Games and Economic Behavior 119 (2020) 197–215

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Strategic cautiousness as an expression of robustness to ambiguity

Gabriel Ziegler [a,*], Peio Zuazo-Garin [b]

[a] Northwestern University, Department of Economics, 2211 Campus Drive, Evanston, IL 60208, USA
[b] University of the Basque Country (UPV/EHU), Department of Foundations of Economic Analysis I, Avenida Lehendakari Agirre 83, 48015, Bilbao, Spain

ARTICLE INFO

Article history:
Received 15 January 2019
Available online 15 November 2019

JEL classification:
C72
D82

Keywords:
Game theory
Decision theory
Ambiguity
Knightian uncertainty
Incomplete preferences
Bayesian rationality
Cautiousness
Iterated admissibility

ABSTRACT

Economic predictions often hinge on two intuitive premises: agents rule out the possibility of others choosing unreasonable strategies ('strategic reasoning'), and prefer strategies that hedge against unexpected behavior ('cautiousness'). These two premises conflict and this undermines the compatibility of usual economic predictions with reasoning-based foundations. This paper proposes a new take on this classical tension by interpreting cautiousness as robustness to ambiguity. We formalize this via a model of incomplete preferences, where (*i*) each player's strategic uncertainty is represented by a possibly non-singleton set of beliefs and (*ii*) a rational player chooses a strategy that is a best-reply to every belief in this set. We show that the interplay between these two features precludes the conflict between strategic reasoning and cautiousness and therefore solves the inclusion-exclusion problem raised by Samuelson (1992). Notably, our approach provides a simple foundation for the iterated elimination of weakly dominated strategies.

1. Introduction

Economists commonly use iterated strategy elimination procedures as solution concepts in games. Such procedures thus constitute one of the cornerstones for modeling agents' behavior in economic theory. The predictive power of iterated elimination procedures is in general lower than that of equilibrium-related notions; however, since the latter requires players to correctly forecast their opponents' behavior (see Aumann and Brandenburger, 1995), the former seems more appropriate in situations of multiple equilibria wherein either the players or the economic analyst lack accurate data about past play or such data appears uninformative about future behavior.[1] For instance, this is the case in many application of auction theory, e.g., wireless spectrum, carbon emission rights and online advertising.[2] Consequently, thorough understanding of the forces

* Corresponding author.
E-mail addresses: ziegler@u.northwestern.edu (G. Ziegler), peio.zuazo@ehu.eus (P. Zuazo-Garin).

[1] In Dekel and Fudenberg's (1990) words (p. 243): "Nash equilibrium and its refinements describe situations with little or no 'strategic uncertainty,' in the sense that each player *knows* and is *correct* about the beliefs of the other players regarding how the game will be played. While this will sometimes be the case, it is also interesting to understand what restrictions on predicted play can be obtained when the players' strategic beliefs may be inconsistent, that is, using only the assumption that it is common knowledge that the players are rational."

[2] See Milgrom (1998), Cramton and Kerr (2002) and Varian (2007), respectively.

https://doi.org/10.1016/j.geb.2019.10.001
0899-8256/

behind iterated elimination is relevant from both a purely theoretical perspective and a more applied point of view, and is key to effective mechanism design and correct identification in empirical analyses.[3]

The conceptual appeal of iterated elimination procedures is that they carry the intuitive game-theoretic appeal of *strategic reasoning*: if a player is certain that some of her opponent's strategies are not going to be played, then she might deem some of her own strategies to be unreasonable.[4] However, as discussed in Samuelson's (1992) classic analysis, strategic reasoning is in conflict with the criterion of *cautiousness*, which dictates that players favor strategies that, ceteris paribus, hedge against unexpected behavior. If players are modeled as expected utility maximizers, the clash seems inescapable: Strategic reasoning requires each player i's beliefs to assign zero probability to some of the strategies of i's, while cautiousness requires player i's decision to be sensitive to those strategies that receive zero probability (and are therefore of negligible importance for the maximization problem). Given that economic modeling often invokes the avoidance of weakly dominated strategies—a specific kind of cautiousness—as a criterion for equilibrium selection,[5] the seemingly mutually exclusive nature of strategic reasoning and cautiousness requires clarification. Such an understanding is desirable in particular in scenarios where behavior is likely to be reasoning-based and cautiousness plays a role.

This paper proposes a new take on this longstanding problem by suggesting a novel theoretical foundation for the interplay between strategic reasoning and cautiousness. The analysis by Samuelson (1992) clearly shows that two ingredients are necessary to overcome this tension: First, multiple beliefs are needed to account for epistemic conditions that would be mutually excluding if required to be satisfied by a single belief. Second, the best-reply needs to be sensitive to all these beliefs. We achieve this within our framework by augmenting the underlying standard decision-theoretic foundation for each player by allowing for incomplete preferences à la Bewley (2002) where: (*i*) Each player's strategic uncertainty is represented by a possibly non-singleton set of beliefs thus allowing for ambiguity, and (*ii*) a rational player chooses a strategy that is a best-reply to every belief in her set, so that the resulting choice is robust to the possible ambiguity faced by the player.[6] Under this set-up, and inspired by Brandenburger et al. (2008), we say that a player *assumes* certain behavior by her opponents if *at least one* of the beliefs in her set has full-support on the collection of states representing such behavior. Consequently, the introduction of ambiguity and the requirement of robustness give great flexibility: It is possible for a player to assume certain behavior and, simultaneously, assume certain more restrictive behavior. If the player is also rational, her choice needs to be a best-reply to both of these beliefs. Hence, in particular, the tension between strategic reasoning and cautiousness is solved: A player can be strategically sophisticated by having one belief that assigns zero probability to her opponents playing dominated strategies, and at the same time cautious by having another belief that assigns positive probability to every strategy of her opponents. Thus, our model overcomes the problem as identified by Samuelson (1992) since it allows precisely for the two necessary ingredients.

Based on the above, we build a framework that provides reasoning-based foundations for iterated admissibility—the iterated elimination of weakly dominated strategies. In Theorem 1 we show that, when type spaces are belief-complete (roughly speaking, *rich* enough to capture any possible belief hierarchy), iterated admissibility characterizes the behavioral implications of rationality, cautiousness, and common assumption thereof. From our characterization, it is easy to see that the foundations of iterated admissibility necessarily require the presence of ambiguity whenever strategic reasoning has any bite. If the elimination procedure consists of multiple rounds, the set of ambiguous beliefs needs to contain a specific belief with full-support on the set of opponents' strategies that survive each round. Theorem 2 provides the analysis for the relaxation of belief-completeness and shows that, in this case, it is *self-admissible* sets à la Brandenburger et al. (2008) which characterize the behavioral implications of rationality, cautiousness and common assumption thereof. Although the main approach in the paper is conceptual and focused on the link between cautiousness in reasoning-based processes and robustness to ambiguity, the results provide a methodological contribution for the use of incomplete preferences in game theory, which is a subject of interest in itself aside from its interpretation as a reflection of ambiguity.[7]

The literature studying the conflict between strategic reasoning and cautiousness is epitomized by the seminal paper by Brandenburger et al. (2008), who shed light on the question by building upon the lexicographic probability system approach by Blume et al. (1991a,b).[8] Lexicographic probability systems represent the uncertainty faced by a decision maker whose preferences depart from standard Bayesian preferences by allowing violations of the continuity axiom. In this setting,

[3] See Bergemann and Morris (2009, 2011); Bergemann et al. (2011) and Aradillas-Lopez and Tamer (2008), respectively.

[4] This is clearly exemplified by the informal argument for competitive prices in Bertrand duopoly models. Consider a market consisting of profitable, identical firms A and B: If A slightly lowers its mark-up it should absorb all the demand and increases its profit; now, this is easy to forecast by B, which might in turn decide to lower its mark-up more than slightly and thus absorb itself all the demand and increase her profit with respect to the losses obtained under A's, hypothetical, initial slight cut. Obviously, this logic leads to the standard zero mark-up conclusion. Sketches of this elementary intuition in modern economic theory can be traced back to Keynes (1936): "It is not a case of choosing those [faces] that, to the best of one's judgment, are really the prettiest, nor even those that average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees."

[5] E.g., Kohlberg and Mertens (1986), Palfrey and Srivastava (1991), Feddersen and Pesendorfer (1997), or Sobel (2017, 2019).

[6] Due to incompleteness such a strategy might not exist for a given set of beliefs. In such case we also say that the player is not rational.

[7] As argued by Aumann (1962): "Of all the axioms of utility theory, the completeness axiom is perhaps the most questionable. [...] [W]e find it hard to accept even from the normative viewpoint. Does 'rationality' demand that an individual make definite preference comparisons [...]" Previous applications of Bewley's (2002) model to game theory include, among others, Lopomo et al. (2011, 2014), who study mechanism design and optimal contracting, respectively.

[8] Early contributions along the same lines include, for example, Brandenburger (1992) and Stahl (1995).

Brandenburger et al. (2008) provide reasoning-based foundations for finitely many iterations of weakly dominated strategy elimination based on rationality and finite-order assumption of rationality, but also present a celebrated impossibility result: under some standard technical conditions and generically in all games, common assumption of rationality cannot be satisfied. This negative result has spurred a line of research concerned with obtaining sound epistemic foundations for iterated admissibility. Keisler and Lee (2015) and Yang (2015) propose answers by tweaking topological properties of the modeling of higher-order beliefs and the notion of assumption, respectively, while Lee (2016) obtains foundations by proposing a modification in the definition of coherence.[9] Catonini and De Vito (2018a) also provide foundations by introducing a weaker notion of the likeliness-ordering of events that characterizes the lexicographic probability system, and via an alternative definition of cautiousness that restricts attention to the payoff-relevant component of the states. In a slightly different direction, Heifetz et al. (2019) propose a new solution concept, *comprehensive rationalizability*, that coincides with iterated admissibility in many settings and admits epistemic foundations. Within a standard Bayesian decision-theoretic model, Barelli and Galanis (2013) provide a characterization for iterated admissibility by introducing an exogenous 'tie breaking' criterion. Robustness to ambiguity is studied by Stauber (2011, 2014) with a different interpretation from ours.

Our paper can be regarded as complementary to the lexicographic probability system approach as standard Bayesian preferences are also abandoned by dropping completeness instead of continuity. Both these relaxations allow for multiple beliefs, but while the former requires a specific order, our model drops the order altogether and allows for multiplicity directly. However, apart from the transparent link between cautiousness and robustness to ambiguity that our framework allows for, the *nice* structure of the sets of ambiguous beliefs representing incomplete preferences has some additional advantages. First, it is easy to show that rationality and common assumption of rationality is a non-empty event and thus, that iterated admissibility is properly founded for *all* games. Second, the definitions and formalism involved do not require departures from the canonical definition of the objects involved: (i) The modeling of higher-order beliefs (i.e., the type structures employed), including the definition of coherence, and the version of *assumption* that we rely on are natural extensions of their counterparts in the realm of standard Bayesian preferences; and (ii) the notion of cautiousness invoked in our theorems is not necessarily restricted to environments where the sets of states have a specific structure (e.g. games).[10] Finally, the presence of ambiguity via incomplete preferences has been shown to be empirically testable by recent work by Cettolin and Riedl (2019).

The rest of the paper is structured as follows. First, Section 2 provides an informal, non-technical overview of the effect of robustness to ambiguity on predictions in games, and specifically, on iterated admissibility as a solution concept. Section 3 reviews both the game-theoretic and the decision-theoretic preliminaries and Section 4 introduces the epistemic framework and the interpretation of strategic cautiousness as a manifestation of robustness of ambiguity. Section 5 the presents the epistemic characterization results. Section 6 concludes. All proofs and purely technical digressions are relegated to the appendices.

2. Non-technical overview

2.1. Examples

To illustrate the intuition behind the usual tension between rationality and cautious behavior and to show how our approach avoids this issue, we present two examples.

Example 1. Consider a two player game with the following payoff matrix[11]:

Ann \ Bob	L	R
T	1, 1	0, 1
D	0, 2	1, 0

Clearly, no action is strictly dominated for either player, so (standard) rationalizability predicts $\{T, D\} \times \{L, R\}$. However, R is weakly dominated by L. Deleting R will therefore make D strictly dominated in the reduced game. Thus, iterated admissibility has a unique prediction in this game: (T, L).

Now assume that one wishes to study how players themselves reason about this game. If Bob is rational and cautious he should play L. Suppose Ann is cautious as well. Therefore her belief has to put positive probability on Bob playing L

[9] Similar to Epstein and Wang (1996), coherency is imposed on the preferences directly, not only on the beliefs that represent the preferences.

[10] Though they are sensitive to topological specifications.

[11] This is the leading example of Brandenburger et al. (2008) and was introduced by Samuelson (1992).

and on Bob playing R. However, if Ann believes that Bob is rational and cautious, then she should rule out Bob playing R. This is the ‘inclusion-exclusion’ problem as identified by Samuelson (1992). On the one hand, Ann should include R in her belief because she is cautious. On the other hand, she should exclude R because she believes that Bob is rational and cautious. ◇

In our framework, we have more flexibility because players are not Bayesian, but are allowed to have a (potentially non-singleton) set of beliefs. To see how this relaxation avoids the tension just described, we provide a slightly more elaborate example, which also explores the reasoning of the players more explicitly.

Example 2. Again, there are two players, Ann and Bob, who play the following game:

Ann \ Bob	A	B	C	D
H	1, 0	1, 4	1, 2	1, 0
M	1, 4	1, 0	1, 2	1, 0
L	0, 2	0, 2	1, 2	3, 2

Now suppose that each player faces ambiguity (as described by Bewley, 2002) about the strategy choice of their opponent. That is, neither player has a unique belief about the opponent’s strategy choice, but rather each has a set of beliefs. In particular, suppose that Ann has a convex closed set of beliefs described by two extreme points. Her first belief is uniform across all of Bob’s strategies, $\mu_A^1(s_B) = 1/4$ for $s_B = A, B, C, D$, and her second (extreme) belief is uniform across A, B, and C only, $\mu_A^2(s_B) = 1/3$ for $s_B = A, B, C$. Similarly, Bob faces uncertainty about Ann’s choice. Consider the following set of beliefs for Bob, which also has two extreme beliefs. The first is uniform across all of Ann’s strategies, $\mu_B^1(s_A) = 1/3$ for $s_A = H, M, L$, and the second belief assigns equal probability to H and M, $\mu_B^2(H) = \mu_B^2(M) = 1/2$.

Let us check what strategies are rational for each player given their beliefs. Preferences à la Bewley (2002) are incomplete, and for incomplete preferences there is no obvious definition of rationality: Optimality is a stronger requirement than maximality for incomplete orders. As stated in the introduction, the solution to the inclusion-exclusion problem requires that a best-reply to be sensitive to all beliefs. Thus, we identify rationality with optimality so that a rational strategy is a best-reply to *all* beliefs, i.e. the choice needs to be robust to the ambiguity faced by the player. In this example this implies that Ann will not rationally choose L since it is not a best-reply that is robust to the ambiguity that she faces. H and M, on the other hand, are best-replies to all beliefs and are therefore rational choices for Ann. For Bob, only D is not rational because it is not a best-reply to any of his beliefs. The three other strategies A, B, and C are rational as they are best-replies to all of his beliefs. Thus, with these sets of beliefs the prediction of the model would correspond to iterated admissibility. This is not a coincidence and foreshadows our results on the characterization iterated admissibility, explained in more detail below, where the strategic reasoning is also made explicit. ◇

2.2. Heuristic treatment of strategic reasoning

In the previous examplesit can be seen that a set of beliefs enables strategic reasoning and cautiousness to be incorporated. To study games in general, players need to be allowed to reason about the reasoning process of other players too. This necessitates the formalizing of infinite sequences of the following form:

a_1: Ann is rational and cautions	b_1: Bob is rational and cautions
a_2: a_1 holds and Ann assumes b_1	b_2: b_1 holds and Bob assumes a_1
a_3: a_1 holds and Ann assumes b_1 & b_2	b_2: b_1 holds and Bob assumes a_1 & a_2
...	...

If this infinite sequence holds, we say that there is *rationality, cautiousness, and common assumption thereof* (RCCARC).

To study these infinite sequences and to see which strategies are played if they hold, (epistemic) types need to be introduced for each player. Accordingly, consider T_A and T_B as type spaces for Ann and Bob, respectively. Usually, each of Ann’s type $t_A \in T_A$ is associated with a belief about Bob’s strategy *and* type, i.e. a probability distribution over $S_B \times T_B$. However, the idea here is to model players who face ambiguity, so each type is associated with a (closed, and convex) set of beliefs about $S_B \times T_B$. Thus, for a strategy-type pair of Ann (s_A, t_A), strategy s_A is said to be *rational* if s_A is a best-reply

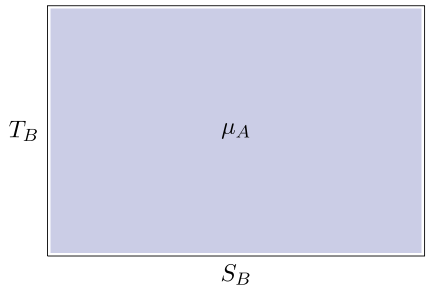

Fig. 1. Cautiousness.

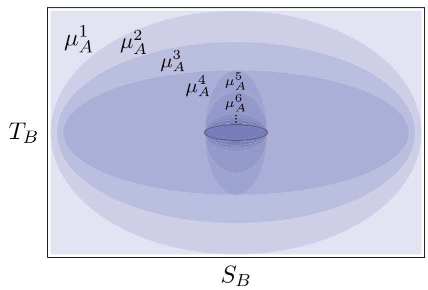

Fig. 2. Rationality, cautiousness, and common assumption thereof.

to *all* of the beliefs associated with t_A. Whether a player is cautious depends only on her beliefs: she thinks everything is possible. That is, one of her beliefs has full support on the full space of uncertainty. Thus, we say that Ann's type t_A is *cautious* if there exists a belief in the associated set of belief which has full support on $S_B \times T_B$.

For example, consider a type of Ann's, t_A, which has only a singleton set of beliefs $\{\mu_A\}$ with support as depicted in Fig. 1. For such a cautious type, the question arises of which strategies are rational. Accordingly, consider the marginal of μ_A on Bob's strategy space S_B. This marginal has full support on S_B and if Ann is rational, her rational choice has to be a best-reply to this marginal. It then follows from Pearce (1984) that she must choose a strategy which is not weakly dominated.

Now, it is possible to study the infinite sequences described above. In this case the picture that emerges looks like Fig. 2. Here the small area with solid boundary corresponds to all strategy-type combinations of Bob satisfying RCCARC. Now, set a strategy-type combination (s_A, t_A) for Ann. Does this type correspond to RCCARC for Ann? i.e. does the type satisfy the sequence $a_1, a_2, \ldots$? It is already known that if a_1 holds there needs to be a belief in the associated set of beliefs which has support as μ_A^1. Next, it is considered that Ann assumes b_1. This rules out some of Bob's strategy-type pairs, but also requires t_A to have a belief which has full support on the remaining pairs. Thus, in the associated set of beliefs there needs to a belief μ_A^2. In the next step, Ann is considered to assume b_1 and b_2. Similar reasoning applies and there needs to be a belief like μ_A^3 in the set of beliefs corresponding to t_A. This procedure can now be iterated (as indicated in the picture) to verify whether the type t_A corresponds to RCCARC for Ann. Only finite games are considered here so at some stage n this iteration no longer rules out any *strategies* for Bob. However, it might be the case that at *every* step there are still some types of Bob's that need to be ruled out. In the worst case there needs to be a different belief for each iteration as the support of each belief is changing over the course of the sequence. However, this does not cause a problem. For each type the set of beliefs could be potentially very large.[12] Since such large sets of beliefs are within the framework under consideration, the event RCCARC is not empty. Thus we do not get a negative result, as Brandenburger et al. (2008) find in a different framework. To illustrate more specifically how this analysis works, types are added explicitly for the example considered above.

Example 2 *(continuing from p. 200).* Consider the following type space $T_i = \{t_i^0, t_i^1, t_i^2, t_i^3\}$ for $i = A, B$ and define (with some abuse of notation) the following beliefs on $S_B \times T_B$:

$$\mu_A^1(s_B, t_B) = 1/16, \text{ for all } (s_B, t_B) \in S_B \times T_B,$$
$$\mu_A^2(s_B, t_B) = 1/9, \text{ for all } (s_B, t_B) \in \{A, B, C\} \times \{t_B^1, t_B^2, t_B^3\}, \text{ and}$$
$$\mu_A^3(s_B, t_B^3) = 1/3, \text{ for all } s_B \in \{A, B, C\}.$$

[12] That is, not finitely generated sets.

Similarly, define the following beliefs on $S_A \times T_A$:

$$\mu_B^1(s_A, t_A) = 1/12 \text{ for all } (s_A, t_A) \in S_A \times T_A,$$

$$\mu_B^2(s_A, t_A^1) = 1/6, \text{ for all } s_A \in S_A, \ \mu_B^2(s_A, t_A) = 1/8, \text{ for all } (s_A, t_A) \in \{H, M\} \times \{t_A^2, t_A^3\}, \text{ and}$$

$$\mu_B^3(s_A, t_A^3) = 1/2, \text{ for all } s_A \in \{H, M\}.$$

Given these beliefs, define the set of beliefs $M_i(t_i)$ for each type as follows: for $i = A, B$ set $M_i(t_i^0) = \{\mu_i^3\}$, $M_i(t_i^1) = \{\mu_i^1\}$, $M_i(t_i^2)$ is the convex hull of μ_i^1 and μ_i^2, and $M_i(t_i^3)$ is the convex hull of μ_i^1, μ_i^2, and μ_i^3

Now, it is possible to analyze the infinite sequences $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ introduced above. a_1 is the event that Ann is rational and cautious, so we must collect all strategy-type pairs which satisfy the full-support requirement (cautiousness) and the requirement that the strategy is a best-reply to all beliefs of the given type (rationality). Here, all types but t_i^0 have at least one belief with full support on $S_{-i} \times T_{-i}$. Together with rationality this gives that the following strategy-type pairs correspond to a_1: $S_A \times \{t_A^1\} \cup \{H, M\} \times \{t_A^2, t_A^3\}$. Similarly, b_1 corresponds to $\{A, B, C\} \times \{t_B^1, t_B^2, t_B^3\}$. So both a_1 and b_1 rule out some strategy-type pairs and in particular the weakly dominated strategy D is ruled out. Next, to get to a_2, we want to find all types of Ann that assume b_1. That is, all types of Ann that have at least one belief with full-support on $\{A, B, C\} \times \{t_B^1, t_B^2, t_B^3\}$. Only t_A^2 and t_A^3 satisfy this requirement, leaving $\{H, M\} \times \{t_A^2, t_A^3\}$ corresponding to a_2. For Bob, it emerges that b_2 corresponds to $\{A, B, C\} \times \{t_B^2, t_B^3\}$. Again, note that in this step the interactive reasoning leads to the ruling out of L, which is weakly dominated after elimination of D. In the next step (i.e. a_3 and b_3), types t_i^2 are ruled out, but no more strategies. This construction, however, would lead to the conclusion that a_4 and b_4 do not correspond to any strategy-type pairs. The solution, and this is the main idea of how to prove one direction of Theorem 1, is to add more types. For each iteration add another type with full support on the previous rounds (similar to types t_i^3). This gives an infinite (but countable) number of types and only the "limiting" type corresponds to RCCARC. This argument shows that the illustration in Fig. 2 is accurate in the sense that for higher order iterations the supported strategies are constant, but only types are removed in each round.

Theorem 2 provides a direct (and hence different) way to construct finite type spaces so that for strategy-type pairs satisfying RCCARC the strategies of iterated admissibility (or those of any other self-admissible set) are obtained. ◇

3. Preliminaries

This section presents the main standard concepts and formalism related to game and decision theory. The object of study is the inclusion-exclusion problem inherent in the iterated elimination of weakly dominated strategies raised by Samuelson (1992). Thus, Section 3.1 recalls the formalization of strategic-form games, iterated admissibility (Luce and Raiffa, 1957; Moulin, 1979) and self-admissible sets (Brandenburger et al., 2008). However, our analysis models players as individual decision makers whose beliefs may display ambiguity via incomplete preferences. Section 3.2 recalls the necessary decision-theoretical toolbox and Bewley's (2002) model of incomplete preferences as formalized by Gilboa et al. (2010), and highlights its key features.[13]

3.1. Games and iterated strategy elimination

A *game* consists of a tuple $G := \langle I, (S_i, u_i)_{i \in I} \rangle$ where I is a finite set of *players*, and for each player i there is a finite set of (pure) *strategies* S_i and a *utility function* $u_i : S \to \mathbb{R}$, where $S := \prod_{i \in I} S_i$ denotes the set of strategy *profiles*. For each player i a randomization of own strategies $\sigma_i \in \Delta(S_i)$ is referred to as a *mixed* strategy,[14] and a probability measure $\mu_i \in \Delta(S_{-i})$, where $S_{-i} := \prod_{j \neq i} S_j$, as a *conjecture*. When necessary, with some abuse of notation, we use s_i to refer to the degenerate mixed strategy that assigns probability one to s_i. Each conjecture μ_i and possibly mixed strategy σ_i naturally induce *expected* utility $U_i(\mu_i; \sigma_i)$ and based on this, each player i's *best-reply* correspondence is defined by assigning to each conjecture μ_i the subset of pure strategies $BR_i(\mu_i)$ that maximize its corresponding expected utility.[15]

Following the duality results of Pearce (1984), we use the best-reply correspondence directly to define iterated admissibility whose foundations are then studied in Section 5. Strategy s_i is *iteratively admissible* if it survives the iterated elimination of weakly dominated strategies; i.e., if it is not weakly dominated given strategy profiles $S_{-i} \times S_i$, it is not weakly dominated given strategy profiles $W_{-i}^1 \times W_i^1$ consisting only of strategies surviving the first elimination round, etc. Thus, formally, strategy s_i is iteratively admissible if $s_i \in W_i^\infty := \bigcap_{n \geq 0} W_i^n$, where $W_i^0 := S_i$ and for any $n \in \mathbb{N}$,

[13] Section A in the appendix provides further details on the decision theoretic foundations and on how to envision games as decision problems, as is standard in the literature since Tan and da Costa Werlang (1988) (see Di Tillio (2008) for a more detailed formulation).

[14] Throughout the paper, for any topological space X, as usual, $\Delta(X)$ denotes the set of probability measures on the Borel σ-algebra of X.

[15] That is, given conjecture μ_i the expected utility is $U_i(\mu_i; \sigma_i) := \sum_{(s_{-i}; s_i) \in S} \mu_i[s_{-i}] \cdot \sigma_i[s_i] \cdot u_i(s_{-i}; s_i)$ for each possibly mixed strategy σ_i, and the set of best-replies is $BR_i(\mu_i) := \arg\max_{s_i \in S_i} U_i(\mu_i; s_i)$.

$$W_i^n := \left\{ s_i \in W_i^{n-1} \;\middle|\; \begin{array}{l} \text{There exists some } \mu_i \in \Delta(S_{-i}) \text{ such that:} \\ (i) \quad \text{supp } \mu_i = \prod_{j \neq i} W_j^{n-1}, \\ (ii) \quad s_i \in BR_i(\mu_i) \end{array} \right\}.$$

Finally, that set of strategy profiles $Q = \prod_{i \in I} Q_i$ is said to be a *self-admissible set* (SAS) if for every player i the following three conditions are satisfied:

(i) No $s_i \in Q_i$ is weakly dominated given $S_{-i} \times S_i$.
(ii) No $s_i \in Q_i$ is weakly dominated given $Q_{-i} \times S_i$.
(iii) For every $s_i \in Q_i$ and every mixed strategy σ_i such that $U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i)$ for every s_{-i}, it holds that supp $\sigma_i \subseteq Q_i$.

The connection between the notions of self-admissibility and iterated admissibility is immediately apparent: the set of iteratively admissible strategy profiles is a self-admissible set of game G, but in general there are other self-admissible sets. For details see Brandenburger and Friedenberg (2010), who also study properties of self-admissible sets for specific (classes of) games.

3.2. Decision problems and incomplete preferences

We follow the reformulation of Anscombe and Aumann's (1963) framework by Fishburn (1970). The decision maker faces *decision environment* (Z, Θ) where: (i) Z is a set of *outcomes*, which can be informally understood as the elements that will ultimately yield direct utility to the decision maker; and (ii) Θ is a set of *states* (of the world) about which the decision maker might face uncertainty, and which may affect how her choices relate to outcomes. We refer to randomizations of outcomes, $\ell \in \Delta(Z)$, as *lotteries*. A *preference* is a binary relation $\succsim$ over the set of *acts*, $\mathcal{F}$, which is the collection of all maps $f : \Theta \to \Delta(Z)$ that assign a lottery to each state. $\mathscr{M}(\Theta)$ denotes the set of closed and convex nonempty subsets of Θ.[16] Throughout the paper we focus on preferences which we call *Bewley* preferences, since they were introduced by Bewley (2002).[17] The main point of departure from the preferences of a standard Bayesian decision maker (i.e., one whose preferences satisfy the axioms by Anscombe and Aumann, 1963) is that completeness of the preferences is dropped. Theorem 1 by Gilboa et al. (2010) provides the following convenient representation for these preferences[18]: $\succsim$ is a Bewley preference if and only if there exist a non-constant utility-function $u : Z \to \mathbb{R}$ and a set of *ambiguous beliefs* $M \in \mathscr{M}(\Theta)$ such that for every pair of acts f, g,[19]

$$f \succsim g \iff \int_\Theta \mathbb{E}_{f(\theta)}[u(z)]\mathrm{d}\mu \geq \int_\Theta \mathbb{E}_{g(\theta)}[u(z)]\mathrm{d}\mu \text{ for every } \mu \in M.$$

A decision maker's epistemic attitude with respect to the source of uncertainty may not be represented by a single belief, as in the standard case, but rather by a possibly non-singleton set of beliefs that reflects the decision maker's possible ambiguity towards that source of uncertainty. As argued extensively in Section 5, this is key to resolving the inclusion-exclusion problem. With such preferences the decision maker is allowed to have beliefs with different supports, but also needs to respond robustly to her ambiguity by best-replying to all of her beliefs: for act f to be regarded as at least as good as another act g, the expected utility for f must be at least as high as the expected utility for g for *every* belief in the set of ambiguous beliefs. Notably, recent work by Cettolin and Riedl (2019) presents experimental tools to test whether the preferences display this form of ambiguity via incompleteness.

4. Reasoning in games

For the rest of the paper we consider game G to be fixed and therefore drop most explicit mentions to it. In this section we present the epistemic framework that we employ below to establish foundations for iterated admissibility in Section 5. Formally, for each player we specify a choice and a representation of her beliefs on her opponents' strategies, her beliefs on her opponents' beliefs over their opponents' strategies, etc. These elements suffice to assess whether under

[16] To be more mathematically precise, Z is assumed to be finite, Θ is compact and metrizable and the elements of $\mathcal{F}$, simple and measurable in the Borel σ-algebra of Θ. Space $\mathscr{M}_i(S_{-i} \times T_{-i})$ is endowed with the topology induced by the Hausdorff metric and is therefore compact and metrizable.

[17] Actually we rely on a more modern version by Gilboa et al. (2010) of Bewley's (2002) original preferences. Bewley's (2002) version requires the decision-maker to have a designated default act always chosen unless ranked strictly lower than some alternative. This is commonly known in the literature as *inertia* (see, Bewley, 2002, or Lopomo et al., 2011). Furthermore, the version of Gilboa et al. (2010) allows for infinite state spaces which are necessary in our framework.

[18] Details about the axioms on preferences and how to map the game theoretic setup to a decision environment can be found in Appendix A.

[19] More precisely, M is non-empty closed and convex. Moreover, M is unique and u is unique up to positive affine transformations.

such specifications, the player is being rational, has preferences that exhibit ambiguity (i.e. multiple beliefs) or has certain higher-order beliefs on her opponents' rationality and the presence of ambiguity in their preferences. The question then is which precise constraints on rationality and higher-order beliefs on opponents' rationality induce the behavior captured by iterated admissibility. Section 5 provides an answer based on the formalism developed in this section.

However, some previous methodological work is required. As seen above, when ambiguity via incomplete preferences is allowed for, the representation of uncertainty may require non-singleton sets of beliefs. It follows that standard type structures as introduced by Harsanyi (1967–1968) and standard belief-hierarchies à la Mertens and Zamir (1985) are not suitable for analyzing strategic reasoning: They fail to capture the possibility of ambiguity. Instead, we rely on a modified version of type structure that accounts for ambiguous beliefs.[20] Thus, in Section 4.1 we first introduce these ambiguous type structures. We build on them and then, in Section 4.2, define the restrictions on behavior and beliefs required for the results in Section 5.

4.1. Ambiguous type structures

The study of strategic reasoning requires an instrument that formalizes players' beliefs about their opponents' choices, players' beliefs about their opponents' beliefs about their opponents' choices and so on. When players have complete preferences this hierarchical uncertainty can easily be represented through type structures. Thus, it is convenient to extend the definition of the latter so that can deal with the possibility of ambiguity. Formally, an *ambiguous type structure* consists of a list $\mathcal{T} := \langle T_i, M_i \rangle_{i \in I}$ where for each player i there is[21]:

(i) A set of (*ambiguous*) *types* T_i.
(ii) An *ambiguous belief map* $M_i : T_i \to \mathscr{M}_i(S_{-i} \times T_{-i})$, where $T_{-i} := \prod_{j \neq i} T_j$, that associates each type with ambiguous beliefs on opponents' strategy-type pairs.

It is easy to see why ambiguous type structures capture the idea of hierarchical reasoning mentioned at the beginning of the paragraph. For any player i's type t_i it is possible to compute the following by recursive marginalization[22]:

(1) *First-order* ambiguous beliefs that represent type t_i's uncertainty about her opponents' strategies, $M_{i,1}(t_i) \in \mathscr{M}_{i,1} := \mathscr{M}_i(S_{-i})$, which is easily obtained by taking the marginals on S_{-i} of the beliefs in $M_i(t_i)$.
(2) *Second-order* ambiguous beliefs that represent type t_i's uncertainty about her opponents' strategy-first-order ambiguous beliefs pairs, $M_{i,2}(t_i) \in \mathscr{M}_{i,2} := \mathscr{M}_i(\prod_{j \neq i}(S_j \times \mathscr{M}_{j,1}))$.
...
(n) *nth-order* ambiguous beliefs that represent type t_i's uncertainty about her opponents' strategy-($n-1$)th-order ambiguous beliefs pairs, $M_{i,n}(t_i) \in \mathscr{M}_{i,n} := \mathscr{M}_i(\prod_{j \neq i}(S_j \times \mathscr{M}_{j,n-1}))$.
...

Ambiguous type structure $\mathcal{T}$ is said to be *complete* if every map M_i is surjective, that is, if for every possible ambiguous beliefs the ambiguous type structure may admit, there exists some type that is mapped to such ambiguous beliefs.[23]

4.2. Behavioral and epistemic conditions

The analysis of each player i's reasoning is focused on strategy-type pairs (s_i, t_i), which specify both player i's choice, and as described above, her ambiguous beliefs on her opponents' choices, her ambiguous beliefs on her opponents' first-order ambiguous beliefs, etc. Thus, each strategy-type pair (s_i, t_i) enables questions such as the following to be addressed: Is player i rational given her beliefs? Do her preferences embody some kind of ambiguity? What are her higher-order beliefs about her opponents' rationality and ambiguity? Next, we first formalize the notion of rationality that we employ (Section 4.2.1). Second, we introduce our formalization of cautiousness as a manifestation of ambiguity (Section 4.2.2). Finally,

[20] These type structures are regarded to Ahn's (2007) *ambiguous hierarchies* what Harsanyi's (1967–1968) type structures are to Mertens and Zamir's (1985) belief hierarchies.

[21] We assume each T_i to be compact and metrizable and each M_i, continuous. See Footnote 40.

[22] The conceptual simplicity that follows contrasts the notational complexity that it requires; technically, for each $n \in \mathbb{N}$ we have:

$$M_{i,n+1}(t_i) = \left\{ \mu_i \in \Delta\left(\prod_{j \neq i}(S_j \times \mathscr{M}_{j,n})\right) \,\middle|\, \begin{array}{l} \text{There exists some } \mu_i' \in M_i(t_i) \text{ such that:} \\ \\ \mu_i[E] = \mu_i'\left[\left(\prod_{j \neq i}(\mathrm{id}_{S_j} \times M_{j,n})\right)^{-1}(E)\right] \\ \\ \text{for every measurable } E \subseteq \prod_{j \neq i} S_j \times \mathscr{M}_{j,n} \end{array} \right\}.$$

.

[23] As shown by Ahn (2007), the answers to the following modified questions in (Dekel and Siniscalchi, 2015, p. 629): "Is there a[n] [ambiguous] type structure that generates *all* [ambiguous] hierarchies of beliefs? Is there a[n] [ambiguous] type structure into which every other [ambiguous] type structure can be embedded?" are yes, and yes. Within a Bayesian framework, Friedenberg (2010) studies such a richness requirement more generally.

we define the appropriate tool to impose restrictions on higher-order beliefs (Section 4.2.3), which is a generalization to Bewley preferences of the usual notion of full-support belief for standard Bayesian preferences.

4.2.1. Rationality

We say that strategy s_i is *rational* for type t_i if s_i is a best-reply to every first-order ambiguous belief induced by t_i; thus, the set of strategy-type pairs in which player i is rational is formalized as follows:

$$R_i := \left\{ (s_i, t_i) \in S_i \times T_i \,\middle|\, s_i \in \bigcap_{\mu_i \in M_i(t_i)} BR_i(\text{marg}_{S_{-i}} \mu_i) \right\}.$$

Note that the definition implicitly requires each type t_i, in order to be eligible for rational behavior, to satisfy that the intersection of the best-replies to the ambiguous first-order beliefs induced by it is non empty.[24] This is a consistency requirement in the vein of Bayesian updating for conditional probability systems in the literature of dynamic games: When a conditional probability system fails to satisfy Bayesian updating it may not admit sequential best-replies.[25]

4.2.2. Cautiousness and ambiguity

We next argue that cautiousness, intuitively thought of as the decision maker considering every state of the world when deciding which choice is best, can be interpreted as a product of ambiguity in the sense that types that exhibit cautiousness tend to represent preferences that also display ambiguity. We first formalize the notion of cautiousness that takes part in the characterizations result in Section 5 and then discuss its link to ambiguity.[26]

Definition 1 *(Cautiousness).* Let G be a game and $\mathcal{T}$, an ambiguous type structure. Then, for any player i and any type t_i we say that type t_i is cautious if at least one belief in $M_i(t_i)$ has full-support on $S_{-i} \times T_{-i}$. We denote the set of player i's strategy-type pairs in which the type is cautious by C_i.

If at an intuitive level cautiousness is seen as the idea that a decision maker takes every possible contingency into account, then that is present in this definition. Cautiousness requires, loosely speaking, that every state is taken into account by the decision maker.[27] The link with ambiguity is easy to see. In principle, it is possible for a type to display cautiousness but not ambiguity. This is the case of every type whose set of ambiguous beliefs consists of a single belief with full-support on $S_{-i} \times T_{-i}$ as in Fig. 1. However, if in addition to cautiousness the type also exhibits some form of strategic sophistication in the sense of having a (different) belief that rules out some proper subset of $S_{-i} \times T_{-i}$, then, necessarily, the type displays ambiguity: The corresponding ambiguous beliefs a fortiori contain at least two different beliefs. Hence, the introduction of ambiguity not only enables strategic reasoning and cautiousness to be made compatible, but is indeed, necessary when strategic reasoning has any bite.

4.2.3. Assumption

Hereafter we refer to measurable subsets $E \subseteq S \times T$ as events. A standard Bayesian decision maker is said to *assume* event E when the unique subjective belief induced by her preference has full-support on E.[28] Some changes are in order if this idea is to be extended to Bewley preferences: The set of ambiguous beliefs may contain beliefs that have different supports. We say that a Bewleyian decision maker assumes event E when *at least one* belief in her set of ambiguous beliefs has full-support on E. Given the inclusion-exclusion problem, it is natural to consider such a weak version of assumption. As discussed in Section 1, it is necessary to have multiple beliefs which have potentially different supports to resolve the tension between strategic reasoning and cautiousness.

Definition 2 *(Assumption).* Let G be a game and $\mathcal{T}$, an ambiguous type structure. For any player i, any type t_i and any event $E_{-i} \subseteq S_{-i} \times T_{-i}$ we say that type t_i assumes E_{-i} if at least one belief in $M_i(t_i)$ has full-support on the topological closure of E_{-i}. We denote the set of player i's strategy-type pairs in which the type assumes E_{-i} by $A_i(E_{-i})$.

[24] We discuss this requirement in detail in Appendix A, where we first separate the condition that ensures non-emptiness of the intersection of best replies (*decisiveness*) from rationality per se, and provide a behavioral characterization for it.

[25] We thank Pierpaolo Battigalli for this observation. This issue, which refers to the distinction between a choice being optimal or undominated, is discussed in further detail in Section A in the appendix, which also provides a behavioral foundation for a non-empty intersection.

[26] We note that all of the following analyses could have been carried out employing a slightly weaker notion of cautiousness than the one introduced in Definition 1. In principle it would suffice to require full support on S_{-i} rather than $S_{-i} \times T_{-i}$. Our reason for opting for the stronger notion is twofold: (*i*) It does not prevent our characterization from dispensing with impossibility issues à la Brandenburger et al. (2008) (see Section 5.3), so it is clear that it is not modifications in the notion of cautiousness that enable for this to be achieved; and (*ii*) since it does not apply only to state spaces with product structure, it has a more general decision-theoretic foundation.

[27] Cautiousness is also present in the analysis by Brandenburger et al. (2008). However, there it is incorporated into the definition of rationality. We find it more transparent to explicitly define the event when a player is cautious.

[28] Technically, we are considering the collapse of the notion of assumption (see Brandenburger et al., 2008 and Dekel et al., 2016) under the lexicographic probability system when the preferences satisfy continuity and the corresponding lexicographic probability system thus collapses to a single belief.

Remark 1. Cautiousness as defined in Definition 1 can be restated in terms of assumption: A type t_i is cautious if it assumes $S_{-i} \times T_{-i}$.

5. Iterated admissibility and ambiguous types

This section presents the main results of the paper. Based on the observation made in the previous section that the presence of ambiguity can reconcile strategic reasoning with cautiousness, we provide foundations for iterated admissibility and self-admissibility in terms of rationality, cautiousness, and certain higher-order assumption constraints. We provide those foundations in Section 5.1. Then, in Section 5.2, we discuss the link between iterated assumption and ambiguity to resolve the inclusion-exclusion problem. Finally, in Section 5.3 we review the seminal impossibility result due to Brandenburger et al. (2008) within the approach in terms of lexicographic probability systems, recall some of the responses in the related literature, and explore the connection with our result.

5.1. Epistemic foundation

As mentioned above, the epistemic foundation of iterated admissibility is to be formulated in terms of rationality, cautiousness, and higher-order assumption restrictions. The set of strategy-type pairs in which player i exhibits common assumption in rationality and cautiousness is given by $CARC_i := \bigcap_{n\geq 0} CARC_{i,n}$, where each $CARC_{i,n}$ is defined recursively by setting:

$$CARC_{i,0} := S_i \times T_i,$$

$$CARC_{i,n} := CARC_{i,n-1} \cap A_i(\prod_{j\neq i} R_j \cap C_j \cap CARC_{j,n-1}),$$

for every $n \in \mathbb{N}$. That is, $CARC_i$ brings together all the strategy-type pairs (s_i, t_i) where player i's type t_i assumes that every player $j \neq i$ is rational, cautious, and assumes that every player $j \neq i$ assumes that every player $k \neq j$ is rational, cautious, and so on. Based on the above[29]:

Theorem 1 *(Foundation of iterated admissibility). Let G be a game. For any player i the following holds:*

(*i*) *For any complete ambiguous type structure, any player i and any strategy-type pair (s_i, t_i), if type t_i is consistent with cautiousness and assumption of rationality and cautiousness and s_i is rational for t_i, then s_i is iteratively admissible; i.e.,*

$$\text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i) \subseteq W_i^{\infty}.$$

(*ii*) *For any player i and any strategy s_i, if s_i is iteratively admissible then there exist a complete ambiguous type structure $\mathcal{T}$ and a type t_i consistent with cautiousness and assumption of rationality and cautiousness for which s_i is rational; i.e.,*

$$W_i^{\infty} \subseteq \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i).$$

Thus, Theorem 1 provides a complete characterization of iterated admissibility. Part (i) is a sufficiency result. It shows that whenever a player chooses in a robust way that maximizes with respect to higher-order assumptions that represent common assumption in rationality and cautiousness, then the resulting strategy is necessarily iteratively admissible. Part (ii) is, partially, the necessity counterpart: while it is not true that every time an iteratively admissible strategy is chosen this is due to the player being rational, cautious, and best-replying to the higher-order assumption restrictions that represent common assumption in rationality and cautiousness, it is true that every iteratively admissible strategy is a rational choice for a type that is consistent with common assumption in rationality and cautiousness. The proof of the theorem is provided by iteration and relies on the slightly stronger result according to which, n rounds of elimination of weakly dominated strategies characterize the behavioral implications of rationality, cautiousness and $n-1$ rounds of assumption in rationality and cautiousness.[30] Notably, from a conceptual perspective, the theorem reveals that whenever the elimination procedure involves more than one round, satisfying the epistemic conditions above requires players' preferences to display ambigu-

[29] The theorem is stated and holds only for a complete type structure because the assumption operator is not monotone. This is similar to, for example, assumption in Brandenburger et al. (2008) or strong belief of Battigalli and Siniscalchi (2002). An example showing why completeness is needed is available upon request.

[30] This statement is properly formalized in Theorem B.1 in Section B. For expositional reasons, we opted here to present the result corresponding to only the full iteration process, so that the result for iterated admissibility and the result for self-admissibility (Theorem 2 below) can be compared straightforwardly.

ity. As the next theorem shows, if the requirement of completeness of the type structure is dropped then the behavioral consequences of rationality, cautiousness and common assumption thereof are captured by self-admissibility:

Theorem 2 *(Foundation of self-admissibility). Let G be a game. Then:*

(i) *For any ambiguous type structure $\mathcal{T}$ the set of strategies consistent with rationality, cautiousness and common assumption of rationality and cautiousness is a self-admissible set; i.e., the following set is self-admissible:*

$$\prod_{i\in I}\mathrm{Proj}_{S_i}(R_i\cap C_i\cap CARC_i).$$

(ii) *For any self-admissible set Q there exists a finite ambiguous type structure $\mathcal{T}$ for which Q characterizes the behavioral implications of rationality, cautiousness and common assumption of rationality and cautiousness; i.e., such that:*

$$\prod_{i\in I}\mathrm{Proj}_{S_i}(R_i\cap C_i\cap CARC_i)=Q\,.$$

The interpretation is analogous to that of Theorem 1. Part (i) states that given an arbitrary ambiguous type structure, not necessarily complete, the set of strategy profiles that are consistent with rationality, cautiousness and common assumption of rationality and cautiousness is a self-admissible set. Part (ii) offers the partial converse: For any given self-admissible set Q there exists an ambiguous type structure $\mathcal{T}$, notably, finite, such that Q is exactly the set of strategy profiles that are consistent with rationality, cautiousness and common assumption of rationality and cautiousness within $\mathcal{T}$. Theorems 1 and 2 are clearly connected because the set of iteratively admissible strategy profiles is itself self-admissible. In particular, for a fixed game this reveals that the set of iteratively admissible strategies can be understood as strategies obtained not only in a very *large* complete type structure, but also under a *smaller* finite one in which, as shown in the proof of Theorem 2, each player i only has as many types as there are iteratively admissible strategies plus one additional dummy type.[31]

5.2. Iterated assumption and ambiguity

The main distinctive feature of assumption with respect to the usual belief for Bayesian agents, and as in the assumption operator of Brandenburger et al. (2008), is the failure of monotonicity.[32] Whenever a Bayesian agent believes in event E, she also believes in every event F such that $E\subseteq F$: The (Bayesian) belief μ_i that assigns probability one to E assigns probability one to F. This is not the case with our notion of assumption. Type t_i might assume event E via some belief $\mu_i\in M_i(t_i)$ that has full-support on E, but she may fail to assume an event F such that $E\subseteq F$[33]; even if t_i assumed such F, it certainly, could not be via μ_i. Thus, when considering a sequence of nested events such as the finite iterations in the common assumption events defined above, a single belief can assign probability one to all the events in the sequence simultaneously, but different beliefs are required in order to assume each of them at the same time. This is exactly why the inclusion-exclusion problem arises within a standard Bayesian framework, but it can be resolved within our framework.

In principle there is no reason to consider that the assumption of an event is an expression of cautiousness; for every type there exists *always* an event that is assumed and this simply relates to which specific states play some role in how preference ranks acts. However, the assumption of different nested events is a non-trivial feature that reveals a cautious attitude: Whenever a type assumes two nested events E and F, the preference represented is crucially sensitive to comparisons at every state in E but also to comparisons at every state in the larger event F, in particular to those outside E. Of course, as mentioned above, the simultaneous assumption of different events necessarily requires belief multiplicity.

5.3. (Non-)Emptiness of common assumption of rationality and cautiousness

The canonical epistemic foundation of iterated admissibility in the literature is due to Brandenburger et al. (2008). Their seminal result shows that m rounds of elimination of non-admissible strategies characterize the behavioral implications of rationality and mth-order mutual assumption of rationality for finite m in a model where players' uncertainty is formalized by type structures where types are mapped to lexicographic probability systems. As shown by Blume et al. (1991a), lexicographic probability systems arise under a variation of Anscombe and Aumann's (1963) preferences in which the axiom of continuity is relaxed (rather than that of completeness, as in Bewley's (2002) variant). However, Brandenburger et al. (2008) also reveal a vexatious feature of the common assumption case: Their celebrated impossibility result shows that for

[31] The proof of Theorem 2 proceeds in a way very similar to the one by Brandenburger et al. (2008) of their characterization result for self-admissible sets (Theorem 8.1). In particular, we need exactly the same number of types for each player.

[32] This is also reminiscent of strong belief as defined and studied by Battigalli and Siniscalchi (1999, 2002).

[33] We are implicitly assuming that the topological closure of F contains that of E.

every generic game, if the type structure is complete and maps types continuously, then common assumption in rationality is empty. Below we also discuss the work by Keisler and Lee (2015), Yang (2015), Lee (2016) and Catonini and De Vito (2018a), who propose changes in the formalism that allow for sound epistemic foundations, and compare their results to ours.

Notice first that within our set-up, and for every game G, common assumption in rationality and cautiousness is never empty in complete ambiguous type structures. The intuition behind the claim is easy to see: For each iteration in player i's reasoning process set a belief $\mu_i^n \in \Delta(S_{-i} \times T_{-i})$ that has full-support on the topological closure of $\prod_{j\neq i} R_j \cap CARC_{j,n}$ (these collections of strategy-type pairs are clearly never empty; thus, the belief μ_i^n always exists). Then, define M_i as the topological closure of the convex hull of $\{\mu_i^n\}_{n\in\mathbb{N}}$, and by virtue of the ambiguous type structure being complete, pick type t_i with ambiguous beliefs M_i.[34] By construction, t_i is a type representing common assumption of rationality and cautiousness and hence, $CARC_i$ is non-empty.

Furthermore, as briefly mentioned in Section 1, the non-emptiness of rationality and common assumption thereof does not follow from specific alterations in the formalism (beyond the different decision-theoretic model underlying the approach). This is easier to visualize by direct comparison with other studies that also provide sound foundations for iterated admissibility. Keisler and Lee (2015) obtain their result by dropping the requirement that types are mapped continuously, Yang (2015) considers a weaker version of assumption than that in Brandenburger et al. (2008) and Lee (2016) explicitly imposes coherence on the preferences, which is usually only checked for the beliefs that represent the preferences. For lexicographic probability systems, which he builds on, this makes a difference. As said, we do not require any of these modifications: Our type structures map types continuously, our notion of assumption is a direct adaption of that in Brandenburger et al. (2008) and Dekel et al. (2016),[35] and the coherence requirement implicit in our type structures resembles the standard one in literature due to Brandenburger and Dekel (1993).[36] Finally, Catonini and De Vito (2018a) consider a weaker version of the likeliness-ordering of events that characterizes the lexicographic probability system and an alternative version of cautiousness where only the payoff-relevant aspect of the states of the world play any role. Again (and despite Theorems 1 and 2 would remain unchanged under this alternative notion of cautiousness), we obtain our non-emptiness result with a standard, purely decision-theoretic notion of cautiousness that does not require any specific structure of the set of states.

To end this section, we present a comparison between lexicographic probability systems and ambiguous beliefs that provides some understanding of the differences between the two approaches with respect to the presence of ambiguity. Remember that a lexicographic probability system consists of a finite sequence beliefs $\{\mu^k\}_{k=1}^n \subseteq \Delta(\Theta)$,[37] where the order of the sequence represents the epistemic priority attached to each element: μ^1 is the decision maker's 'primary' hypothesis, μ^2 is the 'secondary' hypothesis, and so on. This is reflected by the lexicographic consideration, i.e. if act f is better than g for belief μ^1, then the comparison between the two acts for the rest of the beliefs in the sequence is immaterial and the decision maker prefers f to g. The main distinction between lexicographic probability systems and ambiguous beliefs is then clear: Both are composed of multiple beliefs, but the former incorporates a hierarchy in terms of epistemic priority and hence removes any trace of ambiguity. However, as we show above, this hierarchy is not important to overcome the inclusion-exclusion problem; what is important is the *multiplicity* of beliefs.

6. Conclusions

Cautiousness in games is intuitively understood as the idea that even when a player deems some of her opponents' strategies to be completely unlikely (typically on the basis of strategic reasoning), she still prefers to choose strategies that are immune to deviations towards such unexpected strategies. This is at odds with the strategically sophisticated expected utility maximization process representing a standard Bayesian rational decision maker who believes her opponent to be rational too: Every suboptimal strategy of the latter is assigned zero probability by the subjective belief of the former, and cannot therefore affect the decision process.

This paper proposes a new theoretic understanding of cautiousness in interactive settings that reconciles it with strategic sophistication. We interpret cautiousness under strategic sophistication as a manifestation of robustness to ambiguity, which renders more choices as non-optimal. Then we show that the resulting behavioral implications can be obtained as a consequence of rationality and related higher-order assumption constraints. Specifically:

(i) We introduce the possibility of ambiguity in beliefs by allowing players' preferences to be incomplete. This is done by replacing the standard Anscombe and Aumann (1963) decision-theoretic framework behind each player with a model

[34] As shown by Ahn (2007), this assignment can take place in an ambiguous type structure that maps types to ambiguous beliefs continuously.

[35] See also Footnote 28.

[36] The requirement is explicit in the construction by Ahn (2007).

[37] Brandenburger et al. (2008) use lexicographic conditional probability systems, but their result extends to more general lexicographic probability systems as shown by Dekel et al. (2016).

of (possibly) incomplete preferences à la Bewley (2002) so that each player's uncertainty about her opponents' behavior is represented by a possibly non-singleton set of beliefs that reflects the decision maker's possibly ambiguous uncertainty. Our main result implies that for choices that are iteratively admissible the justifying set of beliefs has to be non-singleton for non-trivial games.

(*ii*) We apply the framework described above to study the epistemic (i.e. reasoning-based) foundations of iterated admissibility in belief-complete type structures and find that it characterizes the behavioral implications of rationality, cautiousness, and common assumption thereof (Theorem 1). For non-complete type structures we find that it is self-admissible sets that characterize the behavioral implications of such an event (Theorem 2).

Thus, the main insight is immediately apparent: The inclusion-exclusion problem of Samuelson (1992) can be resolved not only by relaxing continuity of preferences (i.e. through lexicographic probability systems), but also by relaxing completeness (while maintaining continuity). Notably, this enables us to provide a sound epistemic foundation of iterated admissibility—a challenging task within the framework of lexicographic probability systems. Using our approach, it is easy to see that the event of rationality, cautiousness, and common assumption thereof is non-empty across all games—unlike, for instance, the foundations for iterated admissibility under lexicographic probability systems, as found by Brandenburger et al. (2008), and the instruments involved in our characterization (type structures and assumption operators) are straightforward generalizations of those in the realm of standard Bayesian preferences. In addition, the suggested link between ambiguity via incomplete preferences and the presence of cautiousness is potentially testable by applying techniques for the identification of incompleteness of preferences recently developed in the literature on experimental economics (see Cettolin and Riedl, 2019).

Finally, the formalism shows that even with incomplete preferences, an iterative solution concept is valid and well-founded. To elaborate, note that the inclusion-extension problem extends, well-beyond iterated admissibility, to every (non trivial) iterated deletion procedure that incorporates cautiousness. This is apparent in Dekel and Fudenberg's (1990) procedure (the *DF-procedure*; *persistency* in Brandenburger, 2003, and Catonini and De Vito, 2018b), which consists of one round of elimination of weakly dominated strategies followed by the iterated elimination of strictly dominated strategies. Here, the notion of cautiousness behind the first elimination round requires player i's beliefs to assign positive probability to every strategy by her opponents (i.e. to *include* all strategies of the opponents) whereas the iterated elimination that follows requires player i's beliefs to assign zero probability to opponents' strategies that did not survive the first round (i.e. to *exclude* some strategies). Hence, the presence of inclusion-exclusion issues makes understanding the DF-procedure problematic from the standard Bayesian perspective. Unsurprisingly, the tension can again be solved via multiplicity of beliefs resulting from ambiguity. Say that player i *believes* event E if *at least* one belief in her set of ambiguous beliefs assigns probability one to E. It is easy to show then that the DF-procedure characterizes the behavioral implications of rationality, cautiousness (as defined in Section 4.2.2) and *common belief* thereof.[38] In addition, it is immediately possible to replicate, within this framework, the well-known result that rationalizability (the iterated elimination of strictly dominated strategies) characterizes rationality and common belief thereof. A comparison between these two observations and Theorem 1 illustrates the theoretical connection between cautiousness and strategic reasoning on the one hand, and ambiguity on the other: In the absence of cautiousness (i.e. rationalizability) behavior can be explained without appealing to ambiguous beliefs, but the latter becomes a sine qua non condition as soon as the solution concept relies on any notion of cautiousness (i.e. iterated admissibility and the DF-procedure).

Acknowledgments

The authors thank the editors, Pierpaolo Battigali, Giacomo Bonanno and Wiebe van der Hoek, and two anonymous referees. Thanks are also due to Eric Mbakop, Emiliano Catonini, Eddie Dekel, Amanda Friedenberg, Elena Iñarra, Willemien Kets, Peter Klibanoff, Christoph Kuzmics, Jean-Philippe Lefort, Matthew Leisten, Martin Meier, Wojcech Olszweski, Alessandro Pavan, Antonio Penta, Burkhard Schipper, Sebastian Schweighofer-Kodritsch, Lorenzo Stanca, Alexander Theisen, Can Urgun, and very especially to Marciano Siniscalchi, as well as to audiences at Universitat de Barcelona (UB), Universitat Pompeu Fabra (UPF), Universidad Pública de Navarra (UPNA), Royal Holloway-University of London, NOeG Winter Workshop 2017, RUD 2018, GTM 2018, LOFT 2018, the 3rd GRODE Workshop, and SAET Conference 2019 for insightful comments and helpful discussion. Ziegler acknowledges financial support from The Eisner Foundation and Zuazo-Garin, from the Spanish Ministry of Economy and Competitiveness, the Department of Education, Language Policy and Culture of the Basque Government (grants ECO2012-31346 and POS-2016-2-0003 and IT568-13, respectively), and the European Research Council (ERC grant #579424). Zuazo-Garin also expresses his gratitude to Northwestern University and Kellogg School of Management (the Department of Economics, MEDS and the CMS-EMS, particularly) for their great hospitality. The authors are responsible for all errors.

[38] In Ziegler and Zuazo-Garin (2019) we systematically study iterated deletion procedures under cautiousness, distinguishing different notions of cautiousness: *Weak* cautiousness (identifiable with the notion of cautiousness here), and strong cautiousness, which helps understand the iterated elimination of never *strict* best-replies.

Appendix A. Decision theory

A.1. Decision problems and Bewley preferences

Given a decision environment (Z, Θ), a *decision problem* consists of a triplet (Z, Θ, F) where F is a subset of acts that we call *feasible* and represents the acts that are materially available to the decision maker.[39] Bewley preferences satisfy the following axioms:

A1. *Preorder.* $\succsim$ is reflexive and transitive.
A2. *Monotonicity.* For any pair of acts f, g,

$$f(\theta) \succsim g(\theta) \text{ for any } \theta \in \Theta \implies f \succsim g.$$

A3. *Continuity.* For any three acts f, g, h the following two are closed in $[0, 1]$:

$$\left\{\lambda \in [0,1] \,\middle|\, \lambda f + (1-\lambda)g \succsim h\right\} \quad \text{and} \quad \left\{\lambda \in [0,1] \,\middle|\, h \succsim \lambda f + (1-\lambda)g\right\}.$$

A4. *Nontriviality.* There exist two acts f, g such that $f \succsim g$ and not $g \succsim f$.
A5. *Certainty-Completeness.* For any two constant acts f, g either $f \succsim g$ or $g \succsim f$.
A6. *Independence*. For any acts f, g, h and any $\alpha \in (0, 1)$,

$$f \succsim g \iff \alpha f + (1-\alpha)h \succsim \alpha g + (1-\alpha)h.$$

Theorem 1 in Gilboa et al. (2010) shows that these axioms characterize the preferences to get the representation stated in the main text.

A.2. Games as decision problems

Players are envisioned as individual decision makers facing a decision problem where their opponents' strategies are part of the description of the states of the world and strategies are the feasible acts. For obvious reasons, for each player i, game G is a very specific decision problem (Z_i, Θ_i, F_i) consisting of:

- *Outcomes.* In contexts of complete (payoff-relevant) information, player i's utility depends only on the strategy profiles chosen in the game; hence, we identify outcomes with the latter: $Z_i := S$.
- *States.* Player i's primary source of uncertainty (and the only payoff-relevant one) is strategic: it refers to her opponents' behavior (S_{-i}). However, player i's beliefs about her opponents' strategies could be affected by an additional non payoff-relevant unobserved parameters about which she might face uncertainty, say T_{-i}.[40] We identify the set of states of the world with these joint sources of uncertainty: $\Theta_i := S_{-i} \times T_{-i}$.
- *Acts and feasible acts.* Player i's set of acts is $\mathcal{F}_i := \Delta(S)^{S_{-i} \times T_{-i}}$. Notice that within the context of a game this set of acts is not feasible. First, player i cannot make her choice contingent on a parameter t_{-i} that she does not observe. Second, in situations of simultaneous choice, player i cannot make her choice contingent on her opponents' choices. Still, player i might (and typically will) have preferences on modeled but unavailable options. The set of player i's feasible acts is then identified with her mixed strategies:

$$F_i := \left\{ f \in \mathcal{F}_i \,\middle|\, \begin{array}{l} \text{There exists a } \sigma_i \in \Delta(S_i) \text{ such that:} \\ f(s_{-i}, t_{-i})[(s'_{-i}; s'_i)] = \begin{cases} \sigma_i[s'_i] & \text{if } s'_{-i} = s_{-i}, \\ 0 & \text{otherwise,} \end{cases} \\ \text{for any } (s_{-i}, t_{-i}) \in S_{-i} \times T_{-i} \text{ and any } (s'_{-i}; s'_i) \in S \end{array} \right\}.$$

In addition, remember that game G already incorporates utility functions; thus, each player i's set of Bewley preferences under consideration needs to be restricted to those preferences whose risk attitude is represented by utility function u_i. Now, Theorem 1 by Gilboa et al. (2010) implies that for any set of parameters T_{-i}, each Bewley preference for decision environment $(S, S_{-i} \times T_{-i})$ whose risk attitude is represented by u_i is biunivocally associated with ambiguous beliefs $M_i \subseteq \Delta(S_{-i} \times T_{-i})$.[41] Thus, there is no loss of generality in switching the focus from Bewley preferences to ambiguous beliefs, the collection of which we denote by $\mathscr{M}_i(S_{-i} \times T_{-i})$.

---

39 That is, the decision maker may have preference on elements of not only F, but $\mathcal{F}$, which means that she might have preferences on options that are not materially available in the problem under study.

40 To ensure appropriate construction, T_{-i} is assumed to be compact and metrizable.

41 Remember that M_i is non-empty, closed, and convex. Of course, M_i is a subset of $\Delta(S_{-i})$ in cases in which we omit set of parameters T_{-i}.

A.3. Decisiveness

We refer to the types that admit rational choices as *decisive*. The foundation of decisiveness in terms of preferences is provided by Proposition 1 below. Decisive types are those induced by preferences that are possibly incomplete but display completeness *at the top*: the decision maker is indifferent between two acts that are not less preferred than another act.[42]

Proposition 1 *(Behavioral foundation of decisiveness). Let G be a game and $\mathcal{T}$, an ambiguous type structure. Then, any player i's type t_i is decisive if and only if there exists a subset of feasible acts $F_i^* \subseteq F_i$, such that $\succsim_i$, the Bewley preference represented by $(u_i, M_i(t_i))$, satisfies*

$$f \sim_i g \succ_i h,$$

for every $f, g \in F_i^*$ and every $h \in F_i \setminus F_i^*$.

Proof. Fix player i, type t_i and event $E_{-i} \subseteq S_{-i} \times T_{-i}$ and let $\succsim_i$ denote the Bewley preference represented by $(u_i, M_i(t_i))$. The 'if' part is immediate, so we focus on the 'only if' part. To see it simply take $S_i^* := \bigcap_{\mu_i \in M_i(t_i)} BR_i(\text{marg}_{S_{-i}} \mu_i)$ and set:

$$F_i^* := \left\{ f_i \in \mathcal{F}_i \;\middle|\; \begin{array}{l} \text{There exists a } \sigma_i \in \Delta(S_i) \text{ such that:} \\ (i) \quad f_i(s_{-i}, t_{-i})[(s'_{-i}; s'_i)] = \begin{cases} \sigma_i[s'_i] & \text{if } s'_{-i} = s_{-i}, \\ 0 & \text{otherwise,} \end{cases} \\ \qquad \text{for any } (s_{-i}, t_{-i}) \in S_{-i} \times T_{-i} \text{ and any } (s'_{-i}; s'_i) \in S, \\ (ii) \quad \sigma_i[S_i^*] = 1 \end{array} \right\}.$$

Clearly, $F^* \subseteq F$ and $f \sim_i g \succ_i h$ for every $f, g \in F_i^*$ and every $h \in F_i \setminus F_i^*$. □

Notice that in the presence of incomplete preferences 'undomination' (an act not being strictly worse than *some* other act) and 'optimality' (an act being at least as good as *every* other act) are two different concepts, which is not the case under completeness: An optimal act is always undominated but an undominated act might not be optimal; furthermore, every Bewley preference admits undominated acts, but there may not exist optimal ones. Decisiveness ensures the existence of the latter, which in turn, restores the equivalence of undomination and optimality. In consequence, imposing decisiveness on incomplete preferences is similar in spirit to the requirement of Bayesian updating for conditional probability systems in the literature of extensive-form games.[43] As for decisiveness, Bayesian updating guarantees the existence of optimal strategies by forcing them to be equivalent to undominated ones.

Appendix B. Characterization result

As mentioned in the main text, we first prove the result for every finite iteration.

Theorem B.1. *Let G be a game and $\mathcal{T}$ a complete ambiguous type structure. For any $n \in \mathbb{N}$ and every player i the following holds:*

$$\text{Proj}_{S_i}(R_i \cap C_i \cap CARC_{i,n}) = W_i^{n+1}.$$

Proof. For the sake of convenience, for each player i we denote $X_{i,0} := S_i \times T_i$ and for any $n \in \mathbb{N}$, $X_{i,n} := R_i \cap C_i \cap CARC_{i,n-1}$. Now, we proceed by induction on n:

Initial Step ($n = 0$). For the right-hand inclusion, set strategy-type pair $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i$ and denote $\bar{M}_i = M_i(\bar{t}_i)$. Then, since $\bar{t}_i$ is cautious, we know that there exists a belief $\mu_i^1 \in \bar{M}_i$ whose support is $S_{-i} \times T_{-i}$, and since $(\bar{s}_i, \bar{t}_i) \in R_i$, we know that $\bar{s}_i$ is a best-reply for $\text{marg}_{S_{-i}} \mu_i^1$. Thus, $\bar{\mu}_i^1 := \text{marg}_{S_{-i}} \mu_i^1$ is a conjecture in with full-support on S_{-i} for which $\bar{s}_i$ is a best-reply. Hence, $\bar{s}_i \in W_i^1$.

For the left-hand inclusion, set strategy $\bar{s}_i \in W_i^1$ and conjecture $\bar{\mu}_i$ with full-support on S_{-i} for which $\bar{s}_i$ is a best-reply. Then, take arbitrary full-support belief $\eta_i \in \Delta(T_{-i})$ and set $\mu_i^1 := \bar{\mu}_i \times \eta_i$ and $\bar{M}_i := \{\mu_i^1\}$. Since $\mathcal{T}$ is complete, we know that there exists a type $\bar{t}_i \in T_i$ such that $M_i(\bar{t}_i) = \bar{M}_i$. Since μ_i^1 has full-support on $S_{-i} \times T_{-i}$ we know that $\bar{t}_i$ is cautious,

[42] Despite the following characterization relying on an axiom evoking existence, G being a finite game implies that the verification of the condition requires only finitely many bets.

[43] The definition of conditional probability systems (originally due to Renyi, 1955) requires the decision maker to update her beliefs according to the chain rule whenever possible; this requirement is usually referred to as *Bayesian updating*.

and hence, that $(\bar{s}_i, \bar{t}_i) \in C_i$, and as $\bar{s}_i$ is a best-reply to the marginal on S_{-i} induced by the unique belief in $M_i(\bar{t}_i)$ it follows that $(\bar{s}_i, \bar{t}_i) \in R_i$. Thus, it can be concluded that $(\bar{s}_i, \bar{t}_i)$ is a strategy-type pair in $R_i \cap C_i$ that induces $\bar{s}_i$.

Inductive Step. Suppose that $n \geq 0$ is such that the claim holds. We next verify that it also holds for $n+1$. For the right-hand inclusion, set strategy-type pair $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_{i,n+1}$ and denote $\bar{M}_i := M_i(\bar{t}_i)$. Then, since $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_{i,n}$ it is known from the induction hypothesis that $\bar{s}_i \in W_i^{n+1}$, and since $(\bar{s}_i, \bar{t}_i) \in R_i \cap CARC_{i,n+1}$ there must exist a belief $\mu_i^{n+1} \in \bar{M}_i$ whose support is the closure of $X_{-i,n+1} := \prod_{j \neq i} X_{j,n+1}$ and whose marginal on S_{-i} admits $\bar{s}_i$ as a best-reply. It follows from the induction hypothesis and completeness that the support of $\bar{\mu}_i^{n+1} := \text{marg}_{S_{-i}} \mu_i^{n+1}$ is W_{-i}^{n+1} and hence, it can be concluded that $\bar{s}_i \in W_i^{n+2}$.

For the left-hand inclusion, set strategy $\bar{s}_i \in W_i^{n+2}$ and family of conjectures $\{\bar{\mu}_i^k\}_{k=1}^{n+2}$ such that for each $k = 1, \ldots, n+2$: (i) $\bar{\mu}_i^k$ has full-support on W_{-i}^{k-1}, and (ii) $\bar{s}_i$ is a best-reply to $\bar{\mu}_i^k$. Now, set arbitrary $k = 0, \ldots, n+1$ and for any player $j \neq i$ and any strategy $s_j \in W_j^k$ define:

$$Y_{j,k}(s_j) := \text{Proj}_{T_j} \left(\{s_j\} \times T_j \cap X_{j,k} \right),$$

which is known from the induction hypothesis to be non-empty. It is also known from the induction hypothesis that $\{Y_{j,k}(s_j) | s_j \in W_j^k\}$ is a finite cover of $\text{Proj}_{T_j}(X_{j,k})$. Now, for each $s_{-i} \in W_{-i}^k$ pick arbitrary belief $\eta_i^k(s_{-i}) \in \Delta(\prod_{j \neq i} Y_{j,k}(s_j))$ whose support is the closure of $\prod_{j \neq i} Y_{j,k}(s_j)$, and define belief μ_i^k in $\Delta(S_{-i} \times T_{-i})$ as follows:

$$\mu_i^k[E] := \sum_{s_{-i} \in W_{-i}^{k-1}} \bar{\mu}_i^{k+1}[s_{-i}] \cdot \eta_i^k(s_{-i}) \left[E \cap \prod_{j \neq i} \{s_j\} \times Y_{j,k}(s_j) \right].$$

Obviously, μ_i^k is well-defined and its support is exactly the closure of $X_{-i,k} := \prod_{j \neq i} X_{j,k}$.[44] Notice in addition that since the marginal of μ_i^k on S_{-i} is precisely $\bar{\mu}_i^{k+1}$, we know that $\bar{s}_i$ is a best-reply to μ_i^k. Then, let $\bar{M}_i$ be the convex hull of $\{\mu_i^k\}_{k=0}^{n+1}$ and pick type $\bar{t}_i \in T_i$ such that $M_i(\bar{t}_i) = \bar{M}_i$. Clearly, the following two hold:

- $(\bar{s}_i, \bar{t}_i) \in C_i \cap CARC_{i,k}$ for any $k = 0, \ldots, n+1$. To see this, simply note that for any $k = 0, \ldots, n+1$, it holds that $\mu_i^k \in M_i(\bar{t}_i) = \bar{M}_i$. Then, the claim is proven since (as seen above) the support of μ_i^k is exactly the closure of $X_{-i,k}$.
- $(\bar{s}_i, \bar{t}_i) \in R_i$. This follows immediately from—as seen above—$\bar{s}_i$ being a best-reply to the conjecture induced by each belief in $\{\mu_i^k\}_{k=0}^{n+1}$ and thus, also to each belief in $M_i(\bar{t}_i)$.

Thus, it can be concluded that $(\bar{s}_i, \bar{t}_i)$ is a strategy-type pair in $R_i \cap C_i \cap CARC_{i,n+1}$ that induces $\bar{s}_i$. □

Theorem 1 *(Foundation of iterated admissibility). Let G be a game. For any player i the following holds:*

(i) *For any complete ambiguous type structure, any player i and any strategy-type pair (s_i, t_i), if type t_i is consistent with cautiousness and assumption of rationality and cautiousness and s_i is rational for t_i, then s_i is iteratively admissible; i.e.,*

$$\text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i) \subseteq W_i^\infty.$$

(ii) *For any player i and any strategy s_i, if s_i is iteratively admissible then there exist a complete ambiguous type structure $\mathcal{T}$ and a type t_i consistent with cautiousness and assumption of rationality and cautiousness for which s_i is rational; i.e.,*

$$W_i^\infty \subseteq \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i).$$

Proof. For the right-hand inclusion set strategy-type pair $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_i$ and simply notice that since $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_{i,n}$ for any $n \geq 0$, Theorem B.1 reveals that $\bar{s}_i \in W_i^n$ for any $n \geq 1$. Thus, $\bar{s}_i \in W_i^\infty$.

For the left-hand inclusion, set strategy $\bar{s}_i \in W_i^\infty$. Since, in particular, $\bar{s}_i \in W_i^{n+1}$ for any $n \geq 0$, it is known from Theorem B.1 that for any $n \geq 0$ there exists a type $t_i^n \in T_i$ such that $(\bar{s}_i, t_i^n) \in R_i \cap C_i \cap CARC_{i,n}$. Now, let $\bar{M}_i$ denote the closure of the convex-hull of $\bigcup_{n \geq 0} M_i(t_i^n)$ and pick type $\bar{t}_i \in T_i$ such that $M_i(t_i) = \bar{M}_i$. Obviously, $\bar{s}_i$ is a best-reply is to every conjecture induced by the beliefs in $M_i(\bar{t}_i)$ and $\bar{t}_i$ is cautious and is consistent with common assumption in rationality and cautiousness. Thus, $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_i$ and hence, $\bar{s}_i \in \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i)$. □

[44] For the latter, simply note that for any (s_{-i}, t_{-i}), $\mu_i^k[N] > 0$ for any neighborhood N of (s_{-i}, t_{-i}) if and only if $\eta_i^k(s_{-i})[N] > 0$ for any neighborhood N of (s_{-i}, t_{-i}).

For the characterization of self-admissible sets, the first thing needed is the simple observation that the reasoning process about *strategies only* stops after finitely many rounds.

Lemma B.1. *Let G be a game. Set an ambiguous type structure $\mathcal{T}$. There exists a $N \in \mathbb{N}$ such that for all $n \geq N$,*

$$\prod_{i\in I} \mathrm{Proj}_{S_i}\left(R_i \cap C_i \cap CARC_{i,n}\right) = \prod_{i\in I} \mathrm{Proj}_{S_i}\left(R_i \cap C_i \cap CARC_{i,N}\right).$$

Proof. By definition $CARC_{i,n+1} \subseteq CARC_{i,n}$, so that it also holds that $R_i \cap C_i \cap CARC_{i,n+1} \subseteq R_i \cap C_i \cap CARC_{i,n}$. Since S_i is finite there has to be an $N_i \in \mathbb{N}$ such that $n \geq N_i$

$$\mathrm{Proj}_{S_i}\left(R_i \cap C_i \cap CARC_{i,n}\right) = \mathrm{Proj}_{S_i}\left(R_i \cap C_i \cap CARC_{i,N_i}\right).$$

Take $N = \max_i N_i$. □

Theorem 2 *(Foundation of self-admissibility). Let G be a game. Then:*

(i) *For any ambiguous type structure $\mathcal{T}$ the set of strategies consistent with rationality, cautiousness and common assumption of rationality and cautiousness is a self-admissible set; i.e., the following set is self-admissible:*

$$\prod_{i\in I} \mathrm{Proj}_{S_i}(R_i \cap C_i \cap CARC_i).$$

(ii) *For any self-admissible set Q there exists a finite ambiguous type structure $\mathcal{T}$ for which Q characterizes the behavioral implications of rationality, cautiousness and common assumption of rationality and cautiousness; i.e., such that:*

$$\prod_{i\in I} \mathrm{Proj}_{S_i}(R_i \cap C_i \cap CARC_i) = Q.$$

Proof. For the first part set an ambiguous type structure $\mathcal{T}$ and consider

$$Q := \prod_{i\in I} \mathrm{Proj}_{S_i}\left(R_i \cap C_i \cap CARC_i\right).$$

If $Q = \emptyset$ then Q is a SAS. So assume it is non-empty. Set $s_i \in \mathrm{Proj}_{S_i}(R_i \cap C_i \cap CARC_i)$; then there exists a t_i such that $(s_i, t_i) \in R_i \cap C_i \cap CARC_i$. Thus $(s_i, t_i) \in R_i \cap C_i$ implies that condition (i) of SAS is satisfied. Furthermore, with N from Lemma B.1 and because $(s_i, t_i) \in CARC_i \subseteq CARC_{i,N+1}$ there must exist a $\mu_i \in M_i(t_i)$ such that $\mathrm{supp}\, \mu_i = \prod_{j\neq i} R_j \cap C_j \cap CARC_{j,N}$. Then, $\bar{\mu}_i := \mathrm{marg}_{S_{-i}} \mu_i$ is a conjecture with full-support on Q_{-i} (again using Lemma B.1) for which s_i is a best-reply. Hence, condition (ii) of SAS is satisfied. Lastly, consider mixed strategy σ_i such that $U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i)$ for every s_{-i}. Then, by Lemma D.2 of Brandenburger et al. (2008) $\mathrm{supp}\, \sigma_i \subseteq BR_i(\mathrm{marg}_{S_{-i}} \mu_i)$ for every $\mu_i \in M_i(t_i)$ giving $(r_i, t_i) \in R_i$ for all $r_i \in \mathrm{supp}\, \sigma_i$. Then it also holds that $(r_i, t_i) \in R_i \cap C_i \cap CARC_{i,n}$ for every $n \geq 1$, so that $r_i \in \mathrm{Proj}_{S_i}(R_i \cap C_i \cap CARC_i)$ and thus, condition (iii) of SAS is satisfied too.

For the second part set SAS Q. By definition of SAS (and Pearce, 1984), for each $s_i \in Q_i$, there exist a $\mu_i^1(s_i), \mu_i^2(s_i) \in \Delta(S_{-i})$ such that $\mathrm{supp}\, \mu_i^1(s_i) = S_{-i}$ and $\mathrm{supp}\, \mu_i^2(s_i) = Q_{-i}$. By Lemma D.4 of Brandenburger et al. (2008) we choose $\mu_i^1(s_i)$ so $r_i \in BR_i(\mu_i^1(s_i))$ if and only if $r_i \in \mathrm{supp}\, \sigma_i$ for a mixed strategy σ_i with $U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i)$ for every s_{-i}.

Now, consider the set of types $T_i := \{t_i(s_i) | s_i \in Q_i\} \cup \{\star_i\}$; to get an ambiguous type structure define $M_i(\star_i) \subseteq \Delta(S{-}i \times T_{-i})$ such that there is no $\eta_i \in M_i(\star_i)$ with $\mathrm{supp}\, \eta_i = S_{-i} \times T_{-i}$. For $s_i \in Q_i$, define,

$$\begin{aligned} Y_i(s_i) := \{(r_i, t_i(s_i)) : &\text{ either } r_i = s_i \text{ or} \\ &\exists \sigma_i \in \Delta(S_i), \text{ such that } r_i \in \mathrm{supp}\, \sigma_i \text{ and } U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i) \text{ for all } s_{-i} \in S_{-i}\} \end{aligned}$$

and then define two beliefs $\eta_i^1(s_i), \eta_i^2(s_i) \in \Delta(S_{-i} \times T_{-i})$ such that

$$\begin{aligned} &\mathrm{supp}\, \eta_i^1(s_i) = S_{-i} \times T_{-i} &&\text{and} &&\mathrm{marg}_{S_{-i}} \eta_i^1(s_i) = \mu_i^1(s_i), \\ &\mathrm{supp}\, \eta_i^2(s_i) = \prod_{j\neq i} \cup_{s_j \in Q_j} Y_j(s_j) \cap R_j &&\text{and} &&\mathrm{marg}_{S_{-i}} \eta_i^2(s_i) = \mu_i^2(s_i). \end{aligned}$$

To complete the description of the type structure, set $M_i(t_i(s_i))$ to be the convex hull of $\eta_i^1(s_i)$ and $\eta_i^2(s_i)$. Note that R_i only depends on the marginal beliefs on the strategies, so for $\eta_i^2(s_i)$ to be well-defined the following is required:

Claim 1. $\mathrm{Proj}_{S_i} \bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i = Q_i$. If $s_i \in Q_i$, then $(s_i, t_i(s_i)) \in Y_i(s_i)$ and by construction also $(s_i, t_i(s_i)) \in R_i$. Conversely, set $r_i \in \mathrm{Proj}_{S_i} \bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i$. So there exists a $s_i \in Q_i$ such that $(r_i, t_i(s_i)) \in Y_i(s_i) \cap R_i$. If $r_i = s_i$, the proof is complete. If not, then by property (iii) of self-admissible sets (and definition of $Y_i(s_i)$), we it also follows that $r_i \in Q_i$. ★

Next we prove that the type structure satisfies that:

$$Q = \prod_{i \in I} \mathrm{Proj}_{S_i} \left(R_i \cap C_i \cap CARC_i\right).$$

Claim 2. $\bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i = R_i \cap C_i$. Consider $(r_i, t_i) \in \bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i$. Then for some $s_i \in Q_i$ we have $\eta_i^1(s_i)$ and thus $t_i(s_i)$ is cautious. Conversely, for $(r_i, t_i) \in R_i \cap C_i$ it is needed that $t_i \neq \star_i$ since $\star_i$ is not cautious. Thus, there exists a $s_i \in Q_i$ such that $t_i = t_i(s_i)$. If $s_i = r_i$, the proof is complete. If not, then R_i requires that $r_i \in BR_i(\mu_i^1(s_i))$, which holds if and only if (see above) $r_i \in \mathrm{supp}\, \sigma_i$ for a mixed strategy σ_i with $U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i)$ for every s_{-i}. Thus, in either case it holds that $(r_i, t_i(s_i)) \in Y_i(s_i)$. ★

Claim 3. $\bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i = R_i \cap C_i \cap CARC_{i,1}$. $(r_i, t_i) \in \cup_{s_i \in Q_i} Y_i(s_i) \cap R_i$, that is $(r_i, t_i(s_i)) \in Y_i(s_i) \cap R_i$ for some $s_i \in Q_i$. Then, $(r_i, t_i(s_i)) \in CARC_{i,1}$ due to $\eta_i^2(s_i)$. The converse follows from Claim 2. ★

Induction concludes the proof. □

References

Ahn, David S., 2007. Hierarchies of ambiguous beliefs. J. Econ. Theory 136, 286–301.
Anscombe, Frank J., Aumann, Robert J., 1963. A definition of subjective probability. Ann. Math. Stat. 34, 199–205.
Aradillas-Lopez, Andres, Tamer, Elie, 2008. The identification power of equilibrium in simple games. J. Bus. Econ. Stat. 26, 261–310.
Aumann, Robert J., 1962. Utility theory without the completeness axiom. Econometrica 30, 445–462.
Aumann, Robert J., Brandenburger, Adam, 1995. Epistemic conditions for Nash equilibrium. Econometrica 63, 1161–1180.
Barelli, Paulo, Galanis, Spyros, 2013. Admissibility and event rationality. Games Econ. Behav. 77, 21–40.
Battigalli, Pierpaolo, Siniscalchi, Marciano, 1999. Hierarchies of conditional beliefs and interactive epistemology in dynamic games. J. Econ. Theory 88, 188–230.
Battigalli, Pierpaolo, Siniscalchi, Marciano, 2002. Strong belief and forward induction reasoning. J. Econ. Theory 106, 356–391.
Bergemann, Dirk, Morris, Stephen, 2009. Robust implementation in direct mechanisms. Rev. Econ. Stud. 76 (4), 1175–1204.
Bergemann, Dirk, Morris, Stephen, 2011. Robust implementation in general mechanisms. Games Econ. Behav. 71 (2), 261–281.
Bergemann, Dirk, Morris, Stephen, Tercieux, Olivier, 2011. Rationalizable implementation. J. Econ. Theory 146, 1253–1274.
Bewley, Truman F., 2002. Knightian decision theory. Part I. Decis. Econ. Finance 25, 79–110.
Blume, Lawrence, Brandenburger, Adam, Dekel, Eddie, 1991a. Lexicographic probabilities and choice under uncertainty. Econometrica 59, 61–79.
Blume, Lawrence, Brandenburger, Adam, Dekel, Eddie, 1991b. Lexicographic probabilities and equilibrium refinements. Econometrica 59, 81–98.
Brandenburger, Adam, 1992. Economic analysis of markets and games. In: Dasgupta, Partha, Gale, Douglas, Hart, Oliver, Maskin, Eric (Eds.), Chap.: Lexicographic Probabilities and Iterated Admissibility. MIT Press, pp. 282–290.
Brandenburger, Adam, 2003. Cognitive processes and economic behaviour. In: Dimitri, Nikola, Basili, Marcello, Gilboa, Itzhak (Eds.), Chap.: On the Existence of a 'Complete' Possibility Structure. Routledge, pp. 30–34.
Brandenburger, Adam, Friedenberg, Amanda, Keisler, H. Jerome, 2008. Admissibility in games. Econometrica 76, 307–352.
Brandenburger, Adam, Friedenberg, Amanda, 2010. Self-admissible sets. J. Econ. Theory 145 (2), 785–811.
Brandenburger, Adam, Dekel, Eddie, 1993. Hierarchies of beliefs and common knowledge. J. Econ. Theory 59, 189–198.
Catonini, Emiliano, De Vito, Nicodemo, 2018a. Cautious belief and iterated admissibility. Mimeo.
Catonini, Emiliano, De Vito, Nicodemo, 2018b. Weak belief and permissibility. Mimeo.
Cettolin, Elena, Riedl, Arno, 2019. Revealed preferences under uncertainty: incomplete preferences and preferences for randomization. J. Econ. Theory 181, 547–585.
Cramton, Peter, Kerr, Suzi, 2002. Tradeable carbon permit auctions: how and why to auction not grandfather. Energy Policy 30, 333–345.
Dekel, Eddie, Friedenberg, Amanda, Siniscalchi, Marciano, 2016. Lexicographic beliefs and assumption. J. Econ. Theory 163, 955–985.
Dekel, Eddie, Fudenberg, Drew, 1990. Rational behavior with payoff uncertainty. J. Econ. Theory 52, 243–267.
Dekel, Eddie, Siniscalchi, Marciano, 2015. Chapter 12 - Epistemic game theory. In: Young, H. Peyton, Zamir, Shmuel (Eds.), Handbook of Game Theory with Economic Applications, vol. 4. Elsevier, pp. 619–702.
Di Tillio, Alfredo, 2008. Subjective expected utility in games. Theor. Econ. 3, 323–387.
Epstein, Larry G., Wang, Tan, 1996. "Beliefs about beliefs" without probabilities. Econometrica 64, 1343–1373.
Feddersen, Timothy, Pesendorfer, Wolfgang, 1997. Voting behavior and information aggregation in elections with private information. Econometrica 65, 1029–1058.
Fishburn, Peter C., 1970. Utility Theory for Decision Making. John Wiley & Sons.
Friedenberg, Amanda, 2010. When do type structures contain all hierarchies of beliefs? Games Econ. Behav. 68 (1), 108–129.
Gilboa, Itzhak, Maccheroni, Fabio, Marinacci, Massimo, Schmeidler, David, 2010. Objective and subjective rationality in a multiple prior model. Econometrica 78, 755–770.
Harsanyi, John C., 1967–1968. Games with incomplete information played by 'Bayesian' players, I–III. Manag. Sci. 14, 159–182. 320–334, 486–502.
Heifetz, Aviad, Meier, Martin, Schipper, Burkhard C., 2019. Comprehensive rationalizability. Games Econ. Behav. 116, 185–202.
Keisler, Jerome, H., Byong Soo, Lee, 2015. Common assumption of rationality. Mimeo.
Keynes, John M., 1936. The General Theory of Employment, Interest and Money. Harcourt Brace and Co.
Kohlberg, Eilon, Mertens, Jean-François, 1986. On the strategic stability of equilibria. Econometrica 54, 1003–1037.
Lee, Byung Soo, 2016. Admissibility and assumption. J. Econ. Theory 163, 42–72.
Lopomo, Giuseppe, Rigotti, Luca, Shannon, Chris, 2011. Knightian uncertainty and moral hazard. J. Econ. Theory 146, 1148–1172.
Lopomo, Giuseppe, Rigotti, Luca, Shannon, Chris, 2014. Uncertainty in mechanism design. Mimeo.

Luce, R. Duncan, Raiffa, Howard, 1957. Games and Decisions: Introduction and Critical Survey. Dover Books on Mathematics.
Mertens, Jean-François, Zamir, Shmuel, 1985. Formulation of Bayesian analysis for games with incomplete information. Int. J. Game Theory 14, 1–29.
Milgrom, Paul, 1998. Game theory and the spectrum auctions. Eur. Econ. Rev. 42, 771–778.
Moulin, Herve, 1979. Dominance solvable voting schemes. Econometrica 47, 1337–1351.
Palfrey, Thomas R., Srivastava, Sanjay, 1991. Nash implementation using undominated strategies. Econometrica 59, 479–501.
Pearce, David G., 1984. Rationalizable strategic behavior and the problem of perfection. Econometrica 52, 1029–1050.
Renyi, Alfred, 1955. On a new axiomatic theory of probability. Acta Math. Hung. 6, 285–335.
Samuelson, Larry, 1992. Dominated strategies and common knowledge. Games Econ. Behav. 4, 284–313.
Sobel, Joel, 2017. A note on pre-play communication. Games Econ. Behav. 102, 477–486.
Sobel, Joel, 2019. Iterated weak dominance and interval-dominance supermodular games. Theor. Econ. 14 (1), 71–102.
Stahl, Dale O., 1995. Lexicographic rationalizability and iterated admissibility. Econ. Lett. 47 (2), 155–159.
Stauber, Ronald, 2011. Knightian games and robustness to ambiguity. J. Econ. Theory 146, 248–274.
Stauber, Ronald, 2014. A framework for robustness to ambiguity of higher-order beliefs. Int. J. Game Theory 43, 525–550.
Tan, Tommy Chin-Chiu, da Costa Werlang, Sérgio Ribeiro, 1988. The Bayesian foundations of solution concepts of games. J. Econ. Theory 45, 370–391.
Varian, Hal R., 2007. Position auctions. Int. J. Ind. Organ. 25, 1163–1178.
Yang, Chih-Chun, 2015. Weak assumption and iterated admissibility. J. Econ. Theory 158, 87–101.
Ziegler, Gabriel, Zuazo-Garin, Peio, 2019. Strategic knightian uncertainty. Mimeo.