(Bayesian) Coalitional Rationalizability*

Xiao Luo†

Institute of Economics, Academia Sinica, Taipei 115, Taiwan, ROC

and

Chih-Chun Yang

Department of Economics, National Taiwan University, Taipei 100, Taiwan, ROC

May 2006

Abstract

We extend Ambrus's [*QJE*, 2006] concept of "coalitional rationalizability (c-rationalizability)" to situations where, in seeking mutual beneficial interests, players in groups (i) make use of Bayes rule in expectation calculations and (ii) contemplate various deviations – i.e. the validity of deviation is checked against any arbitrary sets of strategies, not only against restricted subsets of strategies. In this paper we offer an alternative notion of c-rationalizability suitable for such complicated interactions. More specifically, following Bernheim's [*Econometrica* **52**(1984), 1007-1028] and Pearce's [*Econometrica* **52**(1984), 1029-1051] approach, we define c-rationalizability by the terminology "coalitional rationalizable set (CRS)". Roughly speaking, a CRS is a product set of pure strategies from which no group of player(s) would like to deviate. We show that this notion of c-rationalizability possesses nice properties similar to those of conventional rationalizability. We also provide its epistemic foundation. *JEL Classification: C70, C72, D81*

---

*This paper was partially done while the first author was visiting Harvard University. We especially thank Attila Ambrus for his encouragement and helpful comments. We also thank Steve Chiu, Satoru Takahashi, and participants in seminars at Harvard University and Taiwan Game Theory Group. Financial support from the National Science Council of Taiwan is gratefully acknowledged. The usual disclaimer applies.

†Corresponding author. Fax:+886-2-2785 3946. *email:* xluo@econ.sinica.edu.tw

1 Introduction

In their seminal paper, Bernheim (1984) and Pearce (1984) proposed the solution concept of rationalizability as the logical implication of common knowledge of Bayesian rationality (see also Tan and Werlang (1988)). In strategic games, the concept of rationalizability can be defined in terms of a best-response (product) set of strategies, which reflects the idea that "rational" behavior should be justified by "rational" beliefs and conversely, "rational" beliefs should be based on "rational" behavior. In finite games, the set of rationalizable strategies can be derived from iterative deletion of never best response strategies (see Bernheim's (1984) Proposition 3.1 and Pearce's (1984) Proposition 2) and, if moreover correlated beliefs are permitted, it is equivalent to iterated (strict) dominance. The concept of rationalizability aims to be weak: it determines not what actions should actually be taken, but what actions could be ruled out with confidence.

However, the concept of rationalizability does not take into account the intriguing and important possibility that groups or coalitions of players would be willing to coordinate their moves, in order to achieve mutual beneficial outcomes. For example, in the following coordination game:

	a	b
a	2,2	0,0
b	0,0	1,1

it is pretty clear that two players can coordinate their actions to achieve the best payoffs (2,2). To capture these important aspects of collective strategic behavior, Ambrus (2006) first offered a solution concept of "coalitional rationalizability" (henceforth, c-rationalizability) by using an iterative procedure of restrictions, in which members of a coalition will confine play to a subset of their strategies if it is in their mutual interest to do so. In this coordination game c-rationalizability predicts the unique payoff-dominant outcome (a,a); in particular, it excludes the strict Nash equilibrium (b,b).

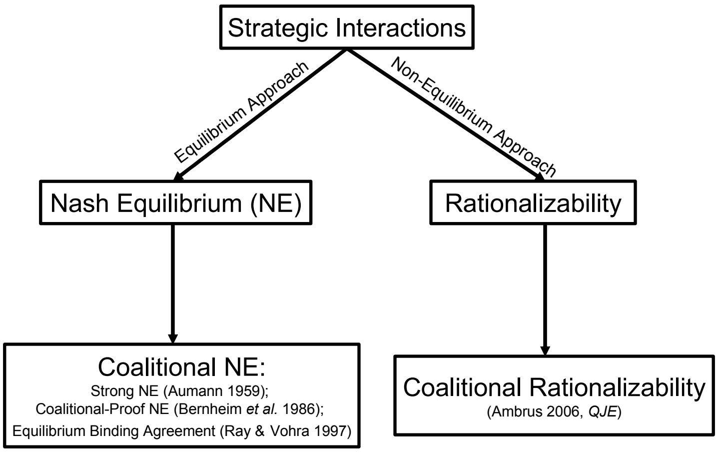

Figure 0

The main purpose of this paper is to further study the logical implication of common knowledge of the fact that every player is Bayesian rational and also aware of arrangements of strategies within conceivable coalitions. We extend Ambrus's (2006) notion of c-rationalizability to situations where, in pursuit of mutual beneficial interests, players in groups (i) make use of Bayes rule in expectation calculations and (ii) contemplate various deviations – i.e. the validity of deviation is checked not only against restricted subsets of strategies, *but also against arbitrary sets of strategies*. That is, this paper is motivated mainly by the following two considerations: First, in the standard view a Bayesian player should update, using Bayes rule, his/her subjective prior beliefs on conditional events. In other words, a rational player who conforms to Savage's (1954) axioms must update his/her prior beliefs according to Bayes rule.[1] Thus, when players consider implicit agreements to confine their collective actions within a set of strategies, each player should make Bayesian updating of the initial priors and calculate his/her resulting expected payoffs of making the implicit agreements. In this paper, we offer an alternative notion of c-rationalizability (Definition 1) to accommodate the effect of Bayesian updating.[2]

[1] The decision-theoretic foundation of this rule is laid by the axiomatization of Savage (1954); see, e.g., Kreps (1988, Chapter 10) and Myerson (1991, Section 1.4). See also Epstein and LeBreton (1993) for more extensive discussions and Ghirardato (2002) for a recent formalization of Bayesian updating.

[2] Ambrus (2006) did not consider Bayesian updating, and he required the marginal distribu-

Second, on a conceptual level, Ambrus's (2006) analysis does not go far enough to account for all the aspects of collective and coalitional "stability;" Ambrus's notion of c-rationalizability requires the validity of deviation check *only* against restricted subsets of strategies when players in groups contemplate deviations – i.e. a valid deviation is a deviation from a product set of strategies to its subset. However, there seems no reason to suppose that coalitional deviations are restricted only to subsets of strategies. In fact, members of deviating coalitions may in general confine, enlarge, or even revise and rearrange play to any arbitrary set of their strategies if doing so is in their mutual interest, and the validity of deviation should go further to check against completely free and unrestrained any sets of strategies.[3] For example, in the aforementioned coordination game, the superior coordination outcome (a,a) is coalitionally stable, but the inferior coordination outcome (b,b) cannot be regarded as a coalitionally stable one (since players can improve their payoffs by jointly moving to (a,a)). The alternative notion of c-rationalizability is proposed here to accommodate the presence of very universal coalition deviations.[4]

In order to easily understand our analysis in this paper, we follow closely Pearce's (1984) and Bernheim's (1984) approach to conventional rationalizability and carry out a *ceteris paribus* study of the coalitional version of rationalizability. All the major features and nice properties of the conventional rationalizability, as a special case of c-rationalizability with the restriction to singleton coalitions only, are essentially preserved. More specifically, we define the notion of c-rationalizability by the "c-rationalizable set (CRS)," which can be viewed as the counterpart of the "best-response set" in the notion of conventional rationalizability (see Definition 1). The central result of this paper is to show that there is a (nonempty) largest CRS (Theorem 1). Moreover, the largest CRS can be derived from an iterative procedure of restrictions to c-best response strategies (Proposition 1); any Pareto-dominant pure Nash equilibrium and, hence, any

tion on the strategies of players outside the coalition to be fixed (see also Footnote 8).

[3]Based on a similar consideration, Kahn and Mookherjee (1994) proposed 'universal coalition-proof equilibrium' to accommodate more general coalition formation in Bernheim et al.'s (1987) 'coalition-proof equilibrium'. Unfortunately, the notion of 'universal coalition-proof equilibrium' may fail to exist; see Section 5(1) for more discussion.

[4]Ambrus's original notion of c-rationalizability cannot be extended to the case of universal coalition deviations; see Section 4(1).

strong pure Nash equilibrium are c-rationalizable (Proposition 2). We next formulate the coalitional version of iterated strict dominance (Definition 2), and show it is equivalent to the notion of c-rationalizability (Theorem 2). We also provide its epistemic foundation (Theorem 3).

The rest of this paper is organized as follows. Section 2 contains some preliminaries. Section 3 offers the definition of c-rationalizability. Subsection 3.1 investigates the existence of c-rationalizability; Subsection 3.2 formulates the notion of "c-dominance" and shows an equivalence theorem; Subsection 3.3 provides an epistemic characterization of c-rationalizability. Section 4 is concluding remarks. To facilitate reading, all the proofs are relegated to Appendix. Throughout this paper, we restrict our attention to pure strategies and, moreover, players are allowed to hold correlated beliefs about the strategies of their opponents.[5]

2 Preliminaries

Let Δ denote the space of the probability distributions on a finite space S of states.[6] For $\mu \in \Delta$, let $\mu(A)$ denote the probability of an arbitrary subset A of states under the distribution μ and, if $\mu(A) \neq 0$, let $\mu|_A$ denote the probability distribution conditional on A. (We decree that $\mu|_\emptyset \equiv \mu$.) We denote the *space of probability distributions conditional on* A by

$$\Delta|_A \equiv \{\mu \in \Delta |\ \mu(A) = 1\}.$$

Let $\Delta^0 \equiv \{\mu \in \Delta |\ \mu(s) > 0\ \forall s \in S\}$.

Let $\mu^n \overset{A}{\rightsquigarrow} \mu$ denote "$\mu^n \to \mu$ in $\Delta|_A$ with $\mu^n \in \Delta^0$" – i.e., a "trembling" sequence $\{\mu^n\}$ of full-support distributions on S converges to distribution μ in $\Delta|_A$. For an arbitrary subset B of states, let $\mu|_B \equiv \lim_{n\to\infty} \mu^n|_B$ if $\mu(B) = 0$. That is, $\mu|$ is a "conditional probability system (CPS)"; see Myerson (1986, 1991). Clearly, if $\mu(B) \neq 0$, $\mu|_B = \lim_{n\to\infty} \mu^n|_B$ for *all* sequence $\mu^n \overset{A}{\rightsquigarrow} \mu$. If $\mu(B) = 0$, $\mu|_B$ could be an arbitrary conditional probability distribution in $\Delta|_B$, depending upon the chosen "trembling" sequence (Lemma 1 in Appendix).

[5] See Aumann (1974, 1987) for extensive discussions on subjectivity and correlation.

[6] Δ is endowed with the weak* topology, and may be regarded as a simplex in a Euclidean space of dimension $|S|$.

3 Coalitional rationalizability

Consider a finite game: $G \equiv (I, \{S_i\}_{i\in I}, \{u_i\}_{i\in I})$, where I is a nonempty finite set of players, S_i is a nonempty finite set of i's strategies, and $u_i : S \equiv \times_{i\in I} S_i \to \Re$ is i's payoff function. For (nonempty) coalition $J \subseteq I$, let $S_J \equiv \times_{j\in J} S_j$, let $S_{-J} \equiv \times_{i\notin J} S_i$, and let $S_{-j} \equiv \times_{i\neq j} S_i$.

Let A and B be subsets of S in product-form. Say a coalition $\mathcal{J}_{AB}$ is a "feasible coalition from A to B" if $B = B_{\mathcal{J}_{AB}} \times A_{-\mathcal{J}_{AB}}$ – i.e. $\mathcal{J}_{AB}$ can be interpreted as a coalition by which set A can be rearranged to set B. Clearly, the coalition $\mathcal{J}_{AB}$ must include a player j with $A_j \neq B_j$. Note that B is not necessarily a subset of A.

Definition 1. A nonempty product subset $\mathcal{R} \subseteq S$ is a *coalitional rationalizable set (CRS)* if $\mathcal{R} \Rrightarrow \mathcal{R}'$ only for $\mathcal{R}' = \mathcal{R}$, where $\mathcal{R} \Rrightarrow \mathcal{R}'$ (via $\mathcal{J}_{\mathcal{R}\mathcal{R}'}$) is defined as: $\forall j \in \mathcal{J}_{\mathcal{R}\mathcal{R}'}$

(1.1) $\forall r_j \in \mathcal{R}_j$, $u_j(r_j, \mu) < u_j(s_j, \mu|_{\mathcal{R}'_{-j}})$ for some $s_j \in S_j$, where $\mu^n|_{\mathcal{R}'_{-j}} \to \mu|_{\mathcal{R}'_{-j}}$ ($\neq \mu$ if $r_j \in \mathcal{R}'_j$) as $\mu^n \overset{\mathcal{R}_{-j}}{\rightsquigarrow} \mu$, and

(1.2) $\forall r'_j \in \mathcal{R}'_j \backslash \mathcal{R}_j$, $u_j(r'_j, \mu|_{\mathcal{R}'_{-j}}) \geq u_j(s_j, \mu|_{\mathcal{R}'_{-j}})$ for all $s_j \in S_j$, for some $\mu^n|_{\mathcal{R}'_{-j}} \to \mu|_{\mathcal{R}'_{-j}}$ as $\mu^n \overset{\mathcal{R}_{-j}}{\rightsquigarrow} \mu$.

In particular, $r_j \in \mathcal{R}_j$ is said to be a *c-rationalizable strategy.*

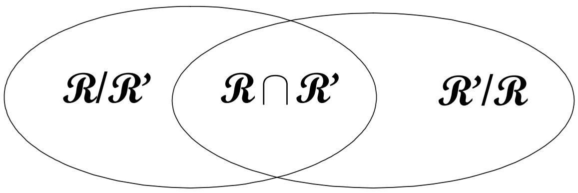

Figure 1. *A typical relationship between* $\mathcal{R}$ *and* $\mathcal{R}'$.

In Definition 1 $u_j(r_j, \mu)$ is j's expected payoff from using r_j given his prior belief μ, and $u_j(s_j, \mu|_{\mathcal{R}'_{-j}})$ is j's expected payoff from using s_j under his posterior belief $\mu|_{\mathcal{R}'_{-j}}$ after moving from $\mathcal{R}$ to $\mathcal{R}'$ via coalition $\mathcal{J}_{\mathcal{R}\mathcal{R}'}$. Conditions (1.1-2) say that: every player j in coalition $\mathcal{J}_{\mathcal{R}\mathcal{R}'}$, whatever prior belief j may hold, can always obtain, by using Bayes rule, a strictly higher expected payoff $u_j(s_j, \mu|_{\mathcal{R}'_{-j}})$ if the 'credible' move from $\mathcal{R}$ to $\mathcal{R}'$ is made. For example, in the Prisoner's Dilemma game, the noncooperative Nash outcome is the unique CRS because the Pareto-optimal cooperative outcome is not "credible." (Technically, the mild 'credible' condition (1.2) purports to overcome the notorious problem of emptiness, typified by Condorcet's paradox, under the 'core-like' blocking arrangements.)[7]

Remark 1. It is easy to see that a CRS $\mathcal{R}$ necessarily satisfies both the "best response" and "closed under rational behavior" properties (see Basu and Weibull 1991), i.e. $\mathcal{R}_i = BR(\mathcal{R}_{-i})\ \forall i$, where

$$BR(\mathcal{R}_{-i}) \equiv \{s_i \in S_i |\ \exists \mu \in \Delta|_{\mathcal{R}_{-i}} \text{ s.t. } u_i(s_i, \mu) \geq u_i(s_i', \mu) \forall s_i' \in S_i\}.$$

Definition 1, with the restriction of $|\mathcal{J}_{\mathcal{R}\mathcal{R}'}| = 1$, is essentially the correlated version of rationalizability (cf. Luo 2001, Section 4.1); every c-rationalizable strategy is rationalizable. It is also straightforward to verify that there is a Nash equilibrium with the support in $\mathcal{R}$.

Remark 2. We follow Selten's (1975) idea of "trembling-hand" to offer a way of updating beliefs in contingencies with zero prior probability in complex coalitional interactions. To see how Bayes rule plays a role, consider a two-person game:

	a	b	c
a	$3,0$	$0,3$	$0,2$
b	$0,3$	$3,0$	$0,0$
c	$2,0$	$0,0$	$1,1$

In this game, $S \Rrightarrow \{a, b\} \times \{a, b\}$ via coalition $\{1, 2\}$. Intuitively, if each player assigns a prior probability of less than 0.5 to a of the opponent and hence the player could achieve an expected payoff of less than 1.5 by using c, then it is beneficial for the two players to move to $\{a, b\} \times \{a, b\}$ because each can guarantee

[7]Definition 1(1.2) requires coalitional deviations be immune to further individual deviations. This requirement is also in the same spirit as Milgrom and Roberts's 'strongly coalition proof equilibria' and Kaplan's 'semistrong equilibria' (see Milgrom and Roberts 1996, p.115). Cf. also Roth's (1976) 'protected' condition.

a higher expected payoff of 1.5; and if each player assigns a prior probability of more than or equal to 0.5 to a, then it is also beneficial for the two players to move to $\{a, b\} \times \{a, b\}$ because each can achieve a higher expected payoff by using a instead of c. (Without using Bayes rule, the player could achieve an expected payoff of 2 by using c, higher than that by using a or b after the deviation. In this game, every strategy is c-rationalizable in Ambrus's (2006) sense.)[8]

3.1 Existence

The central result in this paper is that there is a largest (w.r.t. set inclusion) CRS:

$$\mathcal{R}^* \equiv \bigcup_{\mathcal{R} \text{ is a CRS}} \mathcal{R}$$

that consists of the union of all CRSs. Formally, we have

Theorem 1. *$\mathcal{R}^*$ is a largest CRS.*

The proof of Theorem 1 in Appendix shows that the set of c-rationalizable strategies can be derived from an iterative procedure of restrictions to c-best response strategies. Moreover, every such procedure leads to the same outcome. Let "$A \Rightarrow B$ (via $\mathcal{J}_{AB}$)" denote "$A \Rrightarrow B$ (via $\mathcal{J}_{AB}$) with $B \subseteq A$." Note that $A \Rightarrow A$. Formally, we have[9]

Proposition 1. *$\mathcal{R}^* = \mathcal{D}$ where $\mathcal{D} \equiv \bigcap_{k=0}^{\infty} \mathcal{D}_k$ with $\mathcal{D}^0 = S$, $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$, and $\mathcal{D} \Rightarrow \mathcal{D}'$ only for $\mathcal{D}' = \mathcal{D}$.*

[8] Ambrus's (2006) notion of c-rationalizability requires that the marginal expectation concerning the strategies of players outside the coalition be fixed. In interpreting the 'fixed marginal' requirement for correlated beliefs, players should believe as if the opponents play the strategies recommended by a mediator in conformity with *a prior* distribution; see, e.g., Milgrom and Roberts (1996, Remark on p.118). This sort of updating beliefs is however different from the Bayesian updating rule commonly used in economics. (Ambrus (2006) implicitly used Bayes rule in a situation where an eliminated strategy is not a best response to a belief over the underlying supported restriction.)

[9] See Dufwenberg and Stegeman (2002), Chen et al. (2005), and Apt (2005) for the similar formulation of iterative procedures.

As individual rationality can be considered a special case of c-rationality with respect to a minimal coalition, in general we would not expect c-rationalizable outcomes to be Pareto efficient. For example, in the Prisoner's Dilemma game, the "noncooperation" action remains only the c-rationalizable strategy. Given that c-rationalizability is built on the idea that players try to attain common aspirations, we can, however, establish some relationships to Pareto efficient (Nash) outcomes. Let $\mathcal{G}'=(I, \{S_i'\}, \{u_i\})$ denote the reduced game after iterated elimination of strictly dominated strategies. Formally, we have

Proposition 2. *(2.1) Any Pareto-dominant equilibrium is c-rationalizable. Moreover, any strong Nash equilibrium is c-rationalizable. (2.2) Any Pareto-best strategy profile is a c-rationalizable strategy profile; any strong Pareto-best strategy profile is a unique c-rationalizable strategy profile. Moreover, if there is a (strong) Pareto-best strategy profile s_J^* for a coalition J in the reduced game G' – i.e., for all $j \in J$, $(u_j(s_J^*, s_{-J}') > u_j(s)$ $\forall s, s' \in S'$ with $s_J^* \neq s_J)$ $u_j(s_J^*, s_{-J}) \geq u_j(s)$ $\forall s \in S'$, then s_j^* is a (unique) c-rationalizable strategy for player j. (2.3) In a common interest game where there is a strategy profile which strictly Pareto-dominates all other strategy profiles, the Pareto-best Nash equilibrium is a unique c-rationalizable strategy profile. (2.4) In the class of games with strategic complementarities where the players' payoffs are all monotonic functions of the opponents' strategies as defined in Milgrom and Roberts (1996, p.124), there is a c-rationalizable Pareto-best Nash equilibrium s^* in $\mathcal{G}'$* and if, moreover, *s^* is a strict Nash equilibrium and each player's payoffs are strict monotonic functions of the opponents' strategies at s^*, then s^* is a unique c-rationalizable strategy profile.*

Proposition 2 is simple but useful in the study of c-rationalizable behavior. Since the Nash equilibrium concept does not take into account coalitional rationality, a Nash equilibrium may fail to be c-rationalizable. Proposition 2(2.1) asserts that Pareto-dominant and strong Nash equilibria must be c-rationalizable. In the class of common interest games, the unique c-rationalizable strategy profile is the Pareto-best strong Nash equilibrium even though there are other strict Nash equilibria. We would also like to point out that Ambrus's (2006) motivating example of "Voting with Costly Participation" can be easily analyzed from the perspective of Proposition 2(2.2).[10] Indeed, any Pareto-best element in the set of strategies surviving iterated deletion of dominated strategies is a c-rationalizable strategy profile and, moreover, it is a coalition-proof equilibrium for any admissible coalition communication structure; see Milgrom and Roberts (1996, Theorem 1).

As pointed out by Milgrom and Roberts (1990), many games of economic interest, e.g., the Bertrand pricing game and the macroeconomic coordination games, belong to the class of games with strategic complementarities in Proposition 2(2.3); see also Vives (2005). While there is a wide range of Nash equilibria and rationalizable strategies in this class of games, the notion of c-rationalizability can provide a unique prediction, which is particularly interesting for games with no strong Nash equilibrium. The same is true for games with a unique Nash equilibrium because the games are indeed dominance solvable; see Milgrom and Roberts (1990, Theorem 6). To illustrate Proposition 2(2.3), consider the following parametric game where $\theta \in [1, 2]$:

	a	b
a	2,2	0,θ
b	θ,0	1,1

It is easy to verify that this game with any lattice structure on strategies is a game with strategic complementarities in Proposition 2. The Pareto-best Nash equilibrium (a, a) is a *unique* c-rationalizable strategy profile if $\theta < 2$. (Every strategy is c-rationalizable when $\theta = 2$. In this case, (a, a) is not a strict Nash equilibrium.)

Remark 3. Throughout this paper, we restrict our attention to pure strategies. As far as mixed strategies are concerned, a Pareto-dominant Nash equilibrium may fail to be c-rationalizable. Consider a three-person game (where players 1, 2, and 3 pick the row, column, and matrix, respectively):

[10] In that example it is a strictly dominated strategy for voter 3 to show up and vote for alternative A. Thus, it is c-rationalizable only for voters 1 and 2 to show up and vote for alternative A and, then, voter 3 would choose to stay at home. Indeed, for Ambrus' notion of c-rationalizability, the 'uniqueness' condition in Corollary 2(2.2) can be relaxed as follows: for all $j \in J$, $u_j(s_J^*, s_{-J}) > u_j(s)\ \forall s \in S'$ with $s_J^* \neq s_J$.

	a	b	c
a	$6,6,0$	$2,0,0$	$0,5,6$
b	$0,2,0$	$0,0,0$	$0,0,0$
c	$5,0,6$	$0,0,0$	$1,1,0$

a

	a	b	c
a	$2,2,2$	$2,2,2$	$0,0,0$
b	$2,2,2$	$2,2,2$	$0,0,0$
c	$0,0,0$	$0,0,0$	$0,0,0$

b

In this game, it easy to verify that the largest CRS is $\{a, b\} \times \{a, b\} \times \{\mathbf{b}\}$ in which each player gets a constant payoff of 2. The profile $(0.5a + 0.5c,\ 0.5a + 0.5c, \mathbf{a})$ is a Pareto-dominant mixed Nash equilibrium with an expected payoff of 3 for each player; in particular, player 3's strategy **a** is not c-rationalizable.

3.2 Iterated c-dominance

The iterative procedure in Proposition 1 requires that, whatever prior beliefs are held, players in a coalition exclude "inferior" strategies from their considerations if it is in their mutual interest to do so. This procedure, however, needs complicated calculations of expected utility. To the extent that players are concerned about their decisions, it is convenient to have a similar notion of (strict) dominance under the consideration of strategy arrangements within coalitions.[11] We here formulate the belief-free notion of "c-dominance" and show an equivalence theorem between iterated c-dominance and c-rationalizablity.

Definition 2. A subset $B \subseteq A$ is a *c-dominated restriction from* A, denoted by $A \Downarrow B$ (via $\mathcal{J}_{AB}$), if $\forall j \in \mathcal{J}_{AB}$, $\forall a_j \in A_j$, $\forall c_j \in A_j \backslash B_j$, $\forall b_{-j} \in B_{-j}$, $\forall c_{-j} \in A_{-j} \backslash B_{-j}$,

(2.1) $u_j(c_j, b_{-j}) < u_j(\sigma_j, b_{-j})$ for some $\sigma_j \in \Delta(S_j)$, and

(2.2) $u_j(a_j, c_{-j}) < u_j(\sigma_j, b_{-j})$ for some $\sigma_j \in \Delta(S_j)$,

where $u_j(\sigma_j, b_{-j}) \equiv \sum_{s_j \in S_j} \sigma_j(s_j)\, u_j(s_j, b_{-j})$.

[11] With no coalition considerations, it is well known that a strategy is a never best response to any correlated conjecture concerning the opponents' moves if, and only if, it is strictly dominated possibly by a mixed strategy; see, e.g. Pearce (1984, Proposition 2). See also Shimoji and Watson's (1998) and Shimoji's (2004) related work on extensive-form rationalizability.

In words, from coalition member j's point of view, c-dominated strategy $c_j \in A_j \backslash B_j$ is strictly dominated given set B in the usual sense. Furthermore, given any $-j$'s c-dominated profile $c_{-j} \in A_{-j} \backslash B_{-j}$, j can always do better if coalition members jointly move from A to B, regardless of whatever $b_{-j} \in B_{-j}$ will be eventually used (cf. Figure 2).

	b_{-j}	c_{-j}
a_j	$u_j(a_j, b_{-j})$	$u_j(a_j, c_{-j})$
c_j	$u_j(c_j, b_{-j})$	$u_j(c_j, c_{-j})$

Figure 2: *The problem faced by j.*

The following Theorem 2 shows that the set of c-rationalizable strategies can be derived from an iterative elimination of c-dominated strategies. Moreover, the order of elimination does not matter. Formally, we have

Theorem 2. *$\mathcal{R}^* = \widehat{\mathcal{D}}$ where $\widehat{\mathcal{D}} \equiv \bigcap_{k=0}^{\infty} \widehat{\mathcal{D}}_k$ with $\widehat{\mathcal{D}}^0 = S$, $\widehat{\mathcal{D}}^k \Downarrow \widehat{\mathcal{D}}^{k+1}$, and $\widehat{\mathcal{D}} \Downarrow \widehat{\mathcal{D}}'$ only for $\widehat{\mathcal{D}}' = \widehat{\mathcal{D}}$.*

It is usually helpful to analyze the complex games by using the notion of (iterated) c-dominance. For example, the set of c-rationalizable strategies in Ambrus's (2006) example of "Dollar Division Game with External Reward" can be easily derived by applying one round of elimination of all c-dominated strategies. Let us reconsider the game in Remark 2:

	a	b	c
a	$3, 0$	$0, 3$	$0, 2$
b	$0, 3$	$3, 0$	$0, 0$
c	$2, 0$	$0, 0$	$1, 1$

In this game, $S \Downarrow \{a, b\} \times \{a, b\}$ via coalition $\{1, 2\}$. Intuitively, under restriction $\{a, b\} \times \{a, b\}$, (i) a strictly dominates c (that is ruled out by the restriction); (ii) $0.5a + 0.5b$ can guarantee an expected payoff of 1.5 that is higher than the best payoff of 1 resulting from the opponent's c (that is ruled out by the restriction).

3.3 Epistemic foundation

We now turn to exploring the epistemic foundation for our notion of c-rationalizablity. We offer sufficient/necessary epistemic conditions for the notion of c-rationalizablity, in terms of common knowledge of "c-rationality."

Consider a standard Aumann's semantic model of knowledge for game $\mathcal{G}$

$$\mathcal{M}(\mathcal{G}) \equiv (\Omega, \{P_i(.)\}, \{s_i(.)\}),$$

where Ω is the set of states with typical element $\omega \in \Omega$, $P_i(\omega) \subseteq \Omega$ is player i's information structure at ω, and $s_i(\omega) \in X_i$ is i's strategy at ω (cf. Osborne and Rubinstein 1994, Chapter 5).

For $A \subseteq S$ let $\Omega_A \equiv \{\omega \in \Omega |\ s(\omega) \in A\}$. Let $\widehat{u}_j(\nu|_{B_{-j}}) \equiv \max_{s_j \in S_j} u_j(s_j, \nu|_{B_{-j}})$. Let $\mathcal{J}^*_{AB}$ denote a "credible" coalition from A to B, i.e., $\mathcal{J}^*_{AB}$ is $\mathcal{J}_{AB}$ with the property that $\forall b_j \in B_j \backslash A_j$, $u_j(b_j, \mu|_{B_{-j}}) \geq u_j(s_j, \mu|_{B_{-j}})\ \forall s_j \in S_j$, for some $\mu^n|_{B_{-j}} \to \mu|_{B_{-j}}$ as $\mu^n \overset{A_{-j}}{\rightsquigarrow} \mu$. Let $\boxed{E}_J \subseteq E$ denote a *self-evident event in E among the players in coalition J,* i.e., $\boxed{E} \subseteq K_j \boxed{E} \subseteq E\ \forall j \in J$. Let $\boxed{E} \equiv \boxed{E}_I$ denote a self-evident event in E (among all the players).

Say *i is c-rational at ω* if, whenever $i \in \mathcal{J}^*_{AB}$ and $\omega \in \boxed{\Omega_A}_{\mathcal{J}^*_{AB}}$ such that $\forall j \in \mathcal{J}^*_{AB} \backslash \{i\}$, $\forall \omega' \in P_i(\omega)$, $u_j(s_j(\omega'), \nu) < \widehat{u}_j(\nu|_{B_{-j}})$ for all CPS (i.e. conditional probability system) $\nu|$ with $\nu|_{A_{-j}} = \nu$ in $\Delta|_{s_{-j}(P_j(\omega'))}$ (and $\nu|_{B_{-j}} \neq \nu$ if $s_j(\omega') \in B_j$), for $s_i(\omega) \notin B_i$ there is a CPS $\mu|$ with $\mu|_{A_{-i}} = \mu$ in $\Delta|_{s_{-i}(P_i(\omega))}$ such that

$$u_i(s_i(\omega), \mu) \geq u_i\left(s_i, \mu|_{B_{-i}}\right) \text{ for all } s_i \in S_i.$$

Let $R_i \equiv \{\omega \in \Omega |\ i \text{ is c-rational at } \omega\}$, and let $R \equiv \cap_{i \in N} R_i$. Note that, if $|\mathcal{J}^*_{AB}| = 1$, "c-rationality" is the (individual) rationality. The following theorem shows that c-rationalizablity can be regarded as the logical consequence of common knowledge of c-rationality. Formally, we have

Theorem 3. (3.1) For all $\omega \in \boxed{R}$, $s_i(\omega)$ is a c-rationalizable strategy. (3.2) For every c-rationalizable strategy r_i, there is a model of knowledge such that $r_i = s_i(\omega)$ for $\omega \in \boxed{R}$.

Remark 4. Our semantic framework does not require explicit specification of players' beliefs at an epistemic state; see, for example, Aumann (1995) and Lipman

(1994). We may, however, consider a set of CPS associated with a state. Ahn (2006) provided such a construction of universal type space where players are allowed to have a compact set of beliefs or CPS.

4 Concluding remarks

In many real life situations, groups of individuals often have an incentive to choose, voluntarily and without binding agreement, to coordinate their action choices and make joint decisions in noncooperative environments. Ambrus (2006) took the first step to offer a solution concept of c-rationalizability for situations in which coalitions can plan profitable deviations from an initial proposal to subsets of strategies by using no Bayesian updating. Following this line of research, we have proposed in this paper an alternative notion of c-rationalizability that applies to situations where, in seeking mutual beneficial interests, members in groups (i) make use of Bayes rule in expectation calculations and (ii) contemplate various deviations – i.e. the validity of deviation is checked not only against restricted subsets of strategies, but also against arbitrary sets of strategies.

We have shown that c-rationalizability is a well-defined solution concept that identifies consequences of common knowledge of coalitional rationality and, moreover, possesses similar nice properties of the conventional rationalizability. We have shown that the set of c-rationalizable strategies can be fully characterized by the largest CRS, which can derived from any iterative procedure of restrictions to c-best response strategies. We have formulated the coalitional version of dominance and, then, shown that the set of c-rationalizable strategies can be solved by performing any iterative deletion procedure of c-dominated strategies. Thus, we have offered a coalitional analogue to the connection between conventional rationalizability and iterated strict dominance, which is valuable on both practical and conceptual levels. Within the standard semantic framework, we have also provided an epistemic characterization of c-rationalizability in terms of common knowledge of coalitional rationality. As a consequence, our notion of c-rationalizability can be viewed as a natural extension of rationalizability *a la* Bernheim (1984) and Pearce (1984).

We would like to make some final remarks and conclusions:

1. Various coalitional equilibrium concepts in the literature, e.g. Aumann's (1959) strong Nash equilibrium, Bernheim et al.'s (1987) coalition-proof Nash equilibrium, and Ray and Vohra's (1997, 1999) equilibrium binding agreements, often fail to exist in a natural class of games. Yet, while Ambrus (2006) suggested a well-defined coalitional solution concept, his solution concept fails to be immune against being blocked by coalitions using more complex strategy arrangements. For example, consider Ambrus's (2006, Figure III) example of a three-person game:

	a	b
a	$2, 2, 2$	$0, 0, 0$
b	$0, 0, 0$	$3, 3, 0$

a

	a	b
a	$0, 0, 0$	$0, 0, 0$
b	$0, 0, 0$	$1, 1, 1$

b

Ambrus's notion of c-rationalizability prescribes the outcome $(1, 1, 1)$ which can, however, be improved on by the appealing Pareto-dominant Nash equilibrium outcome $(2, 2, 2)$. In contrast, our notion of c-rationalizability is a logically consistent solution concept that accommodates very complex coalitional reasoning. Each player always has a pure c-rationalizable strategy in our sense; in particular, any Pareto-dominant Nash equilibrium must be c-rationalizable (Proposition 2(2.2)).

It is also worthwhile to mention that Definition 1(1.1) requires that any deviating coalition member j who uses an unaffected strategy r_j in $\mathcal{R}_j \cap \mathcal{R}'_j$ can obtain a strictly higher expected payoff after the deviation takes place, whereas j's posterior belief $\mu|_{\mathcal{R}'_{-j}}$ is changed. The following three-person game shows that this requirement is necessary.

	a	b	c
a	$2, 2, 1$	$1, 1, 0$	$0, 0, 0$
b	$0, 0, 0$	$0, 0, 0$	$0, 0, 0$

a

	a	b	c
a	$1, 5, 0$	$1, 0, 2$	$0, 4, 2$
b	$0, 0, 0$	$0, 5, 2$	$0, 1, 2$

b

If the requirement were not in place, then there would be a logical inconsistency – i.e. there would be no 'c-rationalizable' strategy in this case. The major problem is that the grand coalition could deviate from $\{(a, a, \mathbf{a})\}$ to $\{a, b\} \times \{b, c\} \times \{\mathbf{b}\}$. Clearly, this sort of deviation is not sensible because player 1 has no interest in committing to the deviation.

2. Proposition 1 shows that the set of c-rationalizable strategies can be derived from *any* iterative procedure of restrictions to c-best response strategies. This paper thereby provides an additional rationale for defining a coalitional version of rationalizability directly by an iterative procedure of restrictions as in Ambrus (2006). Ambrus (2005) observed that there are other possible definitions to his formulation of 'supported restriction' and thus studied a general notion of γ-rationalizability by a class of sensible best response correspondences which satisfy four properties. It is easy to verify that our iterative procedures of restrictions satisfy the "closed under rational behavior" and "Pareto optimality" properties (i.e. Properties (i) and (iv) in Ambrus's (2005) definition of a sensible best response correspondence), but the iterative procedures may violate the "individual rationality" and "monotonicity" properties (i.e. Properties (ii) and (iii) in Ambrus's (2005) definition of a sensible best response correspondence). The main reason for these violations is that our iterative procedures are flexible to allow one coalition at a time, but *not all coalitions simultaneously*, to eliminate never c-best response strategies at each stage of an iteration. We believe that this flexibility is of relevance to practical use of c-rationalizability.

3. To analyze collective behavior in social interactions, the notion of c-rationalizablity in this paper is motivated and developed by using the 'core' idea that no group of players who, by rearranging and replotting their strategies, can each expect a higher payoff than that the player can expect to obtain from the original preparations of strategies. This 'core' idea may suffer from a conceptual deficiency: a CRS must be unblocked by any feasible set of strategies, including those that can, in turn, be blocked in the same sense; see Greenberg (1990) for more extensive discussions. Our notion of c-rationalizability can easily be tailored to overcome this deficiency. Specifically, we can define a 'coalition-proof rationalizable set' (CPRS) as a nonempty product subset $\mathcal{R}$ of S satisfying $\mathcal{R} \not\Rightarrow \mathcal{R}'$ only for $\mathcal{R}' = \mathcal{R}$, where $\mathcal{R}'$ is itself a CPRS. That is, a CPRS satisfies an appealing property of "consistency": a CPRS is immune against being blocked only by those that are themselves a CPRS. Apparently, any CRS is itself a CPRS and the existence of CPRS is an immediate implication of Theorem 1. Nevertheless, Milgrom and Roberts (1996, p.115) pointed out, "While this symmetric treatment is mathematically elegant, there is no reason to suppose it is descriptive

of the real possibilities for coalitional deviations." As a strong Nash equilibrium maintains the collective stability stronger than a coalition-proof Nash equilibrium does, we believe that the well-defined notion of c-rationalizability would be more suitable for the description of rational behavior in noncooperative social environments with no binding agreements or commitments.

4. Greenberg (1990) offered an integrative approach to social interactions by using a 'stability' criterion. Within a non-equilibrium framework of coalitional reasoning in strategic games, Greenberg (1990) proposed several coalitional negotiation processes where coalitions openly negotiate to make contingent threats (i.e. "coalitional contingent threats" situation) or to make irrevocable commitments (i.e. "coalitional commitments" situation), and he analyzed these coalitional interactions through his theory of social situations. Within a similar framework in which coalitional moves are publicly observed in social environments, Chwe (1994) and Xue (1998) studied the "stable" outcomes under coalitional interactions where players are farsighted; Herings et al. (2004) analyzed the social environment by using Pearce's (1984) extensive-form rationalizability in the associated multistage game. The main differences of our approach in this paper are that: (1) coalitional moves are secretly conducted and cannot be publicly observed; (2) implicit agreements made by coalitions are in general in the form of constraint sets of strategies to be confined, rather than stringent specifications of a particular course of actions.[12] We would also like to point out that the minimal CRS in Definition 1 can be viewed as a coalitional version of Basu and Weibull's (1991) minimal curb set, which is an interesting set-valued solution for strategic games.[13]

5. Finally, extension of this paper to more general class of games, e.g., extensive games with imperfect or incomplete information and games with general preferences, is an important subject for further research. Extension of this paper to permit nonproduct-set deviations is also an intriguing topic worth further studying.

[12] See Ambrus (2006, Section VII) for extensive discussions on this form of agreements.

[13] Voorneveld (2004, 2005) proposed a set-valued solution of '(minimal) preparation sets', and showed minimal preparation sets are intimately related to Basu and Weibull's (1991) 'minimal curb sets' and Kalai and Samet's (1984) 'persistent retracts'.

Appendix: Proofs

To prove Theorem 1, we need the following Lemmas 1-4.

Lemma 1. *Suppose $A \cap B \neq A$. For any $\mu \in \Delta|_{A\setminus B}$ and $\nu \in \Delta|_B$, there is $\mu^n \overset{A}{\rightsquigarrow} \mu$ such that $\nu = \lim_{n\to\infty} \mu^n|_B$.*

Proof: Consider $\widetilde{\mu}^n \to \mu$ and $\nu^n \to \nu$ such that $\widetilde{\mu}^n \in \Delta|_{S\setminus B}$, $\nu^n \in \Delta|_B$, and $\frac{1}{2}\widetilde{\mu}^n + \frac{1}{2}\nu^n \in \Delta^0$. Define $\mu^n \equiv \left(1 - \frac{1}{n}\right)\widetilde{\mu}^n + \frac{1}{n}\nu^n$. Thus, $\mu^n \overset{A}{\rightsquigarrow} \mu$ and $\nu = \lim_{n\to\infty} \mu^n|_B$.■

Lemma 2. *Suppose $A \Rightarrow B$ with $A \in \mathcal{M}$. Then, $B \neq \emptyset$ and $B \in \mathcal{M}$, where*

$$\mathcal{M} \equiv \{\mathcal{A}|\ BR(\mathcal{A}_{-i}) \subseteq \mathcal{A}_i\ \forall i\}.$$

Proof: Clearly, $A \neq \emptyset$ since $A \Rrightarrow B$. For any fixed i, let $u_i(a^*) \equiv \max_{a\in A} u_i(a)$. Since $A \in \mathcal{M}$, for all $\mu \in \Delta|_{A_{-i}}$,

$$\max_{s_i\in S_i} u_i(s_i, \mu) = \max_{a_i\in A_i} u_i(a_i, \mu) \leq \max_{\mu\in\Delta|_{A_{-i}}} \max_{a_i\in A_i} u_i(a_i, \mu) = u_i(a^*).$$

Since $a^* \in A$ and $A \Rightarrow B$, $a_i^* \in B_i$. Thus, $B \neq \emptyset$.

Let $\mu \in \Delta|_{B_{-i}}$. Since $B \subseteq A$, $\mu \in \Delta|_{A_{-i}}$. Consider $\mu^n \overset{A_{-j}}{\rightsquigarrow} \mu$. Therefore, $\mu = \mu|_{B_{-i}}$. Thus, for all $s_i \in BR(B_{-i})$, there is $\mu \in \Delta|_{B_{-i}}$ such that $u_i(s_i, \mu) \geq u_i(s_i', \mu|_{B_{-i}})$ $\forall s_i' \in S_i$. Let $A \Rightarrow B$ via $J \equiv \mathcal{J}_{AB}$. Since $A \in \mathcal{M}$, $BR(B_{-j}) \subseteq B_j$ $\forall j \in J$. For $i \notin J$, $BR(B_{-i}) \subseteq B_i$ since $B_i = A_i$.■

Lemma 3. *Suppose $\mathcal{A} \subseteq A \Rrightarrow B$, $\mathcal{A} \in \mathcal{M}$, and $\mathcal{A} \cap B = \emptyset$. Then, $\mathcal{A}_{-j} \cap B_{-j} = \emptyset$ $\forall j$.*

Proof: Since $\mathcal{A} \cap B = \emptyset$, $\mathcal{A}_j \cap B_j = \emptyset$ for some j. Since $\mathcal{A} \in \mathcal{M}$, $A_j \supseteq \mathcal{A}_j \supseteq BR(s_{-j})$ for all $s_{-j} \in \mathcal{A}_{-j} \cap B_{-j}$. But, since $A \Rrightarrow B$, by Definition 1(1.1), $B_j \supseteq BR(s_{-j})$ for all $s_{-j} \in \mathcal{A}_{-j} \cap B_{-j}$. Therefore, $\mathcal{A}_{-j} \cap B_{-j} = \emptyset$. That is, $\mathcal{A}_{j'} \cap B_{j'} = \emptyset$ for some $j' \neq j$ and, hence, $\mathcal{A}_{-j} \cap B_{-j} = \emptyset$ $\forall j$.■

Lemma 4. *Suppose $\mathcal{A} \subseteq A \Rrightarrow B$. Then, (4.1) $\mathcal{A} \Rightarrow \mathcal{A} \cap B$ if $\mathcal{A} \cap B \neq \emptyset$;(4.2) $\mathcal{A} \Rrightarrow \mathcal{B} \subseteq B$ if $\mathcal{A} \cap B = \emptyset$ and $\mathcal{A}, B \in \mathcal{M}$.*

Proof: Let $A \Rrightarrow B$ via $J \equiv \mathcal{J}_{AB}$. (4.1) For any j consider $\nu^n \overset{\mathcal{A}_{-j}}{\rightsquigarrow} \nu$ and $\nu^n|_{(\mathcal{A}\cap B)_{-j}} \to \nu|_{(\mathcal{A}\cap B)_{-j}}$. Define

$$\mu^n \equiv \left[\frac{1}{n}\nu^n\left(\mathcal{A}_{-j} \cap B_{-j}\right)\right]\nu^n + \left[1 - \frac{1}{n}\nu^n\left(\mathcal{A}_{-j} \cap B_{-j}\right)\right]\nu^n|_{\mathcal{A}_{-j}}.$$

It is easily verified that $\mu^n \overset{A_{-j}}{\rightsquigarrow} \nu$ and $\mu^n|_{B_{-j}} \to \nu|_{(\mathcal{A}\cap B)_{-j}}$. Since $\mathcal{A} \subseteq A \Rrightarrow B$, $\forall j \in J$, $\forall a_j \in \mathcal{A}_j$, $u_j(a_j, \nu) < u_j(s_j, \nu|_{(\mathcal{A}\cap B)_{-j}})$ for some $s_j \in S_j$, where $\nu|_{(\mathcal{A}\cap B)_{-j}} \neq \nu$ if

$a_j \in \mathcal{A}_j \cap B_j$. But, since $\mathcal{A}_{-J} \subseteq A_{-J} = B_{-J}$, $\mathcal{A} \cap B = (\mathcal{A} \cap B)_J \times \mathcal{A}_{-J}$. Therefore, $\mathcal{A} \Rightarrow \mathcal{A} \cap B$ via J.

(4.2) Define $\widetilde{\mathcal{B}} \equiv B_J \times \mathcal{A}_{-J}$. Since $\mathcal{A}_{-J} \subseteq A_{-J} = B_{-J}$, $\widetilde{\mathcal{B}} \subseteq B$. Since $B \in \mathcal{M}$, $\widetilde{\mathcal{B}}_j \supseteq BR\left(\widetilde{\mathcal{B}}_{-j}\right)$ $\forall j \in J$. Let $\mathcal{B}$ be the (nonempty) set of surviving iterated elimination of never-best responses for all the players in coalition J in the finite subgame restricted on $\widetilde{\mathcal{B}}$. Clearly, $J = \mathcal{J}_{\mathcal{A}\mathcal{B}}$ and $\mathcal{B} \subseteq B$ with $\mathcal{B}_j = BR(\mathcal{B}_{-j})$ $\forall j \in J$. To complete the proof, it remains to verify $\mathcal{A} \Rrightarrow \mathcal{B}$ via J. Since $\mathcal{A} \in \mathcal{M}$, by Lemma 3, $\mathcal{A}_{-j} \cap B_{-j} = \emptyset$.

(1) Consider $\nu^n \overset{\mathcal{A}_{-j}}{\rightsquigarrow} \nu$. Since $\mathcal{A}_{-j} \subseteq A_{-j}$, $\nu^n \overset{A_{-j}}{\rightsquigarrow} \nu$. Since $A \Rrightarrow B \supseteq \mathcal{B}$, by Lemma 1, it follows that $\forall j \in J = \mathcal{J}_{\mathcal{A}\mathcal{B}}$, $\forall a_j \in \mathcal{A}_j \subseteq A_j$,

$$u_j(a_j, \nu) < \; u_j(s_j, \nu') \text{ for some } s_j \in S_j,$$

for all $\nu \in \Delta|_{\mathcal{A}_{-j}}$ and $\nu' \in \Delta|_{\mathcal{B}_{-j}}$. That is, Definition 1(1.1) is satisfied.

(2) Since $\mathcal{B}_j = BR(\mathcal{B}_{-j})$, by Lemma 1, $\forall j \in J = \mathcal{J}_{\mathcal{A}\mathcal{B}}$, $\forall b_j \in \mathcal{B}_j \backslash \mathcal{A}_j \subseteq \mathcal{B}_j$,

$$u_j(b_j, \nu|_{\mathcal{B}_{-j}}) \geq \; u_j(s_j, \nu|_{\mathcal{B}_{-j}}) \text{ for all } s_j \in S_j,$$

for some $\nu^n|_{\mathcal{B}_{-j}} \to \nu|_{\mathcal{B}_{-j}}$ as $\nu^n \overset{\mathcal{A}_{-j}}{\rightsquigarrow} \nu$. That is, Definition 1(1.2) is satisfied.■

Proof of Theorem 1: The proof is split into two parts. First of all, define

$$\mathcal{D} \equiv \bigcap_{k=0}^{\infty} \mathcal{D}^k$$

where $\mathcal{D}^0 = S$, $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$, and $\mathcal{D} \Rightarrow \mathcal{D}'$ only for $\mathcal{D}' = \mathcal{D}$. By Lemma 2, $\mathcal{D} \neq \emptyset$.

Part I: $\mathcal{D}$ is a CRS. Assume, in negation, that $\mathcal{D}$ is not a CRS, i.e., $\mathcal{D} \Rrightarrow \mathcal{D}' \neq \mathcal{D}$. Clearly, $\mathcal{D}' \neq \emptyset$ since $\mathcal{D} \Rrightarrow \emptyset$ implies $\mathcal{D} \Rightarrow \emptyset$. We distinguish two cases.

Case 1.1. $\mathcal{D}' \cap \mathcal{D} \neq \emptyset$. By Lemma 4(4.1), $\mathcal{D} \Rightarrow \mathcal{D}' \cap \mathcal{D}$. Thus, $\mathcal{D}' \cap \mathcal{D} = \mathcal{D}$. By Lemma 2, $\mathcal{D} \in \mathcal{M}$. Therefore, for every $d'_j \in \mathcal{D}'_j \backslash \mathcal{D}_j$, $u_j(d'_j, \mu|_{\mathcal{D}'_{-j}}) < u_j(s_j, \mu|_{\mathcal{D}'_{-j}})$ for some $s_j \in S_j$, for all $\mu^n|_{\mathcal{D}'_{-j}} \to \mu|_{\mathcal{D}'_{-j}}$ as $\mu^n \overset{\mathcal{D}_{-j}}{\rightsquigarrow} \mu$. By Definition 1(1.2), $\mathcal{D} \not\Rrightarrow \mathcal{D}' \supsetneq \mathcal{D}$, a contradiction.

Case 1.2. $\mathcal{D}' \cap \mathcal{D} = \emptyset$. Let $\mathcal{D}^k \supseteq \mathcal{D}'$ and $\mathcal{D}^{k+1} \not\supseteq \mathcal{D}'$ (cf. Figure 3). Thus, $\exists d'_{j^0} \in \mathcal{D}'_{j^0} \backslash \mathcal{D}^{k+1}_{j^0}$ and, hence, $d'_{j^0} \in \mathcal{D}'_{j^0} \backslash \mathcal{D}_{j^0}$. Thus, $j^0 \in \mathcal{J}_{\mathcal{D}\mathcal{D}'} \cap \mathcal{J}_{\mathcal{D}^k\mathcal{D}^{k+1}}$. Since by

Lemma 2, $\mathcal{D}^{k+1} \in \mathcal{M}$, $BR(\mathcal{D}^{k+1}_{-j^0}) \subseteq \mathcal{D}^{k+1}_{j^0}$. Since $\mathcal{D} \Rrightarrow \mathcal{D}'$, by Definition 1(1.2), $d'_{j^0} \in BR(\mathcal{D}'_{-j^0})$. Therefore, $\mathcal{D}'_{-j^0} \nsubseteq \mathcal{D}^{k+1}_{-j^0}$, i.e., $\exists d'_{-j^0} \in \mathcal{D}'_{-j^0} \backslash \mathcal{D}^{k+1}_{-j^0}$. We proceed in two steps.

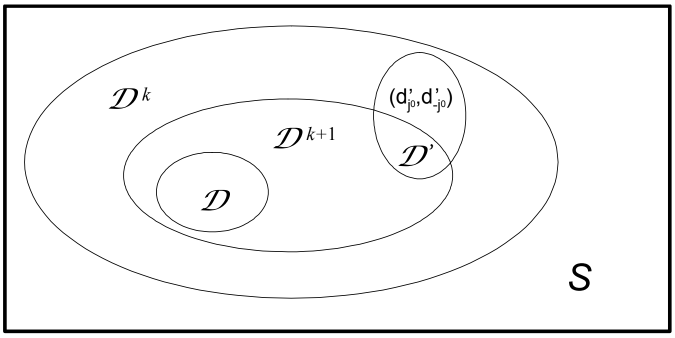

Figure 3.

Step 1. Let $d_{-j^0} \in \mathcal{D}_{-j^0}$. By Lemma 2, $\mathcal{D} \in \mathcal{M}$. Since $\mathcal{D} \Rrightarrow \mathcal{D}'$, by Lemma 3, $d_{-j^0} \notin \mathcal{D}'_{-j^0}$. Consider $\mu^n \overset{\mathcal{D}_{-j0}}{\rightsquigarrow} \mu$ with $\mu(d_{-j^0}) = 1$. Since $\mathcal{D} \Rrightarrow \mathcal{D}'$, by Lemma 1,

$$u_{j^0}(d_{j^0}, d_{-j^0}) = u_{j^0}(d_{j^0}, \mu) < u_{j^0}(s_{j^0}, d'_{-j^0}) \text{ for some } s_{j^0} \in S_{j^0},$$

for all $d_{j^0} \in \mathcal{D}_{j^0}$ and $d'_{-j^0} \in \mathcal{D}'_{-j^0}$. Since $\mathcal{D}' \subseteq \mathcal{D}^k \in \mathcal{M}$,

$$u_{j^0}(d_{j^0}, d_{-j^0}) < u_{j^0}(d'_{j^0}, d'_{-j^0}) \text{ for some } d'_{j^0} \in \mathcal{D}^k_{j^0},$$

for all $d_{j^0} \in \mathcal{D}_{j^0}$, $d_{-j^0} \in \mathcal{D}_{-j^0}$ and $d'_{-j^0} \in \mathcal{D}'_{-j^0}$.

Step 2. Let $d'_{-j^0} \in \mathcal{D}'_{-j^0} \backslash \mathcal{D}^{k+1}_{-j^0}$ (since $\mathcal{D}'_{-j^0} \backslash \mathcal{D}^{k+1}_{-j^0} \neq \emptyset$). Consider $\nu^n \overset{\mathcal{D}^k_{-j0}}{\rightsquigarrow} \nu$ with $\nu\left(d'_{-j^0}\right) = 1$. Since $\mathcal{D}' \subseteq \mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$, by Lemma 1,

$$u_{j^0}(d'_{j^0}, d'_{-j^0}) = u_{j^0}(d'_{j^0}, \nu) < u_{j^0}(s_{j^0}, d_{-j^0}) \text{ for some } s_{j^0} \in S_{j^0},$$

for all $d_{-j^0} \in \mathcal{D}_{-j^0}$ and $d'_{j^0} \in \mathcal{D}^k_{j^0}$. Since $\mathcal{D} \in \mathcal{M}$, there is $d'_{-j^0} \in \mathcal{D}'_{-j^0}$ such that

$$u_{j^0}(d'_{j^0}, d'_{-j^0}) < u_{j^0}(d_{j^0}, d_{-j^0}), \text{ for some } d_{j^0} \in \mathcal{D}_{j^0},$$

for all $d_{-j^0} \in \mathcal{D}_{-j^0}$ and $d'_{j^0} \in \mathcal{D}^k_{j^0}$, contradicting Step 1.

Part II: $\mathcal{R}^* = \mathcal{D}$**.** Assume, in negation, that there is a CRS $\mathcal{R} \nsubseteq \mathcal{D}$. Then, there exists $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$ via $J = \mathcal{J}_{\mathcal{D}^k\mathcal{D}^{k+1}}$ such that $\mathcal{R} \subseteq \mathcal{D}^k$ and $\mathcal{R} \nsubseteq \mathcal{D}^{k+1}$. If $\mathcal{R}\cap$ $\mathcal{D}^{k+1} \neq \emptyset$, by Lemma 4(4.1), $\mathcal{R} \Rrightarrow \mathcal{R} \cap \mathcal{D}^{k+1}$, contradicting that $\mathcal{R}$ is a CRS. If $\mathcal{R}\cap$ $\mathcal{D}^{k+1} = \emptyset$, by Lemma 4(4.2), $\mathcal{R} \Rrightarrow \mathcal{R}'$ with $\mathcal{R} \cap \mathcal{R}' = \emptyset$, contradicting that $\mathcal{R}$ is a CRS.$\blacksquare$

Proof of Proposition 2: (2.1) Let s^* be a Pareto-dominant Nash equilibrium, i.e., s^* Pareto-dominates all other (mixed) Nash equilibria. Assume, in negation, that s^* is not c-rationalizable. By Proposition 1, $s^* \in \mathcal{D}^k \backslash \mathcal{D}^{k+1}$ for some $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$. Since every player i's Nash equilibrium strategy s_i^* is a best response to s_{-i}^*, the deviating coalition$\mathcal{J}_{\mathcal{D}^k\mathcal{D}^{k+1}}$ contains more than two players. Therefore, $s_{-i}^* \notin \mathcal{D}_{-i}^{k+1}$. Note that there is a Nash equilibrium s^{**} with the support in $\mathcal{D}^{k+1}$. Since $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$, by Lemma 1, $u_j(s^*) < u_j(s^{**})$ for all $j \in \mathcal{J}_{\mathcal{D}^k\mathcal{D}^{k+1}}$, contradicting that s^* is a Pareto-dominant Nash equilibrium. Thus, s^* is c-rationalizable.

Now, let s^* be a strong Nash equilibrium (see Aumann 1959), i.e., for each (mixed) strategy profile $s \neq s^*$, there is a player j such that $s_j \neq s_j^*$ and $u_j(s) \leq u_j(s^*)$. Assume, in negation, that s^* is not c-rationalizable. By Proposition 1, $s^* \in \mathcal{D}^k \backslash \mathcal{D}^{k+1}$ for some $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$. Define $J \equiv \left\{ j | \ s_j^* \in \mathcal{D}_j^k \backslash \mathcal{D}_j^{k+1} \right\}$. Consider a subgame restricted on $\mathcal{D}_J^{k+1} \times \left\{ s_{-J}^* \right\}$. There is a Nash equilibrium s^{**} in the subgame. Since by Lemma 2, $\mathcal{D}^{k+1} \in \mathcal{M}$, s_j^{**} is a best response to s_{-j}^{**} $\forall j \in J$. Similar to the proof for the Pareto-dominant Nash equilibrium case, it follows that for all $j \in J$, $u_j(s^*) < u_j(s^{**})$ for some $s_J \in \mathcal{D}_J^{k+1}$, contradicting that s^* is a strong Nash equilibrium. Thus, s^* is c-rationalizable.

(2.2) Let s_J^* be a Pareto-best strategy profile for J in the reduced game $\mathcal{G}'$=$(I, \{S_i'\}, \{u_i\})$, i.e., for all $j \in J$, $u_j(s_J^*, s_{-J}) \geq u_j(s)$ $\forall s \in S'$. Assume, in negation, that s_J^* is not c-rationalizable. By Proposition 1, for some $j^0 \in J$, $s_{j^0}^* \in \mathcal{D}_{j^0}^k \backslash \mathcal{D}_{j^0}^{k+1}$ for some $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$. Consider $s \in \mathcal{D}$. Clearly, $s \in S'$ and $(s_J^*, s_{-J}) \in \mathcal{D}^k$. As any dominated strategy is dominated by an undominated strategy in any finite game (see, e.g., Milgrom and Roberts 1996, Lemma 1), it follows that $s_{j^0}^*$ is a best response to $\left(s_{J\backslash j^0}^*, s_{-J}\right)$. By $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$ and $\mathcal{D} \in \mathcal{M}$, it follows that $u_{j^0}(s_J^*, s_{-J}) < u_{j^0}(s_{j^0}, s_{-j^0})$ for some $s_{j^0} \in \mathcal{D}_{j^0} \subseteq S_{j^0}'$, contradicting that s_J^* is Pareto-best for J. Thus, s_J^* is c-rationalizable.

Now, assume that for all $j \in J$, $u_j\left(s_J^*, s_{-J}'\right) > u_j(s)$ $\forall s, s' \in S'$ with $s_J^* \neq s_J$. Thus, J would be willing to confine their play to s_J^*. Again by Proposition 1, s_J^* is a unique c-rationalizable strategy profile for J. Therefore, a (strong) Pareto-best strategy profile is a (unique) c-rationalizable strategy profile.

(2.3) In a common interest game where there is a strategy profile strictly

Pareto-dominates all other strategy profiles, the Pareto-best strategy profile is a strong Pareto-best Nash equilibrium and, hence, is a unique c-rationalizable strategy profile.

(2.4) Milgrom and Shannon (1994) showed that, if each player's payoffs are always nondecreasing (or nonincreasing) in the opponents' strategies in games with strategic complementarities, then the reduced game $\mathcal{G}'$ has a Pareto-best Nash equilibrium s^* that is given by the largest (or smallest) element in $\mathcal{G}'$. Thus, s^* is c-rationalizable.

Now, let s^* be a strict Nash equilibrium. That is, for all $i \in I$, $u_i(s^*) > u_i\left(s'_i, s^*_{-i}\right)$ $\forall s'_i \neq s^*_i$. By the strict monotonicity of $u_i(s^*_i, .)$ in s_{-i} , $u_i(s^*) > u_i\left(s^*_i, s'_{-i}\right)$ $\forall s'_{-i} \neq s^*_{-i}$. Therefore, $u_i(s^*) > u_i\left(s'_i, s^*_{-i}\right) \geq u_i\left(s'_i, s'_{-i}\right)$ $\forall s'_i \neq s^*_i$ $\forall s'_{-i} \neq s^*_{-i}$. That is, for all $i \in I$, $u_i(s^*) > u_i(s')$ $\forall s' \neq s^*$. By (2.2), s^* is a unique c-rationalizable strategy profile.■

Proof of Theorem 2: The proof follows immediately from the following.
Lemma 5: *$A \Downarrow B$ iff $A \Rightarrow B$.*
Proof: The result is clearly true if $A = B$, so we assume $B \subsetneq A$.

"only if part": Suppose that $A \Downarrow B$ via $J = \mathcal{J}_{AB}$. Let $j \in J$. Consider $\mu^n|_{B_{-j}} \to \mu|_{B_{-j}}$ as $\mu^n \overset{A_{-j}}{\rightsquigarrow} \mu$. We distinguish three cases as follows.

Case (1.1) $\mu(B_{-j}) = 1$. By Definition 2(2.1), $c_j \in A_j \backslash B_j$ is strictly dominated in the subgame restricted on $S_j \times B_{-j}$. Thus, $c_j \in A_j \backslash B_j$ is a never-best response in the subgame restricted on $S_j \times B_{-j}$. Since $\mu(B_{-j}) = 1$, $\mu = \mu|_{B_{-j}}$. Therefore, $u_j(c_j, \mu) < u_j(s_j, \mu|_{B_{-j}})$ for some $s_j \in S_j$.

Case (1.2) $\mu(B_{-j}) = 0$. By Definition 2(2.2), for any $a_j \in A_j$, there is $\sigma_j \in \Delta(S_j)$ such that

$$\begin{aligned} u_j(a_j, \mu) &\leq \max_{c_{-j} \in A_{-j} \backslash B_{-j}} u_j(a_j, c_{-j}) \\ &< \min_{b_{-j} \in B_{-j}} u_j(\sigma_j, b_{-j}) \\ &\leq u_j(\sigma_j, \mu|_{B_{-j}}). \end{aligned}$$

Therefore, $u_j(a_j, \mu) < u_j(s_j, \mu|_{B_{-j}})$ for some $s_j \in S_j$.

Case (1.3) $\mu(B_{-j}) \neq 0$ *or* 1. Clearly, $\mu = q\mu' + (1-q)\mu''$ where $q \equiv \mu(B_{-j})$, $\mu'(B_{-j}) = 1$ and $\mu''(B_{-j}) = 0$. Since $\mu'(B_{-j}) = 1$, for any given $a_j \in A_j$,

$$u_j(a_j, \mu') \leq u_j(s'_j, \mu') = u_j(s'_j, \mu'|_{B_{-j}}) \text{ for some } s'_j \in S_j.$$

Since $\mu''(B_{-j}) = 0$, by Case (1.2),

$$u_j(a_j, \mu'') < u_j(s''_j, \mu'|_{B_{-j}}) \text{ for some } s''_j \in S_j.$$

Therefore,

$$\begin{aligned} u_j(a_j, \mu) &= q u_j(a_j, \mu') + (1-q)\, u_j(a_j, \mu'') \\ &< q u_j(s'_j, \mu'|_{B_{-j}}) + (1-q)\, u_j(s''_j, \mu'|_{B_{-j}}) \\ &= u_j(q s'_j + (1-q)\, s''_j, \mu'|_{B_{-j}}). \end{aligned}$$

But, since $\mu''(B_{-j}) = 0$, $\mu|_{B_{-j}} = \mu'|_{B_{-j}}$. Thus, $u_j(a_j, \mu) < u_j(q s'_j + (1-q)\, s''_j, \mu|_{B_{-j}})$. Hence, $u_j(a_j, \mu) < u_j(s_j, \mu|_{B_{-j}})$ for some $s_j \in S_j$.

"if part": Suppose $A \Rightarrow B$ via $J = \mathcal{J}_{AB}$. Let $j \in J$. Consider $\mu^n|_{B_{-j}} \to \mu|_{B_{-j}}$ as $\mu^n \overset{A_{-j}}{\rightsquigarrow} \mu$. We distinguish two cases.

Case (2.1) $\mu(B_{-j}) = 1$. Since $\mu(B_{-j}) = 1$, $\mu|_{B_{-j}} = \mu$. By Definition 1(1.1), for every $c_j \in A_j \backslash B_j$, $u_j(c_j, \mu) < u_j(s_j, \mu)$ for some $s_j \in S_j$. That is, $c_j \in A_j \backslash B_j$ is a never-best response in the subgame restricted on $S_j \times B_{-j}$. Therefore, $c_j \in A_j \backslash B_j$ is strictly dominated in the subgame restricted on $S_j \times B_{-j}$ (see, e.g., Osborne and Rubinstein's (1995) Lemma 60.1). Thus, Definition 2(2.1) holds.

Case (2.2) $\mu(c_{-j}) = 1$ *for some* $c_{-j} \in A_{-j} \backslash B_{-j}$. Let $\overline{\sigma}_j \in \Delta(S_j)$ be such that

$$\min_{\sigma_{-j} \in \Delta(B_{-j})} u_j(\overline{\sigma}_j, \sigma_{-j}) = \max_{\sigma_j \in \Delta(S_j)} \min_{\sigma_{-j} \in \Delta(B_{-j})} u_j(\sigma_j, \sigma_{-j}).$$

Since $\mu(B_{-j}) = 0$, by Definition 1(1.1) and Lemma 1, $u_j(a_j, \mu) < \max_{\sigma_j \in \Delta(S_j)} u_j(\sigma_j, \sigma_{-j})$ for all $\sigma_{-j} \in \Delta(B_{-j})$. Since $\Delta(B_{-j})$ is compact, by the well-known Maximum Theorem,

$$u_j(a_j, \mu) < \min_{\sigma_{-j} \in \Delta(B_{-j})} \max_{\sigma_j \in \Delta(S_j)} u_j(\sigma_j, \sigma_{-j}).$$

By the Minmax Theorem,

$$\begin{aligned}
u_j(a_j, c_{-j}) &= u_j(a_j, \mu) \\
&< \min_{\sigma_{-j} \in \Delta(B_{-j})} \max_{\sigma_j \in \Delta(S_j)} u_j(\sigma_j, \sigma_{-j}) \\
&= \max_{\sigma_j \in \Delta(S_j)} \min_{\sigma_{-j} \in \Delta(B_{-j})} u_j(\sigma_j, \sigma_{-j}) \\
&= \min_{\sigma_{-j} \in \Delta(B_{-j})} u_j(\overline{\sigma}_j, \sigma_{-j}) \\
&\leq u_j(\overline{\sigma}_j, b_{-j}) \ \forall b_{-j} \in B_{-j}.
\end{aligned}$$

Thus, Definition 2(2.2) holds.■

Theorem 3. (3.1) For all $\omega \in \boxed{R}$, $s_i(\omega)$ is coalitionally rationalizable. (3.2) For every coalitionally rationalizable strategy r_i, there is a model of knowledge such that $r_i = s_i(\omega)$ for $\omega \in \boxed{R}$.

Proof: (3.1) Let $\omega \in \boxed{R}$. Define $\mathcal{R} \equiv \times_{i \in I} \left\{ s_i(\omega') \mid \omega' \in \boxed{R} \right\}$. Clearly, $\boxed{R} \subseteq \Omega_{\mathcal{R}}$. Since $\boxed{R}$ is self-evident we have $P_i(\omega) \subseteq \boxed{R}$, and thus $s_j(P_i(\omega)) \subseteq \mathcal{R}_j$ $\forall j \neq i$. Note that by Myerson's (1986) Theorem 1, a CPS $\mu|$ satisfies $\mu|_{\mathcal{A}_{-i}} = \mu$ in $\Delta|_{\mathcal{A}_{-i}}$ iff there exists a sequence of full-support distributions $\mu^n \to \mu$ in $\Delta|_{\mathcal{A}_{-i}}$ (i.e. $\mu^n \overset{\mathcal{A}_{-i}}{\rightsquigarrow} \mu$) such that $\mu^n|_{B_{-i}} \to \mu|_{B_{-i}}$ $\forall B_{-i}$. But since $\boxed{R} \subseteq R$, $\omega \in R_i$ $\forall i$. Therefore, for all $i \in \mathcal{J}^*_{\mathcal{R}\mathcal{R}'}$ and $s_i(\omega) \notin \mathcal{R}'_i$,

$$u_i(s_i(\omega), \mu) \geq u_i(s_i, \mu|_{\mathcal{R}_{-i}}) \ \forall s_i \in S_i,$$

for some $\mu^n|_{\mathcal{R}'_{-i}} \to \mu|_{\mathcal{R}'_{-i}}$ as $\mu^n \overset{\mathcal{R}_{-i}}{\rightsquigarrow} \mu$ if Definition 1(1.1) holds to be true for all $j \in \mathcal{J}^*_{\mathcal{R}\mathcal{R}'} \backslash \{i\}$. Thus, $\mathcal{R} \not\Rrightarrow \mathcal{R}'$ if $\mathcal{R}_i \backslash \mathcal{R}'_i \neq \emptyset$ for some i. By the following Lemma 6, $\mathcal{R} \subseteq \mathcal{R}^*$. Hence, $s_i(\omega)$ is coalitionally rationalizable.

(3.2) Define Ω as the set of coalitionally rationalizable profiles, i.e., $\Omega = \mathcal{R}^*$ is the largest CRS. For any $\omega = (s_i)_{i \in I}$ in Ω, define $s_i(\omega) = s_i$ $\forall i$ and $P_i(\omega) = \{\omega' \in \Omega |\ s_i(\omega') = s_i(\omega)\}$ $\forall i$. Clearly, $s_j(P_i(\omega)) = \mathcal{R}^*_j$ $\forall j \neq i$. Since $\mathcal{R}^*$ is a CRS, every player i is c-rational at every $\omega \in \Omega$. Therefore, $\Omega = \boxed{R}$ is a self-evident event in R.■

Lemma 6: *If $A \Rrightarrow B$ only if $B \supseteq A$, then $A \subseteq \mathcal{R}^*$.*

Proof. It suffices to show that for $A \neq \emptyset$, $A \subseteq \mathcal{D}$. Assume, in negation, that $A \nsubseteq \mathcal{D}$. That is, there are $\mathcal{D}^k \supseteq A$ and $\mathcal{D}^{k+1} \not\supseteq A$ where $\mathcal{D}^k \Rrightarrow \mathcal{D}^{k+1}$ via $J \equiv \mathcal{J}_{\mathcal{D}^k\mathcal{D}^{k+1}}$. If $A \cap \mathcal{D}^{k+1} \neq \emptyset$, by Lemma 4(4.1), $A \Rrightarrow A \cap \mathcal{D}^{k+1}$, a contradiction.

Now assume that $A \cap \mathcal{D}^{k+1} = \emptyset$. Define $\widetilde{\mathcal{B}} \equiv \mathcal{D}_J^{k+1} \times A_{-J}$. Since $A_{-J} \subseteq \mathcal{D}_{-J}^k = \mathcal{D}_{-J}^{k+1}$, $\widetilde{\mathcal{B}} \subseteq \mathcal{D}^{k+1}$. Since $\mathcal{D}^{k+1} \in \mathcal{M}$, $\widetilde{\mathcal{B}}_j \supseteq BR\left(\widetilde{\mathcal{B}}_{-j}\right)$ $\forall j \in J$. Let $\mathcal{B}$ be the (nonempty) set of surviving iterated elimination of never-best strategies within $\cup_{j \in J}\left(\mathcal{D}_j^{k+1} \backslash A_j\right)$ in the subgame restricted on $\widetilde{\mathcal{B}}$. Clearly, $J = \mathcal{J}_{A\mathcal{B}}$. Moreover, $\mathcal{B} \subseteq \mathcal{D}^{k+1}$ with $A_i \cap \mathcal{D}_i^{k+1} \subseteq \mathcal{B}_i$ $\forall i \in I$ and $\mathcal{B}_j \backslash A_j \subseteq BR(\mathcal{B}_{-j})$ $\forall j \in J$. To complete the proof, it remains to verify $A \Rrightarrow \mathcal{B}$ via J. Let $j \in J$.

(1) Since $A \Rrightarrow B$ only if $B \supseteq A$, $A_j \subseteq BR(A_{-j})$. Therefore, by $\mathcal{D}^{k+1} \in \mathcal{M}$, $\mathcal{D}_j^{k+1} \supseteq A_j$ if $\mathcal{D}_{-j}^{k+1} \supseteq A_{-j}$. But, since $A \cap \mathcal{D}^{k+1} = \emptyset$, $A_{-j} \nsubseteq \mathcal{D}_{-j}^{k+1}$ and, hence, $A_{-j} \nsubseteq \mathcal{B}_{-j}$. By Lemma 1, $\forall j \in J$, $\forall b_j \in \mathcal{B}_j \backslash A_j \subseteq BR(\mathcal{B}_{-j})$,

$$u_j(b_j, \mu|_{\mathcal{B}_{-j}}) \geq u_j(s_j, \mu|_{\mathcal{B}_{-j}}) \text{ for all } s_j \in S_j,$$

for some $\mu^n|_{\mathcal{B}_{-j}} \to \mu|_{\mathcal{B}_{-j}}$ as $\mu^n \overset{A_{-j}}{\rightsquigarrow} \mu$. That is, Definition 1(1.2) is satisfied.

(2) Let $j \in J$. Consider $\mu^n|_{\mathcal{B}_{-j}} \to \mu|_{\mathcal{B}_{-j}}$ as $\mu^n \overset{A_{-j}}{\rightsquigarrow} \mu$. Clearly, $\mu^n \overset{\mathcal{D}_{-j}^k}{\rightsquigarrow} \mu$ since $A_{-j} \subseteq \mathcal{D}_{-j}^k$. To verify Definition 1(1.1), we distinguish three cases.

(2.1) $\mu(\mathcal{B}_{-j}) = 1$. Then, $\mu = \mu|_{\mathcal{B}_{-j}} = \mu|_{\mathcal{D}_{-j}^{k+1}}$. Since $A_j \cap \mathcal{D}_j^{k+1} \subseteq \mathcal{B}_j$, $A_j \backslash \mathcal{B}_j \subseteq \mathcal{D}_j^k \backslash \mathcal{D}_j^{k+1}$. Since $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1}$, by Definition 1(1.1), $\forall a_j \in A_j \backslash \mathcal{B}_j$, $u_j(a_j, \mu) < u_j\left(s_j, \mu|_{\mathcal{B}_{-j}}\right)$ for some $s_j \in S_j$.

(2.2) $\mu(\mathcal{B}_{-j}) = 0$. Since $\mathcal{B}_i \supseteq \mathcal{D}_i^{k+1} \cap A_i$ $\forall i \in I$, $\mathcal{B}_{-j} \supseteq \mathcal{D}_{-j}^{k+1} \cap A_{-j}$. Thus, $\mu\left(\mathcal{D}_{-j}^{k+1}\right) = 0$. Since $\mathcal{D}^k \Rightarrow \mathcal{D}^{k+1} \supseteq \mathcal{B}$, by Definition 1(1.1) and Lemma 1, $\forall a_j \in A_j \subseteq \mathcal{D}_j^k$, $\forall \mu' \in \Delta|_{\mathcal{B}_{-j}}$, $u_j(a_j, \mu) < u_j(s_j, \mu')$ for some $s_j \in S_j$.

(2.3) $\mu(\mathcal{B}_{-j}) \neq 0$ *or* 1. The rest of proof is totally similar to (1.3) in the proof of Theorem 2.■

References

1. D.S. Ahn, Hierarchies of ambiguous beliefs, mimeo, UC_Berkeley, 2006.

2. A. Ambrus, Coalitional rationalizability, forthcoming in: QJE (2006).

3. A. Ambrus, Theories of coalitional rationality, mimeo, Harvard University, 2005.

4. K.R. Apt, Order independence and rationalizability, mimeo, National University of Singapore, 2005.

5. R.J. Aumann, Acceptable points in general cooperative n-person games, in: Contributions to the Theory of Games IV, Princeton University Press, 1959.

6. R.J. Aumann, Subjectivity and correlation in randomized strategies, JME 1(1974), 67-96.

7. R.J. Aumann, Correlated equilibrium as an expression of Bayesian rationality, Econometrica 55(1987), 1-18.

8. R.J. Aumann, Backward induction and common knowledge of rationality, GEB 8(1995), 6-19.

9. K. Basu and J.W. Weibull, Strategy subsets closed under rational behavior, EL 36(1991), 141-146.

10. B.D. Bernheim, Rationalizable strategic behavior, Econometrica 52(1984), 1007-1028.

11. B.D. Bernheim, B. Peleg, and D.Whinston, Coalition-proof Nash equilibria I. Concepts, JET 42(1987), 1-12.

12. Y.C. Chen, N.Y. Long and X. Luo, Iterated strict dominance in general games, mimeo, Academia Sinica, 2005.

13. M.S.Y. Chwe, Farsighted coalitional stability, JET 63(1994), 299-325.

14. M. Dufwenberg and M. Stegeman, Existence and uniqueness of maximal reductions under iterated strict dominance, Econometrica 70(2002), 2007-2023.

15. L. Epstein and M. LeBreton, Dynamically consistent beliefs must be Bayesian, JET 61(1993), 1-22.

16. J. Greenberg, The Theory of Social Situations: An Alternative Game-Theoretic Approach, Cambridge University Press, Cambridge, 1990.

17. C.M. Kahn and D. Mookherjee, Universal Coalition-proof equilibrium: Theory and applications, mimeo, Bell Communications - Economic Research Group, 1994.

18. E. Kalai and D. Samet, Persistent equilibria in strategic games, IJGT 13(1984), 129-144.

19. D. Kreps, Notes on the Theory of Choice, Westview, Bouldler, 1988.

20. B.L. Lipman, A note on the implication of common knowledge of rationality, GEB 6(1994), 114-129.

21. X. Luo, General systems and φ-stable sets – a formal analysis of socioeconomic environments, JME 36(2001), 95-109.

22. P. Milgrom and J. Roberts, Rationalizability, learning, and equilibrium in games with strategic complementarities, Econometrica 58(1990), 1255-1278.

23. P. Milgrom and J. Roberts, Coalition-proofness and correlation with arbitrary communication possibilities, GEB 17(1996), 113-128.

24. P. Milgrom and C. Shannon, Monotone comparative statics, Econometrica 62(1994), 157-180.

25. R. Myerson, Multistage games with communication, Econometrica 54(1986), 323-358.

26. R. Myerson, Game Theory: Analysis of Conflict, Harvard University Press, 1991.

27. M.J. Osborne and A. Rubinstein, A Course in Game Theory, The MIT Press, 1994.

28. D. Pearce, Rationalizable strategic behavior and the problem of perfection, Econometrica 52(1984), 1029-1051.

29. D. Ray and R. Vohra, Equilibrium binding agreements, JET 73(1997), 30-78.

30. D. Ray and R. Vohra, A theory of endogenous coalition structures, GEB 26(1999), 286-336.

31. A. Roth, Subsolutions and the supercore of cooperative games. Mathematics of Operations Research 1 (1976), 43-49.

32. R. Selten, Reexamination of the perfectness concept for equilibrium points in extensive games, IJGT 4(1975), 25-55.

33. M. Shimoji, On the equivalence of weak dominance and sequential best response, GEB 48(2004), 385-402.

34. M. Shimoji and J. Watson, Conditional dominance, rationalizability, and game forms, JET 83(1998), 161-195.

35. T. Tan and S. Werlang, The Bayesian foundations of solution concepts of games, JET 45(1988), 370-391.

36. X. Vives, Complementarities and games: New developments, JEL 43(2005), 437-479.

37. M. Voorneveld, Persistent retracts and preparation, GEB 51(2005), 228-232.

38. M. Voorneveld, Preparation, GEB 48(2004), 403-414.

39. L. Xue, Coalitional stability under perfect foresight, ET 11(1998), 603-627.