Mathematical Social Sciences 91 (2018) 6–16

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/mss

Common belief in approximate rationality

Angie Mounir [a,*], Andrés Perea [a], Elias Tsakas [b]

[a] Maastricht University, EpiCenter and Department of Quantitative Economics, P.O. Box 616, 6200 MD Maastricht, The Netherlands
[b] Maastricht University, EpiCenter and Department of General Economics, P.O. Box 616, 6200 MD Maastricht, The Netherlands

HIGHLIGHTS

- We allow for the possibility that players may err when making a choice.
- Players' error margins are their private information.
- Lower bounds of probabilities assigned to margins of errors are common knowledge.
- Common belief in F-rationality is a generalization of common belief in rationality.

ARTICLE INFO

Article history:
Received 21 July 2016
Received in revised form 20 July 2017
Accepted 8 October 2017
Available online 23 October 2017

ABSTRACT

This paper substitutes the standard rationality assumption with approximate rationality in normal form games. We assume that players believe that their opponents might be ε-rational, i.e. willing to settle for a suboptimal choice, and so give up an amount ε of expected utility, in response to the belief they hold. For every player i and every opponents' degree of rationality ε, we require player i to attach at least probability $F_i(\varepsilon)$ to his opponent being ε-rational, where the functions F_i are assumed to be common knowledge amongst the players. We refer to this event as belief in F-rationality. The notion of Common Belief in F-Rationality ($CBFR$) is then introduced as an approximate rationality counterpart of the established Common Belief in Rationality. Finally, a corresponding recursive procedure is designed that characterizes those beliefs players can hold under $CBFR$.

1. Introduction

Rationality of players in situations of strategic interaction has been a crucial axiom upon which the vast majority of game theoretic concepts are based. A player is rational if he only plays optimal choices, which are those that maximize his expected utility given his beliefs about opponents' choices. Specifying the set of choices any rational player can make in a normal form game has therefore been the central question many existing solution concepts attempted to answer. Rationalizability, epistemically characterized by rationality and Common Belief in Rationality (CBR) (Pearce, 1984; Bernheim, 1984; Brandenburger and Dekel, 1987; Tan and Werlang, 1988), is a crucial one of such concepts. Under rationality and CBR, all players are rational, believe in their opponents' rationality, and so on. Choices that can be made under CBR are those that survive Iterated Elimination of Strictly Dominated Choices (IESDC).

Despite its solid epistemic foundations, experimental findings in certain well-known games came at odds with rationality and CBR. Examples of such games include the Traveler's Dilemma and Guess 2/3 of the Average (Nagel, 1995; Becker et al., 2005). This discrepancy between experimental outcomes and theoretical ones triggered the need to develop new theoretical models of reasoning that provide better theoretical foundations for observed experimental behavior in these games. One such aspect highlighted in experiments is that players may err and/or believe that their opponents might make mistakes.

One way to introduce such mistakes is by replacing the standard notion of rationality with ε-rationality, originally introduced by Radner (1980). The basic underlying idea is that players may settle for suboptimal choices, as long as the utility induced by these choices is sufficiently close to the utility induced by the optimal ones. Formally, a choice is ε-rational given a belief about the opponents' choices, whenever the expected utility of this choice (given the belief) is at most ε away from the optimal expected utility (given the same belief). This idea initially attracted a lot of attention, as it allowed us to explain cooperation in finitely repeated prisoner's dilemma or in finitely repeated principal–agent games (e.g., see Radner, 1981), which is not possible

* Corresponding author.
E-mail addresses: a.mounir@maastrichtuniversity.nl (A. Mounir), a.perea@maastrichtuniversity.nl (A. Perea), e.tsakas@maastrichtuniversity.nl (E. Tsakas).

https://doi.org/10.1016/j.mathsocsci.2017.10.001
0165-4896/

using more standard solution concepts such as Nash equilibrium or Common Belief in Rationality. Still the behavioral foundations of ε-rationality remained in the background.

The recent surge of behavioral economics, as well as work in other disciplines, has provided basis for assuming that players may sometimes be ε-rational. For instance, in their seminal work, March and Simon (1958) introduced the notion of "satisficing" which replaced the usual concept of "optimizing". In our context, this would mean that players may settle for suboptimal choices as long as these are sufficiently close to the optimal one, thus making them satisfactory. Another example is the well-known Weber–Fechner–Stevens set of laws from psychophysics, according to which people may fail to perceive (small) differences when they compare two choices, thus implying that the size of ε represents the player's cognitive constraints. In either case, Baye and Morgan (2004) provide empirical evidence supporting ε-rationality when compared to standard rationality.

The notion of ε-rationality originally appeared in the definition of Radner's ε-equilibrium concept, while more recently, Dekel et al. (2006) introduced the concept of ε-interim correlated rationalizability. Both ε-equilibrium and ε-interim correlated rationalizability assume the value of each player's ε to be transparent with players having common belief therein. This is relaxed in our model, as we assume players to be ε-rational with the value of each player's ε being his private information. Thus players in our model not only account for the possibility of potential mistakes on their opponents' part, but also perceive the exact margin of error their opponents might have as uncertain. Therefore, players form beliefs about the extent of their opponents' potential errors. In an attempt to restrict such uncertainty, players adopt a certain lower bound to the beliefs they could form about their opponents' error margins. Such lower bound is captured by the commonly believed weakly increasing functions $F_i : [0, \infty) \to [0, 1]$. For every player i, the function F_i assigns a minimum bound to the probability that player i would assign to each potential level of the opponents' ε.

Thus the function F_i does not specify the exact belief player i assigns to each of his opponents' possible extents of rationality, but serves only as a lower bound for such beliefs. The lower the extent of irrationality ε, the lower the minimum bound on probability $F_i(\varepsilon)$ each player must assign to the event of his opponents having a margin of error of at most ε. For every possible extent of irrationality ε of player i, there is a set of choices that could be ε-rationally made by player i given some belief about his opponents' choices. The uncertainty about the opponents' extent of irrationality can be translated into restrictions on the probabilities that could be assigned to different *sets* of opponents' choice combinations. This translation of uncertainty into restrictions on probabilities assigned to sets of opponents' choice combinations plays a crucial role throughout the paper.

For every given level ε of the opponents' error margin, a lower bound of F_i means player i assigns at least probability $F_i(\varepsilon)$ to the event of his opponents' willingness to give up an amount of ε or less in terms of expected utility. We refer to this event as belief in F-rationality. Establishing common belief among all players in that event, this paper introduces Common Belief in F-Rationality. Notably, this framework captures the intuitive uncertainty players might have about the extent to which their opponents might deviate from their optimal choices. Moreover, adopting F-rationality offers a generalization to some of the existing concepts in the literature such as ε rationality and, of course, common belief in rationality. It also helps linking some existing concepts in the literature such as Common Belief in ε Rationality and Common p-Belief in Rationality.

Models and concepts in the literature that are in the same spirit as ours, include Rosenthal's t-Solution (Rosenthal, 1989), Quantal Response Equilibria (McKelvey and Palfrey, 1995) and Utility Proportional Beliefs (Bach and Perea, 2014). All of these models recognize that players might assign positive probabilities to opponents' suboptimal choices, none however explicitly defines the underlying reason to be the presence of a potential error margin. CBFR, just like all of the three aforementioned models, allows players to assign positive probabilities to suboptimal choices of their opponents'. However, CBFR differs in that positive probabilities need not be assigned to all opponents' choices. Moreover, in all three of the models mentioned, better choices receive higher probabilities, and in case of utility proportional beliefs and Rosenthal's t-Solution these probabilities are proportional to the utility generated by these choices[1] . In our case, restrictions are imposed on the probabilities assigned to *groups* of choices rather than on the probabilities assigned to each individual choice. Furthermore, unlike CBFR, both Rosenthal's t-Solution and Quantal Response Equilibria are equilibrium concepts.

After the concept is introduced, the recursive procedure of Iterated Elimination of F-Dominated Beliefs (IEFB) is developed, which characterizes exactly those first-order beliefs that can be held under common belief in F-rationality. Furthermore, a fixed point characterization of the concept of CBFR is provided. The paper also considers two special cases. One case is that in which a specific value of ε for each player is given full probability, or Common Belief in ε-rationality. The second special case is where there is common belief that each player assigns at least p to the event of his opponent being 0-rational, or Common Belief in p-Belief in Rationality. Finally, the model is applied to the n-price Traveler's Dilemma game to demonstrate the potential behavioral implications of our epistemic solution concept. The paper is divided into six sections. Section 3 defines Common Belief in F-Rationality. Section 3 presents the recursive procedure of Iterated Elimination of F-Dominated Beliefs. Section 4 covers the two special cases. Section 5 uses a 3-price Traveler's Dilemma game to illustrate how the recursive procedure works and then summarizes some general results for the n-price game. Section 6 is a discussion and conclusion of the paper.

2. Common belief in *F*-rationality

Consider an n-player finite normal form game $\Gamma = (\mathcal{I}, C, U)$, where $\mathcal{I} = \{1, \ldots, n\}$ is the finite set of players, $C = \{C_1, \ldots, C_n\}$ is an n-tuple of finite sets of choices and $U = \{U_1, \ldots, U_n\}$ where $u_i : C_i \times C_{-i} \to \Re$ is the utility function of player i. Let $b_i \in \triangle(C_{-i})$ be a belief player i holds about his opponents' choice combinations, where $\triangle(C_{-i})$ is the set of probability distributions on $(C_1 \times \cdots C_{i-1} \times C_{i+1} \times \cdots C_n)$. Expected utility[2] $u_i(c_i, b_i)$ of the choice $c_i \in C_i$ is then the utility of that choice given the belief b_i player i holds about his opponents' choices, i.e.,

$$u_i(c_i, b_i) := \sum_{c_{-i} \in C_{-i}} b_i(c_{-i}) \times u_i(c_i, c_{-i}).$$

We assume every player to have a margin of error ε of zero or more. Moreover, we assume these error margins to be every player's private information. Let F_i be a weakly increasing function $F_i : [0, \infty) \to [0, 1]$ held by player i and $F = (F_1, \ldots, F_n)$. For every player i, F_i characterizes a lower bound for the belief that player can hold about each of his opponents' potential values of ε. For every given level ε of the opponents' error margin, a lower bound of F_i means player i assigns at least probability $F_i(\varepsilon)$ to the event

1 More precisely, in both utility proportional beliefs and Rosenthal's t-Solution, the differences in probabilities assigned to opponents' choices are proportional to the differences in utilities generated by these choices.

2 Note that we are using the same notation u_i to refer to both the utility of player i and the expected utility of choice c_i of player i given belief b_i. The former has the form $u_i : C_i \times C_{-i} \to \Re$, while the latter is written as $u_i(c_i, b_i)$.

of his opponents' willingness to give up an amount of ε or less in terms of expected utility. Respecting this lower bound, the player is said to believe in his opponents' F-rationality. Common Belief in F-Rationality implies that every player i believes in his opponents' F-rationality, believes his opponents believe in i's F-rationality, and so on. A formal characterization of CBFR, therefore, requires the use of infinite belief hierarchies, which can be intelligently defined within an epistemic model. Let M be a finite epistemic model assigning to every player i a finite set of types T_i. Every type $t_i \in T_i$ holds a belief $b_i(t_i) \in \triangle(C_{-i} \times T_{-i})$ which is a probability distribution over the opponents' choice-type combinations.

As usual, each $t_i \in T_i$ induces an infinite belief hierarchy, with $b_i^1(t_i) \in \triangle(C_{-i})$ being the first-order belief held by t_i (see for instance Heifetz and Samet, 1998 for more details on how an epistemic type induces a belief hierarchy). Moreover, let $u_i(c_i, t_i)$ be the expected utility of choice c_i given the first-order belief held by type t_i, i.e., $u_i(c_i, t_i) := u_i(c_i, b_i^1(t_i))$. A choice-type pair (c_i, t_i) is then ε-rational if c_i is ε-optimal given $b_i^1(t_i)$.

Definition 1. Let $\varepsilon \geq 0$. A choice-type pair (c_i, t_i) of player i is ε-rational, if for all $c_i' \in C_i$,

$$u_i(c_i, b_i^1(t_i)) \geq u_i(c_i', b_i^1(t_i)) - \varepsilon.$$

A type t_i believes in the opponents' F-rationality if for every $\varepsilon \geq 0$ it assigns at least probability $F_i(\varepsilon)$ to opponents' ε-rational choice-type combinations (c_{-i}, t_{-i}). Note that player i's opponents' choice-type combination $(c_{-i}, t_{-i}) = \big((c_1, t_1), \ldots (c_{i-1}, t_{i-1}), (c_{i+1}, t_{i+1}), \ldots, (c_n, t_n)\big)$ is said to be ε-rational, if every individual opponent's choice combination (c_j, t_j) is ε-rational where $j \neq i$. To formally define belief in opponents' F-rationality, let $R_{-i}^{\varepsilon} := \{(c_{-i}, t_{-i}) \in (C_{-i} \times T_{-i}) \mid c_{-i} \text{ is } \varepsilon\text{-optimal for } b_{-i}^1(t_{-i})\}$.

Definition 2. A type t_i believes in the opponents' F-rationality if for every $\varepsilon \geq 0$,

$$b(t_i)(R_{-i}^{\varepsilon}) \geq F_i(\varepsilon).$$

The above definition illustrates how the function F_i serves as a lower bound of the actual belief player i can hold about his opponents' extent of rationality. The distribution over players—i's extent of rationality induced by the belief held by t_i must assign to each of the opponents' possible ε a probability of at least $F_i(\varepsilon)$. For the purpose of defining types expressing k-fold, as well as common belief in F-rationality, let FR_i^k be the set of player i's types expressing k-fold belief in F-rationality, where $k \geq 1$. A type t_i believing in the opponents' F-rationality is said to express 1-fold belief in F-rationality, i.e., $t_i \in FR_i^1$. Moreover, for any $k \geq 2$ type t_i is said to express k-fold belief in F-rationality, written as $t_i \in FR_i^k$, if it only assigns positive probability to opponents' type combinations $t_{-i} \in FR_{-i}^{k-1}$ expressing $(k-1)$-fold belief in F-rationality. Finally, type t_i expresses common belief in F-rationality, denoted by $t_i \in CBFR_i$, if it expresses k-fold belief in F-rationality for all $k \geq 1$.[3]

Definition 3. Sets of types expressing k-fold belief in F-rationality and common belief in F-rationality can be defined recursively by the following sequence:

$$FR_i^1 := \{t_i \in T_i | b_i(t_i)(R_{-i}^{\varepsilon}) \geq F_i(\varepsilon), \ \forall \ \varepsilon \geq 0\};$$

$$FR_i^k := \{t_i \in T_i | b_i(t_i)(C_{-i} \times FR_{-i}^{k-1}) = 1\};$$

and

$$CBFR_i := \bigcap_{k=1}^{\infty} FR_i^k.$$

Note that Common Belief in Rationality (CBR) is a special case of CBFR, with $F = F^*$ where $F^*(\varepsilon) = 1$ for every $\varepsilon \geq 0$.[4] Furthermore, it is worth noting that our model might bear some similarity to the established $\triangle$-rationalizability of Battigalli and Siniscalchi (2003). One main aspect of similarity is that both impose exogenous restrictions on beliefs and establish common belief therein. However, one crucial difference is that the exogenous restrictions in case of $\triangle$-rationalizability are imposed on first-order beliefs, while in our model the restriction is imposed on second-order beliefs. The main restriction in our model is outlined in Definition 2. For a type t_i to believe in the opponents' F-rationality, it has to assign at least $F_i(\varepsilon)$ in probability to the event $R_{-i}(\varepsilon)$ for every $\varepsilon \geq 0$. Thus our main restriction is imposed on the probability player i assigns to his opponents' choice-type combinations. For any $(c_{-i}, t_{-i}) \in R_{-i}(\varepsilon)$, it must be that c_{-i} is ε-optimal for the respective first-order beliefs held by t_{-i}. Hence the restriction imposed by Definition 2 is on the probability player i assigns to sets of combinations of his opponents' choices and first-order beliefs. This core difference between our model and $\triangle$-rationalizability implies that the former cannot be a special case of the latter.

3. Iterated elimination of F-Dominated Beliefs

Now that Common Belief in F-Rationality has been defined, this section develops the recursive procedure of "Iterated Elimination of F-Dominated Beliefs (IEFB)", to characterize those first-order beliefs that can be held by epistemic types expressing common belief in F-rationality. IEFB does so by translating the basic uncertainty each player has about his opponents' rationality, bounded by $(F_i)_{i\in\mathcal{I}}$, into restrictions on the set of potential beliefs about opponents' choices. Thus, IEFB eliminates beliefs rather than choices, with every round k of the procedure resulting in a new restricted feasible belief set B_i^k for each player i and where $B_i^0 = \triangle(C_{-i})$.

Notably, IEFB differs from Common Belief in Rationality (CBR) in that the former eliminates beliefs which may or may not result in the elimination of some choice-combinations, while the latter proceeds directly to eliminating choice-combinations with the elimination of the belief implied. So for some choice-combination $c_{-i} \in C_{-i}$, IEFB could result in eliminating $b_i^1(c_{-i}) > a$ as a potential first-order belief, while maintaining $b_i^1(c_{-i}) \leq a$ as a valid one, where $0 \leq a \leq 1$. For Correlated Rationalizability, on the other hand, assigning positive probability to a choice-combination is either kept feasible for any a or eliminated for all a, i.e., $a \in \{0, 1\}$.

[3] Note that throughout the paper, for notation simplicity and without loss of generality, we consider finite epistemic models. Still, notice that our construction can be generalized to a complete epistemic model, i.e. an epistemic model that induces all belief hierarchies (Friedenberg, 2010). The details of such extension are not presented in this paper, but can nevertheless be provided upon request.

[4] Notice that for every sequence $\{F_k\}_{k=1}^{\infty}$ that satisfies $F_k(\varepsilon) \uparrow 1$ for all $\varepsilon \geq 0$, it is the case that $\bigcap_{k=1}^{\infty} F_k R_i^1 = F^* R_i$, i.e., the sequence of sets of i's types that satisfy one fold belief in F_k-rationality converges to the set of i's beliefs that satisfy one-fold belief in F^*-rationality. This follows from the definition of the set of types expressing one-fold belief in F-rationality formalized in Definitions 2 and 3. Then, we can inductively prove that $\bigcap_{k=1}^{\infty} F_k R_i^m = F^* R_i^m$ for every $m \geq 1$. Indeed, formally, observe that

$$\begin{aligned}\bigcap_{k=1}^{\infty} F_k R_i^2 &= \bigcap_{k=1}^{\infty} \{t_i \in T_i | b_i(t_i)(C_{-i} \times F_k R_{-i}^1) = 1\} \\ &= \{t_i \in T_i | b_i(t_i)(C_{-i} \times \bigcap_{k=1}^{\infty} F_k R_{-i}^1) = 1\} \\ &= \{t_i \in T_i | b_i(t_i)(C_{-i} \times F^* R_{-i}^1) = 1\} \\ &= F^* R_i^2,\end{aligned}$$

and likewise for $m > 2$. Hence, $\bigcap_{k=1}^{\infty} CBF_k R = CBF^* R = CBR$, thus implying that the sequence of beliefs in $\{CBF^k R\}_{k=1}^{\infty}$ converges to the beliefs in CBR.

IEFB can therefore be thought of as a generalization of Correlated Rationalizability, as Theorem 2 shows. Note that Correlated Rationalizability in turn is equivalent to Iterated Elimination of Strictly Dominated Choices (IESDC).

This section is divided into two subsections. Section 3.1 formally introduces the recursive procedure and links it to IESDC. Section 3.2 then considers some practical implementation matters of the procedure.

3.1. The recursive procedure

Before we proceed to describing the steps of the recursive procedure, we define $\Gamma^k = (C, B^k; u)$ to be the belief restricted game resulting from round k of the procedure, i.e. it is the belief restricted game round $k+1$ of the procedure starts from where C and u are as defined above and $B^k = \{B_1^k, \ldots, B_n^k\}$. Recall that $B_i^k \subseteq \triangle(C_{-i})$ is the feasible set of player i after k rounds of the recursive procedure and that $B_i^0 = \triangle(C_{-i})$. Moreover, note that $\Gamma^0 = (C, B^0; u)$ is simply the original game Γ as every $B_i^0 = \triangle(C_{-i})$ by definition. Recall that for player i to believe in his opponents' F-rationality, he must assign a probability of at least $F_i(\varepsilon)$ to the event of his opponents' having a margin of error of at most ε. This can be translated into probability distributions over opponents' choice combinations, by assigning at least $F_i(\varepsilon)$ to opponents' choice combinations c_{-i} that are ε-optimal for some belief. Define the set

$$C_i^\varepsilon(\Gamma^k) = \{c_i \in C_i \mid \exists\, b_i \in B_i^k \text{ s.t. } u_i(c_i, b_i) \geq u_i(c_i', b_i) - \varepsilon \quad \forall c_i' \in C_i\}.$$

The first step of the recursive procedure states that player i must assign a probability of at least $F_i(\varepsilon)$ to the set of choice combinations $C_{-i}^\varepsilon(\Gamma^0)$, for every $\varepsilon \geq 0$. Beliefs satisfying the conditions $b_i(C_{-i}^\varepsilon(\Gamma^0)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$ constitute the feasible belief set $B_i^1 \subseteq \triangle(C_{-i})$. The set B_i^1 contains those first-order beliefs player i can hold while believing in his opponents' F-rationality. Applying this for all players i concludes the first round of the procedure of Iterated Elimination of F-Dominated Beliefs. Under up to 2-fold belief in F-rationality, player i not only believes in his opponents' F-rationality, but also believes that his opponents $-i$ believe in i's F-rationality, making the relevant belief restricted game $\Gamma^1 = (C, B^1; u)$. Consequently, the sets $(C_i^\varepsilon(\Gamma^0))_{i\in\mathcal{I}}$ require an update, as the feasible belief sets of some players are now potentially smaller. In general, $C_i^\varepsilon(\Gamma^1) \subseteq C_i^\varepsilon(\Gamma^0)$ for every player $i \in \mathcal{I}$.

The new set $C_{-i}^\varepsilon(\Gamma^1)$ in turn implies new restrictions on the feasible belief set B_i^1 of player i of the form $b_i(C_{-i}^\varepsilon(\Gamma^1)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$. These restrictions then characterize the new set B_i^2 of first-order beliefs player i can feasibly hold under up to 2-fold belief in F-rationality. The new restricted belief sets B_j^2 of each player $j \neq i$ can be obtained in a similar manner. The recursive procedure proceeds by iteratively eliminating beliefs and stops when $B_i^k = B_i^{k-1}$ for every player i, which may or may not happen after finitely many steps. In the latter case feasible belief sets only stabilize in the limit. The steps of round k of Iterated Elimination of F-Dominated Beliefs are summarized below.

Procedure 1. *Iterated elimination of F-dominated beliefs:*

- *Initial step: Define $\Gamma^0 = (C, B^0; u)$*
- *Inductive step: Assume that $\Gamma^{k-1} = (C, B^{k-1}; u)$ has been defined. Then $\Gamma^k = (C, B^k; u)$ is the game where for each player i*

  $$B_i^k := \{b_i \in \triangle(C_{-i}) \mid b_i(C_{-i}^\varepsilon(\Gamma^{k-1})) \geq F_i(\varepsilon)\ \forall \varepsilon \geq 0\}.$$

Theorem 1 shows that the recursive procedure characterizes exactly those first-order beliefs that can be held by a type expressing common belief in F-rationality.

Theorem 1. *A belief $b_i \in \triangle(C_{-i})$ can be held by a type $t_i \in T_i$ expressing common belief in F-rationality iff it survives all rounds of the recursive procedure of Iterated Elimination of F-Dominated Beliefs.*

It is worth noting that even though the recursive procedure may not stop after finitely many rounds, the resulting sets of feasible beliefs $B_i^\infty = \bigcap_{k\geq 0} B_i^k$ surviving all rounds of the procedures are always nonempty. This is shown in Corollary 1, which is based on Theorem 2 linking our recursive procedure to Correlated Rationalizability. Let C_i^k be the choices of player i surviving k rounds of the procedure of iterated elimination of never-best replies characterizing Correlated Rationalizability.

Theorem 2. *For every $k \geq 0$ and every player $i \in \mathcal{I}$, the following holds:*

$\triangle(C_{-i}^k) \subseteq B_i^k$.

Note that for any finite static game, there must be some $k \geq 0$ for which $\triangle(C_{-i}^k) = \triangle(C_{-i}^{k-1})$. This in turn implies the nonemptiness of the limit set of first-order beliefs surviving the procedure (Corollary 1).

Corollary 1. *The limit of the feasible belief sets B_i^∞ for every player i is always nonempty.*

Theorem 3 provides a fixed point characterization of the limit sets $(B_i^\infty)_{i\in\mathcal{I}}$ of the procedure. Let $\Phi_i \subseteq \triangle(C_{-i})$ be a closed and convex set of first-order beliefs of player i. Moreover, let $C_i^\varepsilon(\Phi_i) = \{c_i \in C_i | \exists \varphi_i \in \Phi_i \ s.t.\ u_i(c_i, \varphi_i) \geq u_i(c_i', \varphi_i) - \varepsilon \ \forall c_i' \in C_i\}$. Note that by definition, if $\Phi_i = B_i^k$ for some round k of the procedure, then $C_i^\varepsilon(\Phi_i) = C_i^\varepsilon(\Gamma^k)$. We say that the collection $(\Phi_i)_{i\in\mathcal{I}}$ is a best-response set if for every player i,

$$\Phi_i \subseteq \{\varphi_i \in \triangle(C_{-i}) | \varphi_i(C_{-i}^\varepsilon(\Phi_{-i})) \geq F_i(\varepsilon)\ \forall \varepsilon \geq 0\}.$$

Theorem 3 shows that the limit sets $(B_i^\infty)_{i\in\mathcal{I}}$ of first-order beliefs formed by the procedure constitute a best-response set.

Theorem 3. *Let $(B_i^\infty)_{i\in\mathcal{I}}$ be the limit set of first-order beliefs resulting from the procedure of Iterated Elimination of F-Dominated Beliefs, then $(B_i^\infty)_{i\in\mathcal{I}}$ is a best response set.*

Furthermore, Theorem 4 shows that the limit set $(B_i^\infty)_{i\in\mathcal{I}}$ resulting from the procedure is also the maximal best-response set. Thus for any best response $(\Phi_i)_{i\in\mathcal{I}}$, with Φ_i as defined above for every i, we have $\Phi_i \subseteq B_i^\infty$ for all i.

Theorem 4. *Let $(\Phi_i)_{i\in\mathcal{I}}$ be a best response set, then $\Phi_i \subseteq B_i^\infty$ for every i.*

3.2. Practical implementation matters

The definition of B_i^k appearing in the inductive step of the recursive procedure involves an infinite number of restrictions. However, these can be reduced to a finite number of restrictions by defining what we call "the critical" ε of choice c_i. Since the choice set of every player is finite, it is possible to characterize for every choice $c_i \in C_i$ within the decision problem Γ^k a critical value $\varepsilon^{c_i}(\Gamma^k)$, where $\varepsilon^{c_i}(\Gamma^k)$ is the minimum value of ε that makes that choice ε-optimal for some belief $b_i \in B_i^k$. Player i believing in his opponents' F-rationality, then should assign at least probability $F_i(\varepsilon')$ to the set of choice combinations c_{-i} that have a critical ε of at most ε', for every $\varepsilon' \geq 0$. These restrictions compose the new restricted belief sets B_i^{k+1} for every player i.

Definition 4. The critical ε of choice c_i in decision problem Γ^k, denoted by $\varepsilon^{c_i}(\Gamma^k)$, is defined as

$$\varepsilon^{c_i}(\Gamma^k) = Min\{\varepsilon \mid \varepsilon \geq 0, \exists\, b_i \in B_i^k \text{ with } u_i(c_i, b_i) \geq u_i(c_i', b_i) - \varepsilon \quad \forall c_i' \in C_i\}.$$

Lemma 1 shows that the critical epsilon $\varepsilon^{c_i}(\Gamma^k)$ can be equivalently defined as the highest ε such that choice c_i is ε-strictly dominated. We say that a choice c_i is ε-strictly dominated if there is an $r_i \in \triangle(C_i)$ such that $u_i(r_i, c_{-i}) \geq u_i(c_i, c_{-i}) + \varepsilon$ for all $c_{-i} \in C_{-i}$, where r_i is a randomization assigning probability $r_i(c_i) \geq 0$ to every choice $c_i \in C_i$.

Lemma 1. *The critical $\varepsilon^{c_i}(\Gamma^k)$ of choice c_i in decision problem Γ^k is defined as*

$$\varepsilon^{c_i}(\Gamma^k) = Max\{\varepsilon \mid \varepsilon \geq 0, \exists\, r_i \text{ with } u_i(r_i, b_i) \geq u_i(c_i, b_i) + \varepsilon \quad \forall b_i \in B_i^k\}.$$

Lemma 2 shows how the notion of critical ε can be used to reduce the infinite set of inequalities characterizing each B_i^k to a finite number of inequalities. Central to such simplification is an ascending ranking $\{\varepsilon_{-i}^1, \ldots, \varepsilon_{-i}^M\} = \{\varepsilon^{c_{-i}}(\Gamma^k) \mid c_{-i} \in C_{-i}\}$ of critical epsilons of the set of choice combinations C_{-i}, where $M = \mid C_{-i} \mid$, $\varepsilon_{-i}^1 = 0$ and $\varepsilon_{-i}^1 \leq \cdots \leq \varepsilon_{-i}^M$.

Lemma 2. *The infinite set of inequalities $b_i(C_{-i}^{\varepsilon}(\Gamma^k)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$ is equivalent to the finite set of inequalities of the form*

$$b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq Lim_{\varepsilon \uparrow \varepsilon_{-i}^{m+1}} F_i(\varepsilon) \quad \forall m \in \{1, \ldots, M-1\}.$$

Because $C_{-i}^{\varepsilon^M} = C_{-i}$, it follows that $b_i(C_{-i}^{\varepsilon^M}) = 1$.

4. Special cases

In this section two special cases are considered, Common Belief in ε-Rationality and Common Belief in p-Belief in Rationality. The former is when one specific value of ε receives full probability, while the latter refers to the case in which the only requirement imposed by functions F_i is that a minimum probability of p must be assigned to the event of the opponent being 0-rational.

4.1. Common belief in ε-rationality

Suppose that there exists for each player $-i$ some ε_{-i}^* such that

$$F_i(\varepsilon) = \begin{cases} 0 & \text{if} \quad \varepsilon < \varepsilon_{-i}^* \\ 1 & \text{if} \quad \varepsilon \geq \varepsilon_{-i}^*. \end{cases}$$

F-rationality in this case reduces to Radner (1980)'s ε-rationality, with $\varepsilon = \varepsilon_i^*$ for every i. It therefore becomes possible for each player i to classify his opponents' choices C_{-i} into ε_{-i}^*-rational and ε_{-i}^*-irrational choices, and so allowing for both the model and the recursive procedure to be simplified. The concept of Common Belief in F-Rationality also reduces to Common Belief in ε-Rationality. An epistemic type $t_i \in T_i$ believes in players $-i$'s ε-rationality, $t_i \in \varepsilon R_i^1$, if it only assigns positive probability to choice-type combinations (c_{-i}, t_{-i}) where c_{-i} is ε_{-i}^*-optimal for t_{-i}. Types expressing $k \geq 2$ fold belief in ε-rationality, $t_i \in \varepsilon R_i^k$, can be defined in a manner analogous to the general case introduced in Definition 3.

Definition 5. Formally, k-fold belief in ε-rationality and common belief in ε-rationality can be defined recursively by the following sequence:

$$\varepsilon R_i^1 := \{t_i \in T_i : b_i(t_i)(c_{-i}, t_{-i}) > 0 \quad \text{implies} \quad c_{-i} \text{ is } \varepsilon_{-i}^* - \text{optimal} \quad \text{for} \quad t_{-i}\};$$

$$\varepsilon R_i^k := \{t_i \in T_i : b_i(t_i)(C_{-i} \times \varepsilon R_{-i}^{k-1}) = 1\}$$

and

$$CB\varepsilon R_i := \bigcap_{k=1}^{\infty} \varepsilon R_i^k.$$

The above specification of F_i reduces the inequalities characterizing the feasible belief set in round k of the recursive procedure to one single equality for each round (by Lemma 2), i.e.,

$$b_i(C_{-i}^{\varepsilon_{-i}^*}(\Gamma^{k-1})) = 1.$$

The recursive procedure can therefore be simplified into one in which for every round choices c_i of every player i that possess a critical epsilon strictly greater than ε_i^* are eliminated. Notably, the recursive procedure now proceeds by eliminating choices rather than beliefs. Since the set of choices for each player is finite, the procedure must stop after a finite number of rounds. Recursive Procedure 2 lists the steps of the now simplified procedure. Corollary 2 then shows the procedure works.

Procedure 2. *Iterated elimination of ε-dominated choices:*

- *Initial step: Define $\Gamma^0 = (C^0; u)$, where $C^0 = \{C_1, \ldots, C_n\}$ and therefore $\Gamma^0 = \Gamma$.*
- *Inductive step: Assume that $\Gamma^{k-1} = (C_i^{k-1}, C_{-i}^{k-1}; u_i, u_{-i})$ has been defined. Then $\Gamma^k = (C_i^k, C_{-i}^k; u_i, u_{-i})$ where for each player i*

  $$C_i^k = \{c_i \in C_i \mid \varepsilon^{c_i}(\Gamma^{k-1}) \leq \varepsilon_i^*\}.$$

Corollary 2. *A choice $c_i \in C_i$ can rationally be made under common belief in ε-rationality if and only if it survives all rounds of the recursive procedure of iterated elimination of ε-dominated choices*

4.2. Common belief in p-belief in rationality

Another special case is that in which F takes the form $F_i(\varepsilon) = p$ for all $\varepsilon \geq 0$ and for all players i. This is the case where each player assigns a probability of at least p to the event of his opponents being 0-rational, while the remaining probability $1-p$ can be assigned to any other degree(s) of rationality. Doing so, the player is said to believe in his opponents' p-rationality, as he assigns at least probability p to his opponents choosing rationally (or being 0-rational). In this respect, a link should be made to an established concept in the literature, namely Common p-Belief in Rationality (Moderer and Samet, 1989; Hu, 2007). Common belief in p-belief in rationality refers to the case where there is common belief in the event that every player p-believes in the opponents' rationality. Common p-Belief in Rationality, on the other hand, refers to the case where player i assigns probability at least p to the event of his opponents being 0-rational, player i assigns probability at least p to the event of every player $-i$ assigning probability at least p to the event of i being 0-rational, and so on. Common Belief in p-Belief in Rationality can thus be considered a strengthening, or a special case, of Common p-Belief in Rationality. Types expressing k fold belief in p-belief in rationality, written as pR_i^k, are defined below.

Definition 6. Formally, k-fold belief in p-belief in rationality and common belief in p-belief in rationality can be defined recursively by the following sequence:

$$pR_i^1 := \{t_i \in T_i : b_i(t_i)(R_{-i}^0) \geq p\};$$

$$pR_i^k := \{t_i \in T_i : b_i(t_i)(C_{-i} \times pR_{-i}^{k-1}) = 1\}$$

$$CBpR_i := \bigcap_{k=1}^{\infty} pR_i^k.$$

Note that the only restriction implied by F_i here is that $F_i(0) = p$. Each player i's choices can therefore be classified into two sets, $C_i^{0,k}$ and $C_i^{\neg 0,k}$. The former includes those $c_i \in C_i$ that are 0-rational given the belief restricted game Γ^k facing player i in round k of the recursive procedure, while the latter groups all choices that are not 0-rational given Γ_i^k. Restriction on the belief set then becomes $b_i(C_{-i}^{0,k}) \geq p$. Note that $C_i^k = C_i$ for all $k \geq 0$ and for all i, where $C_i^k = C_i^{0,k} \cup C_i^{\neg 0,k}$. So the recursive procedure never eliminates any choices, it only redistributes choices between the two above defined sets. Steps of the recursive procedure below are simultaneously applied to each player and repeated until $C_i^{0,k} = C_i^{0,k-1}$ and so $B_i^k = B_i^{k-1}$ for all i.

Procedure 3. *Iterated elimination of p-dominated choices:*

- *Initial Step: Define* $\Gamma^0 = (C^{0,0}, C^{\neg 0,0}, B^0; u)$.
- *Inductive Step: Assume that* $\Gamma^{k-1} = (C^{0,k-1}, C^{\neg 0,k-1}, B^{k-1}; u)$ *has been defined. Then* $\Gamma^k = (C^{0,k}, C^{\neg 0,k}, B^k; u)$ *where for each player i*

$$C_i^{0,k} = \{c_i \in C_i \mid \varepsilon^{c_i}(\Gamma^k) = 0\}$$

$$C_i^{\neg 0,k} = \{c_i \in C_i \mid \varepsilon^{c_i}(\Gamma^k) > 0\}$$

$$B_i^k := \{b_i \in \triangle(C_{-i}) : b_i(C_{-i}^{0,k}) \geq p\}.$$

5. An example: the Traveler's Dilemma

This section uses the Traveler's dilemma of Basu (1994) both to illustrate the recursive procedure and to show how our model could help bring theoretical predictions closer to experimental findings. Section 5.1 uses a three price Traveler's Dilemma game to illustrate the steps of the procedure for two different functions F_i. Section 5.2 then presents some general results for the n-price Traveler's Dilemma under *CBFR*.

5.1. Illustrating examples of the recursive procedure

To show how the recursive procedure works, we use a 3-price Traveler's Dilemma (*TD*) game due to (Basu, 1994), in which both reward and penalty are set equal to 2. We examine the resulting B_j^∞ under two slightly different forms of F_j;

$$\text{Case I} \qquad F_j(\varepsilon) = \begin{cases} \varepsilon & \text{if } \varepsilon \leq 1 \\ 1 & \text{if } \varepsilon > 1 \end{cases}$$

$$\text{Case II} \qquad F_j(\varepsilon) = \begin{cases} \dfrac{\varepsilon}{2} & \text{if } \varepsilon \leq 2 \\ 1 & \text{if } \varepsilon > 2. \end{cases}$$

Case I:

Let $TD = \{P_1, P_2; u_1, u_2\}$, where $P_i = \{1, 2, 3\}$ and

$$u_i(p_i, p_j) = \begin{cases} p_i + 2 & \text{if } p_i < p_j \\ p_i & \text{if } p_i = p_j \\ p_j - 2 & \text{if } p_i > p_j. \end{cases}$$

Table 1 shows our *TD* game in normal form

Let F_j be defined for both players as in Case I above. Start round 1 of the recursive procedure. The relevant belief restricted game is $TD^0 = TD$ or $TD^0 = \{P_1, P_2, B_1^0, B_2^0; u_1, u_2\}$. The critical values of epsilon $\varepsilon^{p_i}(TD^0)$ for each of player i's potential choices are 0, 0 and $\frac{2}{3}$ for $p_i = 1, p_i = 2$ and $p_i = 3$, respectively. Note that these critical values are the same for both players due to the symmetric nature of the game. So $C_i^\varepsilon = \{1, 2\}$ for all $\varepsilon < \frac{2}{3}$ and $C_i^\varepsilon = \{1, 2, 3\}$ for all $\varepsilon \geq \frac{2}{3}$. From Lemma 2, this implies one restriction on player j's belief set

$$b_j(\{1, 2\}) \geq \frac{2}{3} \qquad \text{or} \qquad b_j(\{3\}) \leq \frac{1}{3}.$$

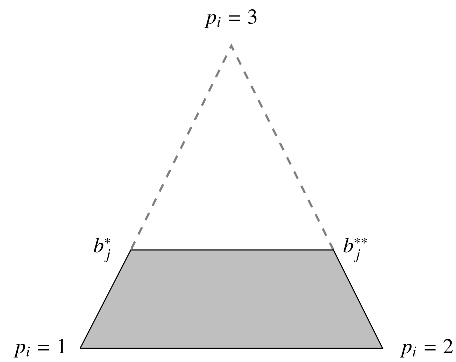

Fig. 1. Feasible belief set B_j^1.

Table 1
Three price Traveler's Dilemma (*TD*)

	$p_2 = 1$	$p_2 = 2$	$p_2 = 3$
$p_1 = 1$	1,1	3,−1	3,−1
$p_1 = 2$	−1,3	2,2	4,0
$p_1 = 3$	−1,3	0,4	3,3

Table 2
Belief restricted game TD^1.

	$p_j = 1$	$p_j = 2$	b_j^*	b_j^{**}
$p_i = 1$	1	3	$\frac{5}{3}$	3
$p_i = 2$	−1	2	$\frac{2}{3}$	$\frac{8}{3}$
$p_i = 3$	−1	0	$\frac{1}{3}$	1

This restriction then defines B_j^1 to be the convex hull of four different points; namely, $B_j^1 = conv(\{(1, 0, 0), (0, 1, 0), b_j^*, b_j^{**}\})$. Note that B_j^1 is also identical for both players due to the symmetry of the game. Fig. 1 is a graphical representation of the feasible belief set B_j^1 after round 1 of the procedure.

Start round 2, with the belief restricted game $TD^1 = \{P_1, P_2, B_1^1, B_2^1; u_1, u_2\}$. Table 2 presents TD^1, where $b_j^* = \frac{1}{3}(p_i = 3) + \frac{2}{3}(p_i = 1)$ and $b_j^{**} = \frac{1}{3}(p_i = 3) + \frac{2}{3}(p_i = 2)$ from Fig. 1. Note that utilities listed in Table 2 are those of player i under TD^1.

The critical values of epsilon for each of both player i's choices under TD^1 become $\varepsilon_i^{p_i=1}(TD^1) = 0$, $\varepsilon_i^{p_i=2}(TD^1) = \frac{1}{3}$ and $\varepsilon_i^{p_i=3}(TD^1) = \frac{4}{3}$. This implies that $C_i^\varepsilon = \{1\}$ for all $\varepsilon < \frac{1}{3}$, $C_i^\varepsilon = \{1, 2\}$ for all $\frac{1}{3} \leq \varepsilon < \frac{4}{3}$ and $C_i^\varepsilon = \{1, 2, 3\}$ for all $\varepsilon > \frac{4}{3}$. Given F_j, restrictions on player j's feasible belief set become

$$b_j(\{1, 2\}) \geq 1 \qquad \text{or} \qquad b_j(\{3\}) = 0$$

$$b_j(\{1\}) \geq (\frac{1}{3}) \qquad \text{or} \qquad b_j(\{2\}) \leq (\frac{2}{3}).$$

The solid line in Fig. 2 represents $B_j^2 = conv(\{(1, 0, 0), b_j^*\})$, or equivalently B_i^2, with $b_j^* = \frac{1}{3}(p_i = 1) + \frac{2}{3}(p_i = 2)$.

Round 3 of the recursive procedure then starts with the belief restricted game $TD^2 = \{P_1, P_2, B_1^2, B_2^2; u_1, u_2\}$ (Table 3). The critical values of epsilon become the following: $\varepsilon_i^{p_i=1}(TD^2) = 0$ and $\varepsilon_i^{p_i=2}(TD^2) = \frac{4}{3}$, implying that $p_i = 2$ receives no positive probability similarly to $p_i = 3$ in round 2. The feasible belief set B_j^3

Table 3
Belief restricted game TD^2.

	$p_j = 1$	b_j^*
$p_i = 1$	1	7/3
$p_i = 2$	-1	1

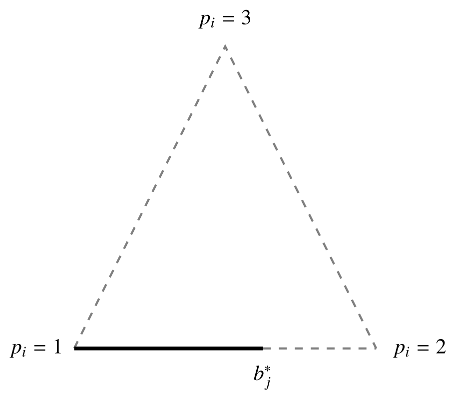

Fig. 2. Feasible belief set B_j^2.

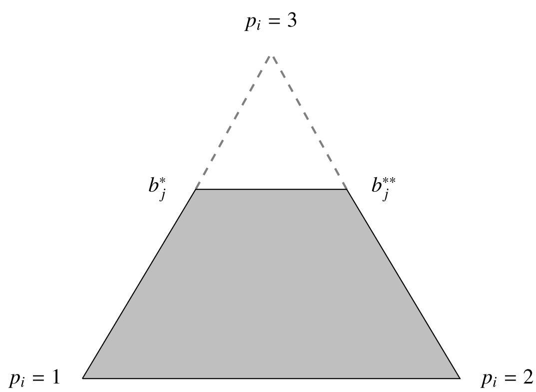

Fig. 3. Feasible belief set B_j^1.

thus collapses to one point assigning probability 1 to the opponent i playing price 1. The procedure stops after 3 rounds and feasible belief sets stabilize to $B_j^3 = B_j^\infty = \{b_j(p_i = 1) = 1\}$ for both players.

Case II:

Let the game TD be as previously defined, and consider the slightly modified form of the F_j function corresponding to Case II. Rounds of the recursive procedure then proceed as follows: Round 1 starts with the same TD^0 as above with the critical epsilons being $\varepsilon_i^{p_i=1}(TD^0) = \varepsilon_i^{p_i=2}(TD^0) = 0$ and $\varepsilon_i^{p_i=3}(TD^0) = \frac{2}{3}$ for both players. The current function $F_j(\varepsilon)$, however, implies different restrictions on the feasible belief set B_j^1, which take the form

$$b_j(\{1,2\}) \geq (\frac{1}{3}) \qquad \text{or} \qquad b_j(\{3\}) \leq (\frac{2}{3}).$$

Fig. 3 represents the feasible belief set B_j^1 after round 1 of the procedure.

Round 2 starts with the belief restricted game $TD^1 = \{P_1, P_2, B_1^1, B_2^1; u_1, u_2\}$. Table 4 presents TD^1, where $b_j^* = (\frac{2}{3})(p_i = 3) + (\frac{1}{3})(p_i = 1)$ and $b_j^{**} = (\frac{2}{3})(p_i = 3) + (\frac{1}{3})(p_i = 2)$ from Fig. 3.

From Table 4, it is evident that the critical epsilons of all three choices remain unchanged. Consequently, the restrictions on player j's feasible belief set, and so the belief set B_j^2 itself remain unchanged as well. By symmetry of the game, the same holds for player i. The recursive procedure stops after two rounds as $B_j^2 = B_j^1$ for both players. The set of first-order beliefs that can be held by a type expressing common belief in F-rationality in the 3-price TD game at hand, thus contains an infinite set of beliefs assigning positive probabilities to both the second and third price.

Table 4
Belief restricted game TD^1.

	$p_j = 1$	$p_j = 2$	b_j^*	b_j^{**}
$p_i = 1$	1	3	$\frac{7}{3}$	3
$p_i = 2$	-1	2	$\frac{7}{3}$	$\frac{10}{3}$
$p_i = 3$	-1	0	$\frac{5}{3}$	2

5.2. CBFR in the Traveler's Dilemma

As indicated earlier, the model we introduce along with the characterizing procedure is purely concerned with players' beliefs and does not by itself say anything about the actual choices players would make in a given game, as the actual ε of each player is his/ her private information. Take for example the case II in the previous section. Although the feasible belief set of each player i contains beliefs assigning probability up to $\frac{2}{3}$ to the opponent's highest price, player i can still react to such a belief by playing a 0-optimal choice. The main message of this section is to show that our model can account for some of the experimental observations in the Traveler's Dilemma even with very little deviation from rationality and common belief in rationality. This is achieved by allowing each player to account for the possibility that his opponent may make a mistake with some probability.

Consider an n-price Traveler's Dilemma where each player i chooses $p_i \in \{1, \ldots, n\}$. The critical epsilon $\varepsilon^n(TD^0)$ of the highest price n is at most $\frac{2}{n}$ while that of any $p_i \neq n$ is 0. To show that $\varepsilon^n(TD^0) = \frac{2}{n}$, note that for an n-price game with reward and punishment set equal to 2, the following holds:

$$u_i(p_i, 1) = \begin{cases} 1 & \text{for } p_i = 1 \\ -1 & \text{otherwise,} \end{cases}$$

and

$$u_i(p_i, n) = \begin{cases} p_i + 2 & \text{for } p_i < n \\ n & \text{for } p_i = n. \end{cases}$$

Consider the belief b_i of player i assigning probability $\frac{2}{n}$ to the opponent's choice $p_j = n$ and $1 - \frac{2}{n}$ to $p_j = 1$. The expected utilities of player i's choices given such belief are as follows: $u_i(1, b_i) = 1 + \frac{4}{n}$, $u_i(n, b_i) = 1 + \frac{2}{n}$ and $u_i(p_i, b_i) = \frac{2}{n}(p_i + 3) - 1$ for prices $1 < p_i < n$. Note that the highest $u_i(p_i, b_i)$ for $1 < p_i < n$ is that of price $n - 1$. It is easy to see that $u_i(n, b_i)$ is exactly $\frac{2}{n}$ lower than the expected utility of the best response choice. Thus there exists a belief b_i for which n is $\frac{2}{n}$-optimal, thereby proving that $\frac{2}{n}$ is at least and upper bound on $\varepsilon^n(TD^0)$. Note that the upper bound on the critical epsilon is a decreasing function in the number of prices. Thus the higher the number of prices the smaller the loss in expected utility a player has to (be believed to) tolerate to choose a higher price. Now assume F_i is defined for players $i = \{1, 2\}$ as follows:

$$F_i(\varepsilon) = \begin{cases} \dfrac{n-2}{2}\varepsilon & \text{if } \varepsilon \leq \dfrac{2}{n-2} \\ 1 & \text{if } \varepsilon > \dfrac{2}{n-2}. \end{cases}$$

Since $\frac{2}{n}$ constitutes an upper limit on the critical epsilon $\varepsilon^n(TD^0)$ of price $p_j = n$, then the restriction imposed on player i's feasible belief set in round 1 of the procedure based on the actual $\varepsilon^n(TD^0)$

must satisfy $b_i(n) \leq \frac{2}{n}$. Given this restriction, one of the extreme points in the belief restricted game TD^1 defined in round 1 for player i becomes $b_i^{1n} = \frac{2}{n}(n) + (1 - \frac{2}{n})(1)$. In TD^1, choice n of player i is $\frac{2}{n}$-optimal for b_i^{1n} as shown above. The critical $\varepsilon^n(TD^1) \leq \frac{2}{n}$. This implies that no further restrictions will be imposed on player i's belief set in round 2 of the procedure.

Assume player $i = \{1, 2\}$ is 0-rational and holds the belief $b_i^{1n} \in B_i^{\infty}$ about his opponent's choices. Player i's best response is then to choose $n - 1$ under CBFR. Thus even if each player is still 0-rational, assigning positive probability to the opponent making mistakes of no more than $\frac{2}{n-2}$ in expected utility in an n-price game, might lead the player to choose for a much higher price than predicted by CBR. Note that the maximum extent of irrationality $\frac{2}{n-2}$ that must be deemed possible is also a decreasing function in n.

6. Discussion and conclusion

By allowing players to be boundedly rational and introducing uncertainty about opponents' extent of rationality/irrationality, this paper develops the concept of Common Belief in F-Rationality. A recursive procedure is also designed to characterize exactly those first-order beliefs that can be held by an epistemic type expressing common belief in F-rationality. The recursive procedure of Iterated Elimination of F-Dominated Beliefs restricts each player's belief set by eliminating those beliefs that are inconsistent with the event of believing in the opponents' F-rationality. Finally, we also consider two special cases of the function F, which correspond to Common Belief in ε-Rationality and Common Belief in p-Belief in Rationality. Both the model and the recursive procedure are illustrated for these special cases.

The model presented in this paper is, however, not free of limitations. One crucial limitation is the common knowledge of functions $(F_i)_{i\in\mathcal{I}}$. Although greatly simplifying the analysis, it is hard to imagine these functions being common knowledge in an actual setting. One could imagine that opponents of a player i could perceive certain forms of the F_i function as more plausible than others, rather than them actually knowing which F_i player i holds. An investigation of such functions requires careful experimental design and could benefit greatly from progress made in the experimental literature on belief elicitation. However, the design of such experiment is beyond the scope of this paper. One other aspect that merits some attention regarding the common knowledge assumption of functions $(F_i)_{i\in\mathcal{I}}$, is what the consequences in a theoretical context of a relaxation of such assumption would be.

Should the common knowledge assumption of functions F_i be relaxed, the game becomes one of incomplete information where players' error margins are private information, i.e. where players' utility functions are private information. Consequently, the model would need to be extended to account for each player i's beliefs about their opponents' functions $(F_j)_{j\neq i}$. One framework that would qualify as a starting point for such extension of our model is that of Bach and Perea (2016). In their model, they provide an incomplete information counterpart of common belief in rationality along with an algorithmic characterization thereof. Players' utility functions in that model are private information and the model considers a finite set of possible utility functions for each player. Dropping the common knowledge assumption of $(F_i)_{i\in\mathcal{I}}$ would give rise to important questions about the conceptual meaning of functions $(F_i)_{i\in\mathcal{I}}$ in such an incomplete information setting. Should these functions be explicitly modeled or should the focus shift from the $(F_i)_{i\in\mathcal{I}}$ functions to the actual error margins since both would then be private information?

Another potential limitation of the model is the underlying assumption that every player i uses the same function F_i to define his/her lower bound of probability for every opponent. Addressing this limitation is more straight forward than the previous one. Although not examined in this paper, we expect the same conclusions to be preserved if every player i is allowed to assign a different lower bound $F_{i,j}$ for each player j. It is worth noting that although for every choice c_i in any given game there exists some $\varepsilon^{c_i} \geq 0$ making that choice ε^{c_i}-optimal for some belief, it is not the case that CBFR allows for any behavior to be justified. The range of beliefs (choices) that players can be reasonably expected to hold (make) under CBFR depends on the domains and shapes of the functions $(F_i)_{i\in\mathcal{I}}$ as well as on the structure of the game at hand.

While the role of the domains and shapes of the functions $(F_i)_{i\in\mathcal{I}}$ are obvious, the implications of the structure of the game might require some clarification. Take for example an n-price Traveler's dilemma. As shown earlier, there are functions $(F_i)_{i\in\mathcal{I}}$ such that some beliefs assigning positive probability to any choice of the opponent up to and including the nth price could be reasonably held by players under CBFR. However, it is not possible to find an F_i such that some reasonable first-order beliefs of player i under CBFR would assign positive probability to any choice of the opponent up to and including the second highest price while no reasonable belief would assign any positive probability to the opponents' nth price. This is due to the fact that the critical epsilon of any price in the Traveler's dilemma is a decreasing function of the number of prices.

Furthermore, it is important to stress that our model remains silent about the way players actually act upon their beliefs. More precisely, our model does not require players to be ε-best responders. In fact, players may actually make perfectly rational choices in response to the beliefs they hold under common belief in F-rationality. Thus the model only assumes that players believe that their opponents might make errors causing them to deviate from their optimal choices. This has been made evident in the Traveler's Dilemma section where players assigning positive probability to the event of their opponent making mistakes up to a certain extent has led even 0-rational players to plausibly choosing prices up to the second highest. This example serves as an indication that some of the observed experimental results, at least in the Traveler's Dilemma game, may be explained by relatively small deviations from standard rationality.

Our model attempted to capture the possibility that a player would make mistakes or believe his opponents' to do so, taking into account that the extent of these mistakes is at best private information. The uncertainty about opponents' potential mistakes is regulated by means of the F-functions assigning to each potential margin of error a maximum level of probability. The game considered was one of complete information. We believe restrictions on beliefs similar to the ones characterized by our model could be obtained had we considered games of incomplete information in which the utility functions of the players are not transparent. Instead of a margin of error, ε would be redefined as the maximum distance between different possible utility functions of the opponent and his true utility function in a manner similar to that of Perea and Roy (2014). The minimum bound on probability $F_i(\varepsilon)$ would then also be reinterpreted accordingly. The set up of the model and procedure of the incomplete information case is expected to bear many similarities to the model presented in this paper. However, the extent of such similarities is beyond the scope of this paper and is a question for further research.

Acknowledgments

We would like to thank participants of Maastricht Lecture Series in Economics (MLSE); Spring Meeting of Young Economists (SMYE) 2012 in Mannheim, Germany; The 10th Conference on Logic and the Foundations of Game and Decision Theory (LOFT 2012) in Seville, Spain; and The 4th World Congress of the Game Theory

Society (Games 2012) in Istanbul, Turkey, for all their questions, discussions and comments. Special thanks are also due to Aviad Heifetz and two anonymous reviewers of LOFT 2012 for valuable comments on the paper.

Appendix

A.1. Theorem 1

Proving Theorem 1 requires proving two statements. First, it must be shown that any first-order belief $b_i^1(t_i)$ held by an epistemic type expressing common belief in F-rationality, i.e. $t_i \in CBFR_i$, must be in the set B_i^∞ of first-order beliefs that survive the iterative procedure (part A). Second, it must be proven that any first-order belief $b_i^1(t_i) \in B_i^\infty$ surviving the iterative procedure can be held by a type expressing common belief in F-rationality (part B). More formally, proving Theorem 1 requires showing that
(A) Any $b_i(t_i)$ with $t_i \in CBFR_i$ has $b_i^1(t_i) \in B_i^\infty$
(B) Any $b_i^1(t_i) \in B_i^\infty$ can be held by a type $t_i \in CBFR_i$.

Proof of part A: We prove by induction on k that $b_i^1(t_i) \in B_i^k$ for all k, whenever $t_i \in CBFR_i$. Let $k = 0$. Take some $t_i \in CBFR_i$, then $b_i^1(t_i) \in B_i^0$, where $B_i^0 = \triangle(C_{-i})$. Let $k \geq 1$. Assume for every $t_i \in CBFR_i$, $b_i^1(t_i) \in B_i^{k-1}$ for both players i, where $B_i^{k-1} = \{b_i \in \triangle(C_{-i}) \mid b_i(C_{-i}^\varepsilon(\Gamma^{k-2})) \geq \varepsilon$ for all $\varepsilon \geq 0\}$.

Take some $t_i \in CBFR_i$, then
(i) $b_i(t_i)(R_{-i}^\varepsilon) \geq F_i(\varepsilon)$ as $t_i \in FR_i^1$ for all $\varepsilon \geq 0$
(ii) $b_i(t_i)(c_{-i}, t_{-i}) > 0$ implies that $t_{-i} \in CBFR_{-i}$.

From the induction assumption above, it follows that $b_{-i}^1(t_{-i}) \in B_{-i}^{k-1}$ whenever $t_{-i} \in CBFR_{-i}$. Let $R_{-i}^{\varepsilon,k-1} := \{(c_{-i}, t_{-i}) \in R_{-i}^\varepsilon(\Gamma^{k-1}) \mid b_{-i}^1(t_{-i}) \in B_{-i}^{k-1}\}$, where $R_{-i}^\varepsilon := \{(c_{-i}, t_{-i}) \mid c_{-i}$ is ε-optimal for $b_{-i}^1(t_{-i})\}$.

Then (i) and (ii) imply, $b_i(t_i)(R_{-i}^{\varepsilon,k-1}) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$. So $b_i(t_i)(Proj_{C_{-i}}R_{-i}^{\varepsilon,k-1} \times T_{-i}) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$, which means $b_i^1(t_i)(Proj_{C_{-i}}R_{-i}^{\varepsilon,k-1}) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$. Since $Proj_{C_{-i}}R_{-i}^{\varepsilon,k-1} = C_{-i}^\varepsilon(\Gamma^{k-1})$, it follows that $b_i^1(t_i)(C_{-i}^\varepsilon(\Gamma^{k-1})) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$, and therefore $b_i^1(t_i) \in B_i^k$.

Proof of part B: Proving this part starts by selecting some subset of the set B_i^∞ of beliefs surviving the iterative procedure. An epistemic model is then defined such that for every selected first-order belief b_i of some player i there is an epistemic type t_i such that the first-order belief of t_i is exactly b_i. We then proceed to show that each one of these constructed types expresses common belief in F-rationality.

The set of beliefs we select for the purpose of this proof consists of two components:

1. For every choice c_i of player i we select a belief $b_i^{c_i} \in B_i^\infty$, where c_i is exactly $\varepsilon^{c_i}(\Gamma^\infty)$-optimal for $b_i^{c_i}$. Thus from Definition 1 belief $b_i^{c_i}$ is such that,

   $$u_i(c_i, b_i^{c_i}) \geq u_i(c_i', b_i^{c_i}) - \varepsilon^{c_i}(\Gamma^\infty), \quad \textit{for all } c_i' \in C_i \quad \textit{and}$$

   $$u_i(c_i, b_i^{c_i}) = u_i(c_i', b_i^{c_i}) - \varepsilon^{c_i}(\Gamma^\infty), \quad \textit{for at least some } c_i' \in C_i,$$

   and where $\varepsilon^{c_i}(\Gamma^\infty)$ is the critical ε of choice c_i in decision problem Γ_i^∞ defined by set B_i^∞ (Definition 4).
   Note that by Theorem 2 B_i^∞ is always non-empty. Characterized by a finite number of linear inequalities (Lemma 2) imposed on the convex, closed and bounded set B_i^0, it is easy to see that set B_i^∞ is also convex, closed and bounded. A belief $b_i^{c_i} \in B_i^\infty$ as defined above therefore always exists.
2. Since we would like to show that the theorem holds for any arbitrary first-order belief surviving the iterative procedure, we also take one arbitrary first-order belief $b_i^* \in B_i^\infty$.

Now that the beliefs have been constructed, the epistemic model can be defined. Consider a finite epistemic model M with a finite set of types T_i for every player i. The set of types T_i is also constructed in two steps:

1. The set of types $T_i^{C_i} = \{t_i^{c_i} \mid c_i \in C_i\}$ where every $t_i^{c_i} \in T_i^{C_i}$ is such that

   $$b_i(t_i^{c_i})(c_{-i}, t_{-i}) = \begin{cases} b_i^{c_i}(c_{-i}) & \text{if} \quad t_{-i} = t_{-i}^{c_{-i}} \\ 0 & Otherwise. \end{cases} \tag{A.1}$$

2. Define the finite set of types $T_i = T_i^{C_i} \cup \{t_i^*\}$ where t_i^* is such that

   $$b_i(t_i^*)(c_{-i}, t_{-i}) = \begin{cases} b_i^*(c_{-i}) & \text{if} \quad t_{-i} = t_{-i}^{c_{-i}} \\ 0 & Otherwise \end{cases} \tag{A.2}$$

   and where $b_i^{c_i}$ and b_i^* are as defined above.

The proof of this part then proceeds in two steps. First we show that any type $t_i^{c_i} \in T_i^{C_i}$ expresses common belief in F-rationality. Once that is established, type $t_i^* \in T_i$ is shown to express common belief in F-rationality completing our proof.

Step 1: Take some $t_i^{c_i} \in T_i^{C_i}$. Since $b_i^1(t_i^{c_i}) = b_i^{c_i}$ and $b_i^{c_i} \in B_i^\infty$ by belief and type construction above, it follows that

$$b_i^1(t_i^{c_i})(C_{-i}^\varepsilon(\Gamma^\infty)) = b_i^{c_i}(C_{-i}^\varepsilon(\Gamma^\infty)) \geq F_i(\varepsilon) \quad \text{for all } \varepsilon \geq 0$$

Recall that

$$C_j^\varepsilon(\Gamma^k) = \{c_j \in C_j \mid \exists\, b_j \in B_j^k \text{ s.t. } u_j(c_j, b_j) \geq u_j(c_j', b_j) - \varepsilon \quad \forall c_j' \in C_j\}.$$

Take some $c_{-i} \in C_{-i}^\varepsilon(\Gamma^\infty)$, then by definition of $C_j^\varepsilon(\Gamma^k)$ it must be that $\varepsilon^{c_j}(\Gamma^\infty) \leq \varepsilon$ for every such c_j with $j \neq i$. Since every c_j in such c_{-i} is $\varepsilon^{c_j}(\Gamma^\infty)$-optimal for the respective $t_j^{c_j}$ by definition of $t_j^{c_j}$, it is also ε-optimal for $t_j^{c_j}$ for any $\varepsilon \geq \varepsilon^{c_j}(\Gamma^\infty)$. So $(c_{-i}, t_{-i}^{c_{-i}}) \in R_{-i}^\varepsilon$, where $R_{-i}^\varepsilon := \{(c_{-i}, t_{-i}) \in (C_{-i} \times T_{-i}) \mid c_j$ is ε-optimal for $b_j^1(t_j)$ *for every* $j \neq i\}$.
Hence

$$b_i(t_i^{c_i})(R_{-i}^\varepsilon) \geq b_i^1(t_i^{c_i})(C_{-i}^\varepsilon(\Gamma^\infty)) \geq F_i(\varepsilon) \quad \text{for all} \quad \varepsilon \geq 0.$$

So $t_i^{c_i} \in FR_i^1$ by Definition 1. Since type $t_i^{c_i}$ by definition only assigns positive probability to types $t_j^{c_j}$ for every $j \neq i$, each of which in turn only assign positive probability to type-combinations $t_{-j}^{c_{-j}}$ and since $t_i^{c_i} \in FR_i^1$ for every player i, it follows that $t_i^{c_i} \in CBFR_i$.

Step 2: Consider $t_i^* \in T_i$. Since $t_i^{c_i} \in CBFR_i$ for every player i (Step 1), the definition of type t_i^* implies that it only assigns positive probability to (c_{-i}, t_{-i}) where $t_{-i} \in CBFR_{-i}$. It therefore suffices to show that $t_i^* \in FR_i^1$ to prove $t_i^* \in CBFR_i$.

Since $b_i^* \in B_i^\infty$,

$$b_i^1(t_i^*)(C_{-i}^\varepsilon(\Gamma^\infty)) \geq F_i(\varepsilon) \quad \text{for all} \quad \varepsilon \geq 0.$$

Since type t_i^* by construction only assigns positive probability to (c_{-i}, t_{-i}) where $t_{-i} = t_{-i}^{c_{-i}}$, the same steps undertaken in Step 1 above can be followed from this point onward leading to the conclusion that $t_i^* \in FR_i^1$. Therefore, $t_i^* \in CBFR_i$

A.2. Theorem 2

This requires proving that for every $k \geq 0$ and every player $i \in \mathcal{I}$, the following holds $B_i^{C_{-i}^k} \subseteq B_i^k$, where $B_i^{C_{-i}^k} = \triangle(C_{-i}^k)$ and where C_i^k is the set of choices of player i surviving k rounds of the procedure of iterated elimination of never-best replies characterizing Correlated Rationalizability. This will be proven by induction.

Induction start: Take some $b_i \in \triangle(C^0_{-i})$ then it is also the case that $b_i \in \triangle(C_{-i})$. By definition of B^0_i, it follows that $b_i \in B^0_i$.

Induction assumption: For any $b_i \in \triangle(C^{k-1}_{-i})$ it is also the case that $b_i \in B^{k-1}_i$.

Induction step: Take some $b_i \in \triangle(C^k_{-i})$. Since $C^k_{-i} \subseteq C^{k-1}_{-i}$ by definition of C^k_{-i}. Then it follows that $b_i \in B^{k-1}_i$. Recall that $C^\varepsilon_i(\Gamma^{k-1}) = \{c_i \in C_i | c_i \text{ is } \varepsilon - \text{optimal for some } b_i \in B^{k-1}_i\}$. Moreover, any $c_i \in C^k_i$ is *0*-optimal for some $b_i \in \triangle(C^{k-1}_{-i})$. From the induction assumption, the latter can be rewritten as follows: any $c_i \in C^k_i$ is *0*-optimal for some $b_i \in B^{k-1}_i$. So $C^k_{-i} \in C^0_{-i}(\Gamma^{k-1})$. Since any $b_i \in \triangle(C^k_{-i})$ is also such that $b_i(C^k_{-i}) = 1$, and since $C^k_{-i} \in C^0_{-i}(\Gamma^{k-1})$, it follows that $b_i(C^0_{-i}(\Gamma^{k-1})) = 1$. So $b_i \in B^k_i$.

A.3. Theorem 3

Let the collection $(B^\infty_i)_{i\in\mathcal{I}}$ be the limit sets of first-order beliefs surviving the procedure of IEFB. Let $(\Phi_i)_{i\in\mathcal{I}}$ be a best-response set where for every player i,

$$\Phi_i \subseteq \{\varphi_i \in \triangle(C_{-i}) | \varphi_i(C^\varepsilon_{-i}(\Phi_{-i})) \geq F_i(\varepsilon)\ \forall\varepsilon \geq 0\}$$

and

$$C^\varepsilon_i(\Phi_i) = \{c_i \in C_i | \exists\varphi_i \in \Phi_i \ s.t.\ u_i(c_i, \varphi_i) \geq u_i(c'_i, \varphi_i) - \varepsilon\ \forall c'_i \in C_i\}.$$

Take some $b_i \in B^\infty_i$. Then by definition of B^k_i, $b_i(C^\varepsilon_{-i}(\Gamma^\infty)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$. By definition, $C^\varepsilon_j(\Gamma^\infty) = C^\varepsilon_j(B^\infty_j)$ for any j. Since $b_i \in B^\infty_i$, it follows that

$$b_i(C^\varepsilon_{-i}(B^\infty_{-i})) \geq F_i(\varepsilon)\ \ \forall\varepsilon \geq 0 \qquad (*).$$

Note that condition (*) is identical to that defining best-response sets $(\Phi_i)_{i\in\mathcal{I}}$. Since condition (*) holds for any $b_i \in B^\infty_i$ and for every player i, it follows that $(B^\infty_i)_{i\in\mathcal{I}}$ is a best-response set.

A.4. Theorem 4

Consider some best response set $(\Phi_i)_{i\in\mathcal{I}}$ where for every player i,

$$\Phi_i \subseteq \{\varphi_i \in \triangle(C_{-i}) | \varphi_i(C^\varepsilon_{-i}(\Phi_{-i})) \geq F_i(\varepsilon)\ \forall\varepsilon \geq 0\}$$

and

$$C^\varepsilon_i(\Phi_i) = \{c_i \in C_i | \exists\varphi_i \in \Phi_i \ s.t.\ u_i(c_i, \varphi_i) \geq u_i(c'_i, \varphi_i) - \varepsilon\ \forall c'_i \in C_i\}.$$

For every round $k \geq 1$ of the procedure, we prove by induction on k that for any best response set $(\Phi_i)_{i\in\mathcal{I}}$, $\Phi_i \subseteq B^k_i$ holds for every i. First take $k = 1$ and show that $\Phi_i \subseteq B^1_i$ for every i. Then by induction on k we show that $\Phi_i \subseteq B^k_i$ for any k and every i. Recall that

$$B^1_i = \{b_i \in \triangle(C_{-i}) | b_i(C^\varepsilon_{-i}(\Gamma^0)) \geq F_i(\varepsilon)\ \ \forall\varepsilon \geq 0\}.$$

By definitions of Γ^0 and Φ_i, it follows for every player j that $C^\varepsilon_j(\Phi_j) \subseteq C^\varepsilon_j(B^0_j) = C^\varepsilon_j(\Gamma^0)$ for every $\varepsilon \geq 0$ and for any Φ_j. Take some $\varphi_i \in \Phi_i$, then by definition

$$\varphi_i(C^\varepsilon_{-i}(\Phi_{-i})) \geq F_i(\varepsilon)\ \ \forall\varepsilon \geq 0.$$

Hence,

$$\varphi_i(C^\varepsilon_{-i}(B^0_{-i})) \geq \varphi_i(C^\varepsilon_{-i}(\Phi_{-i})) \geq F_i(\varepsilon)\ \ \forall\varepsilon \geq 0.$$

Thus, $\Phi_i \subseteq B^1_i$. Assume that $\Phi_i \subseteq B^{k-1}_i$ for all i. Let $k \geq 2$ and show that $\Phi_i \subseteq B^k_i$ for all i and for any k. Proving that requires showing that every $\varphi_i \in \Phi_i$ is such that

$$\varphi_i(C^\varepsilon_{-i}(\Gamma^{k-1})) \geq F_i(\varepsilon)\ \ \forall\varepsilon \geq 0.$$

Since $\Phi_j \subseteq B^{k-1}_j$ for all j by the induction assumption, then for every player j

$$C^\varepsilon_j(\Phi_j) \subseteq C^\varepsilon_j(B^{k-1}_j) = C^\varepsilon_j(\Gamma^{k-1})\ \ \ \forall\varepsilon \geq 0.$$

Hence,

$$\varphi_i(C^\varepsilon_{-i}(B^{k-1}_{-i})) \geq \varphi_i(C^\varepsilon_{-i}(\Phi_{-i})) \geq F_i(\varepsilon)\ \ \forall\varepsilon \geq 0.$$

Thus, $\Phi_i \subseteq B^k_i$ for any $k \geq 2$ for every i. Since this holds for every k, then by definition of B^∞_i, it follows that $\Phi_i \subseteq B^\infty_i$ for every player i.

A.5. Lemma 1

Let $C^\varepsilon_i(\Gamma^k) = \{c_i \in C_i : \varepsilon^{c_i}(\Gamma^k) \leq \varepsilon\}$, and let $B^k_i := \{b_i \in \triangle(C_{-i}) \mid b_i(C^\varepsilon_{-i}(\Gamma^{k-1})) \geq F_i(\varepsilon)\ \forall\varepsilon \geq 0\}$. Since B^k_i can be defined by finitely many linear inequalities (Lemma 2), it is possible to define the set $EB^k_i := \{b_i \in B^k_i \mid b_i \text{ is an extreme point of } B^k_i\}$. Note that since $B^0_i = \triangle(C_{-i})$, $EB^0_i := \{b_i \in \triangle(C_{-i}) \mid b_i(c_{-i}) = 1 \text{ for some } c_{-i} \in C_{-i}\}$. Lemma 1 then requires proving:
(A) For every $c^*_i \notin C^\varepsilon_i(\Gamma^k)$ for some $\varepsilon \geq 0$, there exists $r_i \in \triangle(C_i \setminus \{c^*_i\})$ such that $u_i(r_i, b_i) > u_i(c^*_i, b_i) + \varepsilon$ for all $b_i \in B^k_i$.
(B) For every c^*_i for which there exists $r_i \in \triangle(C_i \setminus \{c^*_i\})$ such that $u_i(r_i, b_i) > u_i(c^*_i, b_i) + \varepsilon$ for all $b_i \in B^k_i$, it is also the case that $c^*_i \notin C^\varepsilon_i(\Gamma^k)$.
(C) For $c^*_i \in C_i$ and for some given k, the critical epsilon $\varepsilon^{c^*_i}(\Gamma^k)$ is defined by

$$\varepsilon^{c^*_i}(\Gamma^k) = Max\{\varepsilon\ |\varepsilon \geq 0, \exists\ r_i \text{ with } u_i(r_i, b_i) \geq u_i(c^*_i, b_i) + \varepsilon\ \ \ \forall b_i \in B^k_i\}.$$

Proof of part A: Let $c^*_i \notin C^\varepsilon_i$ for some $\varepsilon \geq 0$, then for every $b_i \in B^k_i$, there is some c_i such that

$$u_i(c_i, b_i) > u_i(c^*_i, b_i) + \varepsilon$$

Let Γ^k_ε be a modified game of Γ^k with transformed utilities as follows:

For $c^*_i \in C_i$	$u^\varepsilon_i(c^*_i, b_i) = u(c^*_i, b_i) + \varepsilon$	$\forall b_i \in B^k_i$
For every $c_i \in C_i \setminus \{c^*_i\}$	$u^\varepsilon_i(c_i, b_i) = u_i(c_i, b_i)$	$\forall b_i \in B^k_i$.

So for every $b_i \in B^k_i$, there exists some c_i such that

$$u^\varepsilon_i(c_i, b_i) > u^\varepsilon_i(c^*_i, b_i).$$

Lemma 3 in Pearce (1984) then implies, there exists $r_i \in \triangle(C_i \setminus \{c^*_i\})$ such that

$$u^\varepsilon_i(r_i, b_i) > u^\varepsilon_i(c^*_i, b_i)\qquad \forall b_i \in EB^k_i.$$

Or equivalently,

$$u_i(r_i, b_i) > u_i(c^*_i, b_i) + \varepsilon\qquad \forall b_i \in EB^k_i.$$

By the convexity of B^k_i, it follows there is some $r_i \in \triangle(C_i \setminus \{c^*_i\})$ such that

$$u_i(r_i, b_i) > u_i(c^*_i, b_i) + \varepsilon\qquad \forall b_i \in B^k_i.$$

Proof of part B: Consider some c^*_i for which there exists $r_i \in \triangle(C_i)$ such that

$$u_i(r_i, b_i) > u_i(c^*_i, b_i) + \varepsilon\qquad \text{for all}\qquad b_i \in B^k_i.$$

Since $EB^k_i \subset B^k_i$, c^*_i is also such that there is some $r_i \in \triangle(C_i)$ with

$$u_i(r_i, b_i) > u_i(c^*_i, b_i) + \varepsilon\qquad \text{for all}\qquad b_i \in EB^k_i.$$

In Γ_ε^k this becomes,

$$u_i^\varepsilon(r_i, b_i) > u_i^\varepsilon(c_i^*, b_i) \quad \text{for all} \quad b_i \in EB_i^k.$$

Lemma 3 in Pearce (1984) then implies for every $b_i \in B_i^k$ there exists $c_i \in C_i \setminus \{c_i^*\}$ such that

$$u_i^\varepsilon(c_i, b_i) > u_i^\varepsilon(c_i^*, b_i).$$

Or equivalently,

$$u_i(c_i, b_i) > u_i(c_i^*, b_i) + \varepsilon.$$

Thus $c_i^* \notin C_i^\varepsilon(\Gamma^k)$.

Proof of part C: Consider some c_i^* and let $\varepsilon^* = Max\{\varepsilon \mid \varepsilon \geq 0, \exists\, r_i \text{ with } u_i(r_i, b_i) \geq u_i(c_i^*, b_i) + \varepsilon \quad \forall b_i \in B_i^k\}$. Then Show:

1. $\varepsilon^{c_i^*}(\Gamma^k) \leq \varepsilon^*$.
   Assume $\varepsilon^{c_i^*}(\Gamma^k) > \varepsilon^*$ and so $c_i^* \notin C_i^{\varepsilon^*}(\Gamma^k)$. Then part A implies that there is some $r_i \in \triangle(C_i \setminus \{c_i^*\})$ such that

   $$u_i(r_i, b_i) > u_i(c_i^*, b_i) + \varepsilon^* \quad \text{for all} \quad b_i \in B_i^k$$

   This leads to a contradiction by definition of ε^*.
2. $\varepsilon^{c_i^*}(\Gamma^k) \geq \varepsilon^*$.
   Assume $\varepsilon^{c_i^*}(\Gamma^k) < \varepsilon^*$. So by definition of ε^* and for some $r_i \in \triangle(C_i \setminus \{c_i^*\})$,

   $$u_i(r_i, b_i) \geq u_i(c_i^*, b_i) + \varepsilon^* \quad \text{for all} \quad b_i \in B_i^k.$$

   So in particular,

   $$u_i(r_i, b_i) > u_i(c_i^*, b_i) + \varepsilon^{c_i^*}(\Gamma^k) \quad \text{for all} \quad b_i \in B_i^k.$$

   From part b, the latter implies that $c_i^* \notin C_i^{\varepsilon^{c_i^*}}(\Gamma^k)$ which creates a contradiction by definition of $C_i^\varepsilon(\Gamma^k)$. Combining parts 1 and 2 proves that $\varepsilon^{c_i^*}(\Gamma^k) = \varepsilon^*$.

A.6. Lemma 2

Let $\{\varepsilon_{-i}^1, \ldots, \varepsilon_{-i}^M\}$ be an ascending ranking of critical epsilons of the set of choices C_{-i}, where $M = \mid C_{-i} \mid$. Lemma 2 then requires proving two parts:
(A) Any $b_i \in \triangle(C_{-i})$ satisfying the finite set of inequalities $b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq Lim_{\varepsilon \uparrow \varepsilon_{-i}^{m+1}} F_i(\varepsilon)$ for all $m \in \{1, \ldots, M-1\}$, also satisfies the infinite set of inequalities of the form $b_i(C_{-i}^\varepsilon(\Gamma^k)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$.
(B) Any $b_i \in \triangle(C_{-i})$ satisfying the infinite set of inequalities of the form $b_i(C_{-i}^\varepsilon(\Gamma^k)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$, also satisfies the finite set of inequalities $b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq Lim_{\varepsilon \uparrow \varepsilon_{-i}^{m+1}} F_i(\varepsilon)$ for all $m \in \{1, \ldots, M-1\}$.

Proof of part A: Let $b_i \in \triangle(C_{-i})$ be some arbitrary belief satisfying $b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq Lim_{\varepsilon \uparrow \varepsilon_{-i}^{m+1}} F_i(\varepsilon)$ for all $m \in \{1, \ldots, M-1\}$. Consider a subset of the infinite set of inequalities $b_i(C_{-i}^\varepsilon(\Gamma^k)) \geq F_i(\varepsilon)$ for all $\varepsilon \geq 0$ with $\varepsilon \in [\varepsilon^m, \varepsilon^{m+1})$ for some $m \in \{1, \ldots, M-1\}$.

By definition of $C_{-i}^\varepsilon(\Gamma^k)$ and since Γ^k is finite for any $k \geq 0$,

$$C_{-i}^\varepsilon(\Gamma^k) = C_{-i}^{\varepsilon^m}(\Gamma^k) \quad \text{for all} \quad \varepsilon \in [\varepsilon^m, \varepsilon^{m+1}).$$

So

$$b_i(C_{-i}^\varepsilon(\Gamma^k)) \geq F_i(\varepsilon) \quad \text{for all} \quad \varepsilon \in [\varepsilon^m, \varepsilon^{m+1})$$

implies,

$$b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq F_i(\varepsilon) \quad \text{for all} \quad \varepsilon \in [\varepsilon^m, \varepsilon^{m+1}).$$

Since $F_i(\varepsilon)$ is weakly increasing in ε, it follows that $b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq Lim_{\varepsilon \uparrow \varepsilon^{m+1}} F_i(\varepsilon)$ implies $b_i(C_{-i}^{\varepsilon^m}(\Gamma^k)) \geq F_i(\varepsilon)$ for all $\varepsilon \in [\varepsilon^m, \varepsilon^{m+1})$, and so also implies $b_i(C_{-i}^\varepsilon(\Gamma^k)) \geq F_i(\varepsilon)$ for all $\varepsilon \in [\varepsilon^m, \varepsilon^{m+1})$.

Proof of part B: Proving part B is trivial since the finite set of inequalities is by definition of $C_{-i}^\varepsilon(\Gamma^k)$ a subset of the infinite set of inequalities. So satisfying the latter automatically implies satisfying the former.

References

Bach, C., Perea, A., 2014. Utility proportional beliefs. Internat. J. Game Theory 43, 881–902.

Bach, C., Perea, A., 2016. Incomplete information and common belief in rationality, EPICENTER Working Paper No. 7, Maastricht University.

Basu, K., 1994. The traveler's dilemma: Paradoxes of rationality in game theory. Amer. Econ. Rev. 84 (2), 391–394.

Battigalli, P., Siniscalchi, M., 2003. Rationalization and incomplete information. Adv. Theor. Econ. 3 (1), 1–46.

Baye, M.R., Morgan, J., 2004. Price dispersion in the lab and on the Internet: theory and evidence. Rand J. Econ. 35, 449–466.

Becker, T., Carter, M., Naeve, J., 2005. Experts playing the traveler's dilemma, Department of Economics Working Paper Nr. 252/2005, University of Hohenheim.

Bernheim, D., 1984. Rationalizable strategic behaviour. Econometrica 52 (4), 1007–1028.

Brandenburger, A., Dekel, E., 1987. Rationalizability and correlated equilibria. Econometrica 55 (6), 1391–1402.

Dekel, E., Fudenberg, D., Morris, S., 2006. Topologies on types. Theor. Econ. 1, 275–309.

Friedenberg, A., 2010. When do type structures contain all hierarchies of beliefs? Games Econom. Behav. 68, 108–129.

Heifetz, A., Samet, D., 1998. Topology - free typology of beliefs. J. Econom. Theory 82, 324–341.

Hu, T., 2007. On p-rationalizability and approximate common certainty of rationality. J. Econom. Theory 136, 379–391.

March, J.H., Simon, H., 1958. Organizations. John Wiley and Sons.

McKelvey, R., Palfrey, T., 1995. Quantal response equilibria for normal form games. Games Econom. Behav. 10, 6–38.

Moderer, D., Samet, D., 1989. Approximating common knowledge with common beliefs. Games Econom. Behav. 1, 170–190.

Nagel, R., 1995. Unravelling in guessing games: An experimental study. Amer. Econ. Rev. 85, 1313–1326.

Pearce, D., 1984. Rationalizable strategic behaviour and the problem of perfection. Econometrica 52 (4), 1029–1050.

Perea, A., Roy, S., 2014. A New Epistemic Characterization of Proper Rationalizability. Mimeo.

Radner, R., 1980. Collusive behaviour in non-cooperative epsilon equilibria of oligopolies with long but finite lives. J. Econom. Theory 22, 121–157.

Radner, R., 1981. Monitoring cooperative agreements in a repeated principal-agent relationship. Econometrica 49 (5), 1127–1148.

Rosenthal, R., 1989. A bounded-rationality approach to the study of noncooperative games. Internat. J. Game Theory 18, 273–292.

Tan, T., Werlang, S., 1988. The bayesian foundations of solution concepts of games. J. Econom. Theory 45, 370–391.