Strategic Cautiousness as an Expression of Robustness to Ambiguity

Gabriel Ziegler∗ Peio Zuazo-Garin†

October 3, 2019

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.

Keywords: Game theory, decision theory, ambiguity, Knightian uncertainty, incomplete preferences, Bayesian rationality, cautiousness, iterated admissibility.

JEL Classification: C72, D82.

1 Introduction

Economists commonly use iterated strategy elimination procedures as solution concepts in games. Therefore, such procedures 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; still, 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

∗Northwestern University, Department of Economics; 2211 Campus Drive, Evanston, IL 60208, USA; ziegler@u.northwestern.edu.
†University of the Basque Country (UPV/EHU), Department of Foundations of Economic Analysis; Aenida Lehendakari Agirre 83, 48015, Bilbao, Spain; peio.zuazo@ehu.eus.
players or the economic analyst lack accurate data about past play or such data appears uninformative about future behavior.\textsuperscript{1} For instance, this is the case in many application of auction theory, e.g., wireless spectrum, carbon emission rights or online advertising.\textsuperscript{2} Consequently, thorough understanding of the forces behind iterated elimination is both relevant from a purely theoretical perspective and, from a more applied point of view, key to effective mechanism design and correct identification in empirical analyses.\textsuperscript{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.\textsuperscript{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 $i$’s opponents’ strategies, while cautiousness asks for player $i$’s decision to be sensitive to these 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,\textsuperscript{5} the seemingly mutually exclusive nature of strategic reasoning and cautiousness requires clarification. In particular, in scenarios where behavior is likely to be reasoning-based and cautiousness plays some role, this understanding is desirable.

This paper proposes a new take on this longstanding problem by suggesting a novel

\textsuperscript{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.”

\textsuperscript{2}See Milgrom (1998), Cramton and Kerr (2002) and Varian (2007), respectively.

\textsuperscript{3}See Bergemann and Morris (2009, 2011); Bergemann, Morris and Tercieux (2011) and Aradillas-Lopez and Tamer (2008), respectively.

\textsuperscript{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 her mark-up she should absorb all the demand and increases her profit; now, this is easy to forecast by $B$, who might in turn decide to lower her mark-up more than slightly and this way, absorb herself 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.”

theoretical foundation for the interplay between strategic reasoning and cautiousness. The analysis of Samuelson (1992) makes transparent 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 of 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 and 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.\footnote{Due to incompleteness, for given set of beliefs such a strategy might not exist. In this case, we will also say that the player is not rational.}

Under this set-up and inspired by Brandenburger, Friedenberg and Keisler (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 exactly allows 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 as 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. In case 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 of the rounds. Theorem 2 provides the analysis corresponding to the relaxation of belief-completeness and shows that, in this case, it is self-admissible sets à la Brandenburger, Friedenberg and Keisler (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 itself besides its interpretation as a reflection of ambiguity.\footnote{As argued by Aumann (1962): “Of all the axioms of utility theory, the completeness axiom is perhaps the most questionable. [...] We 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, Rigotti and Shannon (2011, 2014), who study mechanism design and optimal contracting, respectively.}

The literature studying the conflict between strategic reasoning and cautiousness is epitomized by the seminal work by Brandenburger, Friedenberg and Keisler (2008), who shed light into the question by building upon the lexicographic probability system approach by Blume, Brandenburger and Dekel (1991a,b).\footnote{Early contributions along the same lines include, for example, Brandenburger (1992) and Stahl (1995).} 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, Brandenburger, Friedenberg and Keisler (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.\footnote{Similar to Epstein and Wang (1996), coherency is imposed on the preferences directly, not only on the beliefs that represent the preferences.} Catonini and De Vito (2018a) also provide foundations by introducing a weaker notion of the likelihood-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, Meier and Schipper (2019) propose a new solution concept, \textit{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 than 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 of these relaxations allow for multiple beliefs, but while the for-
mer 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 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 we rely on are natural extensions of their counterparts in the realm of standard Bayesian preferences, and (ii) our notion of cautiousness we invoke in our theorems does not necessarily restrict itself to environments where the sets of states have a specific structure (e.g., games). 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. Next, we present in Section 5 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 payof matrix:\[11\]

\[
\begin{array}{c|cc}
  & L & R \\
\hline 
T & 1 & 1 \\
D & 0 & 2 & 1 & 0 \\
\end{array}
\]

\[10\] Though they are sensitive to topological specifications.

\[11\] This is the leading example of Brandenburger, Friedenberg and Keisler (2008) and was introduced by Samuelson (1992).
Clearly, no action is strictly dominated for either player and therefore (standard) rationalizability predicts $\{T, D\} \times \{L, R\}$. However, $R$ is weakly dominated by $L$. Deleting $R$ for this reason will make $D$ strictly dominated in the reduced game. Thus, iterated admissibility has a unique prediction in this game: $(T, L)$.

Now, suppose we want 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 $L$ and on Bob playing $R$. However, if Ann believes that Bob is rational and cautious, then she should rule out that Bob is 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, which play the following game:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ann</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>M</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>L</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Now, suppose that each player faces ambiguity (as described by Bewley, 2002) about the strategy choice of their opponent. That is, each player does not have a unique belief about the opponent’s strategy choice, but rather 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, $\mu^1_A(s_B) = 1/4$ for $s_B = A, B, C, D$, and her second (extreme) belief is uniform across $A, B$, and $C$ only, $\mu^2_A(s_B) = 1/3$ for $s_B = A, B, C$. Similarly, Bob faces uncertainty about Ann’s choice. Consider the following set of beliefs for Bob, which has two extreme beliefs too. The first is uniform across all of Ann’s strategies, $\mu^1_B(s_A) = 1/3$ for $s_A = H, M, L$, and the second belief assigns equal probability to $H$ and $M$, $\mu^2_B(H) = \mu^2_B(M) = 1/2$.

Let us check what strategies are rational for each player given their beliefs. However, preferences à la Bewley (2002) are incomplete, and for incomplete preferences rationality does not have an obvious definition: 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 rationally not choose L since it is not a best-reply that is robust to the ambiguity 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 characterization results of iterated admissibility, which will be explained in more detail next, which also makes the strategic reasoning explicit.

2. Non-technical overview

2.2 Heuristic treatment of strategic reasoning

In the previous examples we saw that a set of beliefs allows to incorporate strategic reasoning and cautiousness. To study games in general, we need to allow for players to reason about the reasoning process of other players as well. For this we need to formalize infinite sequences of the following form:

\[ a_1: \text{Ann is rational and cautious} \]
\[ a_2: a_1 \text{ holds and Ann assumes } b_1 \]
\[ a_3: a_1 \text{ holds and Ann assumes } b_1 \& b_2 \]
\[ \ldots \]

\[ b_1: \text{Bob is rational and cautious} \]
\[ b_2: b_1 \text{ holds and Bob assumes } a_1 \]
\[ b_3: b_1 \text{ holds and Bob assumes } a_1 \& a_2 \]
\[ \ldots \]

If this infinite sequence holds, we say there is \textit{rationality, cautiousness, and common assumption thereof} (RCCARC).

To study these infinite sequences and to see which strategies are played if this sequences hold we need to introduce (epistemic) types for each player. For this 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, as we want to model players who face ambiguity, we associate with each type a (closed, and convex) set of beliefs about \(S_B \times T_B\). Then, for a strategy-type pair of Ann \((s_A, t_A)\) we can say \(s_A\) is rational if \(s_A\) is a best-reply to all of the beliefs associated with \(t_A\). If a player is cautious depends only on her beliefs, because it requires that 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 \(t_A\) which has only a singleton set of beliefs \(\{\mu_A\}\) with support as depicted in Figure 1. For such a cautious type, we can ask which strategies are
rational. For this, consider the marginal of $\mu_A$ 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, we can study the infinite sequences described above. In this case we get a picture that looks like Figure 2. Here the small area with solid boundary corresponds to all strategy-type combinations of Bob satisfying RCCARC. Now, fix a strategy-type combination $(s_A, t_A)$ for Ann. Does this type correspond to RCCARC for Ann, that is does the type satisfy the sequence $a_1, a_2, \ldots$? We know already if $a_1$ holds there needs to be a belief in the associated set of beliefs which has support as $\mu^1_A$. Next, we want to consider that Ann assumes $b_1$. This rules out some of Bob’s strategy-type pairs, but also requires that $t_A$ has a belief which has full support on the remaining pairs. Thus, in the associated set of beliefs there needs to a belief $\mu^2_A$. In the next step, we consider Ann to assume $b_1$ and $b_2$. A similar reason applies and there needs to be a belief like $\mu^3_A$ in the set of beliefs corresponding to $t_A$. This procedure can now be iterated (as indicated in the picture) to verify if the type $t_A$
corresponds to RCCARC for Ann. As we are only considering finite games, at some stage \( n \) this iteration would not rule out any strategies for Bob anymore. However, it might be the case that at every step there are still some types of Bob 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 along 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, Friedenberg and Keisler (2008) found in a different framework. To illustrate how this analysis works more concretely, we add types explicitly for the example considered above.

Example 2 (continuing from p. 6). 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 \( S_B \times T_B \):

\[
\mu_A(s_B,t_B) = 1/16, \text{ for all } (s_B,t_B) \in S_B \times T_B,
\]

\[
\mu_A^2(s_B,t_B) = 1/9, \text{ for all } (s_B,t_B) \in \{A,B,C\} \times \{t_B^1, t_B^2, t_B^3\}, \text{ and}
\]

\[
\mu_A^3(s_B,t_B) = 1/3, \text{ for all } s_B \in \{A,B,C\}.
\]

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

\[
M_i(t_i^0) = \{\mu_i^0\}, \ M_i(t_i^1) = \{\mu_i^1\}, \ M_i(t_i^2) \text{ is the convex hull of } \mu_i^1 \text{ and } \mu_i^2, \text{ and } M_i(t_i^3) \text{ is the convex hull of } \mu_i^1, \mu_i^2, \text{ and } \mu_i^3.
\]

Now, we can 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 need to 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 that 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 \( t_A^3 \) and \( t_A^3 \)

\(^{12}\)That is, not finitely generated sets.
satisfy this requirement and we are left with \( \{H, M\} \times \{t^2_A, t^3_A\} \) corresponding to \( a_2 \). For Bob, we get that \( b_2 \) corresponds to \( \{A, B, C\} \times \{t^2_B, t^3_B\} \). Again, note that in this step the interactive reasoning leads to ruling out \( L \), which is weakly dominated after elimination of \( D \). In the next step (i.e. \( a_3 \) and \( b_3 \)), types \( t^2_i \) will be 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 \( t^3_i \)). This gives an infinite (but countable) number of types and only the “limiting” type corresponds to RCCARC. This argument shows that the illustration in Figure 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 we get the strategies of iterated admissibility (or for any other self-admissible set).

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 to the iterated elimination of weakly dominated strategies raised by Samuelson (1992). Thus, in Section 3.1 we recall the formalization of strategic-form games, iterated admissibility (Luce and Raiffa, 1957; Moulin, 1979) and self-admissible sets (Brandenburger, Friedenberg and Keisler, 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, Maccheroni, Marinacci and Schmeidler (2010), and highlights its key features.\(^{13}\)

3.1 Games and iterated strategy elimination

A game consists of a tuple \( G := (I, (S_i, u_i))_{i \in I} \) where \( I \) is a finite set of players, and for each player \( i \) we have 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 \( i \) we refer to a randomization of own strategies \( \sigma_i \in \Delta(S_i) \) as a mixed strategy,\(^{14}\) and to a probability

\(^{13}\)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 \( \sigma \)-algebra of \( X \).
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 puts probability one on \( s_i \). Each conjecture \( \mu_i \) and possibly mixed strategy \( \sigma_i \) 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 \( \mu_i \) the subset of pure strategies \( BR_i(\mu_i) \) that maximize its corresponding expected utility.\(^{15}\)

Due to the duality results of Pearce (1984), we use the best-reply correspondence directly to define iterated admissibility whose foundations we later study 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 \( W_{1-i}^1 \times W_i^1 \) 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 \( n \in \mathbb{N} \),

\[
W_i^n := \begin{cases} 
\{ s_i \in W_{i-1}^{n-1} \mid \text{There exists some } \mu_i \in \Delta(S_{-i}) \text{ such that:} \\
(i) \supp \mu_i = \prod_{j \neq i} W_{j}^{n-1}, \\
(ii) s_i \in BR_i(\mu_i) \}
\end{cases}
\]

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

(i) Every \( s_i \in Q_i \) is not weakly dominated given \( S_{-i} \times S_i \).

(ii) Every \( s_i \in Q_i \) is not weakly dominated given \( Q_{-i} \times S_i \).

(iii) For every \( s_i \in Q_i \) and every mixed strategy \( \sigma_i \) 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 immediate: 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 Anscome and Aumann’s (1963) framework due to Fishburn (1970). The decision maker faces decision environment \( (Z, \Theta) \) where: (i) \( Z \) is a set of outcomes, which can be informally understood as the elements that will ultimately yield direct

\(^{15}\)That is, given conjecture \( \mu_i \), we have expected utility \( U_i(\mu_i; \sigma_i) := \sum_{(s_{-i}, s_i) \in S} \mu_i[s_{-i}] \cdot \sigma_i[s_i] \cdot u_i(s_{-i}; s_i) \) for each possibly mixed strategy \( \sigma_i \), and set of best-replies \( BR_i(\mu_i) := \arg \max_{s_i \in S_i} U_i(\mu_i; s_i) \).
utility to the decision maker, and (ii) $\Theta$ is a set of states (of the world) the decision maker might face uncertainty about, 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. By $\mathcal{M}(\Theta)$ we denote the set of closed and convex nonempty subsets of $\Theta$.\textsuperscript{16} Throughout the paper we focus on preferences which we call Bewley preferences, since they were introduced by Bewley (2002).\textsuperscript{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 dropping completeness of the preferences. Theorem 1 by Gilboa, Maccheroni, Marinacci and Schmeidler (2010) provides the following convenient representation for these preferences:\textsuperscript{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 $M \in \mathcal{M}(\Theta)$ such that for every pair of acts $f, g$,\textsuperscript{19}

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

A decision maker’s epistemic attitude with respect to the source of uncertain may not be represented by a single belief, as in the standard case, but instead, by a possibly non-singleton set of beliefs that reflects the decision maker’s possible ambiguity towards said source of uncertainty. As argued extensively in Section 5 this is key to resolve 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 at least as good as some other act $g$, it must be the case that the expected utility for $f$ is 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 that allow for testing whether the preferences display this form of ambiguity via incompleteness.

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

\textsuperscript{17}Actually we rely on a more modern version by Gilboa, Maccheroni, Marinacci and Schmeidler (2010) of Bewley’s (2002) original preferences. The latter require the decision-maker to have a designated default act always chosen unless ranked strictly lower than some alternative. This is commonly known in literature as inertia (see, Bewley, 2002, or Lopomo, Rigotti and Shannon, 2011). Furthermore, the version of Gilboa, Maccheroni, Marinacci and Schmeidler (2010) allows for infinite state spaces which are necessary in our framework.

\textsuperscript{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.

\textsuperscript{19}More precisely, $M$ is non-empty closed and convex. Moreover, $M$ is unique and $u$ is unique up to positive affine transformations.
4. Reasoning in games

For the remainder 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 we later employ for establishing foundations for iterated admissibility in Section 5. Formally, for each player we specify a choice and a representation of her beliefs over her opponents’ strategies, her beliefs over her opponents’ beliefs over their opponents’ strategies, etc. These elements suffice to assess whether under such specification, the player is being rational, has preferences that exhibit ambiguity (i.e. multiple beliefs) or has certain higher-order beliefs over her opponents’ rationality and 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.

Some previous methodological work is required though. 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 the analysis of 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. Thus, first, in Section 4.1 we introduce these ambiguous type structures. We build on the latter to define next, in Section 4.2, the restrictions on behavior and beliefs necessary for the results in Section 5.

4.1 Ambiguous type structures

The study of strategic reasoning requires some 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 be easily represented through type structures. Thus, it turns out 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 $i$ we have:

1. A set of (ambiguous) types $T_i$.
2. An ambiguous belief map $M_i : T_i \rightarrow \mathcal{M}(S_{-i} \times T_{-i})$, where $T_{-i} := \prod_{j \neq i} T_j$, that associates each type with some ambiguous beliefs over opponents’ strategy-type pairs.

---

20 These type structures can be 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 will assume each $T_i$ to be compact and metrizable and each $M_i$, continuous. See Footnote 40.
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, by recursive marginalization:

(1) First-order ambiguous beliefs that represent type $t_i$’s uncertainty over her opponents’ strategies, $M_{i,1}(t_i) \in \mathcal{M}_{i,1} := \mathcal{M}_i(S_{-i})$, 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 over her opponents’ strategy-first-order ambiguous beliefs pairs, $M_{i,2}(t_i) \in \mathcal{M}_{i,2} := \mathcal{M}_i(\prod_{j \neq i} (S_j \times \mathcal{M}_{j,1}))$.

\[
\vdots
\]

(n) $n^{th}$-order ambiguous beliefs that represent type $t_i$’s uncertainty over her opponents’ strategy-$(n-1)^{th}$-order ambiguous beliefs pairs, $M_{i,n}(t_i) \in \mathcal{M}_{i,n} := \mathcal{M}_i(\prod_{j \neq i} (S_j \times \mathcal{M}_{j,n-1}))$.

\[
\vdots
\]

Ambiguous type structure $\mathcal{T}$ is said to be complete if every map $M_i$ is onto, 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.\textsuperscript{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 over her opponents’ choices, her ambiguous beliefs over her opponents’ first-order ambiguous beliefs, etc. Thus, each strategy-type pair $(s_i, t_i)$ allows for addressing questions such as: Is player $i$ rational given her beliefs? Do her preferences embody some kind of ambiguity? Which are her higher-order beliefs about her opponents’ rationality and ambiguity? Next, we first

\textsuperscript{22}The conceptual simplicity that follows is in contrast with the notational complexity it requires; technically, for each $n \in \mathbb{N}$ we have:

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

\textsuperscript{23}As shown by Ahn (2007), the answers to the following modified questions in (Dekel and Siniscalchi, 2015, p. 629): “Is there an ambiguous type structure that generates all ambiguous hierarchies of beliefs? Is there an 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.
formalize the notion of rationality we employ (Section 4.2.1). Second, we introduce our formalization of cautiousness as a manifestation of ambiguity (Section 4.2.2). Finally, 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:

$$R_i := \left\{ (s_i, t_i) \in S_i \times T_i \mid s_i \in \bigcap_{\mu_i \in M_i(t_i)} \text{BR}_i(\text{marg}_{S_{-i}} \mu_i) \right\}.$$ \hspace{1cm}

Note that the definition implicitly requires that each type $t_i$, in order to be eligible for rational behavior, satisfies that the intersection of the best-replies to the ambiguous first-order beliefs induced by it is non empty.\footnote{We discuss this requirement in detail in Appendix A, where we first disentangle the condition that ensures non-emptiness of the intersection of best replies (decisiveness) from rationality per se, and provide its behavioral characterization.} 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.\footnote{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 in Section A in the appendix, which also provides a behavioral foundation for a non-empty intersection.}

4.2.2 Cautiousness and ambiguity

We argue next that cautiousness, intuitively thought of 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. First, we formalize this notion of cautiousness that take part in the characterizations result in Section 5 and then discuss its relation with ambiguity.\footnote{We note that all of the following analysis 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 $S_{-i}$ instead than $S_{-i} \times T_{-i}$. Our reason for opting for the stronger notion is twofold: (i) it does not preclude our characterization from dispensing with impossibility issues à la Brandenburger, Friedenberg and Keisler (2008) (see Section 5.3), so that it is clear that it is not modifications in the notion of cautiousness what allow for achieving this, and (ii) since it does not apply only to state spaces with product structure, it has a more general decision-theoretic foundation.}
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 we think of cautiousness as the idea that a decision maker takes every possible contingency into account, then this is present in this definition. Cautiousness requires, loosely speaking, that every state is taken into account by the decision maker. The relation with ambiguity is easy to see. In principle, it is possible that a type displays 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 Figure 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, not only does the introduction of ambiguity allow for making strategic reasoning and cautiousness compatible, but indeed, it is necessary when strategic reasoning has any bite.

4.2.3 Assumption

Henceforth, 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 $E$. Some changes are in order if we want to extend this idea 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 support 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 $A_i(E_{-i})$.

---

27 Cautiousness is also present in the analysis of Brandenburger, Friedenberg and Keisler (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, Friedenberg and Keisler, 2008 and Dekel, Friedenberg and Siniscalchi, 2016) under the lexicographic probability system when the preferences satisfy continuity and thus, the corresponding lexicographic probability system collapses to a single belief.
5. Iterated admissibility and ambiguous types

In this section we present 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 such foundation in Section 5.1. Next, in Section 5.2, we discuss the relation 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, Friedenberg and Keisler (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 going 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 $\text{CARC}_i := \bigcap_{n \geq 0} \text{CARC}_{i,n}$, where each $\text{CARC}_{i,n}$ is defined recursively by setting:

$$\text{CARC}_{i,0} := S_i \times T_i,$$

$$\text{CARC}_{i,n} := \text{CARC}_{i,n-1} \cap A_i \left( \prod_{j \neq i} R_j \cap C_j \cap \text{CARC}_{j,n-1} \right),$$

for every $n \in \mathbb{N}$. That is, $\text{CARC}_i$ collects all those 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. Then:

Theorem 1 (Foundation of iterated admissibility). Let $G$ be a game. For any player $i$ it holds that:

(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

---

Remark 1. Cautiousness as defined in Definition 1 can be restated in terms of assumption: a type $t_i$ is cautious if it assumes $S_{-i} \times T_{-i}$.
cautiousness and \( s_i \) is rational for \( t_i \), then \( s_i \) is iteratively admissible; i.e.,

\[
\text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i) \subseteq W^\infty_i.
\]

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

\[
W^\infty_i \subseteq \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i).
\]

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 some type that is consistent with common assumption in rationality and cautiousness. The proof of the theorem is performed 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.\(^3\) 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 ambiguity. 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 \( T \) 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:

\[
\prod_{i \in I} \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i).
\]

\(^3\)This statement is properly formalized in Theorem B1, in Section B. For expositional reasons, we opted for presenting here the result corresponding to only the full iteration process, so that the comparison of the result for iterated admissibility and the result for self-admissibility (Theorem 2 below) is more straightforward.
(ii) For any self-admissible set $Q$ there exists some finite ambiguous type structure $T$ for which $Q$ characterizes the behavioral implications of rationality, cautiousness and common assumption of rationality and cautiousness; i.e., such that:

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

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 some ambiguous type structure $T$, 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 $T$. 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 visible in the proof of Theorem 2, each player $i$ only has as many types as 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 like the assumption operator of Brandenburger, Friedenberg and Keisler (2008), is the failure of monotonicity.\(^{32}\) Whenever a Bayesian agent believes some event $E$, she also believes every event $F$ such that $E \subseteq F$: the (Bayesian) belief $\mu_i$ 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 some $F$ such that $E \subseteq F$;\(^ {33}\) even if $t_i$ assumed such $F$, certainly, it could not be via $\mu_i$. Thus, when considering a sequence of nested events such as the finite iterations in the common assumption events defined above, one 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 can be resolved within our framework, but raises a the problem in the first place.

\(^{31}\)The proof of Theorem 2 proceeds in a very similar way as Brandenburger, Friedenberg and Keisler (2008) prove 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 $E$. 
within a Bayesian framework.

In principle there is no reason to consider that the assumption of some event is an expression of cautiousness; for every type there always exists some event that is assumed and this simply relates to which are the specific states that play some role in the way the preference ranks acts. However, the assumption of different nested events is a non-trivial feature that reveals cautious attitude: whenever a type assumes two nested events \( E \) and \( F \), the represented preference is crucially sensitive to comparisons at every state in \( E \) but in addition, also to comparisons at every state in larger event \( F \), in particular to those outside \( E \). Of course, as mentioned before, 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, Friedenberg and Keisler (2008). Their seminal result shows that \( m \) rounds of elimination of non-admissible strategies characterize the behavioral implications of rationality and \( m^{th} \)-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, Brandenburger and Dekel (1991a), lexicographic probability systems arise under a variation of Anscombe and Aumann’s (1963) preferences in which the axiom of continuity is relaxed (instead of the completeness one, as in Bewley’s (2002) variant). However, Brandenburger, Friedenberg and Keisler (2008) also reveal a vexatious feature of the common assumption case: their celebrated impossibility result shows that for 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 fix a belief \( \mu_i^n \in \Delta(S_{-i} \times T_{-i}) \) that has full-support in the topological closure of \( \prod_{j \neq i} R_j \cap CARC_{j,n} \) (this collections of strategy-type pairs are clearly never empty; thus, such belief \( \mu_i^n \) 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 \( M_i \). By construction, \( t_i \) is a type representing common assumption of rationality.

\[\text{As shown by Ahn (2007), this assignment can take place in an ambiguous type structure that maps types to ambiguous beliefs continuously.}\]
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 works 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, Friedenberg and Keisler (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 assumption in Brandenburger, Friedenberg and Keisler (2008) and Dekel, Friedenberg and Siniscalchi (2016),\footnote{See also Footnote 28.} and the coherence requirement implicit in our type structures resembles the standard one in literature due to Brandenburger and Dekel (1993).\footnote{The requirement is explicit in the construction of Ahn (2007).} 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 some 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 on the set of states.

To end the section, let us present some comparison between lexicographic probability systems and ambiguous beliefs that provides some understanding on the differences between both approaches with respect to the presence of ambiguity. Remember that a lexicographic probability system consists of a finite sequence beliefs $\{\mu^k\}_{k=1}^n \subseteq \Delta(\Theta)$,\footnote{Brandenburger, Friedenberg and Keisler (2008) use lexicographic conditional probability systems, but their result extends to more general lexicographic probability systems as shown by Dekel, Friedenberg and Siniscalchi (2016).} where the order in the sequence represents the epistemic priority attached to each element: $\mu^1$ is the decision maker’s ‘primary’ hypothesis, $\mu^2$ 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 $\mu^1$, then the comparison between both acts for the rest of the beliefs in the sequence is immaterial and the decision maker prefers $f$ over $g$. The main distinction between lexicographic probability systems and ambiguous beliefs is then clear: despite both being composed of multiple beliefs, 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 totally 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 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 characterizes the behavioral implications of said event (Theorem 2).

Thus, the main insight is immediate: 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 allows 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, Friedenberg and Keisler (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 in 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 easily visible in Dekel and Fudenberg’s (1990) procedure (the \textit{DF-procedure}; \textit{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 \textit{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 \textit{exclude} some strategies). Hence, the presence of inclusion-exclusion issues makes understanding of 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.\textsuperscript{38} In addition, it is immediate to replicate, within this framework, the well-known result that rationalizability (the iterated elimination of strictly dominated strategies) characterizes rationality and common belief thereof. The 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: while in the absence of cautiousness (i.e., rationalizability) behavior can be explained without the appeal to ambiguous beliefs, the latter becomes sine qua non as soon as the solution concept relies on any notion of cautiousness (i.e, iterated admissibility and the DF-procedure).

\textsuperscript{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.
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 Inarra, 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 specially 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 ERC Programme (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 immense hospitality. The authors are responsible for all errors.
A Decision Theory

A.1 Decision problems and Bewley preferences

Given a decision environment \((Z, \Theta)\), a decision problem consists of a triplet \((Z, \Theta, F)\) where \(F\) is a subset of acts we call feasible and that 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 \([0, 1]\):

\[
\{ \lambda \in [0, 1] \mid \lambda f + (1 - \lambda)g \succsim h \} \quad \text{and} \quad \{ \lambda \in [0, 1] \mid h \succsim \lambda f + (1 - \lambda)g \}.
\]

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 \(\alpha \in (0, 1)\) we have,

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

Theorem 1 of Gilboa, Maccheroni, Marinacci and Schmeidler (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 \((Z_i, \Theta_i, F_i)\) consisting of:

- Outcomes. In contexts of complete (payoff-relevant) information, player \(i\)'s utility only depends on the strategy profiles chosen in the game; hence, we identify outcomes with the latter: \(Z_i := S\).

\(^{39}\)That is, the decision maker may have preference on elements of not only \(F\), but \(\mathcal{F}\), representing that she might have preferences on options that are not materially available in the problem under study.
- **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 some additional non payoff-relevant unobserved parameter she might face uncertainty about, say $T_{-i}$.\footnote{To ensure appropriate construction, $T_{-i}$ will be assumed to be compact and metrizable.} We identify the set of states of the world with this joint sources of uncertainty: $\Theta_i := S_{-i} \times T_{-i}$.

- **Acts and feasible acts.** Player $i$’s set of acts is $F_i := \Delta(S_{-i}) \times T_{-i}$. Notice that within the context of a game this set of acts is not feasible. First, player $i$ cannot make her choice contingent on parameter $t_{-i}$ 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:

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

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, Maccheroni, Marinacci and Schmeidler (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 some ambiguous beliefs $M_i \subseteq \Delta(S_{-i} \times T_{-i})$.\footnote{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 $T_{-i}$.} Thus, there is no loss of generality in switching the focus from Bewley preferences to ambiguous beliefs, the collection of which we denote by $\mathcal{M}_i(S_{-i} \times T_{-i})$.

### 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 least preferred than some other act.\footnote{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.}
Proposition 1 (Behavioral foundation of decisiveness). Let $G$ be a game and $T$, an ambiguous type structure. Then, any player $i$’s type $t_i$ is decisive if and only if there exists some subset of feasible acts $F_i^* \subseteq F_i$, such that $\succ_i$, 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 $\succ_i$ denote the Bewley preference represented by $(u_i, M_i(t_i))$. The ‘if’ part is immediate, so we focus on the ‘only if’ one. To see it simply take $S_i^* := \bigcap_{\mu_i \in M_i(t_i)} BR_i(\text{marg}_{S_i \setminus i} \mu_i)$ and set:

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

Clearly, $F_i^* \subseteq F$ 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, unlike under completeness, two different concepts: an optimal act is always undominated but an undominated act might not be optimal; furthermore, while every Bewley preference admits undominated acts, optimal ones might not exist. Decisiveness ensures the existence of the latter, which in turn, restores the equivalence of undomination and optimality. In consequence, imposing decisiveness to incomplete preferences is similar in spirit to the requirement of Bayesian updating for conditional probability systems in the literature of extensive-form games. As decisiveness, Bayesian updating guarantees the existence of optimal strategies by forcing them to be equivalent to undominated ones.

**B Characterization result**

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

---

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.
Theorem B1. Let $G$ be a game and $\mathcal{T}$ a complete ambiguous type structure. For any $n \in \mathbb{N}$ and every player $i$ it holds that:

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

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

Initial Step ($n = 0$). For the right-hand inclusion, fix strategy-type pair $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i$ and denote $\tilde{M}_i = M_i(\bar{t}_i)$. Then, since $\bar{t}_i$ is cautious we know that there exists some belief $\mu_1^1 \in \tilde{M}_i$ 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 $\text{marg}_{S_{-i}}\mu_1^1$. Thus, $\bar{\mu}_i^1 := \text{marg}_{S_{-i}}\mu_1^1$ is a conjecture in with full-support on $S_{-i}$ for which $\bar{s}_i$ is a best-reply. Hence, $\bar{s}_i \in W_1^1$.

For the left-hand inclusion, fix strategy $\bar{s}_i \in W_1^1$ 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 $\mu_1^1 := \bar{\mu}_i \times \eta_i$ and $\tilde{M}_i := \{\mu_1^1\}$. Since $\mathcal{T}$ is complete, we know that there exists some type $\bar{t}_i \in T_i$ such that $M_i(\bar{t}_i) = \tilde{M}_i$. Since $\mu_1^1$ has full-support in $S_{-i} \times T_{-i}$ we know that $\bar{t}_i$ is cautious, 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 $M_i(\bar{t}_i)$ we have that $(\bar{s}_i, \bar{t}_i) \in R_i$. Thus, we conclude 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 verify next that so does it for $n + 1$. For the right-hand inclusion, fix strategy-type pair $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap \text{CARC}_{i,n+1}$ and denote $\tilde{M}_i := M_i(\bar{t}_i)$. Then, since $(\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap \text{CARC}_{i,n}$ we know from the induction hypothesis that $\bar{s}_i \in W_i^{n+1}$, and since $(\bar{s}_i, \bar{t}_i) \in R_i \cap \text{CARC}_{i,n+1}$ we know that there exists some belief $\mu_{i}^{n+1} \in \tilde{M}_i$ 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, we conclude that $\bar{s}_i \in W_i^{n+2}$.

For the left-hand inclusion, fix 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, fix arbitrary $k = 0, \ldots, n + 1$ and for any player $j \neq i$ and any strategy $s_j \in W_j^k$ define:

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

which is known from the induction hypothesis to be non-empty. We also know from the induction hypothesis that $\{Y_{j,k}(s_j) \mid s_j \in W_j^k\}$ 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
Thus, we conclude that (\( \bar{s}_i \)) is a best-reply to \( \bar{t}_i \). 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 \( \bar{M}_i(\bar{t}_i) = \bar{M}_i \). 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 it, simply note that for any \( k = 0, \ldots, n + 1 \), we have that \( \mu_i^k \in M_i(\bar{t}_i) = \bar{M}_i \). Then, since—as seen above—the support of \( \mu_i^k \) is exactly the closure of \( X_{-i,k} \), the claim is proven.

- \((\bar{s}_i, \bar{t}_i) \in R_i \). 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, we conclude that \((\bar{s}_i, \bar{t}_i) \) is 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 \( i \) it holds that:

(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 complete ambiguous type structure \( T \) and some 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 fix strategy-type pair \((\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_i \) and simply notice that since \( (\bar{s}_i, \bar{t}_i) \in R_i \cap C_i \cap CARC_i \), for any \( n \geq 0 \), we know from Theorem B1 that \( \bar{s}_i \in W_i^\infty \) for any \( n \geq 1 \). Thus, \( \bar{s}_i \in W_i^\infty \).

\[\text{For the latter, simply notice that for any } (s_{-i}, t_{-i}), \mu_i^k[N] > 0 \text{ for any neighborhood } N \text{ of } (s_{-i}, t_{-i}) \text{ if and only if } \eta_i(s_{-i})[N] > 0 \text{ for any neighborhood } N \text{ of } (s_{-i}, t_{-i}).\]
For the left-hand inclusion, fix strategy $\bar{s}_i \in W_i^\infty$. Since, in particular, $\bar{s}_i \in W_i^{n+1}$ for any $n \geq 0$, we know from Theorem B1 that for any $n \geq 0$ there exists some type $\bar{t}_i^n \in T_i$ such that $(\bar{s}_i, \bar{t}_i^n) \in R_i \cap C_i \cap CARC_{i,n}$. 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 $\bar{M}_i(\bar{t}_i) = \bar{M}_i$. 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, $\bar{s}_i \in \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i)$. ■

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

**Lemma B1.** Let $G$ be a game. Fix an ambiguous type structure $\mathcal{T}$. There exists $N \in \mathbb{N}$ such that for all $n \geq N$, we have

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

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

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

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:

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

(ii) For any self-admissible set $Q$ there exists some 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:

$$\prod_{i \in I} \text{Proj}_{S_i}(R_i \cap C_i \cap CARC_i) = Q.$$
Claim 1: Let \( T \) be the convex hull of \( \bigcup_{i \in I} \text{Proj}_{S_i} (R_i \cap C_i \cap CARC_i) \).

If \( Q = \emptyset \) then \( Q \) is a SAS. So assume it is non-empty. Fix \( s_i \in \text{Proj}_{S_i} (R_i \cap C_i \cap CARC_i) \); then there exists \( t_i \) 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 B1 and because \((s_i, t_i) \in CARC_i \subseteq CARC_{i,n+1}\) there must exist \( \mu_i \in M_i(t_i) \) such that \( \text{supp} \mu_i = \prod_{j \neq i} R_j \cap C_j \cap CARC_{j,N} \). Then, \( \mu_i := \text{marg}_{S_i} \mu_i \) is a conjecture with full-support on \( Q_{-i} \) (using again Lemma B1) for which \( s_i \) is a best-reply. Hence, condition (ii) of SAS is satisfied. Lastly, consider mixed strategy \( \sigma_i \) 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, Friedenberg and Keisler (2008) \( \sigma_i \subseteq BR_i(\text{marg}_{S_i} \mu_i) \) for every \( \mu_i \in M_i(t_i) \) giving us that \((r_i, t_i) \in R_i \) for all \( r_i \in \sigma_i \). Then we also have \((r_i, t_i) \in R_i \cap C_i \cap CARC_{i,n} \) for every \( n \geq 1 \), so that \( r_i \in \text{Proj}_{S_i} (R_i \cap C_i \cap CARC_i) \) and thus, condition (iii) of SAS is satisfied as well.

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

Now, consider the set of types \( T_i := \{ t_i(s_i) | s_i \in Q_i \} \cup \{ * \} \); to get an ambiguous type structure define \( M_i(*_i) \subseteq \Delta(S_{-i} \times T_{-i}) \) such that there does not exist \( \eta_i \in M_i(*_i) \) with \( \text{supp} \eta_i = S_{-i} \times T_{-i} \). For \( s_i \in Q_i \), define,

\[
Y_i(s_i) := \{ (r_i, t_i(s_i)) : \text{either } r_i = s_i \text{ or } \exists \sigma_i \in \Delta(S_{i}), \text{ such that } r_i \in \sigma_i \text{ and } U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i) \text{ for all } s_{-i} \in S_{-i} \}
\]

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

\[
\text{supp} \eta_i^1(s_i) = S_{-i} \times T_{-i} \quad \text{and} \quad \text{marg}_{S_{-i}} \eta_i^1(s_i) = \mu_i^1(s_i),
\]

\[
\text{supp} \eta_i^2(s_i) = \prod_{j \neq i} \bigcup_{s_j \in Q_j} Y_j(s_j) \cap R_j \quad \text{and} \quad \text{marg}_{S_{-i}} \eta_i^2(s_i) = \mu_i^2(s_i).
\]

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 we need the following:

**Claim 1:** \( \text{Proj}_{S_i} \bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i = Q_i \). 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, fix \( r_i \in \text{Proj}_{S_i} \bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i \). So there...
exists $s_i \in Q_i$ such that $(r_i, t_i(s_i)) \in Y_i(s_i) \cap R_i$. If $r_i = s_i$, we are done. If not, then by property $(iii)$ of self-admissible sets (and definition of $Y_i(s_i)$), we also have that $r_i \in Q_i$. ★

We proceed that the type structure satisfies the requirement that

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

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 $(r_i, t_i) \in R_i \cap C_i$ we need $t_i \neq *_i$ since $*_i$ is not cautious. Thus, there exists $s_i \in Q_i$ such that $t_i = t_i(s_i)$. If $s_i = r_i$, we are done. 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) $r_i \in \text{supp} \sigma_i$ for a mixed strategy $\sigma_i$ with $U_i(s_{-i}; \sigma_i) = U_i(s_{-i}; s_i)$ for every $s_{-i}$. Thus, in either case we have $(r_i, t_i(s_i)) \in Y_i(s_i)$. ★

Claim 3: $\bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i = R_i \cap C_i \cap CARC_{i, 1}$. $(r_i, t_i) \in \bigcup_{s_i \in Q_i} Y_i(s_i) \cap R_i$, 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


