2007, Battigalli et al. 2011). The closest paper is Battigalli and Siniscalchi (2003), who consider additional, self-confirming restrictions on (first-order) beliefs. In this paper, I consider self-confirming restrictions that endogenously restrict beliefs at each level of the hierarchy.<sup>11</sup>

In Section 2, I present a motivating example to illustrate the framework in this paper and the difference with Harsanyi's framework. In Sections 3 and 4, I introduce the framework, provide an epistemic definition of equilibrium, and characterize equilibrium for the case of finite strategy and belief spaces. I provide additional examples in Section 5 to illustrate how the framework can be applied to evaluate the robustness of Nash predictions. In Section 6, I discuss the extension to compact and continuous games, and the relationship between RCE and other equilibrium concepts, such as BNE, rationalizability, and self-confirming equilibrium. I conclude in Section 7.

## 2. EXAMPLE: SALES-PITCH GAME

Two sellers simultaneously decide whether to pitch ( $P$ ) their products to a potential client or not to pitch ( $N$ ) and obtain an outside option. Sellers are playing one of two

<sup>9</sup>If nature were viewed as a player, the setup of Jehiel (2005) would allow players to have erroneous perceptions about nature's strategy.

<sup>10</sup>Esponda (2008a) speaks to the second literature by providing a model of bounded rationality based on self-confirming beliefs, but does not allow for higher-order beliefs. In the current paper, I speak to the first literature by allowing for higher-order beliefs, but I do not allow for other aspects of bounded rationality. However, it would be possible to allow for certain types of bounded rationality by relaxing the assumption that a feedback correspondence partitions the state space. All results would go through, except that it would no longer be true that a Nash equilibrium is also an RCE.

<sup>11</sup>Battigalli and Prestipino (forthcoming) provide an epistemic characterization of Battigalli and Siniscalchi's (2003)  $\Delta$  rationalizability. There is also a literature in epistemic game theory that considers restrictions on beliefs that go beyond pure introspection (e.g., Brandenburger et al. 2008, Brandenburger and Friedenberg 2010, and Battigalli and Friedenberg 2012). The type of restrictions that I consider are motivated by the joint consequences of learning and introspection.

	<i>P</i>	<i>N</i>		<i>P</i>	<i>N</i>
<i>P</i>	<i>L, L<sub>a</sub></i>	<i>H, O</i>	<i>P</i>	<i>L, L<sub>b</sub></i>	<i>H, O</i>
<i>N</i>	<i>O<sub>a</sub>, H<sub>a</sub></i>	<i>O<sub>a</sub>, O</i>	<i>N</i>	<i>O<sub>b</sub>, H<sub>b</sub></i>	<i>O<sub>b</sub>, O</i>
	<i>θ<sub>a</sub></i>			<i>θ<sub>b</sub></i>	

FIGURE 1. Payoffs for the sales-pitch game.

possible games,  $\theta_a$  or  $\theta_b$ . In both games, seller 1 gets  $H$  if she is the only one to pitch and gets  $L$  if both players pitch. But seller 1’s outside option differs depending on the game:  $O_a$  in game  $\theta_a$  and  $O_b$  in game  $\theta_b$ . In both games, seller 2 gets an outside option of  $O$ . But seller 2’s payoff from pitching alone is  $H_a$  in game  $\theta_a$  and  $H_b$  in game  $\theta_b$ , while his payoff when both sellers pitch is  $L_a$  and  $L_b$  in each game. Suppose that  $O_a < L < O_b < H$  for seller 1, and that  $L_a < O < H_a$  and  $O < L_b < H_b$  for seller 2. In particular, it is a dominant strategy for seller 1 to pitch in game  $\theta_a$  and for seller 2 to pitch in game  $\theta_b$ . These payoffs are depicted in Figure 1.

2.1 Bayesian Nash equilibrium

In Harsanyi’s framework,  $\theta_a$  and  $\theta_b$  are states chosen by nature and that conveniently represent players’ uncertainty about their own as well as others’ payoffs. The environment is supplemented by specifying a type space, which captures players’ entire hierarchy of beliefs over states. One then proceeds to find a Bayesian Nash equilibrium (BNE) for the fixed type space. In the sales-pitch game, both  $P$  and  $N$  are rationalizable for each player, in the sense of Battigalli and Siniscalchi (2003). It then follows from Battigalli and Siniscalchi (2003) that, by appropriately specifying beliefs over  $\{\theta_a, \theta_b\}$ , any action profile can be induced in a BNE. For example, suppose that one specifies that it is common belief that seller 1 believes the game is  $\theta_a$  and that seller 2 believes the game is  $\theta_b$ , i.e., the players agree to disagree. Then  $(P, P)$  is the unique BNE (because it is dominant for each player to pitch given their beliefs about nature) and is supported by the following higher-order beliefs: It is common belief that player 1 believes  $(P, P, \theta_a)$  and that player 2 believes  $(P, P, \theta_b)$ . While players have no incentives to deviate from  $(P, P)$  given their beliefs, is it reasonable that players will maintain such misperceptions if they repeatedly play  $(P, P)$ ? One cannot answer this question in Harsanyi’s framework. To answer this question, two elements must be added to the environment. First, to determine whether a player has a misperception, one must specify the *objective* game, i.e., is the true game being played  $\theta_a$  or  $\theta_b$ ? Second, one must specify which types of misperceptions are unreasonable.

For example, suppose that the objective game is  $\theta_b$ . In equilibrium, both players pitch, seller 1 gets  $L$ , and seller 2 gets  $L_b$ . While having misperceptions, seller 1 does correctly anticipate her own equilibrium payoff: She incorrectly believes the equilibrium outcome is  $(P, P, \theta_a)$ , but since she would also get  $L$  if that were the case, she does correctly anticipate her payoff. However, seller 1 believes that seller 2 believes  $(P, P, \theta_b)$ , so that seller 1 believes that seller 2 believes that his own payoff is  $L_b$ . But since seller 1 believes  $(P, P, \theta_a)$ , then seller 1 believes seller 2’s payoff to be  $L_a < L_b$ . Therefore, in this



BNE, seller 1 believes that seller 2 does not correctly anticipate his own payoff (or does not optimize).<sup>12</sup>

The framework in this paper allows a researcher to obtain predictions about equilibrium play as a function of transparent restrictions on equilibrium beliefs, which include beliefs about the game being played and beliefs about other players' strategies. For example, suppose that in the game above, a desirable restriction to impose *in equilibrium* is that (i) each player is rational (i.e., her strategy is optimal given her belief) and has a correct belief about her own equilibrium payoff, and (ii) it is common belief that each player is rational and has a correct belief about her own equilibrium payoff. The question is then, "What strategy profiles are optimal given beliefs that satisfy the previous conditions?" The answer is provided by the notion of a rationalizable conjectural equilibrium (RCE).

## 2.2 Rationalizable conjectural equilibrium

The following three variants of the sales-pitch game illustrate the framework introduced in this paper. In the first two examples, the objective game is either  $\theta_a$  or  $\theta_b$ . In the first example, the sellers do not know the objective game; in the second example, seller 1 does not know the objective game but seller 2 knows it. In the third example, the objective game is random and seller 2 knows the realization. In all examples, players are restricted to have correct beliefs about their own equilibrium payoffs. The goal is to find, for each example, the set of RCE's, which is defined in Section 3 to be the set of strategy profiles that can be supported when the above description of the game (including the restrictions on beliefs) is common belief.<sup>13</sup>

**EXAMPLE 1** (Deterministic game; both players uninformed). Let  $\Omega = \{P, N\} \times \{P, N\} \times \{\theta_a, \theta_b\}$  denote the *state space* of the game, over which beliefs are defined. A state  $\omega = (x_1, x_2, \theta)$  represents a strategy profile  $(x_1, x_2)$  and a game  $\theta$ . In addition, the description of the game now includes a partition of  $\Omega$  for each player with the property that two states belong to the same element of the partition if and only if they share the same strategy and yield the same own payoff. For example,  $(P, P, \theta_a)$  and  $(P, P, \theta_b)$  share the property that seller 1 plays  $P$  and obtains payoff  $L$ ; therefore, these two states belong to the same element of seller 1's partition. **Theorem 1** in **Section 4** shows that the set of RCE's can be obtained by the following iterative procedure over players' partitions.

*Step 0.* For each player, keep a state if its associated strategy is a best response to some belief over states that assigns probability 1 to the element of the partition that contains the state. Below, the (degenerate) beliefs that support each element of the partition are underlined and the states for which no such beliefs exist are crossed out:

$$\begin{array}{c} \text{Seller 1} \\ \{(\underline{P, P, \theta_a}), (P, P, \theta_b)\}, (\underline{P, N, \theta_a}), (\underline{P, N, \theta_b})\}, \\ \{(\underline{N, P, \theta_a}), (\underline{N, N, \theta_a})\}, \{(\underline{N, P, \theta_b}), (\underline{N, N, \theta_b})\} \end{array}$$

<sup>12</sup>If the objective game is  $\theta_a$ , it is easy to see that seller 2 misperceives his *own* equilibrium payoff.

<sup>13</sup>*Mutual* rather than common belief of these restrictions suffices for these examples.

$$\begin{aligned} & \text{Seller 2} \\ & \{(\underline{P}, \underline{P}, \underline{\theta_a})\}, \{(\underline{N}, \underline{P}, \underline{\theta_a})\}, \{(\underline{P}, \underline{P}, \underline{\theta_b})\}, \{(\underline{N}, \underline{P}, \underline{\theta_b})\}, \\ & \{(\underline{P}, \underline{N}, \underline{\theta_a})\}, (P, N, \theta_b), (N, N, \theta_a), (N, N, \theta_b)\}. \end{aligned}$$

For example,  $\{(P, P, \theta_a), (P, P, \theta_b)\}$  are not eliminated for seller 1 because her choice of  $P$  is a best response to beliefs that the state is  $(P, P, \theta_a)$ . Finally, those states that are eliminated for either of the sellers are not carried forward to the next step. The states that are carried forward (i.e., the set  $\Gamma(\Omega)$  in Section 4) are

$$(P, P, \theta_b), (P, N, \theta_a), (P, N, \theta_b), (N, P, \theta_b), (N, N, \theta_b).$$

Step 0 captures the self-confirming equilibrium restriction that requires each player's strategy to be optimal given a belief that is consistent with her feedback partition.

*Step 1.* This step captures the additional requirement that each player *believes* that the other player chooses a strategy that is a best response to a belief that is consistent with feedback. The previous elimination process is applied to those states remaining after Step 0:

$$\begin{aligned} & \text{Seller 1} \\ & \{(\underline{P}, \underline{P}, \underline{\theta_b})\}, \{(\underline{P}, \underline{N}, \underline{\theta_a})\}, \{(\underline{P}, \underline{N}, \underline{\theta_b})\}, \{(\underline{N}, \underline{P}, \underline{\theta_b})\}, (N, N, \theta_b)\} \\ & \text{Seller 2} \\ & \{(\underline{P}, \underline{P}, \underline{\theta_b})\}, \{(\underline{N}, \underline{P}, \underline{\theta_b})\}, \{(\underline{P}, \underline{N}, \underline{\theta_a})\}, (P, N, \theta_b), (N, N, \theta_b)\}. \end{aligned}$$

In particular,  $(P, P, \theta_b)$  is now eliminated for seller 1, since she can no longer justify playing  $P$  by believing in the previously eliminated state  $(P, P, \theta_a)$ . For this example, no more states can be eliminated and the procedure stops at Step 1.

The set of RCE's for this example depends on the objective game: a strategy profile  $(x_1, x_2)$  is an RCE of game  $\theta$  if the state  $(x_1, x_2, \theta)$  survives the iterative elimination procedure. If the objective game is  $\theta_a$ , then  $(P, N)$  is the unique RCE, which coincides with the unique Nash equilibrium. In contrast, there are three RCE's if the objective game is  $\theta_b$ :  $(N, P)$ ,  $(P, N)$ , and  $(N, N)$ . While the first RCE is also a Nash equilibrium, the remaining RCE's involve misperceptions. In particular, it is possible that both players decide not to pitch,  $(N, N)$ , which relies on seller 2 misperceiving both  $\theta$  and seller 1's action as well as seller 1 misperceiving seller 2's action. However, the profile  $(P, P)$ , which was previously found to be a BNE outcome for a certain type space, is not an RCE.<sup>14</sup>  $\diamond$

Of course, one could directly impose restrictions on beliefs that yield the same results, such as, for example, requiring players to have correct beliefs if the objective game

<sup>14</sup>A nonstandard feature of the iterative procedure is that its outcome lacks a product structure: If the objective game is  $\theta_b$ , the predicted outcomes are  $(P, N)$ ,  $(N, P)$ , and  $(N, N)$ , but *not*  $(P, P)$ . The lack of a product structure is due to the feedback restrictions. For example,  $(P, P, \theta_b)$  would survive Step 1 if no feedback restriction was assumed:  $P$  is seller 2's best response to  $P, \theta_b$  and  $P$  is seller 1's best response to  $N, \theta_b$ . But with feedback, player 1 is not allowed to believe  $(P, N, \theta_b)$  conditional on the outcome being  $(P, P, \theta_b)$ .



is  $\theta_a$ . The benefit of the proposed framework is to offer a systematic and transparent way to restrict equilibrium beliefs. One first exogenously fixes certain restrictions on beliefs. These restrictions (e.g., it is common belief that each player has a correct belief about her equilibrium payoff) should be motivated by the environment (e.g., players obtain payoff feedback while repeatedly playing the game). One then endogenously derives (nonobvious) conclusions about the strategy profiles that can be supported given these restrictions on beliefs.

**EXAMPLE 2** (Deterministic game; seller 2 informed). Are the RCE's in [Example 1](#) robust to the additional restriction that (it is common belief that) seller 2 knows whether the objective game is  $\theta_a$  or  $\theta_b$ ? This additional restriction is represented by refining the partition of seller 2 so that two states in the same element of his partition share the same  $\theta$ . Importantly, it is *not* the case that a strategy of seller 2 is a mapping from  $\{\theta_a, \theta_b\}$  to  $\{P, N\}$ . It is now important to distinguish between seller 1's belief about how seller 2 would play as a function of  $\theta$  and the actual strategy of seller 2 (which is either  $P$  or  $N$ , since the game is either  $\theta_a$  or  $\theta_b$ , not both). The reason is that equilibrium beliefs are restricted by players' actual, and not hypothetical, choices.

Because seller 2 is now informed of  $\theta$ , Step 0 in [Example 1](#) is modified by replacing the element of seller 2's partition where he plays  $N$  by

$$\{(P, N, \theta_a), (N, N, \theta_a)\}, \{(\cancel{P, N, \theta_b}), (\cancel{N, N, \theta_b})\}.$$

Seller 2 can no longer misperceive the state and, therefore, he must play his dominant strategy of pitching under  $\theta_b$ , thus eliminating states  $(P, N, \theta_b)$  and  $(N, N, \theta_b)$ . As in [Example 1](#), the state  $(P, P, \theta_b)$  is further eliminated in Step 1, so that both players must have correct beliefs in equilibrium. Therefore, RCE coincides with Nash equilibrium:  $(P, N)$  in game  $\theta_a$  and  $(N, P)$  in game  $\theta_b$ .  $\diamond$

**EXAMPLE 3** (Random game; seller 2 informed). Suppose that the objective game can be random, in the sense that  $\Pr(\theta_a) = p^0$  and  $\Pr(\theta_b) = 1 - p^0$ . The interpretation (which is standard in games of asymmetric information) is that the sellers repeatedly play the stage game in [Figure 1](#), where  $\theta$  is independent and identically distributed across periods. Suppose, in addition, that seller 2 observes the realization of  $\theta$  before choosing her action, that  $p^0 \in [0, 1]$ , and that these events are common belief. Now a strategy of seller 2 is a mapping from  $\{\theta_a, \theta_b\}$  to  $\{P, N\}$ , so that the set of strategies for seller 2 is  $X_2 = \{NN, NP, PN, PP\}$ , where the first action is contingent on observing  $\theta_a$  and the second action is contingent on  $\theta_b$ . The state space is  $\Omega = \{P, N\} \times X_2 \times [0, 1]$ .

Because it is dominant for seller 2 to pitch in game  $\theta_b$ , then for  $p^0 < 1$ ,  $PN$  and  $NN$  are dominated strategies for seller 2 and do not survive Step 0. Seller 2's partition for the remaining states is

Seller 2

$$\{(\cancel{P, PP, p}), (N, PP, p)\}, \{(P, NP, p), (\cancel{N, NP, p})\}, \{(P, *P, 0)\}, \{(N, *P, 0)\} \\ \{(\cancel{P, P*, 1}), (N, P*, 1)\}, \{(\cancel{P, N*, 1}), (N, N*, 1)\}$$

for all  $p \in (0, 1)$  and  $* \in \{P, N\}$ . For example, if the state is  $(P, PP, p)$ , then seller 2's feedback about her own payoff is  $L_a$  with probability  $p$  and  $L_b$  with probability  $1 - p$ . It is easy to check that any other state where seller 2 plays  $PP$  generates a different payoff distribution for seller 2.<sup>15</sup> In addition, for seller 2, playing  $P$  is not a best response to  $P$  and playing  $N$  is not a best response to  $N$  in game  $\theta_a$ , so that states  $(P, PP, p)$  and  $(N, NP, p)$  are also eliminated. Restricted to the remaining states, seller 1's partition is

$$\begin{aligned} & \text{Seller 1} \\ & \{(\underline{N, PP, p_1}), \{(\underline{N, PP, p'_1}), \{(P, NP, p_2)\}, \{(\underline{P, NP, p'_2}), \{(\underline{P, *P, 0}), \\ & \{(\underline{N, *P, 0}), \{(P, N*, 1)\}, \{(\underline{N, P*, 1}), \{(\underline{N, N*, 1})\} \end{aligned}$$

for all  $p_1 \in (0, p_H]$ ,  $p'_1 \in (p_H, 1)$ ,  $p_2 \in [p_L, 1)$ , and  $p'_2 \in (0, p_L)$ , where  $p_L \equiv (O_b - L)/(H - L + O_b - O_a)$  and  $p_H \equiv (O_b - L)/(O_b - O_a)$ . In particular, if the objective game is not random (i.e.,  $p^0 = 0$  or  $p^0 = 1$ , as in [Example 2](#)), then RCE coincides with Nash equilibrium and the finding that players must play Nash equilibrium in [Example 2](#) is therefore robust to the introduction of player 1's uncertainty about whether the game is random.<sup>16</sup>  $\diamond$

### 3. THE FRAMEWORK

*Primitives.* To simplify issues regarding measurability, the main exposition assumes that all sets are finite, endowed with the discrete  $\sigma$ -field (see [Section 6.1](#) for extensions).

**DEFINITION 1.** A (simultaneous-move) *game*  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  consists of

- a finite set of players  $I$
- a collection of nonempty strategy sets  $\{X_i\}_{i \in I}$ , where  $X = \times_{i \in I} X_i$
- a set of fundamentals  $\Theta$ , with the true fundamental  $\theta \in \Theta$
- a utility function  $u_i: X \times \Theta \rightarrow \mathbb{R}$  for each  $i \in I$
- a state space  $\Omega = X \times \Theta$
- a collection of feedback correspondences  $\mathcal{P} = \{P_i\}_{i \in I}$ , where  $P_i: \Omega \rightarrow \Omega$  is such that for all  $\omega, \omega' \in \Omega$ , (i)  $\omega \in P_i(\omega)$  and (ii) if  $\omega' \in P_i(\omega)$ , then  $P_i(\omega') = P_i(\omega)$  and  $x_i(\omega') = x_i(\omega)$ .<sup>17</sup>

The corresponding *objective game*  $\mathcal{G}^o(\theta)$  consists of the first four items, where  $\Theta = \{\theta\}$ .

The definition of an objective game is standard and its most common solution concept is [Nash \(1951\)](#) equilibrium.

<sup>15</sup>Each seller's belief about the *distribution* of her equilibrium payoff is assumed to be correct.

<sup>16</sup>More generally, RCE also coincides with Nash equilibrium for all  $p^0 \notin \{0, 1\}$ :  $(N, PP)$  for  $p^0 \leq p_H$  and  $(P, NP)$  for  $p^0 \geq p_L$ .

<sup>17</sup>Throughout the paper, for a given state  $\omega = (x_1, \dots, x_I, \theta)$ , let  $x_i(\omega) = x_i$  and  $\theta(\omega) = \theta$ . Note that [Definition 1](#) requires  $P_i$  to be a partition and  $x_i(\cdot)$  to be  $P_i$ -measurable.

DEFINITION 2. A strategy profile  $x \in X$  is a *Nash equilibrium* of the objective game  $\mathcal{G}^o(\theta)$  if, for every  $i \in I$ ,

$$x_i \in \arg \max_{x'_i \in X_i} u_i(x'_i, x_{-i}, \theta).$$

The notion of a game is more general than meets the eye, and allows for mixed strategies and the existence of asymmetric information by an appropriate relabeling of the primitives. These extensions require an infinite strategy space, and Section 6.1 shows that the results extend to infinite spaces.

EXAMPLE 4 (Objective auction game; e.g., Vickrey 1961). Consider a private-value auction among players  $\{1, \dots, I\}$ , where private values  $(v_1, \dots, v_I) \in V = \times_{i \in I} V_i$  are drawn according to a probability measure  $\theta \in \Delta(V)$ . Each player  $i \in \{1, \dots, I\}$  privately observes  $v_i \in V_i$  and simultaneously submits a bid  $b_i$ . Payoffs are  $U_i(b_1, \dots, b_I, v_i)$ . This situation can be cast as an objective game by letting  $X_i$  be the set of bidding functions  $x_i: V_i \rightarrow \mathbb{R}_+$  and letting  $u_i(x_1, \dots, x_I, \theta) = E_\theta U_i(x_1(v_1), \dots, x_I(v_I), v_i)$ , where the expectation  $E_\theta$  is taken with respect to  $\theta \in \Delta(V)$ .  $\diamond$

Following Fudenberg and Levine (1993a), it is natural to interpret Nash equilibrium as requiring (i) *rationality*—player  $i$  plays a best response  $x_i$  to her (first-order) beliefs about  $(x_{-i}, \theta)$ —and (ii) *correct beliefs*—player  $i$ 's beliefs about  $(x_{-i}, \theta)$  are correct in the sense that  $x_{-i}$  is the equilibrium strategy profile of other players and  $\theta$  is the true fundamental. This paper extends Nash's definition of equilibrium by allowing players to hold higher-order beliefs (e.g., beliefs about what other players believe) and by relaxing the assumption that equilibrium strategies must be supported by beliefs that are correct. These extensions are complementary: It is interesting to allow for higher-order beliefs precisely because players may now hold one of many (possibly incorrect) beliefs, and introspection may help eliminate some incorrect beliefs. Formally, this extension is accomplished by adding two elements to the definition of an objective game (i.e., the last two items in Definition 1).

The first additional element is the state space  $\Omega = X \times \Theta$ , which captures the elements of the game over which players are allowed to be uncertain and, therefore, over which higher-order beliefs will be defined. The second element is a collection of partitional feedback correspondences  $\mathcal{P}$ , which provide a flexible way to capture restrictions on beliefs, either motivated by players' a priori information about the primitives or by "feedback" that results from repeatedly playing a particular strategy profile.<sup>18</sup> The set  $P_i((x, \theta))$ , which is required to include  $(x, \theta)$ , represents the restricted support of player  $i$ 's beliefs when players play  $x$  and the true fundamental is  $\theta$ . It can be interpreted as the set of states that are observationally equivalent to  $(x, \theta)$  for player  $i$ . The default restriction on  $P_i$  that requires player  $i$ 's strategy to be measurable with respect to the partition of  $\Omega$  generated by  $P_i$  captures the assumption that player  $i$  observes, at the very least, her own strategy. As mentioned above and illustrated through examples,  $P_i$

<sup>18</sup>From the point of view of RCE analysis, beliefs are restricted by all of the information possessed by players, and whether information is obtained ex post or ex ante does not make a difference.

can capture further restrictions on beliefs, and the equilibrium set will depend on these exogenously-specified partitions.

**EXAMPLE 2 CONTINUED.** Formally,  $I = \{1, 2\}$ ,  $X_1 = X_2 = \{P, N\}$ , and  $\theta \in \Theta = \{\theta_a, \theta_b\}$  is the true (i.e., objective) fundamental, and  $u_1$  and  $u_2$  are given by Figure 1. Uncertainty is defined over strategy profiles and the fundamental, so that  $\Omega = \{P, N\} \times \{P, N\} \times \{\theta_a, \theta_b\}$ . The restriction that seller 1 observes her own payoff is captured by letting

$$P_1(\omega) = \{\omega' : u_1(\omega) = u_1(\omega'), x_1(\omega) = x_1(\omega')\}.$$

Finally, the assumption that player 2 observes his own payoff and is informed of the true fundamental  $\theta$  is captured by

$$P_2(\omega) = \{\omega' : u_2(\omega) = u_2(\omega'), \theta(\omega) = \theta(\omega'), x_2(\omega) = x_2(\omega')\}. \quad \diamond$$

The goal is to formalize the idea that a solution to a game is a strategy profile  $x$  that satisfies rationality, consistency of beliefs (with respect to feedback correspondences  $\mathcal{P}$ ), and common belief of rationality and consistency. The remainder of this section introduces the machinery to represent and restrict higher-order beliefs, and formally defines RCE.

*Belief spaces.* I adopt machinery from epistemic game theory so as to formalize players' higher-order beliefs. The novelty here is the particular way in which restrictions are imposed on players' entire hierarchy of beliefs. A belief space (also known as an epistemic type space) formalizes players' beliefs about the space of primitive uncertainty, their beliefs about other players' beliefs, and so on. A belief space is not part of the primitives of the game; instead, beliefs (about both strategies and fundamentals) are jointly determined in equilibrium. Players' beliefs over the space of primitive uncertainty or state space  $\Omega$  are represented by a *belief space*

$$\mathcal{B} = \langle \Omega, T, \xi, \{\lambda_i\}_{i \in I} N \rangle,$$

where  $T$  is a finite set, and  $\xi : T \rightarrow \Omega$  and  $\lambda_i : T \rightarrow \Delta(T)$  are functions. As usual,  $\Delta(T)$  is the set of all probability measures on  $T$ . Corresponding to a belief space  $\mathcal{B}$ , there is a set of epistemic states (i.e., profiles of types and states of nature)  $T$ , and each epistemic state  $t \in T$  is associated with<sup>19</sup>

- a state of primitive uncertainty  $\xi(t) = (x_1(\xi(t)), \dots, x_n(\xi(t)), \theta(\xi(t))) \in \Omega$  (i.e., a strategy profile and a fundamental)

<sup>19</sup>An equivalent representation is  $T = T_0 \times T_1 \times \dots \times T_I$ , where  $T_0 = \Omega$ ,  $T_i$  is player  $i$ 's type space, and  $m_i : T_i \rightarrow \Delta(T)$  captures players' beliefs. Applications of BNE follow this product representation (where  $T_0 = \Theta$ ) and define player  $i$ 's strategy as a mapping from  $T_i$  to  $X_i$ . In our setting, in contrast, a strategy is simply an element of  $X_i$ , which is fixed before the notion of a type space is introduced, and this alternative representation would add unnecessary notation.

- a probability measure  $\lambda_i(t) \in \Delta(T)$  that represents player  $i$ 's beliefs over epistemic states (i.e., over strategy profiles, structural uncertainty, and beliefs of all players).<sup>20</sup>

An event is a set  $E \subseteq T$ , and  $\lambda_i(t)[E]$  denotes the probability that player  $i$  assigns to event  $E$  at  $t$ .

*Rationality.* Let  $\Delta(\Omega)$  denote the set of probability distributions over  $\Omega$  and, for any  $\delta \in \Delta(\Omega)$ , let  $\text{marg}_{X_{-i} \times \Theta} \delta$  denote the marginal distribution over  $X_{-i} \times \Theta$ . The set

$$\Phi_i(\delta) \equiv \arg \max_{x'_i \in X_i} \int_{X_{-i} \times \Theta} u_i(x'_i, x_{-i}, \theta) d\text{marg}_{X_{-i} \times \Theta} \delta$$

is the set of (perceived) best responses of player  $i$  who believes that strategies and fundamentals are distributed according to  $\delta \in \Delta(\Omega)$ . Note that player  $i$ 's first-order belief over  $\Omega$  at epistemic state  $t$  is given by  $\lambda_i(t) \circ \xi^{-1} \in \Delta(\Omega)$ . Then

$$R_i^{\mathcal{B}} \equiv \{t : x_i(\xi(t)) \in \Phi_i(\lambda_i(t) \circ \xi^{-1})\}$$

is defined to be the event that player  $i$  is rational given a belief space  $\mathcal{B}$ . Player  $i$  is said to be *rational at  $t$*  if  $t \in R_i^{\mathcal{B}}$ .<sup>21</sup>

*Consistency of beliefs.* Given a belief space  $\mathcal{B}$ , player  $i$ 's beliefs are  *$P_i$ -consistent at  $t$*  if  $\lambda_i(t)[\xi^{-1}(P_i(\xi(t)))] = 1$ ; i.e., player  $i$  puts zero probability on states that are not in  $P_i(\xi(t))$ . Let

$$C_i^{\mathcal{B}, \mathcal{P}} \equiv \{t : \lambda_i(t)[\xi^{-1}(P_i(\xi(t)))] = 1\}$$

denote the event that player  $i$  has consistent beliefs.

Let  $RC_i^{\mathcal{B}, \mathcal{P}} = R_i^{\mathcal{B}} \cap C_i^{\mathcal{B}, \mathcal{P}}$  denote the event that player  $i$  is rational and has consistent beliefs, and let  $RC^{\mathcal{B}, \mathcal{P}} = \bigcap_{i \in I} RC_i^{\mathcal{B}, \mathcal{P}}$  denote the event that all players are rational and have consistent beliefs.

*Restrictions on  $k$ th order beliefs.* Fix any belief space  $\mathcal{B}$ . For an event  $E \subseteq T$ , let

$$\mathbf{B}_i E = \{t \in T : \lambda_i(t)[E] = 1\}$$

denote the event that player  $i$  believes  $E$ , that is, the set of epistemic states at which player  $i$  assigns probability 1 to event  $E$ . Let  $\mathbf{B}E = \bigcap_{i \in I} \mathbf{B}_i E$  denote the event that all players believe  $E$ .

<sup>20</sup>Note that players are allowed to believe that other players' strategies are correlated with each other and with nature; see Aumann (1974) and Brandenburger and Friedenberg (2008) for different approaches that formalize the source of these correlations.

<sup>21</sup>As pointed out by a referee, it can be conceptually problematic to let  $\delta$  be any element of  $\Delta(\Omega)$ , e.g., how do we interpret a best response when a player believes that her own strategy is correlated with the strategy of other players? This issue can be avoided by letting  $\Delta_i(\Omega)$  be the set of  $\delta \in \Delta(\Omega)$  with the property that  $\delta[\{x_i\} \times X_{-i} \times \Theta] = 1$  for some  $x_i \in X_i$  and then letting  $\Delta_i(\Omega)$  be the domain of  $\Phi_i$ . The set  $R_i^{\mathcal{B}}$  is then defined as above with the additional requirement that  $\lambda_i(t) \circ \xi^{-1} \in \Delta_i(\Omega)$ . Because  $P_i$  is  $x_i$ -measurable, player  $i$  indeed is restricted to beliefs in  $\Delta_i(\Omega)$  in equilibrium, and all the results in this paper hold under this alternative definition of rationality.

Define recursively  $(RC^{\mathcal{B}, \mathcal{P}})^1 = RC^{\mathcal{B}, \mathcal{P}}$  and

$$(RC^{\mathcal{B}, \mathcal{P}})^{j+1} = (RC^{\mathcal{B}, \mathcal{P}})^j \cap \mathbf{B}(RC^{\mathcal{B}, \mathcal{P}})^j$$

for all  $j = 1, 2, \dots$ . Then,  $(RC^{\mathcal{B}, \mathcal{P}})^{k+1}$  denotes the event that players are rational and have consistent beliefs, and that there is  $k$ th-order belief of rationality and consistency. Moreover,

$$RC^{\mathcal{B}, \mathcal{P}} \mathbf{CB} RC^{\mathcal{B}, \mathcal{P}} = \bigcap_{k=0}^{\infty} (RC^{\mathcal{B}, \mathcal{P}})^{k+1}$$

denotes the event that there is rationality, consistency, and common belief of rationality and consistency.

*Definition of equilibrium.* A strategy profile is a  $k$ -rationalizable conjectural equilibrium if there exists a belief space such that (i) players' strategies maximize their expected utility given their beliefs (rationality), (ii) beliefs are  $\mathcal{P}$ -consistent (consistency), and (iii) there is  $k$ th-order belief of rationality and consistency.

**DEFINITION 3.** A strategy profile  $x \in X$  is a  *$k$ -rationalizable conjectural equilibrium* ( $k$ -RCE) of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  if there exists a belief space  $\mathcal{B} = \langle \Omega, T, \xi, \{\lambda_i\}_{i \in I} \rangle$  and some  $t \in T$  such that  $\xi(t) = (x, \theta)$  and  $t \in (RC^{\mathcal{B}, \mathcal{P}})^{k+1}$ .

**DEFINITION 4.** A strategy profile  $x \in X$  is a *rationalizable conjectural equilibrium* (RCE) of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  if there exists a belief space  $\mathcal{B} = \langle \Omega, T, \xi, \{\lambda_i\}_{i \in I} \rangle$  and some  $t \in T$  such that  $\xi(t) = (x, \theta)$  and  $t \in RC^{\mathcal{B}, \mathcal{P}} \mathbf{CB} RC^{\mathcal{B}, \mathcal{P}}$ .

Throughout the paper, explicit reference to  $\mathcal{B}$  and  $\mathcal{P}$  is omitted, unless it is not obvious from the context.

*Existence and Nash equilibrium.* The following simple result establishes that a Nash equilibrium is also an RCE, implying that an RCE exists whenever a Nash equilibrium exists.

**CLAIM 1.** If  $x$  is a Nash equilibrium of the objective game  $\mathcal{G}^o(\theta)$  corresponding to the game  $\mathcal{G}(\theta, \Omega, \mathcal{P})$ , then  $x$  is also an RCE of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$ .

**PROOF.** Fix a belief space  $\mathcal{B} = \langle \Omega, T, \xi, \{\lambda_i\}_{i \in I} \rangle$  with  $T = \{t\}$  and  $\xi(t) = (x, \theta) \in \Omega$  (in particular, this is possible because of the assumption that  $\theta \in \Theta$ ). Since  $x$  is a Nash equilibrium, then each player maximizes utility given correct beliefs  $(x, \theta)$ . Since all players have these beliefs at  $t$ , then all players are rational at  $t$ . In addition, by definition of  $P_i$ ,  $\xi(t) \in P_i(\xi(t))$  for all  $i$ , so that all players have consistent beliefs at  $t$ . Therefore,  $T = RC$ . Finally, since  $\mathbf{B}T = T$ , then  $T \mathbf{CB} T = T$ . Hence,  $t \in RCC \mathbf{CB} RC$ .  $\square$

A converse of [Claim 1](#) does not hold in general.<sup>22</sup> But an obvious sufficient condition for an RCE to be a Nash equilibrium is that players get perfect feedback.

<sup>22</sup>It is well known that a 0-RCE (which is a self-confirming equilibrium; see [Section 6.3](#)) need not be a Nash equilibrium (e.g., [Fudenberg and Kreps 1988](#), [Fudenberg and Levine 1993a](#), [Dekel et al. 2004](#)).



CLAIM 2. *If  $x$  is a  $k$ -RCE of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  for some  $k$  and  $P_i(\omega) = \{\omega\}$  for all  $\omega \in \Omega$  and  $i \in I$ , then  $x$  is a Nash equilibrium of the objective game  $\mathcal{G}^o(\theta)$  corresponding to  $\mathcal{G}(\theta, \Omega, \mathcal{P})$ .*

PROOF. By assumption there exists  $\mathcal{B} = \langle \Omega, T, \xi, \{\lambda_i\}_{i \in I} \rangle$  and  $t \in T$  such that  $\xi(t) = (x, \theta)$  and  $t \in (RC)^{k+1}$ . In particular,  $t \in RC$ . Therefore,  $x_i \in \Phi_i(\lambda_i(t))$  and  $\lambda_i(t)[\{x, \theta\}] = 1$  for all  $i$ , implying that  $x_i \in \arg \max_{x'_i \in X_i} u_i(x'_i, x_{-i}, \theta)$  for all  $i$ .  $\square$

Section 6.3 discusses the relationship between the definition of RCE and other, non-Nash, equilibrium concepts in the literature.

#### 4. CHARACTERIZATION OF EQUILIBRIUM

Throughout this section, fix a game  $\mathcal{G}(\theta, \Omega, \mathcal{P})$ . For  $k = 0, 1, \dots$ , let  $\mathcal{E}^k$  denote the set of states  $\omega = (x, \theta) \in \Omega$  such that  $x$  is a  $k$ -RCE of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$ . Let  $\mathcal{E}$  denote the set of states  $\omega = (x, \theta)$  such that  $x$  is a RCE of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$ . Theorems 1 and 2 characterize  $\mathcal{E}^k$  and  $\mathcal{E}$  via an iterative procedure over the set of states  $\Omega$ ; hence, the characterization does not involve the notion of belief spaces introduced in Section 3.

Let  $\Gamma: 2^\Omega \rightarrow 2^\Omega$  be the *equilibrium operator* defined by

$$\Gamma(A) = \left\{ \omega \in \Omega : \forall i \in I \exists \delta_i \in \Delta(\Omega) \text{ such that: } \begin{array}{l} x_i(\omega) \in \Phi_i(\delta_i) \\ \delta_i(P_i(\omega) \cap A) = 1 \end{array} \right\} \quad (1)$$

for all  $A \subseteq \Omega$ . In words,  $\Gamma(A)$  is the set of states where the strategies of all players can be rationalized by beliefs that are consistent and have support in  $A$ . Define recursively, for  $j \in 1, 2, \dots$ ,  $\Gamma^j(\cdot) = \Gamma(\Gamma^{j-1}(\cdot))$ , where  $\Gamma^0(A) = A$  for all  $A \subseteq \Omega$ . It is immediate from (1) that  $\Gamma$  is monotonic ( $A' \subseteq A$  implies  $\Gamma(A') \subseteq \Gamma(A)$ ) so that  $\Gamma^j(\Omega) \subseteq \Gamma^{j-1}(\Omega)$  for all  $j$ .

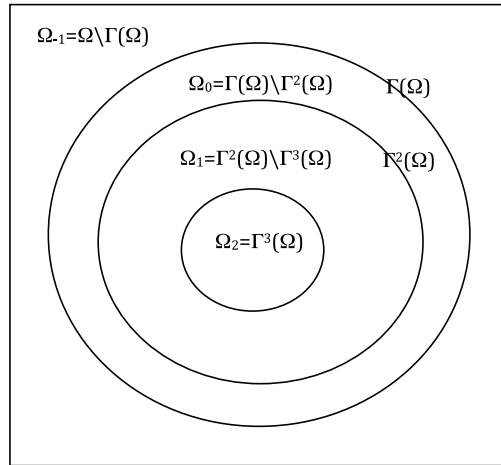
THEOREM 1 (Characterization of  $k$ -RCE). *The set of  $k$ -RCE is obtained by iterative application of  $\Gamma$ : For all  $k = 0, 1, 2, \dots$ ,*

$$\mathcal{E}^k = \Gamma^{k+1}(\Omega).$$

Theorem 1 shows that  $k$ -RCE can be characterized by a procedure that iteratively removes those states  $\omega$  that are not rationalizable *given* the states that belong to  $P_i(\omega)$  and that remain from the previous iteration. The procedure in Theorem 1 generalizes the procedure of iterated deletion of strategies that are never a best response that characterizes the set of (correlated) rationalizable strategies. The generalization accounts for *equilibrium* restrictions on beliefs, which explains why rationalizability is with respect to remaining states that are indistinguishable to each other (for a certain  $\mathcal{P}$ ), and not with respect to the entire set of remaining states.

Theorem 1 follows from Lemmas 1 and 2 below.

LEMMA 1. *For all  $k = 0, 1, \dots$ ,  $\mathcal{E}^k \subseteq \Gamma^{k+1}(\Omega)$ .*

FIGURE 2. Basic belief system,  $k = 2$ .

PROOF. The proof establishes that  $\mathcal{E}^k \subseteq \Gamma(\mathcal{E}^{k-1})$  for all  $k = 0, 1, \dots$  (where  $\mathcal{E}^{-1} \equiv \Omega$ ); the lemma then follows by monotonicity of  $\Gamma$ . Fix  $k$  and let  $\omega \in \mathcal{E}^k$ . By definition of equilibrium, there exist a belief space  $\mathcal{B} = \langle \Omega, T, \xi, \{\lambda_i\}_{i \in I} \rangle$  and  $t \in T$  such that  $\xi(t) = \omega$  and  $t \in (RC)^{k+1}$ . In particular,  $t \in RC$  and  $t \in \mathbf{B}(RC)^k$ . Fix a player  $i$  and let  $\delta_i \in \Delta(\Omega)$  be such that for all  $A \subseteq \Omega$ ,

$$\delta_i[A] = \lambda_i(t)[\xi^{-1}(A)].$$

Since  $t \in R_i$ , then  $x_i(\omega) \in \Phi_i(\delta_i)$ . Since  $t \in C_i$ , then

$$\begin{aligned} \delta_i[P_i(\omega)] &= \lambda_i(t)[\xi^{-1}(P_i(\xi(t)))] \\ &= 1. \end{aligned}$$

Finally,

$$\begin{aligned} \delta_i[\mathcal{E}^{k-1}] &= \lambda_i(t)[\xi^{-1}(\mathcal{E}^{k-1})] \\ &\geq \lambda_i(t)[(RC)^k] \\ &= 1, \end{aligned}$$

where the second line follows from the fact that  $\mathcal{E}^{k-1} \supseteq \xi((RC)^k)$ , which implies that  $\xi^{-1}(\mathcal{E}^{k-1}) \supseteq \xi^{-1}\xi((RC)^k) \supseteq (RC)^k$ , and the third line follows because  $t \in \mathbf{B}(RC)^k$ . Since the argument holds for every  $i$ , then  $\omega \in \Gamma(\mathcal{E}^{k-1})$ .  $\square$

The remainder of the proof uses the equilibrium operator  $\Gamma$  to construct a particular belief space and establishes that  $\Gamma^{k+1}(\Omega) \subseteq \mathcal{E}^k$ . For all  $k \in \{0, 1, \dots\}$ , define the sets  $\Omega_l^k \equiv \Gamma^{l+1}(\Omega) \setminus \Gamma^{l+2}(\Omega)$  for  $l = -1, 0, 1, 2, \dots, k-1$  and  $\Omega_k^k = \Gamma^{k+1}(\Omega)$ . By monotonicity of  $\Gamma$ ,  $(\Omega_l^k)_l$  is a collection of disjoint sets (some of which may be empty) that partitions  $\Omega$ . Figure 2 illustrates an example for  $k = 2$ .

**DEFINITION 5.** A belief space  $\mathcal{B}^k = \langle \Omega, T, \xi, \{\lambda_i^k\}_{i \in I} \rangle$  is *k-basic* if (i)  $T = \Omega$ , (ii)  $\xi$  is the identity function, and (iii) for all  $i$ ,  $\lambda_i^k$  is measurable with respect to  $i$ 's partition,  $P_i$ , and for each  $l = 0, \dots, k$  and all  $\omega \in \Omega_l^k$ ,  $\lambda_i^k(\omega)$  is an element of

$$\{\delta_i \in \Delta(P_i(\omega)) : x_i(\omega) \in \Phi_i(\delta_i)\} \cap \Delta(\Gamma^l(\Omega)). \quad (2)$$

By construction, a *k*-basic belief system always exists for all  $k$ .<sup>23</sup>

**LEMMA 2.** Fix any  $k = 0, 1, \dots$  and any *k*-basic belief space  $\mathcal{B}^k$ . Then  $(RC^{\mathcal{B}^k})^{k+1} = \Gamma^{k+1}(\Omega) \subseteq \mathcal{E}^k$ .

**PROOF.** *Step 1* ( $RC^{\mathcal{B}^k} = \Gamma^1(\Omega)$ ). Let  $\omega \in RC^{\mathcal{B}^k}$ . Then  $x_i(\omega) \in \Phi_i(\lambda_i^k(\omega))$  and  $\lambda_i^k(\omega)[P_i(\omega)] = 1$  for all  $i$ , so that  $\omega \in \Gamma^1(\Omega)$ . Next suppose that  $\omega \in \Gamma^1(\Omega)$ . Then  $\omega \in \Omega_l^k$  for some  $l = 0, 1, \dots, k$ . By (2),  $x_i(\omega) \in \Phi_i(\lambda_i^k(\omega))$  and  $\lambda_i^k(\omega)[P_i(\omega)] = 1$  for all  $i$ , so that  $\omega \in RC^{\mathcal{B}^k}$ .

*Step 2* ( $\Gamma^j(\Omega) \cap \mathbf{B}^{\mathcal{B}^k} \Gamma^j(\Omega) = \Gamma^{j+1}(\Omega)$  for all  $j = 1, 2, \dots, k$ ). Let  $\omega \in \Gamma^j(\Omega) \cap \mathbf{B}^{\mathcal{B}^k} \Gamma^j(\Omega)$ . Since  $\omega \in \Gamma^j(\Omega)$ , there exists  $l \geq j - 1$  such that  $\omega \in \Omega_l^k$ . By (2),  $x_i(\omega) \in \Phi_i(\lambda_i^k(\omega))$  and  $\lambda_i^k(\omega)[P_i(\omega)] = 1$  for all  $i$ . In addition,  $\omega \in \mathbf{B}^{\mathcal{B}^k} \Gamma^j(\Omega)$  implies  $\lambda_i^k(\omega)[\Gamma^j(\Omega)] = 1$  for all  $i$ . Therefore,  $\omega \in \Gamma(\Gamma^j(\Omega)) = \Gamma^{j+1}(\Omega)$ . For the other direction, let  $\omega \in \Gamma^{j+1}(\Omega)$ . Since  $\Gamma$  is monotonic,  $\omega \in \Gamma^j(\Omega)$ . In addition, there exists  $l \geq j$  such that  $\omega \in \Omega_l^k$ . By (2),  $\lambda_i^k(\omega)[\Gamma^j(\Omega)] = 1$  for all  $i$  and, therefore,  $\omega \in \mathbf{B}^{\mathcal{B}^k} \Gamma^j(\Omega)$ .

By Steps 1 and 2, it follows that  $(RC^{\mathcal{B}^k})^j = \Gamma^j(\Omega)$  for  $j = 1, \dots, k + 1$ ; in particular,  $(RC^{\mathcal{B}^k})^{k+1} = \Gamma^{k+1}(\Omega)$ . The desired result then follows from the fact that  $(RC^{\mathcal{B}^k})^{k+1} \subseteq \mathcal{E}^k$ , since  $\mathcal{B}^k$  is one of many possible belief spaces that support equilibrium.  $\square$

The next result establishes that a strategy profile is an RCE if and only if it is a *k*-RCE for all  $k$ .

**THEOREM 2** (Characterization of RCE). We have  $\mathcal{E} = \bigcap_{k=0}^{\infty} \mathcal{E}^k = \bigcap_{k=0}^{\infty} \Gamma^{k+1}(\Omega)$ .

**PROOF.** By definition of equilibrium,  $\mathcal{E} \subseteq \bigcap_{k=0}^{\infty} \mathcal{E}^k$ . By Theorem 1,  $\bigcap_{k=0}^{\infty} \mathcal{E}^k = \bigcap_{k=1}^{\infty} \Gamma^k(\Omega)$ . Therefore, it remains to establish that  $\bigcap_{k=1}^{\infty} \Gamma^k(\Omega) \subseteq \mathcal{E}$ . First, let  $k^* < \infty$  be the (finite) number of elements in  $\Omega$ . Then  $\Gamma^j(\Omega) = \Gamma^{k^*}(\Omega)$  for all  $j \geq k^*$  and, by monotonicity of  $\Gamma$ ,  $\bigcap_{k=1}^{\infty} \Gamma^k(\Omega) \subseteq \Gamma(\bigcap_{k=1}^{\infty} \Gamma^k(\Omega))$ . Second, consider the belief space  $\mathcal{B}^{\infty} = \langle \Omega, T, \xi, \{\lambda_i^{\infty}\}_{i \in I} \rangle$  where (i)  $T = \bigcap_{k=1}^{\infty} \Gamma^k(\Omega)$ , (ii)  $\xi$  is the identity function, and (iii) for each  $\omega \in T$ ,  $\lambda_i^{\infty}(\omega)$  is an (arbitrary) element of

$$\{\delta_i \in \Delta(P_i(\omega)) : x_i(\omega) \in \Phi_i(\delta_i)\} \cap \Delta(T)$$

(which is possible because, by the first step,  $T \subseteq \Gamma(T)$ ). Therefore,  $T = RC^{\mathcal{B}^{\infty}}$ . Finally, since  $\mathbf{B}T = T$ , then  $T = RC^{\mathcal{B}^{\infty}} \mathbf{B}RC^{\mathcal{B}^{\infty}} \subseteq \mathcal{E}$ , where the last inclusion follows

<sup>23</sup>The restriction imposed by (2) is feasible by construction of the sets  $\Omega_l^k$ . Measurability of  $\lambda_i^k$  (which is not needed for the result, but is conceptually appealing, since it implies that each player has a correct belief of her own belief) is feasible due to the assumption that each player's strategy is measurable with respect to her partition.

because  $\mathcal{B}^\infty$  is one of possibly many belief spaces that may support RCE. Therefore,  $T = \bigcap_{k=1}^\infty \Gamma^k(\Omega) \subseteq \mathcal{E}$ .  $\square$

There are two additional implications of these characterization results. First, a corollary is that a  $k$ -basic belief space  $\mathcal{B}^k$  supports the entire set of  $k$ -RCE, i.e.,  $(RC^{\mathcal{B}^k})^{k+1} = \mathcal{E}^k$ . Hence, for RCE analysis, it is without loss of generality to restrict attention to *basic* belief spaces. Such basic belief spaces are quite restrictive: e.g., if two players have the same first-order beliefs, then all their higher-order beliefs must coincide. Moreover, the approach followed here differs from Harsanyi's approach, which fixes a (nonepistemic) type space and defines a solution concept for such a fixed type space. In Harsanyi's case, a natural concern is that a particular type space is one of many that may capture the desired restrictions on hierarchies of beliefs, thus prompting the question, "What are the appropriate type spaces to consider?"<sup>24</sup> In contrast, in this paper an epistemic type space (or belief space) is not part of the primitives of the game, but, rather, it is used to describe the beliefs that support an equilibrium. Therefore, it is a result of the analysis that attention can be restricted to certain belief spaces.

Second, the construction of a basic belief space provides a convenient way to find higher-order beliefs that support a particular equilibrium. This point is illustrated next for the sales-pitch game in Section 2.

**EXAMPLE 1 CONTINUED.** As observed in Section 2, at the end of the iterative procedure, the remaining states and the corresponding partitions are given by

$$\begin{array}{c} \text{Seller 1} \\ \{(P, \underline{N}, \theta_a), (P, N, \theta_b)\}, \{(\underline{N}, P, \theta_b), (N, N, \theta_b)\} \\ \text{Seller 2} \\ \{(\underline{N}, P, \theta_b)\}, \{(\underline{P}, \underline{N}, \theta_a), (P, N, \theta_b), (N, N, \theta_b)\}. \end{array}$$

Suppose that the true fundamental is  $\theta_b$  and consider the RCE  $(N, N)$ . The underlined states represent the (degenerate) beliefs held for each player for each of the elements of her partition (i.e., an example of the  $\lambda_i^k$  function in a  $k$ -basic belief system). While the true state is  $(N, N, \theta_b)$ , player 1 incorrectly believes  $(N, P, \theta_b)$ , player 1 believes that player 2 believes  $(N, P, \theta_b)$ , and so on. Therefore, player 1 believes that  $(N, P, \theta_b)$  is common belief. Similarly, player 2 believes that  $(P, N, \theta_a)$  is common belief. These higher-order beliefs support the equilibrium  $(N, N)$ .  $\diamond$

## 5. EXAMPLES: ROBUST EQUILIBRIUM PREDICTIONS

The following examples illustrate how the framework can be applied to test the robustness of Nash predictions by allowing players to be uncertain about certain aspects of the game.

<sup>24</sup>See Ely and Pęski (2006), Dekel et al. (2007), Liu (2009), and Sadzik (2011) as well as the discussions by Battigalli et al. (2011) and Friedenberg and Meier (2012).

REGION (payoff in parentheses)										
		1	2	3		4	5			
		(1)	(1)	$(0 < b < 1)$		(1)	(1)			
Payoffs		2-player game (other player)				3-player game (other two players)				
		2	3	4	22	23	24	33	34	44
Player 1	2	$2 + b/2$	2	$2 + b/2$	$(4 + b)/3$	1	$1 + b/3$	2	2	$2 + b/3$
	3	$2 + b$	$2 + b/2$	$2 + b$	$2 + b$	$1 + b/2$	$b$	$(4 + b)/3$	$1 + b/2$	$2 + b$
	4	$2 + b/2$	2	$2 + b/2$	$2 + b/3$	2	$1 + b/3$	2	1	$4 + b/3$

FIGURE 3. Payoffs for location game.

5.1 Location game with uncertainty about the number of opponents

Each of  $n$  players must simultaneously choose to locate in one of five regions,  $\{1, \dots, 5\}$ . If a player captures region 3, her payoff is  $0 < b < 1$ ; if she captures any other region, her payoff is 1 per region (see Figure 3). A region is captured by the player(s) whose location is closest to the region; in the case of a tie, the payoff is equally shared among all closest players. In particular, in the two-player version of this game, both players choose region 3 in the unique rationalizable outcome. The new feature being modeled is that there are, in fact, two players in this game, but players are uncertain whether there are actually two or three players in the game. Is the choice of region 3 robust to this additional uncertainty?

Formally, let  $I = \{1, 2, 3\}$  and  $X_i = \{2, 3, 4\}$  for  $i \in I$ . (For simplicity, I assume that players are restricted to locate in regions 2–4.<sup>25</sup>) Let

$$\Theta = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

represent the set of players in the game, and let  $\Omega = \bigcup_{\theta \in \Theta} \times_{i \in \theta} X_i \times \{\theta\}$  be the underlying state space. The (objective) game is given by  $\theta^0 = \{1, 2\}$ ; i.e., players 1 and 2 are the actual players in this game (due to symmetry, nothing changes if the objective game is either  $\{1, 3\}$  or  $\{2, 3\}$ ). Figure 3 shows the payoff of a particular player as a function of her chosen location, the location of the other player(s), and the number of players. Feedback correspondences are such that, for all  $\omega$  with  $i \in \theta(\omega)$ ,

$$P_i(\omega) = \{\omega' \in \Omega : u_i(\omega') = u_i(\omega), i \in \theta(\omega')\},$$

i.e., player  $i$  knows that she is a player in the game and she has correct beliefs about her equilibrium payoff. However, players do not get feedback about the location chosen by other players.<sup>26</sup>

Due to symmetry, it is convenient to carry out the analysis from the point of view of player 1. In addition, because a player does not care which of the other players chooses a particular location, states  $(x_1, x_2, x_3, \{1, 2, 3\})$  and  $(x_1, x_3, x_2, \{1, 2, 3\})$  are simply

<sup>25</sup>It is easy to show that players do not locate in regions 1 and 5 in equilibrium.  
<sup>26</sup>For example, players can be firms with covert sales operations in a particular region.

represented as  $(x_1, x_2x_3)$ , and similarly for all other states. The partition for player 1 then becomes (ignoring states where player 1 does not play)

$$\begin{aligned} & \{(\underline{2}, \underline{2}), (\underline{2}, \underline{4})\}, \{(2, 3), (\underline{2}, \underline{33}), (\underline{2}, \underline{34})\}, \{(\underline{2}, \underline{22})\}, \{(\underline{2}, \underline{23})\}, \{(\underline{2}, \underline{24})\}, \{(\underline{2}, \underline{44})\} \\ & \{(\underline{3}, \underline{2}), (\underline{3}, \underline{4}), (\underline{3}, \underline{22}), (\underline{3}, \underline{44})\}, \{(\underline{3}, \underline{3}), \{(\underline{3}, \underline{23}), (\underline{3}, \underline{34})\}, \{(\underline{3}, \underline{24})\}, \{(\underline{3}, \underline{33})\} \\ & \{(\underline{4}, \underline{2}), (\underline{4}, \underline{4})\}, \{(4, 3), (\underline{4}, \underline{23}), (\underline{4}, \underline{33})\}, \{(\underline{4}, \underline{22})\}, \{(\underline{4}, \underline{24})\}, \{(\underline{4}, \underline{34})\}, \{(\underline{4}, \underline{44})\}. \end{aligned}$$

As in Section 2, an underlined state represents a degenerate belief that supports a state (and, therefore, all states in the same element of the partition). A state is crossed out if there is no such belief.<sup>27</sup> If the same procedure is followed for players 2 and 3, then additional states are eliminated. For example, state  $(2, 33)$  is eliminated because, viewing player 2 as player 1, it was established above that  $(3, 23)$  is eliminated. At the end of this step ( $k = 0$ ), the remaining states (i.e., the set  $\Gamma(\Omega)$ ) are given by  $(2, 3)$ ,  $(3, 2)$ ,  $(4, 3)$ ,  $(3, 4)$ , and  $(3, 3)$ .

Next, consider step  $k = 1$ . For the remaining states, player 1's partition becomes

$$\{(\underline{2}, \underline{3}), \{(\underline{4}, \underline{3}), \{(\underline{3}, \underline{2}), (\underline{3}, \underline{4})\}, \{(\underline{3}, \underline{3})\}.$$

Now, states  $(2, 3)$  and  $(4, 3)$  cannot be justified by incorrect beliefs (since these incorrect beliefs violate the requirement that player 1 believes that the other player is rational and consistent) and must be eliminated. A symmetric argument for the other player eliminates  $(3, 2)$  and  $(3, 4)$ . Then  $\Gamma^2(\Omega) = \{(3, 3)\}$  and  $(3, 3)$  is the unique RCE in the two-player location game. Thus, the prediction that both players choose region 3 is robust to uncertainty about the number of players.<sup>28</sup>

### 5.2 Robust comparative statics under Cournot competition

There are  $n$  firms competing by simultaneously choosing quantities  $x_i \in \mathbb{R}_+$  to maximize profits  $u_i(x, \theta) = (p(x, \theta) - c) \times x_i$ , where  $x \in \mathbb{R}_+^n$  is the vector of firm quantities,  $c \in (0, 1)$  is marginal cost,

$$p(x, \theta) = \max \left\{ 1 - \frac{1}{\theta} \sum_{i=1}^n x_i, 0 \right\}$$

is the inverse demand function, and  $\theta \in [1, 2]$  is the demand parameter.

Firm  $i$ 's best response is given by

$$\text{BR}_i(x_{-i}, \theta) = \max \left\{ \frac{1}{2} \left( (1 - c)\theta - \sum_{j \neq i} x_j \right), 0 \right\},$$

<sup>27</sup>Note that states  $(3, 23)$  and  $(3, 34)$  are not crossed out if  $b = 1$ , since in that case a belief that puts equal probability on each state makes player 1 indifferent between locations 2, 3, and 4.

<sup>28</sup>There are no states remaining for the three-player game, which illustrates the well known fact that when attention is restricted to pure strategies, even the set of Nash equilibria may be empty. Allowing for mixed strategies, however, does not change the result that  $(3, 3)$  is the unique RCE in the two-player game, provided that players have correct beliefs about their payoff from each of the actions in the support of their mixed strategy.



and it is easy to obtain the unique Nash equilibrium quantities  $x_i^{\text{NE}} = (1 - c)\theta/(n + 1)$  for all  $i$  and price  $p^{\text{NE}} = 1 - n(1 - c)/(n + 1)$ .

Two Nash predictions are noteworthy. First, as  $n \rightarrow \infty$ , price converges to marginal cost and profits vanish. Hence, the Cournot environment provides a foundation for competitive equilibrium. Second, an increase in marginal cost (say, due to higher taxes) leads to an increase in prices. These Nash predictions implicitly rely on the assumption that firms know the demand parameter and have correct beliefs about each others' equilibrium quantities. The objective is to evaluate these predictions when firms are uncertain about demand and where only beliefs about own profits (or, equivalently, market price) are restricted to be correct.

A simple observation is that the competitive foundation result must hold when beliefs about equilibrium prices are correct: if price is observed to remain sufficiently above marginal cost with a large number of firms, there is at least one firm that is producing sufficiently little and will realize that it can increase its profit by producing slightly more.<sup>29</sup> Therefore, I focus on the less obvious question of comparative statics with respect to cost. For simplicity, suppose that  $n = 2$ , so that there are only two firms in the industry.

In this case, it is well known that the Nash equilibrium is the unique rationalizable outcome. However, this result relies on the assumption that the true demand parameter  $\theta^0$  is common belief. In contrast, I assume that it is common belief that  $\theta^0 \in [1, 2]$  and then apply the RCE framework to find equilibria that can be supported by beliefs that satisfy this restriction.

Formally,  $I = \{1, 2\}$ ,  $X_1 = X_2 = \mathbb{R}_+$ ,  $\Theta = [1, 2]$ ,  $\theta^0 \in \Theta$  denotes the true demand parameter,  $u_i$  is provided above, and  $\Omega = X_1 \times X_2 \times \Theta$  is the state space. The choice of  $\Omega$  implies that players are uncertain about strategies and demand, but the rest of the game (e.g., the cost  $c$ ) is "common knowledge." As shown below, it is without loss of generality to restrict the strategy spaces to  $X_i = [0, 2]$ .<sup>30</sup>

Finally, for each  $\omega \in \Omega$ ,

$$P_i(\omega) = \{\omega' \in \Omega : x_i(\omega) = x_i(\omega'), p(\omega') = p(\omega)\}.$$

Therefore, players' beliefs about their own quantities and the market price are restricted to be correct in equilibrium. Still, player  $i$  cannot disentangle  $x_{-i}$  from  $\theta$  given  $x_i$  and  $p$ . However, she may refine her beliefs by using the additional information that the other player is rational and consistent, and so on. The objective is to find the set of  $k$ -RCE and RCE strategy profiles and prices as a function of the true fundamental  $\theta^0$  and cost  $c$ . In particular, do equilibrium prices increase with  $c$ ?

*Iteration over state space.* Each cell of firm  $i$ 's partition over  $\Omega$  is uniquely identified by the firm's quantity and the market price  $(x_i, p)$ . So that cell  $(x_i, p)$  survives a step of

<sup>29</sup>Formally, fix any  $n$  with RCE price  $p^n$  and quantity  $x^n$ . Suppose that  $x_i^n < p^n - c$  for some  $i$ . Then  $i$  knows that increasing production by  $y$  increases profits by  $((p^n - c) - x_i^n/\theta)y - y^2/\theta$ . Since the increase is positive for sufficiently small  $y > 0$  (for any belief over  $\theta$ ), then  $x_i^n \geq p^n - c$  for all  $i$ . But then, replacing in the demand function,  $p^n \leq 1 - n(p^n - c)/\theta$ ; hence  $p^n \leq (1 + nc/\theta)/(1 + n/\theta) \leq (1 + 0.5nc)/(1 + 0.5n) \leq (1 + 0.5n_\varepsilon c)/(1 + 0.5n_\varepsilon) = c + \varepsilon$  for all  $n \geq n_\varepsilon \equiv 2(1 - c - \varepsilon)/\varepsilon$ .

<sup>30</sup>The game is then compact and continuous, and the RCE characterization applies (see Section 6.1).

the iterative procedure, firm  $i$  must believe  $(x_{-i}^*, \theta^*)$  such that<sup>31</sup> (i)  $x_i$  is a best response,

$$x_i = \frac{1}{2}((1-c)\theta^* - x_{-i}^*), \quad (3)$$

and (ii) beliefs are consistent with price  $p$ ,

$$p = 1 - \frac{1}{\theta^*}(x_i + x_{-i}^*). \quad (4)$$

Equations (3) and (4) can be solved to obtain the unique belief  $(x_{-i}^*, \theta^*)$  that can rationalize the cell  $(x_i, p)$ :

$$\theta^*(x_i, p) = \frac{x_i}{p-c} \quad (5)$$

and

$$x_{-i}^*(x_i, p) = x_i \left( \frac{1+c-2p}{p-c} \right). \quad (6)$$

In the previous step of the iterative procedure, suppose that the cells with market price  $p$  that have survived (for both players, since the game is symmetric) are those where quantity belongs to  $[\underline{x}, \bar{x}]$ . It then follows that in the current step, the quantities that survive for price  $p$  are those  $x_i$  such that  $\theta^*(x_i, p) \in \Theta = [1, 2]$  and  $x_{-i}^*(x_i, p) \in [\underline{x}, \bar{x}]$ . Equations (5) and (6) then imply that surviving  $x_i$ 's satisfy  $x_i \in [p-c, 2(p-c)]$  and  $x_i \in [\underline{x}, \bar{x}] \times (p-c)/(1+c-2p)$ .<sup>32</sup> The first of these conditions is constant for each step, while the second depends on the set of  $x_i$ 's that survive the previous step.

For example, at step  $k=0$ , the interval  $[\underline{x}, \bar{x}]$  in the second condition is the set of all quantities  $[0, +\infty)$  (or  $[0, 2]$  in the compactified game). Because quantities are positive, only cells with prices  $c < p < p^M$  survive this step, where  $p^M = (1+c)/2$  is the monopoly price. The first condition implies that the quantities that survive for each of these  $p$  are  $[p-c, 2(p-c)]$  and the second condition provides no additional restriction given that  $\bar{x} \geq 2$ .<sup>33</sup> These surviving quantities are then used in the next iteration to replace  $[\underline{x}, \bar{x}]$  in the second condition.

Let  $\Phi^k(p; c)$  denote the set of quantities  $x_i$  such that the cells indexed by  $(x_i, p)$  survive the  $k$ th iteration when cost is  $c$ . By iterating the process started above from the set of quantities  $[0, +\infty)$  (or  $[0, 2]$ ), it follows that

$$\begin{aligned} \Phi^k(p; c) &= [p-c, 2(p-c)] \cap [p-c, 2(p-c)] \times \frac{p-c}{1+c-2p} \cap \dots \\ &\quad \cap [p-c, 2(p-c)] \times \left( \frac{p-c}{1+c-2p} \right)^k. \end{aligned} \quad (7)$$

In particular,  $\Phi^k(p; c)$  may be empty, meaning that no cells with price  $p$  survive step  $k$ .

<sup>31</sup>The arguments by [Moulin \(1979\)](#) are easily adapted to this context to show that it is without loss of generality to restrict attention to degenerate beliefs.

<sup>32</sup>It is understood that  $[a, b] \times c \equiv [a \times c, b \times c]$ .

<sup>33</sup>The second condition requires  $x_i \leq \bar{x}((p-c)/(1+c-2p))$ , which is not binding because  $c < p < (1+c)/2$  and  $\bar{x} \geq 2$  imply that  $2(p-c) < \bar{x}((p-c)/(1+c-2p))$ .

*RCE.* Recall that  $p^{\text{NE}} = (1 + 2c)/3$  is the unique Nash equilibrium price. By characterization of RCE,  $p$  is a  $k$ -RCE price of the game with true parameter  $\theta^0$  if and only if there exist  $x_1$  and  $x_2$  such that

$$p = 1 - \frac{1}{\theta^0}(x_1 + x_2) \quad \text{and} \quad x_i \in \Phi^k(p; c) \quad \text{for } i = 1, 2. \quad (8)$$

First, consider the case of RCE, which, since the game is compact and continuous, is obtained as the limit  $k \rightarrow \infty$  (see [Section 6.1](#)). Then  $\Phi^k(p)$  becomes empty for all  $p \neq p^{\text{NE}}$  and  $\Phi^k(p^{\text{NE}}; c) = [p^{\text{NE}} - c, 2(p^{\text{NE}} - c)] = [(1/3)(1 - c), (2/3)(1 - c)]$  for all  $k$ . Therefore, the Nash equilibrium price is also the unique RCE price.<sup>34</sup>

Second, consider the case where firms engage in only a finite number of levels of introspection  $k$ . Equation (7) can be written as  $\Phi^k(p; c) = [l(p; c), h(p; c)]$ , where

$$l(p; c) = \begin{cases} p - c & \text{if } p \leq p^{\text{NE}} \\ (p - c) \left( \frac{p - c}{1 + c - 2p} \right)^k & \text{if } p \geq p^{\text{NE}} \end{cases}$$

and

$$h(p; c) = \begin{cases} 2(p - c) \left( \frac{p - c}{1 + c - 2p} \right)^k & \text{if } p \leq p^{\text{NE}} \\ 2(p - c) & \text{if } p \geq p^{\text{NE}} \end{cases}$$

By (8), it follows that  $p$  is a  $k$ -RCE price of the game with true parameter  $\theta^0$  if and only if

$$p = 1 - \frac{1}{\theta^0}(x_1 + x_2) \in \left[ 1 - \frac{1}{\theta^0}2h(p; c), 1 - \frac{1}{\theta^0}2l(p; c) \right].$$

Since the extreme points in the above interval are continuous and decreasing in  $p$ , the set of equilibrium prices is given by a compact interval  $[p_L(c, \theta^0, k), p_H(c, \theta^0, k)]$ . Since  $p^{\text{NE}}$  is always an equilibrium price, then the lowest and highest equilibrium prices are below and above  $p^{\text{NE}}$ , respectively. Therefore, the lowest equilibrium price solves

$$p_L = 1 - \frac{1}{\theta^0}4(p_L - c) \left( \frac{p_L - c}{1 + c - 2p_L} \right)^k$$

and the highest solves

$$p_H = 1 - \frac{1}{\theta^0}2(p_H - c) \left( \frac{p_H - c}{1 + c - 2p_H} \right)^k.$$

Finally, by the implicit function theorem, both  $p_L(\cdot)$  and  $p_H(\cdot)$  are increasing in  $c$  for all  $\theta^0$  and  $k$ ; therefore, equilibrium prices are increasing (in the strong set order) in a  $k$ -RCE. Thus, for any level  $k$  of introspection, the standard comparative statics prediction is robust to relaxing the Nash assumption as long as firms are required to have correct beliefs about equilibrium price.

<sup>34</sup>Note, however, that RCE quantities need not be Nash: of course, as established by [Claim 1](#), Nash quantities are also RCE quantities.

### 5.3 Uncertainty about precision of information

Consider the version of the sales-pitch game given in [Example 3](#) of [Section 2](#). In this game, the payoff parameter  $\theta$  is randomly drawn with probability  $\Pr(\theta_a) = p^0$ , seller 2 knows the realization, and seller 1 is uninformed. The Nash assumption implicitly requires players to know this information structure. Is the Nash prediction of this game robust to seller 1's uncertainty about seller 2's information about  $\theta$ ?

Formally, let  $I = \{1, 2\}$  be the set of players and suppose that after  $\theta$  is realized, seller 2 observes a private signal  $s \in \{s_a, s_b\}$ , where  $\Pr(s_a|\theta_a) = \Pr(s_b|\theta_b) = q$ . Let  $X_1 = [0, 1]$  and  $X_2 = [0, 1] \times [0, 1]$  be the sets of strategies for each player, where  $x_1 \in X_1$  denotes the probability that seller 1 pitches and  $x_2 = (x_{2a}, x_{2b}) \in X_2$ , where  $x_{2s}$  denotes the probability that seller 2 pitches after observing signal  $s$ . In particular, players are allowed to play mixed strategies in this example.<sup>35</sup>

A fundamental is now a pair  $(p, q) \in \Theta^* \subseteq [0, 1]^2$ , where  $p$  is the probability of  $\theta_a$  and  $q$  is the precision of seller 2's information. The state space is then  $\Omega = X_1 \times X_2 \times \Theta^*$ . This representation captures the restriction that it is common belief that  $(p, q) \in \Theta^*$ , thus potentially relaxing the assumption that the information structure is common knowledge. Utility functions are extended in a straightforward manner to the domain  $\Omega$ , and  $u_i^*(\omega)$  denotes the random payoff of player  $i$  in state  $\omega$ . Each player is assumed to have correct beliefs about her own payoff distribution and player 2 is assumed to know the precision of her signals by letting

$$P_1(\omega) = \{\omega' : u_1^*(\omega) \stackrel{D}{=} u_1^*(\omega'), x_1(\omega) = x_1(\omega')\}$$

and

$$P_2(\omega) = \{\omega' : u_2^*(\omega) \stackrel{D}{=} u_2^*(\omega'), q(\omega) = q(\omega'), x_2(\omega) = x_2(\omega')\}$$

be the information partitions, where  $\stackrel{D}{=}$  indicates that two random variables have the same distribution.<sup>36</sup>

For concreteness, suppose that

$$O_a = -1 < L = 0 < O_b = 1 < H = 2$$

for seller 1 and

$$L_a = -1 < O = 0 < H_a = 1 < L_b = 2 < H_b = 3$$

<sup>35</sup>More generally, mixed strategies can be handled in the following way. If  $A_i$  is the set of player  $i$ 's pure strategies and  $U_i : \times_{i \in I} A_i \rightarrow \mathbb{R}$  is her utility function, then let  $X_i$  be the set of all probability measures on  $A_i$  and define  $u_i : X \rightarrow \mathbb{R}$  to be the expected value of  $U_i$  under the probability measure over  $\times_{i \in I} A_i$  induced by  $x \in X$ . As a referee mentioned, this interpretation of a mixed strategy as a pure strategy in an enlarged strategy space may be controversial, since it requires players to have beliefs about the *actual mixing* of their opponents. An alternative approach is to follow [Harsanyi \(1973\)](#) and introduce (small) payoff perturbations to purify the mixed strategies.

<sup>36</sup>The results are equivalent under the stronger (and perhaps more realistic) assumption that each seller has correct beliefs about her payoffs from playing each of the actions in the support of her strategy.

for seller 2. In addition, suppose that the true fundamental is given by

$$0 < p^0 < 1/4, \quad q^0 = 1;$$

in particular, seller 2 receives a perfect signal. Finally, let  $\Theta^* = [0, 0.5] \times [0.5, 1]$ , so that it is common belief that  $p \in [0, 0.5]$  and  $q \in [0.5, 1]$ .

*Nash equilibrium.* The objective game, with fundamental  $(p^0, q^0)$ , has a unique Nash equilibrium,  $x_1^{\text{NE}} = 0$ ,  $x_2^{\text{NE}} = (1, 1)$ , where seller 1 never pitches and seller 2 always pitches. To see this claim, first note that  $x_{2b}^{\text{NE}} = 1$  because it is a dominant strategy for seller 2 to pitch under  $\theta_b$ . Then seller 1's payoff is

$$p^0(x_{2a}L + (1 - x_{2a})H) + (1 - p^0)L = 2p^0(1 - x_{2a})$$

if she pitches and

$$p^0O_a + (1 - p^0)O_b = 1 - 2p^0$$

if she does not. Not pitching is a best response for all  $x_{2a}$  because  $p^0 < 1/4$ . Finally, seller 2's best response to  $x_1^{\text{NE}} = 0$  under  $\theta_a$  is  $x_{2a}^{\text{NE}} = 1$ .

*RCE.* Some simple shortcuts simplify the iterative procedure that characterizes RCE. Let  $(x_1^*, (x_{2a}^*, x_{2b}^*), (p^0, q^0))$  be a state that survives the iterative procedure, thus implying that  $(x_1^*, (x_{2a}^*, x_{2b}^*))$  is an RCE of the current game. Because seller 2 knows her information precision ( $q^0 = 1$ ), then  $x_{2b}^* = 1$ . Because seller 1 has correct beliefs about her own payoff from following  $x_1^*$  and because her realized payoffs from pitching and not pitching are different, she can then identify her true payoff from each of her actions (hence, best respond to the true state) provided that she plays each action with positive probability, i.e.,  $x_1^* \in (0, 1)$ . But, as argued above when obtaining the Nash equilibrium, mixing is not a correct best response to any  $x_{2a}^*$  when  $x_{2b}^* = 1$ . Therefore,  $x_1^* \in \{0, 1\}$ . Finally, since seller 2 always pitches under  $\theta_b$ , then he can identify  $x_1^*$  by observing whether he gets  $L_b$  or  $H_b$ . Therefore, seller 2 must play a best response:  $x_{2a}^* = 1$  if  $x_1^* = 0$  and  $x_{2a}^* = 0$  if  $x_1^* = 1$ . Consequently, the only non-Nash profile that could potentially be an RCE is

$$x_1^* = 1, \quad x_2^* = (0, 1),$$

where only seller 1 has misperceptions.

Next, let  $(1, (x_{2a}, x_{2b}), (p, q))$  be a state that represents seller 1's beliefs and supports  $x^* = (x_1^*, x_2^*)$  as a non-Nash RCE. By characterization of RCE, such a state exists if and only if  $(x_{2a}, x_{2b})$  and  $(p, q)$  satisfy the three conditions

$$\text{1C: } p(qx_{2a} + (1 - q)x_{2b}) + (1 - p)((1 - q)x_{2a} + qx_{2b}) = 1 - p^0$$

$$\text{1R: } p \geq 1/2 - p^0$$

$$\text{2BR: } x_{2s} = \begin{cases} 0 \\ \text{anything} \\ 1 \end{cases} \quad \text{if } \Pr(\theta_a|s) \begin{cases} > \\ = \\ < \end{cases} 2/3,$$

where

$$\Pr(\theta_a|s_a) = \frac{qp}{qp + (1 - q)(1 - p)} \quad \text{and} \quad \Pr(\theta_a|s_b) = \frac{(1 - q)p}{(1 - q)p + q(1 - p)}.$$

Condition 1C requires consistency of beliefs for seller 1. Since  $x_1^* = 1$ , seller 1's payoff is a random variable taking values  $L$  and  $H$ . Under  $(x^*, (p^0, q^0))$ , the probability of  $L$  is  $1 - p^0$ , and 1C requires the probability of  $L$  under  $(1, (x_{2a}, x_{2b}), (p, q))$  to be the same. Condition 1R requires  $x_1^* = 1$  to be a best response for seller 1 to  $(1, (x_{2a}, x_{2b}), (p, q))$ . Seller 1 thinks she would obtain  $pO_a + (1 - p)O_b$  from not pitching, which must be lower than what she obtains by pitching under state  $(x^*, (p^0, q^0))$  (and, by consistency, for any state  $(1, (x_{2a}, x_{2b}), (p, q))$  satisfying 1C),  $p^0H + (1 - p^0)L$ .<sup>37</sup> Finally, 2BR captures the restriction, argued above, that surviving states must be such that seller 2 best responds to the true state.<sup>38</sup>

Condition 1C rules out  $x_{2a} = x_{2b} = 0$  and  $x_{2a} = x_{2b} = 1$ . Condition 2BR implies that  $x_{2a} \leq x_{2b}$  but rules out the case  $x_{2a} = x_{2b} \in (0, 1)$ , because that case would imply  $q = 1/2$  and  $p = 2/3$ , which contradicts  $p \in [0, 0.5]$ . Therefore, there are three remaining cases to consider.

First, suppose that  $x_2 = (0, 1)$ . Then 1C becomes

$$p(1 - q) + (1 - p)q = 1 - p^0.$$

The highest possible  $p \leq 1/2$  that satisfies this condition is  $p = p^0$  (by setting  $q = 1$ ). But then 1R cannot be satisfied, because  $p^0 < 1/2 - p^0$  given that  $p^0 < 1/4$ . Similarly, suppose that  $x_2 = (0, \alpha)$ , with  $\alpha \in (0, 1)$ . Then the highest possible  $p \leq 1/2$  that satisfies 1C is (strictly) lower than  $p^0$ , thus leading again to a contradiction when 1R is required.

Finally, suppose that  $x_2 = (\beta, 1)$ , with  $\beta \in (0, 1)$ . Then it is easy to see that  $p = 1/2$ ,  $q = 2/3$ , and  $\beta = 1 - 2p^0$  satisfy conditions 1C, 1R, and 2BR. Therefore,  $x_1^* = 1$ ,  $x_2^* = (0, 1)$  is a non-Nash RCE for the game where  $0 < p^0 < 1/4$  and  $q^0 = 1$ . Thus, in this example, the Nash prediction is not robust to seller 1's uncertainty about the precision of seller 2's information. A final observation is that the Nash prediction would have been robust if mixed strategies had not been allowed. Therefore, whether mixed strategies are permitted may affect whether a non-Nash *pure-strategy* profile is an RCE.

## 6. DISCUSSION

### 6.1 Compact and continuous games

Sections 3 and 4 restricted attention to finite spaces for simplicity, but, as shown next, all the results extend in a straightforward manner to the class of compact and continuous games (all examples in this paper belong to this class).

The framework in Section 3 is now required to satisfy the following conditions: To allow for infinite belief spaces, let  $T$  be a topological space and let  $\xi$  be a measurable

<sup>37</sup>The condition becomes  $p \geq (O_b - (p^0H + (1 - p^0)L))/(O_b - O_a)$ , which reduces to 1R after appropriate substitutions.

<sup>38</sup>Conditional on observing  $s$ , seller 2 obtains  $\Pr(\theta_a|s)L_a + \Pr(\theta_b|s)L_b = 2 - 3\Pr(\theta_a|s)$  from pitching and  $O = 0$  from not pitching.



function.<sup>39</sup> Correspondingly, the definition of equilibrium is modified to require  $(RC^{\mathcal{B}, \mathcal{P}})^j \subseteq T$  to be measurable for all  $j \leq k$  (for  $k$ -RCE) and for all  $j$  (for RCE).<sup>40</sup> Throughout this section, attention is restricted to the following class of games.

**DEFINITION 6.** The game  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  is *compact and continuous* if  $(X_i)_{i \in I}$  and  $\Theta$  are compact, metrizable topological spaces,  $u_i: X \times \Theta \rightarrow \mathbb{R}$  is jointly continuous for each  $i$ , and  $P_i: \Omega \rightarrow \Omega$  has the closed-graph property for each  $i$ .

It is convenient to define the correspondences  $C_i: \Omega \rightarrow \Delta(\Omega)$ , where

$$C_i(\omega) = \{\delta_i \in \Delta(\Omega) : x_i(\omega) \in \Phi_i(\delta_i), \delta_i[P_i(\omega)] = 1\}.$$

In words,  $C_i(\omega)$  is the set of beliefs over  $\Omega$  that are consistent with feedback arising from play of  $\omega$  and that make  $x_i(\omega)$  an optimal strategy. For a compact (hence Borel-measurable) set  $A$ , the equilibrium operator defined in Section 4 can be written as

$$\Gamma(A) = \{\omega \in \Omega : \forall i \exists \delta_i \in C_i(\omega) \cap \Delta(A)\}. \quad (9)$$

**LEMMA 3.** *Let  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  be a compact and continuous game. Then (i) the correspondence  $C_i$  has the closed-graph property and (ii) if  $A \subseteq \Omega$  is a compact set, then  $\Gamma(A)$  is a compact set.*

**PROOF.** (i) Let  $\delta_i^k \in \Delta(\Omega)$  and  $\omega^k \in \Omega$  be convergent sequences with  $\delta_i^k \rightarrow \delta_i$  and  $\omega^k \rightarrow \omega$  such that  $\delta_i^k \in C_i(\omega^k)$  for all  $k$ . The proof shows that  $\delta_i \in C_i(\omega)$ . First, note that  $\delta_i^k \in \Delta(P_i(\omega^k))$  for all  $k$ . Since, by assumption, the correspondence  $P_i(\cdot)$  has the closed-graph property, then the fact that  $\Omega$  is compact and Theorem 17.13 in Aliprantis and Border (2006) imply that the correspondence  $\Delta(P_i(\cdot))$  also has the closed-graph property. Therefore,  $\delta_i \in \Delta(P_i(\omega))$ . Second, note that  $x_i(\omega^k) \in \Phi_i(\delta_i^k)$  for all  $k$ . Since  $(x_i, \delta) \mapsto \int_{X_{-i} \times \mathcal{M}} u_i(x_i, x_{-i}, \mu) d\text{marg}_{X_{-i} \times \mathcal{M}} \delta$  is continuous (by continuity of  $u_i$  and the weak\* topology; see, e.g., Aliprantis et al. 2006) and since  $X_i$  is compact, the theorem of the maximum implies that  $\Phi_i$  has the closed-graph property. Therefore,  $x_i(\omega) \in \Phi_i(\delta_i)$ . By the two previous arguments,  $\delta_i \in C_i(\omega)$ .

(ii) The proof shows that  $\Gamma(A)$  is a closed set; hence also compact. Let  $\omega^k \in \Gamma(A)$ ,  $\omega^k \rightarrow \omega$  be a convergent sequence in  $\Gamma(A)$ . Fix any player  $i$ . Then by (9), there exists  $\delta_i^k \in C_i(\omega^k) \cap \Delta(A)$ . Since  $A$  is compact, then  $\Delta(A)$  is also compact (see Theorem 15.11 in Aliprantis and Border 2006). Since the sequence  $\delta_i^k$  lives in a compact space, then there exists a convergent subsequence  $\delta_i^{k_m} \rightarrow \delta_i$ . Since  $\delta_i^{k_m} \in \Delta(A)$  for all  $k_m$ , then the fact that  $\Delta(A)$  is compact implies that  $\delta_i \in \Delta(A)$ . Finally, since by part (i),  $C_i$  has the closed-graph property, then  $\delta_i \in C_i(\omega)$ . Since the above is true for all  $i$ , then  $\omega \in \Gamma(A)$ .  $\square$

<sup>39</sup>For a topological space  $Z$ , its  $\sigma$ -field is the Borel  $\sigma$ -field and  $\Delta(Z)$  is the set of all Borel probability measures, endowed with the weak\* topology.

<sup>40</sup>As shown in the working paper version of the paper, the results also hold without these additional measurability restrictions.

Since compact sets are Borel-measurable, an immediate corollary of Lemma 3(ii) is that  $\Gamma^j(\Omega)$  is measurable for all  $j$ . In particular, the proof of Theorem 1 carries through. The proof of Theorem 2 also carries through, provided that  $\bigcap_{k=1}^{\infty} \Gamma^k(\Omega) \subseteq \Gamma(\bigcap_{k=1}^{\infty} \Gamma^k(\Omega))$  still holds in this more general setting. As shown below, this result follows by adapting Bernheim's (1984) arguments for the set of rationalizable strategies.

LEMMA 4. *Let  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  be a compact and continuous game. Then  $\bigcap_{k=1}^{\infty} \Gamma^k(\Omega) \subseteq \Gamma(\bigcap_{k=1}^{\infty} \Gamma^k(\Omega))$ .*

PROOF. Let  $\omega \in \bigcap_{k=1}^{\infty} \Gamma^k(\Omega)$  and fix any player  $i$ . From (9), for each  $k$ , there exists  $\delta_i^k \in C_i(\omega) \cap \Delta(\Gamma^k(\Omega))$ . By Lemma 3(i),  $C_i(\omega) \subseteq \Omega$  is a closed set; hence  $C_i(\omega)$  is a compact set because  $\Omega$  is compact. By Lemma 3(ii),  $\Gamma^k(\Omega)$  is a compact set for all  $k$ ; hence  $\Delta(\Gamma^k(\Omega))$  is a compact set for all  $k$ . In addition,  $\{\Gamma^k(\Omega)\}_k$  is a sequence of nested sets, implying that  $\Delta(\Gamma^k(\Omega))$  is a sequence of compact, nested sets. Putting these results together,  $\{C_i(\omega) \cap \Delta(\Gamma^k(\Omega))\}_k$  is a sequence of compact, nested, nonempty sets; hence, their intersection is nonempty, so that there exists  $\delta_i$  such that  $\delta_i \in C_i(\omega)$  and  $\delta_i \in \bigcap_{k=1}^{\infty} \Delta(\Gamma^k(\Omega)) = \Delta(\bigcap_{k=1}^{\infty} \Gamma^k(\Omega))$ . Since such  $\delta_i$  exists for all  $i$ , then  $\omega \in \Gamma(\bigcap_{k=1}^{\infty} \Gamma^k(\Omega))$ .  $\square$

As shown by Lipman (1994) for common belief of rationality, Theorem 2 does not hold for more general classes of games, where transfinite iterations may be required to characterize the solution. In the working paper version, I show that for any game,  $\mathcal{E} = \bigcap_{\alpha \in \text{ON}} \Gamma^\alpha(\Omega)$ , where ON denotes the ordinals.

## 6.2 Comparison to Harsanyi's BNE

The standard way to capture player's uncertainty about the game,  $\Theta$ , was first formalized by Harsanyi. In his framework, the description of the game is completed by specifying a Harsanyi type space  $\langle T^H, \{h_i\}_{i=1, \dots, n} \rangle$ , where  $T^H = \Theta \times T_1^H \times \dots \times T_n^H$  is the set of states of nature and profiles of types, and  $h_i: T_i^H \rightarrow \Delta(T^H)$ . A Harsanyi type space represents an entire hierarchy of beliefs *over the fundamental space*  $\Theta$ . These beliefs are exogenously specified, in the sense that they are not determined in equilibrium, but rather are fixed while conducting equilibrium analysis.<sup>41</sup> A Harsanyi strategy for player  $i$  is a mapping  $\sigma_i$  from own types  $T_i^H$  to the action set  $X_i$ . A strategy profile  $\sigma = (\sigma_1, \dots, \sigma_n)$  is then a BNE if each type best responds to others' equilibrium strategies given its beliefs about the fundamentals and the types of other players, and a BNE prediction of play for type  $t_i^H$  is given by  $\sigma_i(t_i^H)$ .<sup>42</sup> As illustrated by the example in Section 2, the equilibrium hierarchy of beliefs need not satisfy certain restrictions, such as requiring that type  $t_i^H$  has correct beliefs about the equilibrium payoff she obtains if everyone plays according to equilibrium.

<sup>41</sup>In the example of Section 2, the assumption that it is commonly believed that player 1 believes  $\theta_a$  and player 2 believes  $\theta_b$  is formally captured by the type space:  $T_1^H = \{t_1^H\}$ ,  $T_2^H = \{t_2^H\}$  and  $h_1(t_1^H)(\{\theta_a, t_1^H, t_2^H\}) = h_2(t_2^H)(\{\theta_b, t_1^H, t_2^H\}) = 1$ .

<sup>42</sup>Of course, only the first-order beliefs matter for optimality, but, for example, second-order beliefs may matter to justify first-order beliefs and so on.

The objective of this paper is to introduce a framework that can incorporate *desirable restrictions* on the beliefs that are allowed to support equilibrium strategies, motivated by viewing equilibrium as the steady state of a learning process. In principle, one could use Harsanyi's framework to answer this question as follows: Fix a Harsanyi type space and find the set of BNE's. For each BNE and each type profile, there is a unique prediction over play of the game and there is a corresponding hierarchy of beliefs supporting this prediction. Next check whether the hierarchy of beliefs satisfies the desired restrictions. If it does, then the prediction is said to be an *equilibrium* (in the sense defined in this paper). Follow this procedure for every possible type space and the end result is a set of RCE profiles.<sup>43</sup>

The purpose of the procedure outlined in the previous paragraph is to relate the objective of this paper to the familiar notion of BNE. However, while the main question can be posed in the context of Harsanyi's framework, the above procedure goes through unnecessary steps. In particular, the step of finding a BNE is not very useful if one still needs to check whether the resulting hierarchy of beliefs satisfies the desired conditions.

Alternatively, I follow a more direct approach. Hierarchies of beliefs are defined not just on the space of fundamentals, as in Harsanyi's framework, but actually over the products of strategy spaces and the space of fundamentals,  $X \times \Theta$ , as in epistemic models. A type space is now an *epistemic* type space  $\langle T, \{g_i\}_{i=1,2} \rangle$ , where  $T = X \times \Theta \times T_1 \times \dots \times T_n$  is the set of epistemic states and  $g_i: T_i \rightarrow \Delta(T)$ .<sup>44</sup> The procedure for finding the set of equilibria can now be described as follows: Fix an epistemic type space  $\langle T, \{g_i\}_{i=1,2} \rangle$  (which is not a Harsanyi type space) and for each  $t = (x_1, \dots, x_n, \theta, t_1, \dots, t_n)$ , ask whether the hierarchies of beliefs given by  $(t_1, \dots, t_n)$  satisfy the desired equilibrium conditions when actual equilibrium play is given by  $(x_1, \dots, x_n)$  and the true fundamental is  $\theta$ . If the equilibrium conditions are satisfied, then  $(x_1, \dots, x_n)$  is said to be an equilibrium of game  $\theta$ . Then follow this procedure over all possible epistemic type spaces.

A key feature of this paper is to operationalize the meaning of *desirable restrictions* on the entire hierarchy of beliefs in a manner that is transparent, practical, and relates to a learning interpretation of equilibrium. The latter objective is achieved by requiring beliefs to be restricted as a function of *actual* equilibrium outcomes. Transparency is obtained by being explicit about which equilibrium outcomes players must have correct beliefs about. And practicality is achieved by requiring such endogenous restrictions to be mutually believed by the players, therefore providing a natural extension of these restrictions to higher-order beliefs.

To conclude, a few differences with Harsanyi's BNE are noteworthy. First, in contrast to Harsanyi's BNE, the (epistemic) type space is not part of the description of a game. The type space is only a tool that is used to formalize the desired equilibrium restrictions. Equilibrium is defined in terms of type spaces, but equilibrium is *not* defined for a *given* type space.

<sup>43</sup>This procedure works for the RCE outcomes, but misses the  $k$ -RCE outcomes that fail rationalizability, because all BNE outcomes are rationalizable.

<sup>44</sup>This epistemic type space, which is a particular kind of belief space, is chosen to facilitate comparison to Harsanyi's type space.

Second, it is important to distinguish between a belief type and a private information signal. In particular, the set of strategies is fixed as a primitive of the game and, unlike BNE, strategies are *not* mappings from types to action sets.<sup>45</sup> The reason is that under BNE, a strategy captures two features: (i) players' beliefs about other players' actions (or private-information contingent plan of actions in games with asymmetric information) for each possible game (i.e., fundamental) and (ii) player's actual actions for a particular true fundamental. Under RCE, both of these features are also present. However, it is now important to distinguish these two features, because only the second feature is used to restrict beliefs given a feedback partition. Thus, a strategy now represents only the second feature and a belief system is used to formalize the first feature.

Third, the description of the game exogenously specifies a *set* of feasible beliefs, but the equilibrium beliefs about *both* strategies and fundamentals are jointly determined in equilibrium. In contrast, in a BNE, beliefs about fundamentals are fixed at the outset by specifying a particular type space.

### 6.3 Relationship to other solution concepts

Rationalizability and rationalizable conjectural equilibrium are often defined or characterized in terms of a “best response” set. A similar approach can be carried out in this paper.

**DEFINITION 7.** A set  $\hat{\Omega} \subseteq \Omega$  is  $\mathcal{P}$ -rationalizable if, for all  $i \in I$  and  $\omega \in \hat{\Omega}$ , there exists  $\delta_{i\omega} \in \Delta(\Omega)$  such that (i)  $\delta_{i\omega}(P_i(\omega)) = 1$ , (ii)  $x_i(\omega) \in \Phi_i(\delta_{i\omega})$ , and (iii)  $\delta_{i\omega}(\hat{\Omega}) = 1$ .

**CLAIM 3.** A strategy profile  $x \in X$  is an RCE of  $\mathcal{G}(\theta, \Omega, \mathcal{P})$  if and only if there exists a  $\mathcal{P}$ -rationalizable set  $\hat{\Omega}$  that contains  $(x, \theta)$ .

**Claim 3** is a straightforward consequence of the characterization of equilibrium in **Section 4** and facilitates the comparison to other solution concepts. First, consider the case where  $\Theta = \{\theta\}$ , so that players face no structural uncertainty. If  $P_i(\omega) = \Omega$  for all  $i \in I$  and  $\omega \in \Omega$ , then condition (i) in **Definition 7** is nonbinding and the set of RCE's is equivalent to the set of (correlated) rationalizable outcomes. For more general partitions, the definition of RCE coincides with **Rubinstein and Wolinsky's (1994)** definition of RCE. Finally, if condition (iii) from **Definition 7** is eliminated, then the solution concept coincides with the simultaneous-move game version of conjectural or self-confirming equilibrium (**Battigalli 1987, Fudenberg and Levine 1993a**) or, equivalently, 0-RCE in this paper.

Second, consider the case where players do face structural uncertainty—the main contribution of this paper. **Dekel et al. (2004)** extend self-confirming equilibrium to games with asymmetric information, but they do not consider restrictions on higher-order beliefs. **Battigalli (2003)** and **Battigalli and Siniscalchi (2003)** define rationalizability in games of incomplete information without a Harsanyi type space, while

<sup>45</sup>In games with asymmetric information, strategies *are* mappings from private information to actions; see Examples 2 and 3 in Section 2.

Ely and Pęski (2006) and Dekel et al. (2007) define rationalizability in games with a Harsanyi type space (see Battigalli et al. 2011 for the relationship between these approaches). In this paper, I follow the former approach in dispensing with the notion of a Harsanyi type space and, in addition, incorporate feedback restrictions on equilibrium beliefs that refine rationalizability. I propose to capture games with strategic and structural uncertainty by defining a particular notion of a game (and its corresponding objective version) that allows the modeler to weaken several standard assumptions on players' beliefs. I then extend and characterize Rubinstein and Wolinsky's (1994) notion of RCE to games with structural uncertainty.

Finally, Claim 3 could have been the starting point to define RCE. However, since higher-order beliefs are not formally expressed, it is not clear whether this is the "right" definition. Clarity is an advantage of the epistemic framework used to define equilibrium in Section 3.

## 7. CONCLUSION

This paper provides an equilibrium framework for games with structural and strategic uncertainty that relaxes the Nash assumption that players have correct beliefs and integrates learning from both feedback (in the spirit of self-confirming equilibrium) and introspection (in the spirit of rationalizability). The main difference with respect to Harsanyi's framework is that the distinct treatment of strategic and structural uncertainty is eliminated, and beliefs about both of these elements are jointly restricted in equilibrium. The characterization result shows that the RCE framework, while grounded on abstract epistemic foundations, is applicable by following a straightforward iterative procedure over the product of the set of strategies and structural uncertainty. Finally, since some of the restrictions on equilibrium beliefs imposed by RCE are motivated by a learning story, further work should explore explicit learning foundations.<sup>46</sup>

## REFERENCES

- Aliprantis, Charalambos D. and Kim C. Border (2006), *Infinite Dimensional Analysis*, third edition. Springer-Verlag, Berlin. [493]
- Aliprantis, Charalambos D., Dionysius Glycopantis, and Daniela Puzzello (2006), "The joint continuity of the expected payoff functions." *Journal of Mathematical Economics*, 42, 121–130. [493]
- Aumann, Robert J. (1974), "Subjectivity and correlation in randomized strategies." *Journal of Mathematical Economics*, 1, 67–96. [479]
- Aumann, Robert J. (1987), "Correlated equilibrium as an expression of Bayesian rationality." *Econometrica*, 55, 1–18. [468]

<sup>46</sup>Bernheim (1984) shows that Cournot dynamics lead to rationalizable strategies (see also Moulin 1984) and Fudenberg and Levine (1993b) and Fudenberg and Kreps (1995) provide learning foundations for self-confirming equilibrium in games without structural uncertainty.

- Aumann, Robert J. and Adam Brandenburger (1995), “Epistemic conditions for Nash equilibrium.” *Econometrica*, 63, 1161–1180. [468]
- Battigalli, Pierpaolo (1987), *Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali*. Unpublished thesis, Università Bocconi. [468, 496]
- Battigalli, Pierpaolo (1999), “A comment on non-Nash equilibria.” Unpublished paper. [469, 470]
- Battigalli, Pierpaolo (2003), “Rationalizability in infinite, dynamic games with incomplete information.” *Research in Economics*, 57, 1–38. [470, 471, 496]
- Battigalli, Pierpaolo, Alfredo Di Tillio, Edoardo Grillo, and Antonio Penta (2011), “Interactive epistemology and solution concepts for games with asymmetric information.” *The B.E. Journal of Theoretical Economics*, 11(1). [470, 471, 484, 497]
- Battigalli, Pierpaolo and Amanda Friedenberg (2012), “Forward induction reasoning revisited.” *Theoretical Economics*, 7, 57–98. [471]
- Battigalli, Pierpaolo and Danilo Guitoli (1997), “Conjectural equilibria and rationalizability in a game with incomplete information.” In *Decision, Games and Markets* (Pierpaolo Battigalli, Aldo Montesano, and Fausto Panunzi, eds.), 97–124, Kluwer, Dordrecht. [470]
- Battigalli, Pierpaolo and Andrea Prestipino (forthcoming), “Transparent restrictions on beliefs and forward-induction reasoning in games with asymmetric information.” *The B.E. Journal of Theoretical Economics*. [471]
- Battigalli, Pierpaolo and Marciano Siniscalchi (2003), “Rationalization and incomplete information.” *The B.E. Journal of Theoretical Economics*, 3(1). [469, 470, 471, 472, 496]
- Battigalli, Pierpaolo and Marciano Siniscalchi (2007), “Interactive epistemology in games with payoff uncertainty.” *Research in Economics*, 61, 165–184. [471]
- Bergemann, Dirk and Stephen Morris (2005), “Robust mechanism design.” *Econometrica*, 73, 1771–1813. [471]
- Bergemann, Dirk and Stephen Morris (2007), “Belief free incomplete information games.” Discussion Paper 1629, Cowles Foundation for Research in Economics, Yale University. [471]
- Bernheim, B. Douglas (1984), “Rationalizable strategic behavior.” *Econometrica*, 52, 1007–1028. [469, 494, 497]
- Brandenburger, Adam and Eddie Dekel (1987), “Rationalizability and correlated equilibria.” *Econometrica*, 55, 1391–1402. [468, 469, 470]
- Brandenburger, Adam and Eddie Dekel (1993), “Hierarchies of beliefs and common knowledge.” *Journal of Economic Theory*, 59, 189–198. [469]
- Brandenburger, Adam and Amanda Friedenberg (2008), “Intrinsic correlation in games.” *Journal of Economic Theory*, 141, 28–67. [479]



- Brandenburger, Adam and Amanda Friedenberg (2010), "Self-admissible sets." *Journal of Economic Theory*, 145, 785–811. [471]
- Brandenburger, Adam, Amanda Friedenberg, and H. Jerome Keisler (2008), "Admissibility in games." *Econometrica*, 76, 307–352. [471]
- Carlsson, Hans and Eric van Damme (1993), "Global games and equilibrium selection." *Econometrica*, 61, 989–1018. [471]
- Dekel, Eddie, Drew Fudenberg, and David K. Levine (1999), "Payoff information and self-confirming equilibrium." *Journal of Economic Theory*, 89, 165–185. [470]
- Dekel, Eddie, Drew Fudenberg, and David K. Levine (2004), "Learning to play Bayesian games." *Games and Economic Behavior*, 46, 282–303. [468, 470, 471, 480, 496]
- Dekel, Eddie, Drew Fudenberg, and Stephen Morris (2007), "Interim correlated rationalizability." *Theoretical Economics*, 2, 15–40. [484, 497]
- Ely, Jeffrey C. and Marcin Pęski (2006), "Hierarchies of belief and interim rationalizability." *Theoretical Economics*, 1, 19–65. [484, 497]
- Esponda, Ignacio (2008a), "Behavioral equilibrium in economies with adverse selection." *American Economic Review*, 98, 1269–1291. [468, 471]
- Esponda, Ignacio (2008b), "Information feedback in first price auctions." *RAND Journal of Economics*, 39, 491–508. [468]
- Eyster, Erik and Matthew Rabin (2005), "Cursed equilibrium." *Econometrica*, 73, 1623–1672. [471]
- Feinberg, Yossi and Andrzej Skrzypacz (2005), "Uncertainty about uncertainty and delay in bargaining." *Econometrica*, 73, 69–91. [471]
- Friedenberg, Amanda and Martin Meier (2012), "The context of the game." Unpublished paper, Arizona State University. [484]
- Fudenberg, Drew (2006), "Advancing beyond 'Advances in behavioral economics'." *Journal of Economic Literature*, 44, 694–711. [471]
- Fudenberg, Drew and Yuichiro Kamada (2011), "Rationalizable partition-confirmed equilibrium." Unpublished paper. [470]
- Fudenberg, Drew and David M. Kreps (1988), "A theory of learning, experimentation, and equilibrium in games." Unpublished paper. [480]
- Fudenberg, Drew and David M. Kreps (1995), "Learning in extensive-form games I. Self-confirming equilibria." *Games and Economic Behavior*, 8, 20–55. [497]
- Fudenberg, Drew and David K. Levine (1993a), "Self-confirming equilibrium." *Econometrica*, 61, 523–545. [468, 477, 480, 496]
- Fudenberg, Drew and David K. Levine (1993b), "Steady state learning and Nash equilibrium." *Econometrica*, 61, 547–573. [497]

- Gale, David (1953), "A theory of  $n$ -person games with perfect information." *Proceedings of the National Academy of Sciences of the United States of America*, 39, 496–501. [469]
- Gilli, Mario (1999), "On non-Nash equilibria." *Games and Economic Behavior*, 27, 184–203. [469]
- Harsanyi, John C. (1967–1968), "Games with incomplete information played by 'Bayesian' players, parts I, II, and III." *Management Science*, 14, 159–182, 320–334, and 486–502. [467, 469]
- Harsanyi, John C. (1973), "Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points." *International Journal of Game Theory*, 2, 1–23. [490]
- Heifetz, Aviad and Zvika Neeman (2006), "On the generic (im)possibility of full surplus extraction in mechanism design." *Econometrica*, 74, 213–233. [471]
- Heifetz, Aviad and Dov Samet (1998), "Topology-free typology of beliefs." *Journal of Economic Theory*, 82, 324–341. [469]
- Jehiel, Philippe (2005), "Analogy-based expectation equilibrium." *Journal of Economic Theory*, 123, 81–104. [471]
- Jehiel, Philippe (2011), "Manipulative auction design." *Theoretical Economics*, 6, 185–217. [468]
- Jehiel, Philippe and Frédéric Koessler (2008), "Revisiting games of incomplete information with analogy-based expectations." *Games and Economic Behavior*, 62, 533–557. [471]
- Lipman, Barton L. (1994), "A note on the implications of common knowledge of rationality." *Games and Economic Behavior*, 6, 114–129. [494]
- Liu, Qingmin (2009), "On redundant types and Bayesian formulation of incomplete information." *Journal of Economic Theory*, 144, 2115–2145. [484]
- Luce, R. Duncan and Howard Raiffa (1957), *Games and Decisions*. Wiley, New York. [469]
- Mertens, Jean-François and Shmuel Zamir (1985), "Formulation of Bayesian analysis for games with incomplete information." *International Journal of Game Theory*, 14, 1–29. [469]
- Molinari, M. Cristina (1991), *Equilibrio, informazione e apprendimento nella teoria economica*. Unpublished thesis, Università Bocconi. [469]
- Morris, Stephen, Andrew Postlewaite, and Hyun Song Shin (1995), "Depth of knowledge and the effect of higher order uncertainty." *Economic Theory*, 6, 453–467. [471]
- Morris, Stephen and Hyun Song Shin (1998), "Unique equilibrium in a model of self-fulfilling currency attacks." *American Economic Review*, 88, 587–597. [471]
- Moulin, Hervé (1979), "Dominance solvable voting schemes." *Econometrica*, 47, 1337–1351. [469, 488]

- Moulin, Hervé (1984), “Dominance solvability and Cournot stability.” *Mathematical Social Sciences*, 7, 83–102. [497]
- Nash, John F. (1951), “Non-cooperative games.” *The Annals of Mathematics*, 54, 286–295. [467, 476]
- Neeman, Zvika (2004), “The relevance of private information in mechanism design.” *Journal of Economic Theory*, 117, 55–77. [471]
- Pearce, David G. (1984), “Rationalizable strategic behavior and the problem of perfection.” *Econometrica*, 52, 1029–1050. [469]
- Rubinstein, Ariel (1989), “The electronic mail game: Strategic behavior under ‘almost common knowledge’.” *American Economic Review*, 79, 385–391. [470]
- Rubinstein, Ariel and Asher Wolinsky (1994), “Rationalizable conjectural equilibrium: Between Nash and rationalizability.” *Games and Economic Behavior*, 6, 299–311. [468, 470, 496, 497]
- Sadzik, Tomasz (2011), “Beliefs revealed in Bayesian-Nash equilibrium.” Unpublished paper. [484]
- Tan, Tommy C.-C. and Sérgio Ribeiro da Costa Werlang (1988), “The Bayesian foundations of solution concepts of games.” *Journal of Economic Theory*, 45, 370–391. [468]
- Vickrey, William (1961), “Counterspeculation, auctions, and competitive sealed tenders.” *Journal of Finance*, 16, 8–37. [477]
- Weinstein, Jonathan and Muhamet Yildiz (2007a), “Impact of higher-order uncertainty.” *Games and Economic Behavior*, 60, 200–212. [471]
- Weinstein, Jonathan and Muhamet Yildiz (2007b), “A structure theorem for rationalizability with application to robust predictions of refinements.” *Econometrica*, 75, 365–400. [471]