IKERLANAK

UNCERTAIN INFORMATION STRUCTURES AND BACKWARD INDUCTION

by

Peio Zuazo-Garin

2014

Working Paper Series: IL. 79/14

Departamento de Fundamentos del Análisis Económico I

Ekonomi Analisiaren Oinarriak I Saila

eman ta zabal zazu

UPV EHU

University of the Basque Country

Uncertain Information Structures and Backward Induction[*]

Peio Zuazo-Garin[†]

March 25, 2014

Abstract

In everyday economic interactions, it is not clear whether sequential choices are visible or not to other participants: agents might be deluded about opponents' capacity to acquire, interpret or keep track of data, or might simply unexpectedly forget what they previously observed (but not chose). Following this idea, this paper drops the assumption that the information structure of extensive-form games is commonly known; that is, it introduces uncertainty into players' capacity to observe each others' past choices. Using this approach, our main result provides the following epistemic characterisation: if players (i) are rational, (ii) have strong belief in both opponents' rationality and opponents' capacity to observe others' choices, and (iii) have common belief in both opponents' future rationality and opponents' future capacity to observe others' choices, then the backward induction outcome obtains. Consequently, we do not require perfect information, and players observing each others' choices is often irrelevant from a strategic point of view. The analysis extends –from generic games with perfect information to games with not necessarily perfect information– the work by Battigalli and Siniscalchi (2002) and Perea (2014), who provide different sufficient epistemic conditions for the backward induction outcome.

Keywords: Perfect Information, Incomplete Information, Backward Induction, Rationality, Strong Belief, Common Belief. *JEL Classification*: C72, D82, D83.

1 Introduction

1.1 Uncertainty on the information structure: an example

Assumptions regarding common knowledge of the information structure of an economic model can significantly impact predictions. Take for instance the sequential Battle of Sexes with perfect

[*] Financial support (lower than promised) from the (now extinct) Spanish Ministry of Science and Technology (grant ECO2009-11213) is acknowledged. The author thanks Elena Iñarra, Annick Laruelle and Fabrizio Germano for inestimable help and advice, and Javier Arin, Bob Aumann, Christian W. Bach, Gilad Bavly, Eddie Dekel, Federico Echenique, Aviad Heifetz, Ziv Hellman, Nagore Iriberri, Andrés Perea and Dov Samet for discussion and valuable comment, as well as participants in seminars at the University of the Basque Country, the University of Tel Aviv, the Center for the Study of Rationality and the University of Valladolid, and in the 24th Stony Brook Game Theory Festival and the 2nd International Workshop in the University of Tsukuba for helpful feedback. Gratitude for the immense hospitality from both the Center for the Study of Rationality at the Hebrew University of Jerusalem and Eilon Solan and the Game Theory Lab at Tel Aviv University, where the vast majority of the paper was written during scholar visits is expressed.

[†] *BRiDGE* group, University of the Basque Country, Department of Foundations of Economic Analysis I, Avenida Lehendakari Aguirre 83, 48015, Bilbao, Spain; peio.zuazo@ehu.es.

information represented in Figure 1. Two players, Alexei Ivanovich (A) and Polina Alexandrovna (P) choose first and second respectively between actions *left* and *right*, and obtain utility depending on each history of actions according to the numbers depicted at the bottom of the tree in the picture. By information structure we refer to whether Polina chooses having observed Alexei's previous choice or not, which she does in this case of perfect information. The game is played just once, so punishment and reinforcement issues are assumed to be negligible. This description is common knowledge among the players, and we additionally assume that both of them are rational, and that Alexei believes Polina to be rational. It then seems reasonable to predict that the players' choices will lead to the unique backward induction outcome: $(2, 1)$; since Polina is rational and observes Alexei's choice, she will mimic it regardless of whether it is *left* or *right*. Alexei believes all the above, so since he himself is rational too, he will move *left*.

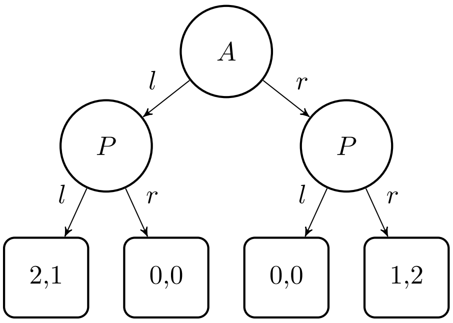

Figure 1: A game with perfect information.

Turn now to a commonly known imperfect information situation (Figure 2): consider the alternative information structure according to which, when her turn arrives, Polina will not have observed Alexei's previous move. Thus, Polina is uncertain of the outcome her choice will induce. Even if we additionally assume that Polina believes both that Alexei is rational and that Alexei believes she is rational, it is easy to see that the previous argument justifying outcome $(2, 1)$ finds no defence this time; and that indeed, depending on reciprocal beliefs concerning opponents' choices, every outcome is consistent with rationality and with any assumption about iterated mutual beliefs in rationality.

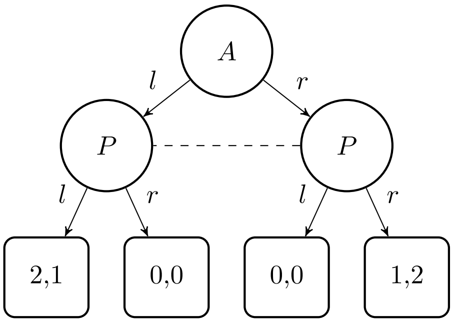

Figure 2: A game w/o perfect information.

Consider finally an imperfect information case such as the one represented in Figure 2, with the following variation: Alexei believes himself to be in a situation like the one in Figure 1; and Polina believes that Alexei believes himself to be in that situation of perfect information. That is, the information structure of the game is not commonly known this time and, actually, Alexei happens to be deluded about it. When it is her turn to choose, despite Polina not observing Alexei's previous move, she can infer that since Alexei believes himself to be in a situation with perfect information, he also believes *left* to be followed by *left* and *right* by *right*, and will therefore choose *left*. Hence, despite not observing Alexei's previous move, Polina believes that Alexei has chosen left and consequently she chooses *left*.

As the example above illustrates, assumptions regarding common knowledge of the information structure of an economic model can significantly impact predictions. The beliefs players hold about the information structure prove to be more relevant in terms of strategic behaviour than players' capacity to observe each others' past choices: that is, more relevant than the factual information structure itself. Consequently, establishing the distinction and exploring the differences concerning strategic implications between notions such as *perfect information*, *incomplete imperfect information* or *common knowledge of perfect information*, which not only refer to the way information flows but also to players' epistemic state concerning this flow, becomes interesting from a game theoretical perspective. In particular, as the comparison between the first and the last situations in the example above suggests, this language allows for extending the class of games for which the backward induction outcome can be considered as a reasonable prediction to the more general setting of contexts with not necessarily perfect information.

1.2 Epistemic sufficiency for backward induction: abandoning perfect information

The literature related to the study of the epistemic assumptions leading to backward induction in extensive-form games with perfect information is prolific and has been abundant in recent years. Despite the apparent simplicity and intuitive appeal of backward induction, and similarly as happens with strategic-form games and Nash equilibria, some discomfort concerning the not-so-explicit epistemic aspects of the solution concept leads to an attempt to deepen our understanding of backward induction. The source of such discomfort lies in this case in the fact that backward induction reasoning seems unable to capture a crucial aspect of sequential playing: the capacity to update beliefs, and in particular, to question the plausibility of a player who showed erratic behaviour in the past actually behaving rationally in the future. Focusing on this apparent weakness of backward induction reasoning, Reny (1992) presents an example of a finite extensive-form game whose unique *extensive-form rationalisable* (EFR, Pearce, 1984) profile does not coincide with its backward induction profile, and Ben Porath (1997) shows that in Rosenthal's centipede the backward induction outcome is not the only one consistent with initial belief in rationality. Still, in Reny's example, the outcome induced by both the EFR profile and the backward induction profile is the same, and Battigalli (1997) generalises this coincidence to the point of proving that for generic finite extensive-form games EFR profiles always lead to the unique backward induction outcome.[1]

A series of results follow the identity above: Battigalli and Siniscalchi (2002) introduce the notion of *strong belief* to represent the idea of forward induction reasoning, and prove that rationality and common strong belief in rationality induce EFR profiles when the type structure is complete (*i.e.*, when it is able to represent any possible belief a player might hold), and hence, lead to the backward induction outcome. Battigalli and Friedenberg (2012) define a new solution concept, *extensive-form best reply sets* (EFBRSs), and prove that rationality and common strong belief in rationality induce profiles included in these sets regardless of whether

[1] While Battigalli's original proof relies on rather intricate mathematics, Heifetz and Perea (2013) present a more intuitive proof that clarifies the logic relating to both outcomes.

the type structure is complete or not. However, they present examples where the outcomes induced by profiles in EFBRSs are not the backward induction outcome, so sufficient epistemic conditions for the backward induction outcome for arbitrary (not necessarily complete) type structures remain unclear. Penta (2011) and Perea (2014), exploit the notion of *future belief in opponents' rationality* and present such sufficient conditions for extensive-form games with perfect information and arbitrary type structures by proving that rationality and common belief in opponents' future rationality induces the backward induction outcome.[2]

A different approach for epistemic analysis in games with perfect information is adopted by Aumann (1995, 1998), who makes use of static partition models that, unlike the models explained above, do not include explicit belief revision. Aumann (1995) proves that *ex ante* common knowledge of rationality induces the backward induction profile. Samet (2013) modifies this result substituting common knowledge by common belief, and defining rationality in terms of beliefs rather than in terms of knowledge, as done by Aumann, in terms of knowledge.[3] Bonanno (2013) also proves that common belief in rationality induces the backward inductive outcome using belief frames that allow for belief revision and by assuming something analogous to belief in opponents' future rationality. Previously, Samet (1996) approached the problem with very rich models that deal with knowledge rather than beliefs, but allow the modelling of hypothetical counterfactual information updates. Arieli and Aumann (2013) provide a novel and interesting an epistemic characterisation of backward induction for games of perfect information similar to that by Battigalli and Siniscalchi (2002), that, unlike the latter work and all the literature in epistemic game theory referenced so far, is performed *via* a syntatic approach rather than the standard semantic one.[4]

The present paper drops the assumptions that the game has perfect information and that this feature is commonly known, and extends the analysis regarding sufficient epistemic assumptions for the backward induction outcome to a broader class of extensive-form games. In order to do so, we introduce uncertainty in what we call the *information structure* of the extensive-form game. By information structure we refer to how each player's set of histories (*i.e.* the histories in which it is the player's turn to make a choice) is partitioned into information sets. The information structure can be regarded as the players' capacity to observe others' past choices, so the uncertainty we introduce can be read as a lack of certainty about whether each player is able to observe or remember her opponents' past choices prior to her turn at making one. Following this approach we prove in Theorem 1 that for arbitrary type structures, under the assumptions

[2] Both Penta and Perea's work is actually more general: Perea (resp. Penta) proves that in generic extensive-form games with not necessarily perfect information, rationality and common belief in opponents' future rationality (resp. and common belief in opponents' future rationality and in Bayesian updating) induce what he defines as strategy profiles surviving the *backward dominance procedure* (resp. the *backwards rationalisability procedure*), which in games with perfect information coincide exactly with the backward induction profile. In addition, Penta proves that in his characterisation result, the assumptions above can be substituted by common certainty of *full rationality* and *belief persistence*. A non probabilistic version of future belief in opponents' rationality can be found in Baltag *et al.* (2009).

[3] The present paper is a variation of the original one by Zuazo-Garin (2013) that applies the idea of uncertainty on information structures and how it epistemically relates to backward induction by adopting the framework introduced by Samet (2013).

[4] For further references on epistemic game theory focused on extensive-form games, see Perea (2007) or Section 7 in Dekel and Siniscalchi (2013).

that: (*i*) players are rational, (*ii*) players strongly believe that opponents were rational and had perfect information, and (*iii*) there is common belief in opponents' future rationality and opponents' future perfect information, the backward induction outcome obtains. Note that we do not assume perfect information but, rather, that even when it is the case that a player has not observed any of her opponents' past choices, she believes that others have, and will do so in the future. Together with assumptions about rationality, these beliefs help the player infer what happened in the past, establish beliefs about future behaviour and, consequently, also choose *the* action that happens to be strictly the best for her. In Theorem 2 we prove that for any extensive-form game and any given information structure, it is possible to construct a type structure such that there is some state at which our assumptions are satisfied and are indeed compatible with the given information structure.

Previous literature on implications of incomplete information in extensive-form games include Battigalli and Siniscalchi (2007) and Penta (2011, 2012) among others. However, these works focus on *payoff uncertainty*, meaning that a history of actions itself does not determine payoffs unless some other certain payoff-relevant parameter is also considered so that issues regarding beliefs on the information structure of the game are not covered.

The rest of the paper is structured as follows: Section 2 describes economic scenarios in which uncertainty about the information structure might be present and heavily influence expected behaviour. Section 3 and 4 detail our formalisations of extensive-form games and information structures, and the epistemic framework and notions needed to perform the analysis, respectively. Section 5 presents our main results in Theorem 1 and Theorem 2 and their respective proofs, and we finish with some remarks and discussion in Section 6.

2 Brief discussion on the economic relevance of uncertainty on the information structure

The capacity of reciprocal pre-choice observation by agents involved in some interaction context is often obvious. It might be obvious that there is perfect information, as in the case of a potential robber at a clothing store who knows that the anti-theft device reveals to the store owner whether he decided to steal or not. But it might alternatively be obvious that there is no perfect information: this is the case in a used car emporium where the seller offers the buyer a car whose quality the latter cannot observe. This distinction leads to the canonical classification of extensive-form games in those with perfect information and those with imperfect information, in which it is common knowledge that there is perfect information and that there is not perfect information, respectively.

Now, it turns out that this apparent dichotomy between games with perfect information and games with imperfect information is a non-exhaustive classification, and we can think of many situations in which it is not obvious that there is perfect information, or it is not obvious that there is no perfect information: in the above example of a game with perfect information, the anti-theft device could just be a cheap fake put there by the owner to fool potential robbers, while in the example of a game with imperfect information, it might be the case that the buyer is an expert whose just needs a brief inspection of the car offered to determine whether it is a good car or not.

We see then that the expected flow of information is sensitive to many aspects surrounding the context of interaction, and it is not clear why agents should not just agree, but *commonly* agree in their appreciation of these aspects and their influence. It is not the aim of this paper to propose some heuristic mechanism that endogenises the rising of different beliefs about the information structure but rather to point out the possibility of the latter being uncertain, to highlight the relevance of such uncertainty, and to provide conditions in which the assumption of perfect information being commonly known can be dropped with no significant strategic consequences. The present section deals with the first two objectives by illustrating the situation in Example 1. Section 3 formalises what exactly we mean by a player's capacity to observe past actions, Section 4 considers beliefs of players in an uncertain information structure such as exogenous parameters, and Section 5 deals with the final objective by establishing under which kind of beliefs about the information structure a player chooses in a case analogous to that of perfect information.

Example 1 (Deluded reputation and private information disclosure). A classic approach to the impact of reputation in agents' behaviour by Milgom and Roberts (1982) and Kreps and Wilson (1982) and recently revisited by Ely and Valimaki (2003) and Ely *et al.* (2008) considers the establishment of reputation as a strategic device an agent might rationally decide to commit to in order to condition potential opponents' beliefs regarding her actions, in case she expects to obtain profit this way in the long run. This is not the only way reputation, which can be interpreted in a broad sense, can be crucial in determining agents' behaviour though.

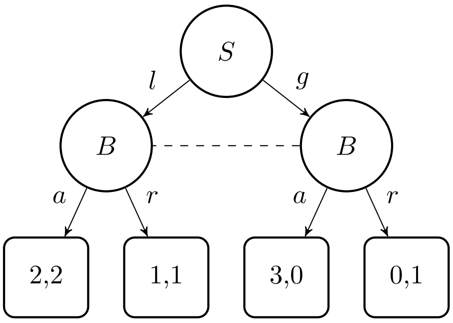

Figure 3: A situation with private information.

Consider again a used car emporium where the seller (S) offers the buyer (B) a car, the quality of which is private information of her own, for some fixed price.[5] More precisely, we assume a context as the one depicted in Figure 3: the seller can offer a good car or a lemon (g and l, respectively), and after a brief inspection, the buyer can accept the offer (a), pay the fixed price and get the car offered, or reject it (r). The true information structure, obviously known to the buyer, corresponds to that of imperfect information: she is unable to determine whether the car the is examining is a good one or a lemon. The preference structure of the game is represented by the numbers at the bottom of the figure: the seller's preferred option is having the lemon accepted, and she prefers selling something than seeing the buyer leave empty-hand, while the buyer's preferred option is getting the good car and the worst preferred one is getting

[5] Which might just be the expected price in a "Lemon Market" (Akerlof, 1970).

the lemon.

As we saw in paragraph 1.1, assumptions on beliefs about the information structure, and not just the specification of a belief structure is required in order to make common belief in rationality concluding. In this situation this is achieved *via* a *reputation of expertise*, where *reputation* serves as an exogenous mechanism that induces certain beliefs about the information structure, and by *expertise* we refer to the ability to discern the quality of a car with a brief examination. That is, in this context, reputation of expertise is actually referring to which the information structure of the game is:[6]

Assumption 1. *There is commonly (and delusively) believed reputation of expertise in auto mechanics.*

Note the implications: the seller believes the buyer to be an expert, *i.e.*, to be able to distinguish a good car from a lemon; since she additionally believes her customer to be rational, the conjectures that a good car will be accepted while a lemon will be rejected, and being herself rational too, concludes that offering a good car is the best option for her. Since both the expertise reputation and the seller's rationality are commonly believed, the buyer is able to induce her opponent's reasoning, and infer that it is the good car the one she is being offered. Thus, since she is rational too, she decides to accept the offer. This way, reputation leads to some kind of private information disclosure that moreover, yields the backward induction outcome of the game. Note also the sharp contrast with situations corresponding to (*i*) the absence of any reputation regarding expertise: if the seller does not take the buyer for an expert, she might try to take advance of her private information and sell a lemon, or (*ii*) the reputation of expertise not being believed by the buyer: despite being offered a good car, the buyer is unable to infer the quality of the car.

The example above illustrates how reputation can induce uncertainty on the information structure. But this uncertainty can be a consequence of a variety of factors: an agent might not be certain of how a deliberate action chosen in order to serve as a signal will be interpreted by subsequent agents (think of education in Spence (1973), for instance), the presence of different element such as a camera or a (possibly one-way) mirror might induce a feeling of surveillance (*i.e.*, of perfect information), etc.... And indeed, as both the example above and the situations described in paragraph 1.1 suggest, this uncertainty can significantly impact expected rational behaviour.

3 Games with uncertain information structure

We consider extensive-form games with incomplete information regarding the information structure of the game. In order to do so we formalise two objects: (*i*) a game tree similar to the extensive-form games with perfect information in Osborne and Rubinstein (1994, Sect. 6.1), which is assumed to be common knowledge, and (*ii*) the set of possible information structures on the given game tree, which is the part of the description of the game that is uncertain. In the

[6] And hence, iterated beliefs about the reputation refer to iterated beliefs about the information structure.

last paragraph we detail the role of strategies in this context, and how they relate to uncertainty about the information structure and outcomes. So, we have:

3.1 Game trees

A (finite) *game tree* is a tuple $\Gamma = \left\langle I, (A_i)_{i \in I}, \mathcal{H}, \mathcal{Z}, (u_i)_{i \in I} \right\rangle$ where:

- I is a finite set of *players.*
- For each player i, A_i is a finite set of *actions.* The set of *possible actions*[7] is denoted $A = \bigcup_{i \in I} A_i$, and we refer to a finite concatenation of elements in $\{\emptyset\} \cup A$, as a *history.*
- $\mathcal{H}$ and $\mathcal{Z}$ are finite and disjoint sets of histories. We assume that the union $\mathcal{H} \cup \mathcal{Z}$ is a rooted and oriented tree with terminal nodes $\mathcal{Z}$. Histories in $\mathcal{H}$ and $\mathcal{Z}$ are called *partial* and *terminal* respectively. For any player i and partial history h, let $A_i(h) = \{a_i \in A_i \,|(h, a_i) \in \mathcal{H} \cup \mathcal{Z}\}$ the set of *actions available to i at h.* We say that player i is *active* at h if $A_i(h) \neq \emptyset$. We also assume that exactly one player is active at each history, that a player is never active twice in a row, and that whenever a player is active, at least two actions are available to her.[8] Additionally, for each player i we define $\mathcal{H}_i = \{h \in \mathcal{H} \,| A_i(h) \neq \emptyset\}$, the set of *partial stories in which player i is active*, and denote $\mathcal{H}_{-i} = \bigcup_{j \neq i} \mathcal{H}_j$. For any pair of histories h and h', we write $h < h'$ when h' *follows* h; that is, when there exists some finite concatenation of actions $(a^n)_{n \leq N}$ such that $h' = \left(h, (a^n)_{n \leq N}\right)$.
- For each player i, $u_i : \mathcal{Z} \to \mathbb{R}$ is player i's *payoff function.* Following Battigalli (1997), we assume that the game has *no relevant ties*, that is, that for any player i and any $h \in \mathcal{H}_i$, function u_i is injective when restricted to the set of terminal histories that follow h.

3.2 Information sequences

For each player i and subset of her histories $v_i \subseteq \mathcal{H}_i$, we say that v_i is an *information set* if none of its elements follow each other, and exactly the same actions are available at all of them.[9] Let V_i be a partition of $\mathcal{H}_i$; we say that V_i is an *information partition for player i* if its cells are information sets and it satisfies perfect-recall.[10] Note that we can then denote $A_i(v_i)$ as the actions available at information set v_i with no ambiguity. An *information structure* is then a profile $V = (V_i)_{i \in I}$ of information partitions. For each player i, we denote by $\mathcal{V}_i$ the set of player i's information partitions. $\mathcal{V}$ denotes the set of information structures.

[7] Not to be confused with the set of action *profiles.*

[8] Formally, the requirements are: (i) $\emptyset \in \mathcal{H}$, (ii) for any $(h, a) \in \mathcal{H} \cup \mathcal{Z}$, $h \in \mathcal{H}$, (iii) for any $h \in \mathcal{Z}$, $(h, a) \notin \mathcal{Z}$ for any $a \in A$, (iv) for any $h \in \mathcal{H}$, $i \in I$ and $a_i \in A_i$ such that $(h, a_i) \in \mathcal{H} \cup \mathcal{Z}$, it holds that if $(h, a) \in \mathcal{H} \cup \mathcal{Z}$ for some $a \in A$, then $a \in A_i$, (v) for any $h \in \mathcal{H}$, any $i \in I$, any $a_i \in A_i$ and $a \in A$ such that $(h, a_i, a) \in \mathcal{H}$, $a \notin A_i$,and (vi) for any $h \in \mathcal{H}$, if $A_i(h) \neq \emptyset$, then $|A_i(h)| \geq 2$.

[9] That is, for any $h, h' \in v_i$, $h \not< h$, and $A_i(h) = A_i(h')$.

[10] Perfect recall is satisfied if: (i) for any $h, h' \in \mathcal{H}_i$ such that $h' \notin V_i(h)$ and $(h, a_i) < h'$ for some $a_i \in A_i$, for any $h''' \in V_i(h')$ there is some $h'' \in V_i(h)$ such that $(h'', a_i) < h'''$, and (ii) for any $h, h', h'', h''' \in \mathcal{H}_i$ such that $h < h''$, $h' < h'''$ and $V_i(h) \neq V_i(h')$, $V_i(h'') \neq V_i(h''')$.

We allow for uncertainty about the information structure of the game. Consequently, a player's information set does not tell just by itself what information sets the player previously went through during the game. Therefore, at a certain information set, the information available to the player is not only the information set itself, but also the previous information sets, if any, in which she previously made some choice. The concept of *information set* needs then to be somehow broadened in order to incorporate not only possible indistinguishability between histories, but also histories of information sets.

For player i, information partition of hers $V_i \in \mathcal{V}_i$ and information sets $v_i, v_i' \in V_i$, we write $v_i < v_i'$ if there is some action $a_i \in A_i(v_i)$ such that for any $h' \in v_i'$ there is some $h \in v_i$ such that $(h, a_i) < h'$. We say that information set v_i is *minimal*, if $v_i \nless v_i'$ for any $v_i' \in V_i$ and any $V_i \in \mathcal{V}_i$. Then, we expand the notion of information set the following way:

Definition 1 (Information sequence). *Let game tree* Γ. *An* information sequence *for player* i *is a concatenation of consecutive information sets of some information partition, with a minimal element; i.e., a sequence* $(v_i^n)_{n \leq N} \subseteq V_i$, *where* $V_i \in \mathcal{V}_i$, v_i^1 *is minimal, and* $v_i^n < v_i^m$ *for any* $n, m \leq N, n < m$. *We denote the set of information sequences for player* i *by* Σ_i, *and for information sequence* $\sigma_i = (v_i^n)_{n \leq N}$, *we denote by* $v_i(\sigma_i) = v_i^N$.

For each player i, each $h \in \mathcal{H}_{-i}$ and each $\sigma_i \in \Sigma_i$ we write $h < \sigma_i$ (resp. $\sigma_i < h$) if there is some $h' \in v_i(\sigma_i)$ such that $h < h'$ (resp. $h' < h$), and for each $j \neq i$ and each $\sigma_i \in \Sigma_i$ and $\sigma_j \in \Sigma_j$, we write $\sigma_i < \sigma_j$ if there is some $h \in v_i(\sigma_i)$ such that $h < \sigma_j$.

3.3 Strategies and terminal histories

In this context, a strategy is not a description of what action to choose at each history or information set, as in the standard cases of *commonly known* perfect or imperfect information respectively, but rather, of what action to choose after any possible information sequence. That is, for each player i, a strategy is a list $s_i \in S_i = \prod_{\sigma_i \in \Sigma_i} A_i(\sigma_i)$, where $A_i(\sigma_i) = A_i(v_i(\sigma_i))$ for any $\sigma_i \in \Sigma_i$. We write $S_{-i} = \prod_{j \neq i} S_j$ to represent the set of player i's opponents' strategies. Note that a strategy profile itself does not induce any terminal history, but a pair (s_i, V_i), induces a *strategy in terms of histories* s_{V_i} given by $h \mapsto s_{v_i(\sigma_i)}$, where σ_i is indeed the unique information sequence such that $\sigma_i \subseteq V_i$ and $v_i(\sigma_i) = V_i(h)$.[11] For profiles s and V we denote $s_V = (s_{V_i})_{i \in I}$. Then, since strategies in terms of histories do induce conditional terminal histories at each partial history $h \in \mathcal{H}$, so does a pair (s, V). This is formally described as,

$$z(s_V | h) = \left\{ z \in \mathcal{Z} \,\middle|\, \begin{array}{c} \text{for any } i \in I \text{ and any } h' \in \mathcal{H}_i \text{ such that} \\ h \leq h' \text{ and } h' < z, \left(h', s_{V_i(h')}\right) \leq z \end{array} \right\},$$

so that each player's conditional payoffs are naturally determined by conditional terminal histories as follows: $u_i(s, V | h) = u_i(z(s_V | h))$.

[11] It is not accurate to speak of strategy in these terms, since a player cannot make her choice contingent on a history if it is the case that she cannot distinguish this history from a different one. It still serves as a description of actions chosen at different histories.

Note that any combination of a strategy profile and an information structure precludes certain information sets being reached, so it is useful to write the following: let player i and $\sigma_i \in \Sigma_i$; then, (i) by $(S_{-i} \times \mathcal{V})(\sigma_i)$ we denote the set of opponents' strategies and information structures such that σ_i might be reached, and (ii) by $S_i(\sigma_i)$, the set of player i's strategies such that σ_i may be reached.[12] Finally, for any $s_i \in S_i$, we define $\Sigma_i(s_i) = \{\sigma_i \in \sigma_i \,|s_i \in S(\sigma_i)\}$, the set of player i's information sequences whose terminal information set might be reached when she plays strategy s_i.

4 Epistemic framework

The epistemic analysis is carried out in a construction following the work by Ben Porath (1997) and Battigalli and Siniscalchi (1999, 2002), among others. First we define the general environment whose central elements are type structures, and then, we formalise the main notions that complete the epistemic language. So first, players' beliefs are modelled with type structures. For the sake of brevity and comprehension, we restrict our attention to these structures and do not detail their relation with belief hierarchies; still, this issue if briefly addressed in paragraph G of Section 6. A *type structure* is defined as follows:

Definition 2 (Type structure). *Let game tree* Γ. *A* type structure *for* Γ *is a list* $\mathrm{T} = \langle E_i, b_i \rangle_{i \in I}$ *where for each player* i*,*

(i) E_i *is a metric and compact* epistemic type space*.*

(ii) $b_i : E_i \to \prod_{\sigma_i \in \Sigma_i} \Delta(E_{-i} \times (S_{-i} \times \mathcal{V})(\sigma_i))$, *where* $E_{-i} = \prod_{j \neq i} E_j$, *is a continuous* conditional belief map*.*[13]

We say that T is *complete* if every b_i is surjective. Type structure T induces set of *states of the world* $\Omega = E \times S \times \mathcal{V}$, where $E = \prod_{i \in I} E_i$. Each element $\omega \in \Omega$ is a description of: (i) the information structure of the game, (ii) each player's strategy and (iii) each player's beliefs about all possible uncertainties.[14] For any state ω we denote $v(\omega) = \mathrm{Proj}_{\mathcal{V}}\omega$, and for each player i, $e_i(\omega) = \mathrm{Proj}_{E_i}\omega$ and $s_i(\omega) = \mathrm{Proj}_{S_i}\omega$.

An *event* is a set of states $W \subseteq \Omega$. Note that some events and information sequences are mutually *belief-inconsistent*: for each player i and information sequence σ_i, let $W^{\sigma_i} = \left(\mathrm{Proj}_{E_{-i} \times S_{-i} \times \mathcal{V}} W\right) \cap (E_{-i} \times (S_{-i} \times \mathcal{V})(\sigma_i))$; then, if $W^{\sigma_i} = \emptyset$, player i will *always*[15] assign

[12] These are respectively described by $\left\{(s_{-i}, V) \in S_{-i} \times \mathcal{V} \,\middle|\sigma_i \subseteq V_i \text{ and } z\left((s_{-i}; s_i)_V\right) > \sigma_i \text{ for some } s_i \in S_i\right\}$ and $\left\{s_i \in S_i \,\middle|\sigma_i < z\left((s_{-i}; s_i)_V\right) \text{ for some } (s_{-i}, V) \in (S_{-i} \times \mathcal{V})(\sigma_i)\right\}$. When necessary, for each $v_i \in V_i \subseteq \mathcal{V}_i$, we denote with no ambiguity $(S_{-i} \times \mathcal{V})(v_i) = (S_{-i} \times \mathcal{V})(\sigma_i)$ where σ_i is such that $v_i(\sigma_i) = v_i$.

[13] Assuming that each $\Delta(E_{-i} \times (S_{-i} \times \mathcal{V})(\sigma_i))$ is endowed with the weak$*$ topology, and their product, with the Tychonoff topology.

[14] Following Perea (2014), we opted for a notationally simpler definition of type structure than usual; to fully adhere to standard notation, we should have defined $\mathrm{T} = \left\langle I, (\mathcal{C}_{-i}, E_i, b_i)_{i \in I}\right\rangle$, where for any $i \in I$, $\mathcal{C}_{-i} = \{E_{-i} \times (S_{-i} \times \mathcal{V})(\sigma_i) \,|\sigma_i \in \Sigma_i\}$ and $b_i : E_i \to \Delta^{\mathcal{C}_{-i}}(E_{-i} \times S_{-i} \times \mathcal{V})$, being $\Delta^{\mathcal{C}_{-i}}(E_{-i} \times S_{-i} \times \mathcal{V})$ the set of conditional probability systems (CPS, see Renyi, 1955) definable over the measurable space composed by $E_{-i} \times S_{-i} \times \mathcal{V}$ and its corresponding Borel σ-algebra, together with set of conditioning events $\mathcal{C}_{-i}$. However, note that our definition does not impose *Bayesian updating*, so the concept of CPS turns out to be too restrictive for our purposes. This last aspect is discussed in paragraph C of Section 6.

[15] That is, no matter what her beliefs are.

null probability to event W at σ_i. This way, for player i and event W, we define *player i's set of information sequences consistent with W* as $\Sigma_i(W) = \{\sigma_i \in \Sigma_i \,|W^{\sigma_i} \neq \emptyset\}$. This set represents i's information sequences in which it might be the case that i assigns positive probability to W. We can now proceed to introduce the main epistemic notions needed for analysis.

4.1 Rationality

Player i's *conditional expected payoff* when she plays s_i and her epistemic type is e_i after information sequence $\sigma_i \in \Sigma_i(s_i)$ is given by,

$$u_i(e_i, s_i \,|\sigma_i) = \sum_{h \in v_i(\sigma_i)} \sum_{(s_{-i},V) \in (S_{-i} \times \mathcal{V})(\{h\})} b_i(e_i, \sigma_i)\left[E_{-i} \times \{(s_{-i}, V)\}\right] u_i\left((s_{-i}; s_i), V \,|h\right).$$

A player is conditionally rational after an information sequence whenever her strategy is not strictly dominated by another in terms of her conditional expected payoff after the information sequence. Thus, the event that *player i is conditionally rational at information sequence σ_i* is defined as,

$$R_{\sigma_i} = \left\{\omega \in \Omega \,\middle|\, s_i(\omega) \in \text{argmax }_{s_i \in S_i(\sigma_i)} u_i\left(e_i(\omega), s_i \,|\sigma_i\right)\right\}.$$

We say that player i is rational if she is conditionally rational after any of her information sequences, so the event that *player i is rational* is defined as $R_i = \left\{\omega \in \Omega \,\middle|\, \omega \in \bigcap_{\sigma_i \in \Sigma_i(s_i(\omega))} R_{\sigma_i}\right\}$, [16] and the event that *players are rational*, as $R = \bigcap_{i \in I} R_i$. Following standards, for player i, we denote $R_{-i} = \bigcap_{j \neq i} R_j$.

Following Baltag *et al.* (2009) and Perea (2014), the hypothesis regarding opponents being rational in the future regardless of their past behaviour is an essential aspect of the present work. Thus, to make this feature explicit, for player i and information sequence $\sigma \in \bigcup_{j \neq i} \Sigma_j$, we define the event that *player i is future rational from σ*, as,

$$FR_i(\sigma) = \{\omega \in \Omega \,|\omega \in R_{\sigma_i} \text{ for any } \sigma_i \in \Sigma_i(s_i(\omega)) \text{ such that } \sigma < \sigma_i\}.$$

For each player i and information sequence σ_i we denote $FR_{-i}(\sigma_i) = \bigcap_{j \neq i} FR_j(\sigma_i)$.

4.2 Perfect information

We say that a player has perfect information at some stage of the game, when it is her turn to make a choice and she knows what her opponents previously chose. Obviously, this can only be so when the information set she finds herself at is a singleton. Thus, for player i

[16] A remark regarding the *one-shot deviation principle* (OSDP) and dynamic inconsistency issues is necessary at this point. Note the following two facts: (i) since the type structure is, in a meta-sense, commonly known, each player *knows* after any information sequence which her beliefs will be at any future information sequence she finds herself at as the plays goes on, and (ii) despite our definition of rationality being provided in terms of *ex ante* strategies, we impose optimality for any information sequence that is not precluded by the strategy itself. Then, from (i) and (ii), we can conclude that despite rationality being defined in terms of *ex ante* strategies, since the beliefs with respect to any choice after any information sequence is evaluated in terms of the beliefs corresponding to that information sequence, the OSDP is not violated and dynamic inconsistency is avoided. Recent work by Battigalli *et al.* (2013) studies these issues in detail.

and her history $h \in \mathcal{H}_i$, the event that *player i has perfect information at h* is defined as $PI_h = \{\omega \in \Omega \,|v_i(\omega)(h) = \{h\}\}$ and the event that *player i has perfect information*, as $PI_i = \bigcap_{h \in \mathcal{H}_i} PI_h$. The event that *there is perfect information* is then $PI = \bigcap_{i \in I} PI_i$. For each player i, we denote $PI_{-i} = \bigcap_{j \neq i} PI_j$, and by V_i^{PI}, the information partition corresponding to the case in which i has perfect information, that is, $\{\{h\} \,|h \in \mathcal{H}_i\}$.

As in the previous paragraph, we are interested in making the presence of perfect information following any given information sequence explicit. This way, for player i and information sequence $\sigma \in \bigcup_{j \neq i} \Sigma_j$, the event that *player i has future perfect information from σ* is defined as,

$$FPI_i(\sigma) = \{\omega \in \Omega \,|\omega \in PI_h \text{ for any } h \in \mathcal{H}_i \text{ such that } \sigma < h\},$$

and for each player i and information sequence σ_i, we denote $FPI_{-i}(\sigma_i) = \bigcap_{j \neq i} FPI_j(\sigma_i)$.

4.3 Beliefs about opponents

Beliefs about opponents' behaviour and opponents' belief hierarchies is a central element of strategic planning. In order to formalise this idea, for player i we define player i's conditional belief operator as the association of each event W with the event that player i conditionally believes W after a given information sequence with probability 1. Formally, for information sequence σ_i, player i's *conditional belief operator at σ_i* is given by,

$$W \mapsto B_i(W \,|\sigma_i) = \{\omega \in \Omega \,|b_i(e_i(\omega), \sigma_i)[W^{\sigma_i}] = 1\}, \text{ for any } W \subseteq \Omega.$$

4.3.1 A strong belief about opponents

Battigalli and Siniscalchi (2002) introduce the concept of strong belief that formalises the notion of forward induction. That is, the ability of players to rationalise, as long as possible, opponents' past behaviour, and conjecture about their beliefs and future behaviour in a way consistent with the rationalisation. They define the strong belief operator, which in our context associates each event W with the event that player i conditionally believes W with probability 1 after any information sequence not belief-inconsistent with W. That is, formally, player i's *strong belief operator* is given by,

$$W \mapsto SB_i(W) = \bigcap_{\sigma_i \in \Sigma_i(W)} B_i(W \,|\sigma_i), \text{ for any } W \subseteq \Omega.$$

So $SB_i(W)$ should be read as the event that player i maintains the hypothesis that W is true as long as it is not contradicted by evidence. In this paper we are interested in the working hypothesis that players believe that their opponents are rational, have perfect information, and commonly believe in their opponents' future rationality and perfect information. Formally, this is represented by the following: let player i; we define the event that *player i strongly believes in* *both opponents' rationality and opponents' perfect information*, as,

$$SBORPI_i = SB_i \left(R_{-i} \cap PI_{-i} \right),$$

and we define the event that *there is strong belief in both opponents' rationality and opponents' perfect information*, as $SBORPI = \bigcap_{i \in I} SBORPI_i$.

4.3.2 A common belief about opponents

In the spirit of Baltag *et al.* (2009) and Perea (2014), for any player i we define the event that *player i believes in both opponents' future rationality and opponents' future perfect information* as,

$$BOFRPI_i = \bigcap_{\sigma_i \in \Sigma_i} B_i \left(FR_{-i} \left(\sigma_i \right) \cap FPI_{-i} \left(\sigma_i \right) \left| \sigma_i \right. \right).$$

That is, this event somehow represents the fact that player i believes that any evidence contradicting opponents' past rationality and perfect information is due to a mistake and should therefore be disregarded. Now, for each player i, let:

$$\begin{aligned} CBOFRPI_i^0 &= BOFRPI_i, \\ CBOFRPI_i^n &= \bigcap_{\sigma_i \in \Sigma_i} B_i \left(CBOFRPI_{-i}^{n-1} \middle| \, \sigma_i \right), \end{aligned}$$

where $CBOFRPI_{-i}^{n-1} = \bigcap_{j \neq i} CBOFRPI_j^{n-1}$ for any $n \in \mathbb{N}$. Then, the event that *there is common belief in both opponents' future rationality and opponents' future perfect information for player i* is defined as $CBOFRPI_i = \bigcap_{n \geq 0} CBOFRPI_i^n$, and the event that *there is common belief in both opponents' future rationality and opponents' future perfect information*, as $CBOFRPI = \bigcap_{i \in I} CBOFRPI_i$. As it will eventually be useful in the proof of Theorem 1, it is easy to check that for any player i and her information sequence σ_i, $CBOFRPR_i \subseteq B_i \left(CBOFRPI_{-i} \left| \sigma_i \right. \right)$, where $CBOFRPI_{-i} = \bigcap_{j \neq i} CBOFRPI_j$.

4.3.3 A remark on perfect information

Note that the event that there is perfect information as defined above, does not imply any belief assumption about other players having perfect information at any history, so this terminology does not exactly coincide with the standard notion of perfect information in games with commonly known information structures. In such games, the statement *the game has perfect information* should be read as the event that there is *perfect information and common belief of perfect information*, which should not be confused with the event that *there is perfect information*, PI, defined above.

Formally, for each player i we define $PICBPI_i^0 = PI_i$ and for any $n \in \mathbb{N}$, $PICBPI_i^n = PICBPI_i^{n-1} \cap \bigcap_{\sigma_i \in \Sigma_i} B_i \left(PICBPI_{-i}^{n-1} \left| \sigma_i \right. \right)$, where $PICBPI_{-i}^{n-1} = \bigcap_{j \neq i} PICBPI_j^{n-1}$. Under this notation, the usual notion of the game having perfect information is formalised by $PICBPI = \bigcap_{i \in I} PICBPI_i$, where $PICBPI_i = \bigcap_{n \geq 0}$, which is precisely the event that there is perfect information and common belief in perfect information.

5 Uncertain information structure and backward induction

Perea (2014) proves that for extensive-form games without uncertainty in the information structure, rationality and common belief in opponents' future rationality[17] induce strategy profiles that survive what he defines as the *backward dominance procedure*, which in the case of games with perfect information and no relevant ties, are outcome equivalent with the unique backward induction profile. Since we deal with uncertainty in the information structure, and thus require a richer description of the definition of strategies, we cannot generalise Perea's result to our set-up in a very straightforward way, without any kind of assumption regarding the information structure.

Recall that a strategy profile in terms of histories $\left((s_h)_{h\in\mathcal{H}_i}\right)_{i\in I}$ is called *inductive* if it satisfies that

$$s_h \in \underset{a_h\in A_i(h)}{\operatorname{argmax}}\; u_i\left(z\left(s_{-i};(s_i,a_h)\,|h\right)\right)$$

for any $h\in\mathcal{H}_i$ and any player i. For a tree with no relevant ties this profile is unique, and we denote it by β. We define the inductive outcome of Γ as $z_{\mathcal{I}} = z\left(\beta\,|\emptyset\right)$ and the event that *the inductive outcome obtains*, as $BIO = \bigcap_{h<z_{\mathcal{I}}}[s_h=\beta_h]$. For each $h\in\mathcal{H}$ we refer to β_h as the *inductive choice at history* h. As said above, it is known that for games with perfect information rationality and common belief in opponents' future rationality induce the backward induction outcome. As seen informally in paragraph 1.1, this result seems to break down when we drop the assumption that the game has perfect information. Theorem 1 shows that under some assumptions, when there is uncertainty about the information structure, perfect information is not necessary:

Theorem 1 (Sufficiency for arbitrary type structures). *Let game tree with no relevant ties* Γ *and arbitrary type structure* T. *Then, if players are rational, there is strong belief in both opponents' rationality and opponents' perfect information, and there is common belief in both opponents' future rationality and opponents' future perfect information, the backward induction outcome obtains; i.e.,*

$$R\cap SBORPI\cap CBOFRPI\subseteq BIO.$$

Proof. In the following proof, we denote: (i) for any $h\in\mathcal{H}$, the event that *history* h *is reached*: $[h]=\left\{\omega\in\Omega\,\middle|\,h<z\left(s_{v(\omega)}(\omega)\right)\right\}$, and (ii) for any $i\in I$, player i's *set of pre-terminal histories*, $\mathcal{Z}_i=\left\{h\in\mathcal{H}_i\,|(h,a_i)\in\mathcal{Z}\text{ for any }a_i\in A_i(h)\right\}$, and $\mathcal{Z}_{-i}=\bigcup_{j\neq i}\mathcal{Z}_j$. Now we proceed in three steps:

A small lemma. Let $i\in I$, $\sigma_i\in\Sigma_i$ and $h\in v_i(\sigma)$. Consider now a state $\omega\in R_{\sigma_i}\cap B_i\left([h]\cap\bigcap_{h'>h,h'\in\mathcal{H}_{-i}}[s_{h'}=\beta_{h'}]\,\middle|\,\sigma_i\right)$. Note that for any strategy $s_i\in S_i(\sigma_i)$, it holds that $u_i\left(e_i(\omega),s_i\,|\sigma_i\right)=u_i\left(z\left(\beta_{-i},s_{v_i(\omega)}\,|h\right)\right)$, and thus, since $\omega\in R_{\sigma_i}$, $s_{\sigma_i}(\omega)=\beta_h$. This way, we

[17] *Common belief in opponents' future rationality*, $CBOFR$, is defined in a way analogous to $CBOFRPI$, by just replacing each $FR_{-i}(\sigma)\cap FPI_{-i}(\sigma_i)$ by the corresponding $FR_{-i}(\sigma_i)$.

conclude that for any $i \in I$, any $\sigma_i \in \Sigma_i$ and any $h \in v_i(\sigma_i)$,

$$R_{\sigma_i} \cap B_i \left([h] \cap \bigcap_{h' > h, h' \in \mathcal{H}_{-i}} [s_{h'} = \beta_{h'}] \middle| \sigma_i \right) \subseteq [s_{\sigma_i} = \beta_h] .$$

A BACKWARD FLOW. Let's proceed by induction: let $i \in I$, $h \in \mathcal{Z}_i$ and $\sigma_i \in \Sigma_i$ such that $v_i(\sigma_i) = \{h\}$. Then, we have that $u_i(e_i, s_i | \sigma) = u_i((h, s_{\sigma_i}))$ for any $e_i \in E_i$ and $s_i \in S_i$, and therefore, that $R_h \subseteq [s_{\sigma_i} = \beta_h]$. Thus, $BOFRPI_i \subseteq \bigcap_{\sigma_i \in \Sigma_i} B_i \left(\bigcap_{h > \sigma_i, h \in \mathcal{Z}_{-i}} [s_h = \beta_h] | \sigma_i \right)$ for any $i \in I$ and $\sigma_i \in \Sigma_i$, and consequently, $CBOFRPI_i \subseteq \bigcap_{\sigma_i \in \Sigma_i} B_i \left(\bigcap_{h > \sigma_i, h \in \mathcal{Z}_{-i}} [s_h = \beta_h] | \sigma_i \right)$.

Now, let $i \in I$ and $\sigma_i \in \Sigma_i$ such that for any $j \in I$ and any $\sigma_j \in \Sigma_j$ such that $\sigma_j > \sigma_i$ it holds that $CBOFRPI_j \subseteq \bigcap_{h > \sigma_j, h \in \mathcal{H}_{-j}} B_j(s_h = \beta_h | \sigma_j)$.[18] Then, since $CBOFRPI_i \subseteq B_i(CBOFRPI_{-i} | \sigma_i)$, from the induction hypothesis we get that

$$CBOFRPI_i \subseteq B_i \left(\bigcap_{j \neq i} \bigcap_{\sigma_j \in \Sigma_j} R_{\sigma_j} \cap PI_j \cap B_j \left(\bigcap_{h > \sigma_j, h \in \mathcal{H}_{-j}} [s_h = \beta_h] \middle| \sigma_j \right) \middle| \sigma_i \right),$$

and therefore, because of the small lemma,[19] $CBOFRPI_i \subseteq B_i \left(\bigcap_{h > \sigma_i, h \in \mathcal{H}_{-i}} [s_h = \beta_h] \middle| \sigma_i \right)$. This way, we conclude that $CBOFRPI_i \subseteq \bigcap_{\sigma_i \in \Sigma_i} B_i \left(\bigcap_{h > \sigma_i, h \in \mathcal{H}_{-i}} [s_h = \beta_h] \middle| \sigma_i \right)$ for any $i \in I$.

A FORWARD FLOW. Note first that from all the above, for any $i \in I$ and $\sigma_i \in \Sigma_i$ such that $v_i(\sigma_i) = \{h\}$ for some $h \in \mathcal{H}_i$,

$$R_{\sigma_i} \cap PI_h \cap CBOFRPI_i \subseteq R_{\sigma_i} \cap B_i \left([h] \cap \bigcap_{h' > \sigma_i, h' \in \mathcal{H}_{-i}} [s_{h'} = \beta_{h'}] \middle| \sigma_i \right) \subseteq [s_h = \beta_h] .$$

Let $z_{\mathcal{I}} = \left(\emptyset, \left(\beta^k \right)_{k=0}^n \right)$, $h^0 = \emptyset \in \mathcal{H}_{i_0}$, and for any $k = 1 \ldots n$, $h^k = \left(h^{k-1}, \beta^{k-1} \right) \in \mathcal{H}_{i_k}$. Since $PI_\emptyset = \Omega$, it is immediate that $R_{i_0} \cap CBOFRPI_{i_0} \subseteq \left[s_{h^0} = \beta^0 \right]$.

Now, let $k \geq n$ such that for any $l < k$, $R_{i_l} \cap SBORPI_{i_l} \cap CBOFRPI_{i_l} \subseteq \left[s_{h^l} = \beta^l \right]$. Let $\sigma_{i_k} \in \Sigma_{i_k}$ such that $h^k \in v_{i_k}(\sigma_{i_k})$. Since by definition we have that $SBORPI_{i_k} \subseteq B_i \left(\bigcap_{l=0,\ldots,k, i_l \neq i_k} \bigcap_{\sigma_{i_l} \in \Sigma_{i_l}, h_l \in v_{i_l}(\sigma_{i_l})} R_{\sigma_{i_l}} \cap PI_{h^l} \middle| \sigma_{i_k} \right)$, and it is known that $CBOFRPI_{i_k} \subseteq B_i \left(\bigcap_{l=0,\ldots,k, i_l \neq i_k} CBOFRPI_l \middle| \sigma_{i_k} \right)$, it is easy to see that $R_{i_k} \cap SBORPI_{i_k} \cap CBOFRPI_{i_k} \subseteq B_{i_k} \left(\left[h^k \right] | \sigma_{i_k} \right)$. In addition, since $CBOFRPI_{i_k} \subseteq B_i \left(\bigcap_{h > h^k, h \in \mathcal{H}_{-i_k}} [s_h = \beta_h] \middle| \sigma_{i_k} \right)$, we obtain that $R_{i_k} \cap SBORPI_{i_k} \cap CBOFRPI_{i_k} \subseteq \left[s_{h^k} = \beta^k \right]$.

Thus, we conclude that $R \cap SBORPI \cap CBOFRPI \subseteq \bigcap_{h < z_{\mathcal{I}}} [s_h = \beta_h] = BIO$. □

[18] The case above ensures the existence of such.

[19] Just note that for any $i \in I$, $\sigma_i \in \Sigma_i$ and $h \in \mathcal{H}_i$ such that $v_i(\sigma_i) = \{h\}$, we have both that (i) $B_i([h] \cap W | \sigma_i) = B_i(W | \sigma_i)$ for any $W \subseteq \Omega$, and (ii) $PI_h \cap [s_{\sigma_i} = \beta_h] \subseteq [s_h = \beta_h]$.

The intuition behind the assumptions in Theorem 1 and its proof can be briefly explained as follows:

(*i*) For each player i, $CBOFRPI_i$ serves as a mechanism to conjecture opponents' behaviour. The fact that she believes that after her choice the game will take the shape of one with perfect information and common belief in rationality leads i to the hypothesis that after her choice everybody will choose inductively. Note though, that if the information set she finds herself at is not a singleton, her own rationality and having a deterministic conjecture about her opponents' future behaviour is still not enough for her to make a choice: since she does not know which history she finds herself at, she does not know where each of her actions will lead. But...

(*ii*) ...note the following: since $CBOFPR_i$ implies that she believes in $CBOFRPI_j$ for any j choosing previous to her, she has beliefs on what j expects to happen in the future, because of $SBORPI_i$, she believes that j's information set is just a singleton, so that j is able to evaluate where each of her actions lead, and because of $SBORPI_i$ again, i is able to conjecture what j actually chose previous to her. This kind of reasoning, as long as evidence against $R_{-i} \cap PI_{-i}$ is not found, enables i to infer what everyone previous to her chose, and thus, conjecture what unique history she finds herself at inside her information set. Thus, she can...

(*iii*) ...predict, because she has beliefs in her opponents' future behaviour at each of their histories, what outcome each of her available actions leads to. Thus, since the fact game has no relevant ties determines a unique optimal choice, rationality implies only one choice for her.

(*iv*) Furthermore, note that the assumptions in Theorem 1 do not imply perfect information at all; in principle, it might be the case that each player is trapped in a *black box* so that she does not observe any other players' choices. Still, her beliefs about opponents' rationality and perfect information together with her own rationality induce inductive behaviour upon her.

In particular, if we assume that the information structure is not uncertain and it corresponds to the case of perfect information, Theorem 1 yields the following corollary:

Corollary 1 (cf. Theorem 5.4 in Perea (2014)). *Let game tree with no relevant ties* Γ *and type structure* T *such that the game has perfect information at every state, that is, such that* $PICBPI = \Omega$. *Then, if players are rational and there is common belief in opponents' future rationality, the backward induction outcome obtains; i.e.,*

$$R \cap CBOFR \subseteq BIO.$$

We omit the proof of the corollary, which is immediate given Theorem 1. Just note that whenever uncertainty about the information structure of the game is removed and players observe each others' choices, we get exactly the assumption and result attained by Perea (2014) for games with perfect information. Now, it is pertinent to wonder whether the assumptions in Theorem 1 are non trivial; that is, if it might be the case that there is some game tree Γ in which BIO fails to obtain under our assumptions due to the fact that $R \cap SBORPI \cap CBOFRPI$ is indeed empty for any type structure T. The following theorem shows that we can always construct a type structure such that this is not the case:

Theorem 2 (Non vacuity). *For any game tree with no relevant ties Γ and any information structure V, there exists some type structure T such that the event that players are rational, there is strong belief in both opponents' rationality and opponents' perfect information, there is common belief in both opponents' future rationality and opponents' future perfect information, and V obtains is not empty; i.e., such that,*

$$R \cap SBORPI \cap CBOFRPI \cap [v = V] \neq \emptyset.$$

Proof. We proceed by construction. First, for each $i \in I$ and $\sigma_i \in \Sigma_i$, let $h_{\sigma_i} \in v_i(\sigma_i)$ such that if $\sigma_i < z_{\mathcal{I}}$, then $h_{\sigma_i} < z_{\mathcal{I}}$, and set $\alpha_{\sigma_i} = \beta_{h_{\sigma_i}}$ for any $\sigma_i \in \Sigma_i$. Now, for each $i \in I$ take $(b_i(\sigma_i))_{\sigma_i \in \Sigma_i} \in \prod_{\sigma_i \in \Sigma_i} \Delta\left(\left(S_{-i} \times \mathcal{V}\right)(\sigma_i)\right)$ such that:

(*i*) $b_i(\sigma_i)\left[\left(S_{-i} \times \left\{V_{-i}^{PI}\right\} \times \mathcal{V}_i\right)(\sigma_i)\right] = 1$ for any $\sigma_i \in \Sigma_i$.

(*ii*) $b_i(\sigma_i)\left[\left(\left\{s_{-i} \in S_{-i} \left| s_{\sigma_j} = \alpha_{\sigma_j} \text{ for any } j \neq i, \sigma_j > \sigma_i \right.\right\} \times \mathcal{V}\right)(\sigma_i)\right] = 1$, for any $\sigma_i \in \Sigma_i$.

(*iii*) $b_i(\sigma_i)\left[\left(\left\{s_{-i} \in S_{-i} \left| h_{\sigma_i} < z\left(s_{V_{-i}^{PI}}, s_{V_i}\right) \text{ for some } s_i \in S_i, V_i \in \mathcal{V}_i \right.\right\} \times \mathcal{V}\right)(\sigma_i)\right] = 1$, for any $\sigma_i \in \Sigma_i$.

(*iv*) $b_i(\sigma_i)\left[\left(\{\alpha_{-i}\} \times \mathcal{V}\right)(\sigma_i)\right] = 1$ for any $\sigma_i \in \Sigma_i$ such that $\left(\{\alpha_{-i}\} \times \mathcal{V}\right)(\sigma_i) \neq \emptyset$.

Now, for any $i \in I$, let epistemic type space $E_i = \{e_i\}$, and conditional belief system map b_i where $b_i(e_i, \sigma_i)\left[(e_{-i}, s_{-i}, V)\right] = b_i(\sigma_i)\left[(s_{-i}, V)\right]$ for any $\sigma_i \in \Sigma_i$ and any $(s_{-i}, V) \in \left(S_{-i} \times \mathcal{V}\right)(\sigma_i)$. Let type structure $\mathrm{T} = \langle E_i, b_i \rangle_{i \in I}$.[20] Note that for any $i \in I$,

(*i*) $B_i\left(PI_{-i} \left| \sigma_i \right.\right) = \Omega$ for any $\sigma_i \in \Sigma_i$.

(*ii*) $B_i\left(\bigcap_{j \neq i} \bigcap_{\sigma_j \in \Sigma_j, \sigma_j > \sigma_i} \left[s_{\sigma_j} = \alpha_{\sigma_j}\right] \middle| \sigma_i\right) = \Omega$ for any $\sigma_i \in \Sigma_i$.

(*iii*) $B_i\left(\left[h_{\sigma_i}\right] \left| \sigma_i \right.\right) = \Omega$ for any $\sigma_i \in \Sigma_i$.

(*iv*) $B_i\left(s_{-i} = \alpha_{-i} \middle| \sigma_i\right) = \Omega$ for any $\sigma_i \in \Sigma_i$ such that $\left(\{\alpha_{-i}\} \times \mathcal{V}\right)(\sigma_i) \neq \emptyset$.

Note in addition that since for any $i \in I$, $\omega \in \Omega$ and $\sigma_i \in \Sigma_i(s_i(\omega))$, $u_i\left(e_i(\omega), s_i(\omega) \left| \sigma_i \right.\right) = u_i\left(z\left(\alpha_{V_{-i}^{PI}}, s_{V_i(\omega)}(\omega) \left| h_{\sigma_i} \right.\right)\right)$, then $\omega \in R_i$ if and only if $\omega \in [s_i = \alpha_i]$. That is, $R_i = [s_i = \alpha_i]$. Now, we want to prove that $\{(e, \alpha)\} \times \mathcal{V} \subseteq R \cap SBORPI \cap CBOFRPI \cap$. Let $\omega \in \{(e, \alpha)\} \times \mathcal{V}$. Then:

- We already checked that $\omega \in R$.

[20] Required topological assumptions are trivially satisfied due to E_i being finite.

- Since $\bigcap_{\sigma_i \in \Sigma_i} \left(B_i \left(PI_{-i} \,|\sigma_i \right) \cap B_i \left(\bigcap_{j \neq i} \bigcap_{\sigma_j \in \Sigma_j, \sigma_j > \sigma_i} \left[s_{\sigma_j} = \alpha_{\sigma_j} \right] \middle| \, \sigma_i \right) \right) = \Omega$, then we have that $BOFRPI_i = \Omega$, and in consequence, $CBOFRPI = \Omega$. Thus, $\omega \in CBOFRPI$.
- Let $\sigma_i \in \Sigma_i \left(R_{-i} \cap PI_{-i} \cap CBOFRPI_{-i} \right) = \Sigma_i \left(R_{-i} \cap PI_{-i} \right)$. This is so, if and only if $\sigma_i \in \Sigma_i \left(\left[s_{-i} = \alpha_{-i}, \mathcal{V}_{-i} = V_{-i}^{PI} \right] \right)$. Then, $\left(\{\alpha_{-i}\} \times \mathcal{V} \right) (\sigma_i) \neq \emptyset$, and therefore, it holds that $B_i \left(s_{-i} = \alpha_{-i} \,|\sigma_i \right) = \Omega$. Since $B_i \left(PI_{-i} \,|\sigma_i \right) = \Omega$, we conclude that $\omega \in SBORPI$, because indeed, $SBORPI = \Omega$.

Thus, T is a type structure for Γ such that $R \cap SBORPI \cap CBOFRPI \cap [v = V] \neq \emptyset$ for any $V \in \mathcal{V}$. □

6 Final Remarks

A. Summary. The present work tries to extend the analysis of sufficient epistemic conditions for the backward induction outcome of generic extensive-form games, from the perfect information case to that with not necessarily perfect information. The main features and conclusions are:

(*i*) We introduce uncertainty about the information structure of the game. Issues concerning players' ability to observe each others' choices in general, and perfect information in particular, are approached *via* an epistemic framework based on type spaces that relies on standard tools in the fields of both in epistemic game theory and analysis of Bayesian games.

(*ii*) Theorem 1 shows that, if players are rational (R), they strongly believe in both opponents' rationality and opponents' perfect information ($SBORPI$), and they commonly believe in both opponents' future rationality and opponents' future perfect information ($CBOFRPI$), then neither common knowledge of perfect information, nor even perfect information is required to obtain the backward induction outcome. In particular, the backward inductive outcome is obtained under these assumptions, even if it is the case that every player is trapped in a *black box* and does not observe any of their opponents' choices.

(*iii*) Theorem 2 shows that the epistemic requirements for the backward inductive outcome to obtain are not trivial: it is always possible to construct a type structure such that $R \cap SBORPI \cap CBOFRPI$ is non empty, or, in words, such that the three conditions are simultaneously satisfied at some state. Moreover, the assumptions are consistent with any information structure the game might happen to have, not just perfect information.

(*iv*) The type structure that defines the epistemic model is not assumed to be complete, so any implicit assumption about players' beliefs about other players' beliefs is consistent with the results.

B. (Non-)Robustness of backward induction. Theorem 1 introduces sufficient epistemic conditions for the backward induction outcome for any game tree, regardless of the factual information structure the game happens to have as it is played. These epistemic assumptions are not very far from those assumed for the case of perfect information and common belief in opponents' perfect information ($PICBPI$), which is what we call perfect information in standard contexts of lack of uncertainty about the information structure. Thus, Theorem 1 addresses robustness properties of the backward induction outcome in two somewhat opposite ways: (i) it proves that the backward induction outcome is robust to shocks in the information structure of the game, as long as these shocks do not affect players' beliefs on their opponents' information structures, and, (ii) it suggests that players believing in opponents' perfect information plays a crucial role in the backward induction outcome being obtained, so that the latter is found very sensitive, *i.e.*, non-robust, to changes in beliefs.

C. Bayesian updating. It is not assumed that as the game progresses, players update their beliefs by Bayesian conditioning, that is, by conditioning previous beliefs to newly unveiled information. This is an assumption we drop from the standard definition of conditional belief systems that can be found on Renyi (1955) or Ben Porath (1997), for instance. Belief revision procedure is free of constraints in the present set-up. Still, were we to impose the condition that beliefs are revised following a Bayesian updating procedure, this would be done by assuming that type structure T is such that the following is satisfies for any player i, any epistemic type e_i and any information sequence σ_i,

$$b_i\left(e_i, \sigma_i'\right)\left[\left(e_{-i}, s_{-i}, V\right)\right] = \frac{b_i\left(e_i, \sigma_i\right)\left[\left(e_{-i}, s_{-i}, V\right)\right]}{b_i\left(e_i, \sigma_i\right)\left[E_{-i} \times \left(S_{-i} \times \mathcal{V}\right)\left(\sigma_i'\right)\right]},$$

for any $\left(e_{-i}, s_{-i}, V\right) \in E_{-i} \times \left(S_{-i} \times \mathcal{V}\right)\left(\sigma_i'\right)$ and any information sequence σ_i' that follows σ_i,[21] and such that $b_i\left(e_i, \sigma_i\right)\left[E_{-i} \times \left(S_{-i} \times \mathcal{V}\right)\left(\sigma_i'\right)\right] > 0$. In any case, results of Theorem 1 and 2 would remain unaffected.

D. Priors, delusion and belief-consistency. Since it is assumed that R holds, there are no delusion issues involved when we assume that there is strong belief in opponents' rationality and there is common belief in opponents' future rationality. But since we do not assume PI to be satisfied, when we ask for strong belief in opponents' perfect information and common belief in opponents' future perfect information, it might be the case that players hold wrong beliefs; that is, *when PI is not satisfied, our epistemic assumptions are not consistent with the Truth Axiom*, which in the current framework is equivalent to players assigning non null probability to the true state of the world.

All the analysis is carried out at interim level and each player's beliefs are specified only at information sets of her own. Since Bayesian updating is not assumed, it could be the case that a player holds mutually inconsistent beliefs at two of her information sets, one following the other; and it might also be the case that if we allow for comparing different players' beliefs at each of their own different information sets, these beliefs are inconsistent. This inconsistency is not fully structural, though: it is possible to construct type structures where $R \cap SBORPI \cap CBOFRPI \neq$

[21] Formally this is means that (i) $v_i\left(\sigma_i'\right) > v_i\left(\sigma_i\right)$, and ($ii$) for any $V_i \in \mathcal{V}_i$, $\sigma_i' \subseteq V_i$ implies $\sigma_i \subseteq V_i$.

$\emptyset$ and players' beliefs are derived from common lexicographic priors.

E. Minimal epistemic condition. It is not clear to us whether some other condition, less restrictive than each player i holding beliefs related to all the rest of players having perfect information, could be enough to generally imply, together with the rest of assumptions about rationality, the backward induction outcome. It is easy to check that if so, this less restrictive condition must of course, be more restrictive than just perfect information (read as *perfect information but* NOT *common belief in opponents' perfect information*).

F. Absence of relevant ties. The fact that the game trees under consideration have no relevant ties is crucial for Theorem 1. Indeed, if the game tree had more than just one backward induction outcome, it would be impossible to infer opponents' past choices at some non-singleton information sets and consequently, induce which choice is actually the inductive one. Consider an uncertain imperfect information game such as in Figure 2 and modify the payoffs so that every non-null payoff is exactly 1. Assume in addition that the conditions in Theorem 1 are satisfied; in this case these reduce to: rationality (R), Alexei's belief in Polina's future rationality and perfect information (implicit in $CBOFRPI_A$) and both Polina's strong belief in Alexei's rationality (implicit in $SBORPI_P$) and her belief in $BOFRPI_A$ (implicit in $CBOFRPI_P$). Since Alexei believes both that Polina is rational and has perfect information, he believes that any choice of his yields him 1. He is therefore indifferent and may choose either left or *right*. Thus, if it is the case that Alexei is deluded and Polina has no perfect information, there is nothing she can infer from $SBORPI_P \cap CBOFRPI_P$, and despite being rational, finds no reason to expect *left* or *right* yielding her a higher payoff than the alternative.

G. Epistemic types and belief hierarchies. The relation between the two fundamental ways of encoding interactive beliefs, namely type spaces and belief hierarchies, has stood as one of the foundational concerns of epistemic game theory since the pioneering work by Harsanyi (1967–1968), and posterior development by Armbruster and Böge (1979), Böge and Eisele (1979), Mertens and Zamir (1985) or Brandenburger and Dekel (1993) among others. Covering such fundamental issues is out of the scope of the present paper; however, we present a brief sketch in what follows for the purposes of completion and self-containment.

First, for any player i and any information sequence σ_i, we set basic conditional uncertainty space $X_i(\sigma_i^0) = (S_{-i} \times \mathcal{V})(\sigma_i)$; higher order uncertainty spaces are recursively obtained as follows: $X_i^n(\sigma_i) = X_i^{n-1}(\sigma_i) \times \prod_{j \neq i} \prod_{\sigma_j > \sigma_i} \Delta\left(X_j^{n-1}(\sigma_j)\right)$ for any $n \in \mathbb{N}$. Then, given type structure $\mathrm{T} = \langle E_i, b_i \rangle_{i \in I}$, type e_i's *conditional belief hierarchy* after information sequence σ_i can be recursively defined the following way: let first order belief $\varepsilon_{i,0}\left(e_i \left| \sigma_i \right.\right) = \mathrm{marg}_{(S_{-i} \times \mathcal{V})(\sigma_i)} b_i\left(e_i, \sigma_i\right)$, and for each $n \in \mathbb{N}$, let higher uncertainty spaces and higher order beliefs:

$$\varepsilon_{i,n}\left(e_i \left| \sigma_i \right.\right)\left[\left(\left(\mu_{-i,k}\right)_{k=0}^{n-1}, \left(s_{-i}, V\right)\right)\right] = b_i\left(e_i, \sigma_i\right)\left[\left\{\left(s_{-i}, V\right)\right\} \times \prod_{j \neq i} \bigcap_{\sigma_j > \sigma_i} \bigcap_{k=0}^{n-1} \varepsilon_{j,k}^{-1}\left(\mu_{\sigma_j,k} \left| \sigma_j \right.\right)\right],$$

for any and $\left(\left(\left(\mu_{\sigma_j,k}\right)_{k=0}^{n-1}\right)_{\sigma_j>\sigma_i}, (s_{-i}, V)\right) \in \Delta\left(X_i^n\right)$. This way, we obtain conditional belief hierarchy $\varepsilon_i(e_i\,|\sigma_i\,) = (\varepsilon_{i,n}(e_i\,|\sigma_i\,))_{n\geq 0}$. Furthermore, we conjecture that following this procedure and employing techniques similar to those by Brandenburger and Dekel (1993) and Battigalli and Siniscalchi (1999), it is possible to give a proper definition of a universal $(S_{-i} \times \mathcal{V})_{i\in I}$-*based type space* $(\mathcal{E}_i, \varphi_i)_{i\in I}$,[22] where for any player i and any information sequence σ_i, $\mathcal{E}_i^0(\sigma_i) = \prod_{n\geq 0} \Delta\left(X_i^n\right)$ is homeomorphic to $\Delta\left(\prod_{j\neq i}\prod_{\sigma_j>\sigma_i} \mathcal{E}_j^0(\sigma_j) \times (S_{-i} \times \mathcal{V})\,(\sigma_i)\right)$, and where standard assumptions such as common belief in coherency or Bayesian updating can be imposed.

References

Akerlof, George A. (1970). "The market for "lemons": quality uncertainty and the market mechanism". *The Quarterly Journal of Economics* **84**, 488–500.

Arieli, Itai and Robert J. Aumann (2013). "The logic of backward induction". Center for the Study of Rationality, Discussion Paper # 652.

Armbruster, W. and W. Böge (1979). "bayesian Game Theory". In: *Game theory and related topics.* North Holland, Amsterdam.

Aumann, Robert J. (1995). "Backwards induction and common knowledge of rationality". *Games and Economic Behavior* **8**, 6–19.

Aumann, Robert J. (1998). "On the centipede game". *Games and Economic Behavior* **23**, 97–105.

Baltag, Alexandru, Sonja Smets and Jonathan A. Zvesper (2009). "Keep hoping for rationality: a solution for the backward induction paradox". *Synthese* **169**, 201–303.

Battigalli, Pierpaolo (1997). "On rationalizability in extensive form games". *Journal of Economic Theory* **74**, 40–61.

Battigalli, Pierpaolo, Alfredo Di Tillio and Dov Samet (2013). *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congres,, Volume I, Economic Theory.* Chap. "Strategies and interactive beliefs in dynamic games". Cambridge University Press.

Battigalli, Pierpaolo and Amanda Friedenberg (2012). "Forward induction reasoning revisited". *Theoretical Economics* **7**, 57–98.

Battigalli, Pierpaolo and Marciano Siniscalchi (1999). "Hierarchies of conditional beliefs and interactive epistemology in dynamic games". *Journal of Economic Theory,* **88**, 188–230.

Battigalli, Pierpaolo and Marciano Siniscalchi (2002). "Strong belief and forward induction reasoning". *Journal of Economic Theory* **106**, 356–391.

[22] In a topological sense: this is the reason why in Definition 2 we considered compact type spaces and continuous conditional belief maps.

Battigalli, Pierpaolo and Marciano Siniscalchi (2007). "Interactive epistemology in games with payoff uncertainty". *Research in Economics* **61**, 165–184.

Ben Porath, Elchanan (1997). "Rationality, Nash equilibrium and backward induction in perfect information games". *The Review of Economic Studies* **64**, 22–46.

Böge, W. and Th. Eisele (1979). "On solutions of Bayesian games". *The International Journal of Game Theory* **8**, 193–215.

Bonanno, Giacomo (2013). "A dynamic epistemic characterization of backward induction without counterfactuals". *Games and Economic Behavior* **78**, 31–43.

Brandenburger, Adam and Eddie Dekel (1993). "Hierarchies of Beliefs and Common Knowledge". *Journal of Economic Theory* **59**, 189–198.

Dekel, Eddie and Marciano Siniscalchi (2013). "Epistemic game theory". Mimeo.

Ely, Jeffrey C. and Juuso Valimaki (2003). "Bad reputation". *The Quarterly Journal of Economics* **118**, 785–814.

Ely, Jeffrey C., Drew Fudenberg and David K. Levine (2008). "When is reputation bad?". *Games and Economic Behavior* **63**, 498–526.

Harsanyi, John C. (1967–1968). "Games with incomplete information played by 'Bayesian' players, I–III". *Management Science* **14**, 159–182, 320–334, 486–502.

Heifetz, Aviad and Andrés Perea (2013). "On the outcome equivalence of backward induction and extensive form rationalizability". Mimeo.

Kreps, David M. and Robert Wilson (1982). "Reputation and imperfect information". *Journal of Economic Theory* **27**, 253–279.

Mertens, Jean-François and Shmuel Zamir (1985). "Formulation of Bayesian analysis for games with incomplete information". *International Journal of Game Theory* **14**(1), 1–29.

Milgom, Paul and John Roberts (1982). "Predation, reputation and entry deterrence". *Journal of Economic Theory* **27**, 280–312.

Osborne, Martin J. and Ariel Rubinstein (1994). *A course in game theory*. MIT Press Books.

Pearce, David G. (1984). "Rationalizable strategic behavior and the problem of perfection". *Econometrica* **52**, 1029–1050.

Penta, Antonio (2011). "Backward induction reasoning in games with incomplete information". Mimeo. University of Winsonsin-Madison.

Penta, Antonio (2012). "Higher order uncertainty and information: static and dynamic games". *Econometrica* **80**, 631–660.

Perea, Andres (2007). "Epistemic foundations for backward induction: an overview". In: *Interactive Logic. Proceedings of the 7th Augustus de Morgan Workshop, Volume 1 of Texts in Logic and Games* (J. van Benthem, D. Gabbay and B. Löwe, Eds.). pp. 159–193. Amsterdam University Press.

Perea, Andrés (2014). "Belief in the opponent's future rationality". *Games and Economic Behavior* **81**, 235–254.

Reny, Philipp J. (1992). "Backward induction, normal form perfection and explicable equilibria". *Econometrica* **60**, 627–649.

Renyi, Alfred (1955). "On a new axiomatic theory of probability". *Acta Mathematica Hungarica* **6**, 285–335.

Rosenthal, Robert W. (1981). "Games of perfect information, predatory pricing and chain-store paradox". *Journal of Economic Theory* **25**, 92–100.

Samet, Dov (1996). "Hypothetical knowledge and games with perfect information". *Games and Economic Behavior* **17**, 230–251.

Samet, Dov (2013). "Common belief of rationality in games with perfect information". *Games and Economic Behavior* **79**, 192–200.

Spence, Michael (1973). "Job market signalling". *The Quarterly Journal of Economics* **87**, 355–374.

Zuazo-Garin, Peio (2013). "Incomplete imperfect information and backward induction". Mimeo.