

PRUDENT RATIONALIZABILITY IN GENERALIZED EXTENSIVE-FORM GAMES WITH UNAWARENESS*

Aviad Heifetz[†] Martin Meier[‡] Burkhard C. Schipper[§]

This Version: January 24, 2019

First Version: March 30, 2011

Abstract

We define a cautious version of extensive-form rationalizability for generalized extensive-form games with unawareness that we call prudent rationalizability. It is an extensive-form analogue of iterated admissibility. In each round of the procedure, for each tree and each information set of a player a surviving strategy of hers is required to be rational vis-a-vis a belief system with a *full-support* belief on the opponents' previously surviving strategies that reach that information set. We demonstrate the applicability of prudent rationalizability. In games of disclosure of verifiable information, we show that prudent rationalizability yields unraveling under full awareness but unraveling might fail under unawareness. We compare prudent rationalizability to extensive-form rationalizability. We show that prudent rationalizability may not refine extensive-form rationalizability strategies but conjecture that the paths induced by prudent rationalizable strategy profiles (weakly) refine the set of paths induced by extensive-form rationalizable strategies.

Keywords: Caution, extensive-form rationalizability, unawareness, disclosure, verifiable information, persuasion games, iterated admissibility, common strong cautious belief in rationality.

JEL-Classifications: C72, D83.

*The first part of the project was published in "Dynamic Unawareness and Rationalizable Behavior", *Games and Economic Behavior*, 2013. We are grateful to Pierpaolo Battigalli for numerous insightful comments and suggestions and to Elchanan Ben Porath for helpful discussions and exchange of ideas. We thank Paulo Barelli, Andy McLennan, Yoram Halevy as well as seminar participants at Barcelona, Bocconi, Caltech, Maryland, Maastricht, Pittsburgh, USC, UC Davis, Tel Aviv, Stony Brook 2007, LOFT 2008, Games 2008 and NSF/NBER/CEME 2009. Aviad is grateful for financial support from the Open University of Israel's Research Fund grant no. 46106. Martin was supported by the Spanish Ministerio de Educación y Ciencia via a Ramon y Cajal Fellowship (IAE-CSIC) and a Research Grant (SEJ 2006-02079). Burkhard is grateful for financial support from the NSF SES-0647811.

[†]The Economics and Management Department, The Open University of Israel. Email: aviadhe@openu.ac.il

[‡]Department of Economics, University of Bath, and Institut für Höhere Studien, Wien. Email: m.meier@bath.ac.uk

[§]Department of Economics, University of California, Davis. Email: bcschipper@ucdavis.edu

1 Introduction

Recently various frameworks for modeling dynamic games with unawareness have been introduced (Halpern and Rego, 2014, Rego and Halpern, 2012, Feinberg, 2012, Grant and Quiggin, 2013, Heifetz, Meier, and Schipper, 2013, Schipper, 2018a, b; for a non-technical survey, see Schipper, 2014). While all of these frameworks are capable of modeling asymmetric awareness among players and endogenous dynamic change of awareness during play, the solution concepts proposed for these frameworks and thus the implicit behavioral assumptions under unawareness differ. The solution concepts can be roughly divided into equilibrium notions (Halpern and Rego, 2014, Rego and Halpern, 2012, Feinberg, 2012, Grant and Quiggin, 2013, Ozbay, 2007, Filiz-Ozbay, 2012, Meier and Schipper, 2014) and rationalizability (Heifetz, Meier, and Schipper, 2013, Guarino, 2017, Schipper and Woo, 2019, Li and Schipper, 2018b, Schipper, 2018b, 2016, Perea, 2018a). As pointed out in Heifetz, Meier, and Schipper (2013) and Schipper (2018b), equilibrium notions that are mere extensions of the mathematical definitions of equilibrium in standard games to the more sophisticated frameworks with unawareness are problematic. Since awareness may change along the “equilibrium” path, behavior cannot be interpreted as a steady state. This is at odds with the common implicit equilibrium assumption of mutual knowledge of play. Where should mutual knowledge of play come from when it could not have been learned previously? In contrast, rationalizability does not feature this implicit assumption. It seeks to answer the question about what behavioral constraints are implied by higher order belief in the rationality.

Heifetz, Meier, and Schipper (2013) extended extensive-form rationalizability of Pearce (1984) and Battigalli (1997) to generalized extensive-form games with unawareness. It entails a notion of forward induction, a feature of special interest in dynamic games with unawareness: If a player raises another player’s awareness of an action, then it is plausible to assume that latter player reasons why the other player made her aware of that action and what it would imply for future behavior. While in many games, extensive-form rationalizability yields interesting restrictions on behavior, in some applications (see for instance Section 5) it is too permissive on beliefs that players can hold. In particular, those beliefs do not feature any form of “caution” or “prudence”. Sometimes this lack of constraints on beliefs does not permit ruling out strategies that could not be justified anymore even if only some tiny bit of caution is assumed.

In this paper, we extend a notion of extensive-form rationalizability featuring some form caution to generalized extensive-form games with unawareness. We call this solution concept prudent rationalizability. Like extensive-form rationalizability, at every information set, the active player looks for a best rationalization for the way this information set has been reached, and replies optimally to a belief over these best-rationalizable strategies. Yet, prudent rationalizability requires additionally that this belief has *full-support* on the opponents’ (recursively defined) prudently-best-rationalizable strategies. In Theorem 1 we prove that prudent ratio-

nalizable strategies exist in every finite dynamic game with perfect recall, including generalized extensive-form games with unawareness as introduced by Heifetz, Meier, and Schipper (2013). The significance of having a rationalizability concept featuring caution for games with unawareness is that it facilitates interesting applications in which strategic revelation of awareness and information plays an important role.

In Section 5 we exemplify the attractiveness of prudent rationalizability in games with disclosure of verifiable information (see Grossman, 1981, Milgrom, 1981, Grossman and Hart, 1980, and Milgrom and Roberts, 1986; see Milgrom, 2008, for a review). Information unravels in these models. Typically, this is proved using sequential equilibrium while an informal intuition is provided with an inductive argument. We show that prudent rationalizability formalizes the inductive intuition and yields the unraveling result in the limit. Yet, since prudent rationalizability is an inductive definition, it also provides predictions for every finite level mutual prudent belief in rationality. This has been useful in experiments. Li and Schipper (2018a) partially identify levels of rationing in experimental disclosure games using prudent rationalizability.

Full unraveling of information is somewhat unrealistic, though. In Section 5.2 we show that when unawareness is introduced into disclosure games, prudent rationalizability need not necessarily entail full information unraveling. Li and Schipper (2018b) find some support for this theoretical prediction in experimental disclosure games with unawareness. It demonstrates that in some games the presence or absence of unawareness may alter predictions dramatically. We also analyze a sender-receiver game with unawareness introduced by Ozbay (2007), and show that prudent rationalizability delivers the same prediction as does his equilibrium notion which incorporates forward-induction reasoning.

We believe that these two applications demonstrate the attractiveness of prudent rationalizability in dynamic games with unawareness. Indeed, in many games with unawareness prudent rationalizability rules out implausible extensive-form rationalizable strategies, with which a player makes an opponent aware of an action which the player would actually like the opponent to *avoid*, just because the player has a firm belief that the opponent would not take it (even if the opponent is indifferent between the revealed action and another one, of which she was aware also before); prudent rationalizability rules out such imprudent behavior.

Schipper and Woo (2019) apply theoretical insights of prudent rationalizability in disclosure games to electoral campaigning in which voters may not be aware of all political issues spanning the multi-dimensional policy space. Candidates campaign for votes by targeting messages to voters that raise political issues and disclose some information about the candidate's agenda. They show that political competition, microtargeting of voters, negative campaigning, and the political reasoning capabilities of voters entailed in forward induction of prudent rationalizability all play a role for unraveling in electoral campaigns.

Prudent rationalizability is not really new. Already Pearce (1984) presents a notion of

cautious extensive-form rationalizability that differs from ours in several respects and which he himself did not find satisfactory; see Section 4.3 for an example and further discussion. Brandenburger and Friedenberg (2007) study iterated weak dominance on the tree, a notion of conditional dominance à la Shimoji and Watson (1994) but featuring admissibility. They consider the associated normal-form and normal-form information sets associated with information sets in the tree. Strategies are eliminated iteratively when they are weakly dominated conditional on these normal-form information sets. They observe that iterated conditional weak dominance is equivalent to iterated (unconditional) weak dominance in the associated normal-form game. Prudent rationalizability is equivalent to iterated conditional weak dominance if instead of our local notion of optimizing over actions at information sets, optimizing over strategy-replacements at information sets is used. Meier and Schipper (2012) show these results formally for generalized extensive-form games with unawareness. They also show that in games with unawareness the relationship between prudent rationalizability to iterated weak dominance is more subtle. Games with unawareness consist of several normal-forms, one for each awareness level. In order to consider the awareness levels of players when eliminating strategies it is still necessary to condition on the appropriate normal-form when eliminating strategies by weak dominance.

We choose to define prudent rationalizability rather than using iterative elimination of weakly dominated strategies in the associated normal-form because our primary purpose is to apply it to dynamic games with unawareness for which - as we just mentioned - the associated normal-form is more involved. It also allows us to compare prudent rationalizability directly to extensive-form rationalizability. Finally, prudent rationalizability is a solution concept for dynamic games while iterative admissibility applies to the normal-form, a static concept. For “aesthetic” reasons, we prefer dynamic solution concepts for dynamic games. Moreover, we simply do not know yet to what extent it is justified to invoke the invariance axiom suggested in the literature on strategic stability for generalized games with unawareness.

In static games, iterated admissibility is a refinement of rationalizability. In the latter solution concept, at every round of elimination a player’s strategy survives only if it is a best response to some belief over the opponents’ strategies which survived the previous rounds, while in the former a strategy survives only if it is a best response to such a *full-support* belief, which does not completely exclude any strategy of the other players that has not been thus far eliminated. This follows from Pearce (1984, Lemma 4). In this paper we are concerned with the connection between the counterparts of these two notions in *dynamic games*, namely extensive-form rationalizability and prudent rationalizability.

Extensive-form rationalizability is particularly interesting because it may be used to refine the notion of sequential equilibrium (Pearce, 1984). Moreover, in generic perfect-information games, extensive-form rationalizability induces the unique backward-induction path (Reny,

1992, Battigalli, 1997, Chen and Micali, 2013, Heifetz and Perea, 2015, Perea, 2018b), even though the extensive-form rationalizable strategies may be distinct from the backward-induction strategies (Reny, 1992). Moreover, extensive-form rationalizability has a transparent epistemic characterization by common strong belief in rationality (Battigalli and Siniscalchi, 2002), a characterization that has recently been extended to dynamic games with unawareness (Guarino, 2017).

While normal-form games every iterative admissible strategy is also rationalizable, a similar inclusion does not obtain in dynamic games: In Section 4 we present an example of a game in which a player’s set of prudent rationalizable strategies is not a subset of her extensive-form rationalizable strategies but is rather *disjoint* from it. Nevertheless, we conjecture that inclusion does obtain in terms of *outcomes*. That is, we conjecture that the set of paths induced by prudent rationalizable strategy profiles is always contained in the set of paths induced by extensive-form rationalizable strategy profiles.

The paper is organized as follows: After outlining in Section 2 generalized extensive-form games with unawareness as introduced in Heifetz, Meier, and Schipper (2013), we present the definition of prudent rationalizability in Section 3, prove existence, and also demonstrate its refining power with an example. Next, we explore in Section 4 the properties of prudent rationalizability by comparing it with extensive-form rationalizability. In Section 5 we demonstrate the applicability of prudent rationalizability by applying it to games of disclosure of verifiable information both with full awareness and under unawareness. Finally, we conclude with further discussions in Section 6. All proofs are relegated to the appendix.

2 Generalized Extensive-form Games with Unawareness

In this section, we outline generalized extensive-form games with unawareness as introduced in Heifetz, Meier, and Schipper (2013). To define a generalized extensive-form game Γ , consider first, as a building block, a finite game tree with perfect information and simultaneous moves with a set of players I , a set of decision nodes N_0 , active players I_n at node n with finite action sets A_n^i of player $i \in I_n$ (for $n \in N_0$), chance nodes C_0 , and terminal nodes Z_0 with a payoff vector $(p_i^z)_{i \in I} \in \mathbb{R}^I$ for the players for every $z \in Z_0$. The nodes $\bar{N}_0 = N_0 \cup C_0 \cup Z_0$ constitute a tree, i.e., they are partially ordered by a precedence relation \prec with which (\bar{N}_0, \prec) is an arborescence (that is, the predecessors of each node in \bar{N}_0 are totally ordered by \prec), for each decision node $n \in N_0$ there is a bijection ψ_n between the action profiles $\prod_{i \in I_n} A_n^i$ at n and n ’s immediate successors, and there is a unique node in \bar{N}_0 with no predecessors – the root of the tree.

Consider now a family \mathbf{T} of subtrees of \bar{N}_0 . A subtree is defined by a subset of nodes $\bar{N}'_0 \subseteq \bar{N}_0$ for which (\bar{N}'_0, \prec) is also a tree (i.e., an arborescence in which a unique node has no

predecessors). For two subtrees $T', T'' \in \mathbf{T}$ we write $T' \preceq T''$ to signify that the nodes of T' constitute a subset of the nodes of T'' .

One of the trees $T_1 \in \mathbf{T}$ is meant to represent the modeler's view of the paths of play that are *objectively* feasible. Each other tree $T \in \mathbf{T}$ represents the feasible paths of play as *subjectively* viewed by some player at some node in T_1 ; or as the frame of mind attributed to the player at some node of T_1 by another player (or even by the same player at a later stage of the game, after her awareness regarding the feasible paths has evolved), whose own frame of mind regarding the feasible paths is represented by yet another $T' \in \mathbf{T}$; and so forth.

Denote by N_i^T the set of nodes in which player $i \in I$ is active in the tree $T \in \mathbf{T}$, and by $N_i = \bigcup_{T \in \mathbf{T}} N_i^T$.

We require three properties:

1. All the terminal nodes in each tree $T \in \mathbf{T}$ are in Z_0 .
2. For every tree $T \in \mathbf{T}$, every node $n \in T$, and every active player $i \in I_n$ there exists a nonempty subset of actions $A_n^{i,T} \subseteq A_n^i$ such that ψ_n maps the action profiles $A_n^T = \prod_{i \in I_n} A_n^{i,T}$ bijectively onto n 's successors in T .
3. For any player $i \in I$, if for two decision nodes $n, n' \in N_i^T$ (i.e., $i \in I_n \cap I_{n'}$) it is the case that $A_n^i \cap A_{n'}^i \neq \emptyset$, then $A_n^i = A_{n'}^i$.

See Heifetz, Meier, and Schipper (2013) for a discussion of these properties.

Within the family \mathbf{T} of subtrees of \bar{N}_0 , some nodes n appear in several trees $T \in \mathbf{T}$. In what follows, we will need to designate explicitly these different appearances of such nodes n as distinct entities. To this effect, in each tree $T \in \mathbf{T}$ label by n_T the copy in T of the node $n \in \bar{N}_0$ whenever the copy of n is part of the tree T , with the caveat that if the move $a_n \in A_n^T$ leads from n to n' , then a_n leads also from the copy n_T to the copy n'_T . Denote by N the union of all decision nodes in all trees $T \in \mathbf{T}$, by C the union of all chance nodes, by Z the union of terminal nodes, and by $\bar{N} = N \cup C \cup Z$ (copies n_T of a given node n in different subtrees T are distinct from one another, so that \bar{N} is a disjoint union of sets of nodes).

In what follows, when referring to a node in \bar{N} we will typically avoid the subscript T when no confusion may arise. For a node $n \in \bar{N}$ we denote by T_n the tree containing n .

Information and awareness of players is modelled with information sets. Different from standard extensive-form games, we allow that the information set at node n in tree T may be in a lower tree. Formally, for each decision node $n \in N$, define for each active player $i \in I_n$ a nonempty information set $\pi_i(n)$ with the following properties:

- I0 Confinement: $\pi_i(n) \subseteq T$ for some tree T .

- I1 No-delusion given the awareness level: If $\pi_i(n) \subseteq T_n$, then $n \in \pi_i(n)$.
- I2 Introspection: If $n' \in \pi_i(n)$, then $\pi_i(n') = \pi_i(n)$.
- I3 No divining of currently unimaginable paths, no expectation to forget currently conceivable paths: If $n' \in \pi_i(n) \subseteq T'$ (where $T' \in \mathbf{T}$ is a tree) and there is a path $n', \dots, n'' \in T'$ such that $i \in I_{n'} \cap I_{n''}$, then $\pi_i(n'') \subseteq T'$.
- I4 No imaginary actions: If $n' \in \pi_i(n)$, then $A_{n'}^i \subseteq A_n^i$.
- I5 Distinct action names in disjoint information sets: For a subtree T , if $n, n' \in T$ and $A_n^i = A_{n'}^i$, then $\pi_i(n') = \pi_i(n)$.
- I6 Perfect recall: Suppose that player i is active in two distinct nodes n_1 and n_k , and there is a path n_1, n_2, \dots, n_k such that at n_1 player i takes the action a_i . If $n' \in \pi_i(n_k)$, then there exists a node $n'_1 \neq n'$ and a path $n'_1, n'_2, \dots, n'_\ell = n'$ such that $\pi_i(n'_1) = \pi_i(n_1)$ and at n'_1 player i takes the action a_i .

We refer to Heifetz, Meier, and Schipper (2013) for illustrations and discussions of these properties.

We denote by H_i the set of i 's information sets in all trees. For an information set $h_i \in H_i$, we denote by T_{h_i} the tree containing h_i . For two information sets h_i, h'_i in a given tree T , we say that h_i precedes h'_i (or that h'_i succeeds h_i) if for every $n' \in h'_i$ there is a path n, \dots, n' in T such that $n \in h_i$. We denote $h_i \rightsquigarrow h'_i$.

If $n \in h_i$ we write also A_{h_i} for A_n^i .

Above properties (in particular the perfect recall property I6) guarantee that with the precedence relation \rightsquigarrow player i 's information sets H_i form an arborescence: For every information set $h'_i \in H_i$, the information sets preceding it $\{h_i \in H_i : h_i \rightsquigarrow h'_i\}$ are totally ordered by \rightsquigarrow .

For trees $T, T' \in \mathbf{T}$ we denote $T \succrightarrow T'$ whenever for some node $n \in T$ and some player $i \in I_n$ it is the case that $\pi_i(n) \subseteq T'$. Denote by \hookrightarrow the transitive closure of \succrightarrow . That is, $T \leftrightarrow T''$ if and only if there is a sequence of trees $T, T', \dots, T'' \in \mathbf{T}$ satisfying $T \succrightarrow T' \succrightarrow \dots \succrightarrow T''$.

A *generalized extensive-form game* Γ consists of a partially ordered set \mathbf{T} of subtrees of a tree \bar{N}_0 satisfying properties 1-3 above, along with information sets $\pi_i(n)$ for every $n \in T$, $T \in \mathbf{T}$ and $i \in I_n$, satisfying properties I0-I6 above.

For every tree $T \in \mathbf{T}$, the *T -partial game* is the partially ordered set of trees including T and all trees T' in Γ satisfying $T \leftrightarrow T'$, with information sets as defined in Γ . A T -partial game is a generalized game, i.e., it satisfies all properties 1-3 and I0-I6.

We denote by H_i^T the set of i 's information sets in the T -partial game.

A (pure) strategy

$$s_i \in S_i := \prod_{h_i \in H_i} A_{h_i}$$

for player i specifies an action of player i at each of her information sets $h_i \in H_i$. Denote by

$$S = \prod_{j \in I} S_j \quad \text{and} \quad S_{-i} = \prod_{j \in I \setminus \{i\}} S_j$$

the set of strategy profiles and the set of strategy profile of player i 's opponents, respectively, in the generalized extensive-form game.

If $s_i = (a_{h_i})_{h_i \in H_i} \in S_i$, we denote by

$$s_i(h_i) = a_{h_i}$$

the player's action at the information set h_i .

With the strategy s_i , at node $n \in N_i$ define the player's action at n to be $s_i(\pi_i(n))$. Thus, the strategy s_i specifies what player i does at each of her active nodes $n \in N_i$, both in case $n \in \pi_i(n)$ and in case $\pi_i(n)$ is a subset of nodes of a tree which is distinct from the tree T_n to which n belongs.

For a strategy $s_i \in S_i$ and a tree $T \in \mathbf{T}$, we denote by s_i^T the strategy in the T -partial game induced by s_i (i.e., $s_i^T(h_i) = s_i(h_i)$ for every information set h_i of player i in the T -partial game). If $R_i \subseteq S_i$ is some set of strategies of player i , denote by R_i^T the set of strategies induced by R_i in the T -partial game. The set of i 's strategies in the T -partial game is thus denoted by S_i^T . Denote by $S^T = \prod_{j \in I} S_j^T$ the set of strategy profiles in the T -partial game. Similarly, denote by $S_{-i}^T = \prod_{j \in I \setminus \{i\}} S_j^T$ the set of strategy profiles of player i 's opponents in the T -partial game.

We say that a strategy profile $s = (s_j)_{j \in I} \in S$ reaches a node $n \in T$ if the players' actions $s_j^T(\pi_j(n'))_{j \in I_{n'}}$ and nature's moves in the nodes $n' \in T$ lead to n with a positive probability. Notice that by property (I4) ("no imaginary actions"), $s_j^T(\pi_j(n'))_{j \in I}$ is indeed well defined: even if $\pi_j(n') \notin T$ for some $n' \in T$, the action profile $s_j^T(\pi_j(n'))_{j \in I_{n'}}$ is an action profile which is actually available in T to the active players $j \in I_{n'}$ at n' . We say that a strategy profile $s \in S$ reaches the information set $h_i \in H_i$ if s reaches some node $n \in h_i$. We say that the strategy $s_i \in S_i$ reaches the information set h_i if there is a strategy profile $s_{-i} \in S_{-i}$ of the other players such that the strategy profile (s_i, s_{-i}) reaches h_i . Otherwise, we say that the information set h_i is excluded by the strategy s_i . Similarly, we say that the strategy profile $s_{-i} \in S_{-i}$ reaches the information set h_i if there exists a strategy $s_i \in S_i$ such that the strategy profile (s_i, s_{-i}) reaches h_i .

As is the case also in standard games, for every given node, a given strategy profile of the

players induces a distribution over terminal nodes in each tree, and hence an expected payoff for each player in the tree.

So far, we outlined generalized extensive-form games with unawareness. See Heifetz, Meier, and Schipper (2013) and Schipper (2018a) for further discussions of the underlying modelling assumptions.

3 Prudent Rationalizability

In this section, we present the solution concept, prudent rationalizability.

A *belief system* of player i

$$b_i = (b_i(h_i))_{h_i \in H_i} \in \prod_{h_i \in H_i} \Delta(S_{-i}^{T_{h_i}})$$

is a profile of beliefs - a belief $b_i(h_i) \in \Delta(S_{-i}^{T_{h_i}})$ about the other players' strategies in the T_{h_i} -partial game, for each information set $h_i \in H_i$, with the following properties

- $b_i(h_i)$ reaches h_i , i.e., $b_i(h_i)$ assigns probability 1 to the set of strategy profiles of the other players that reach h_i .
- If h_i precedes h'_i ($h_i \rightsquigarrow h'_i$) then $b_i(h'_i)$ is derived from $b_i(h_i)$ by conditioning whenever possible on the set of opponents' (partial) strategy profiles that reach h'_i .

Denote by B_i the set of player i 's belief systems.

For a belief system $b_i \in B_i$, a strategy $s_i \in S_i$ and an information set $h_i \in H_i$, define player i 's expected payoff *at* h_i to be the expected payoff for player i in T_{h_i} given $b_i(h_i)$, the actions prescribed by s_i at h_i and its successors, assuming that h_i has been reached.

We say that with the belief system b_i and the strategy s_i player i is *rational* at the information set $h_i \in H_i$ if there exists no action $a'_{h_i} \in A_{h_i}$ such that only replacing the action $s_i(h_i)$ by a'_{h_i} results in a new strategy s'_i which yields player i a higher expected payoff at h_i given the belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$. Note that this is a "local" notion of optimizing over actions available at the information rather than optimizing over entire continuation strategies (i.e., replacement strategies) at the information set. See Meier and Schipper (2012) for a discussion of these two notions and an extension of the one-shot deviation principle to generalized extensive-form games with unawareness.

Definition 1 (Prudent rationalizability in generalized extensive-form games) *Let*

$$\bar{S}_i^0 = S_i$$

For $k \geq 1$ define inductively

$$\bar{B}_i^k = \left\{ b_i \in B_i : \begin{array}{l} \text{for every information set } h_i, \text{ if there exists some profile} \\ s_{-i} \in \bar{S}_{-i}^{k-1} = \prod_{j \neq i} \bar{S}_j^{k-1} \text{ of the other players' strategies} \\ \text{such that } s_{-i} \text{ reaches } h_i \text{ in the tree } T_{h_i}, \text{ then the support} \\ \text{of } b_i(h_i) \text{ is the set of strategy profiles } s_{-i} \in \bar{S}_{-i}^{k-1, T_{h_i}} \text{ that reach } h_i \end{array} \right\}$$

$$\bar{S}_i^k = \left\{ s_i \in \bar{S}_i^{k-1} : \begin{array}{l} \text{there exists } b_i \in \bar{B}_i^k \text{ such that for all } h_i \in H_i \text{ player } i \\ \text{with strategy } s_i \text{ is rational at } h_i \end{array} \right\}$$

The set of prudent rationalizable strategies of player i is

$$\bar{S}_i^\infty = \bigcap_{k=1}^{\infty} \bar{S}_i^k$$

At each level, each player, each tree, and each information set of the player, she forms full support beliefs over the opponents' strategies remaining from the previous level and reaching this information set. The player retains any of her previous level strategies for which there exists such a belief such that the strategy is rational at all information sets.

In the context of finite generalized extensive-form games with unawareness (and perfect recall), every player's set of prudent rationalizable strategies is non-empty.

Theorem 1 *The set of player i 's prudent rationalizable strategies is non-empty.*

The proof appears in the appendix.

3.1 An Example by Ozbay (2007)

In order to demonstrate the extra power of prudent rationalizability, consider the following example of dynamic interaction with unawareness, which is a variant of example 3 in Ozbay (2007). There are three states of nature, $\omega_1, \omega_2, \omega_3$. A chance move chooses one out of four potential distributions over the states of nature:

$$\begin{aligned} \delta_1 &= (1, 0, 0) \\ \delta_2 &= (0, 1, 0) \\ \delta_3 &= (0, 0, 1) \\ \delta_4 &= \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right) \end{aligned}$$

An Announcer gets to know the distribution (but not the realization of the state of nature). A Decision Maker (DM) is initially aware only of the state ω_1 (and hence the DM is certain that ω_1 will be realized with certainty). However, before the DM chooses what to do, the Announcer can choose to make the DM aware of either ω_2, ω_3 , none of them or both of them. Increased awareness makes the DM aware of the relevant marginals of the distributions. For instance, if the Announcer makes the DM aware of ω_2 , the DM becomes aware of the set of distributions

$$\begin{aligned}\delta_{1|\{\omega_1, \omega_2\}} &= (1, 0) \\ \delta_{2|\{\omega_1, \omega_2\}} &= (0, 1) \\ \delta_{4|\{\omega_1, \omega_2\}} &= \left(\frac{1}{2}, \frac{1}{2}\right)\end{aligned}$$

and also becomes certain that the Announcer knows which of these is the true distribution.¹

Subsequently, the DM should choose one out of three possible actions – left, middle or right. The payoffs to the players as a function of the chosen action and the state of nature appear in the following table:

	left	middle	right
ω_1	3, 3	0, 0	2, 2
ω_2	0, 0	5, 5	2, 2
ω_3	2, 2	0, 0	2, 2

The game is thus described in Figure 1.

It is obvious that if the Announcer announces nothing, and hence the DM is certain that ω_1 prevails, the DM will choose ‘left’.

What happens if the Announcer makes the DM aware of ω_2 ? The information set of the DM becomes

$$\left\{ \delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}} \right\}$$

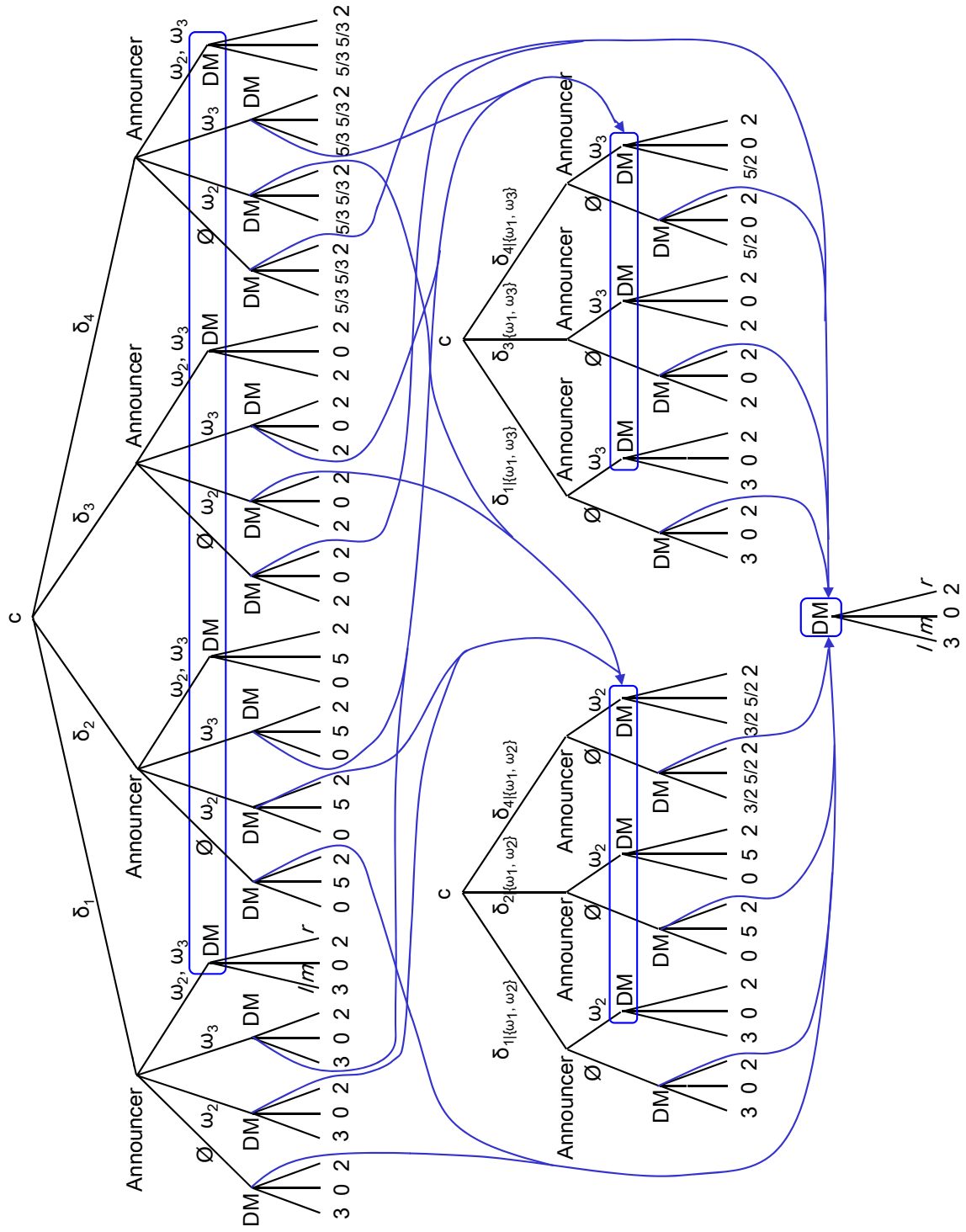
The DM may then assign a high probability to $\delta_{1|\{\omega_1, \omega_2\}}$,² and this will lead the DM to choose ‘left’. Hence, assuming such a belief by the DM, it is rationalizable for the Announcer to make

¹In Ozbay’s example and in what follows the DM’s beliefs about these marginal distributions will not be necessarily related to the prior probabilities with which the distributions were chosen by the chance move. That’s why we do not even bother to specify the probabilities with which the chance move chooses the different distributions.

Put differently, instead of describing this game by a partially ordered set of trees, one for each level of awareness as in Figure 1, we could have replaced each tree with an arborescence in which the initial chance move is erased. Allowing for arborescences instead of trees in the framework for dynamic unawareness is straightforward, but for the sake of clarity of the exposition we avoid this explicit generalization in the body of the paper.

²That is, the DM may assign a high probability to strategies of the Announcer by which the Announcer announces ω_2 (and cause the DM’s information set to become $\left\{ \delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}} \right\}$) when the Announcer has learned that the true distribution is δ_1 .

Figure 1:



the DM aware of ω_2 when the Announcer knows that the true distribution is δ_1 (i.e., when the Announcer knows that ω_1 will be realized with probability 1).

This is not very sensible, though. After all, the Announcer can ensure that the DM chooses ‘left’ by not announcing any new state. When the Announcer likes the DM to choose ‘left’, it makes no sense on the Announcer’s part to announce ω_2 and thus face the risk that the DM assigns a low probability to $\delta_{1|\{\omega_1, \omega_2\}}$ and consequently choose ‘middle’. This idea is captured by Ozbay’s reasoning refinement to his awareness equilibrium notion. Moreover, it is captured even without assuming equilibrium by prudent rationalizability.

Proposition 1 *The DM has a unique prudent rationalizable strategy. With this strategy the DM chooses ‘left’ when no new state is announced, ‘middle’ when only ω_2 is announced, ‘left’ when only ω_3 is announced, and ‘right’ when both ω_2, ω_3 are announced.*

The proof is contained in the appendix.

Caution embodied in the definition of prudent rationalizability is crucial for the argument. In particular, a solution concept like extensive-form rationalizability, which is analogous to prudent rationalizability but without the requirement of full-support beliefs on strategies remaining from the previous levels of the elimination procedure, would not allow us to show the result.

4 Comparison to Extensive-form Rationalizability

In order to explore properties of prudent rationalizability, it is useful to compare it to extensive-form rationalizability. Pearce (1984) defined extensive-form (correlated) rationalizable strategies by a procedure of an iterative elimination of strategies. The inductive definition below is an extension of Battigalli’s (1997) version to generalized extensive-form games with unawareness as introduced in Heifetz, Meier, and Schipper (2013). Battigalli’s (1997) definition differs from Pearce’s (1984) in at least two respects. First, Battigalli allows for correlated beliefs over opponents’ strategies. Second, Battigalli’s definition defines a procedure of an iterative elimination of beliefs. This allows for a more intuitive interpretation as a reasoning procedure. Nevertheless, Battigalli shows that when one allows for correlation in Pearce’s original definition, then both procedures are equivalent. Below Definition 2 differs in two respects from Battigalli (1997). First, it applies to generalized extensive-form games with unawareness. Second, it uses local optimization over actions at the information rather than optimization over replacements of strategies at the information. With respect to the last issue, below definition differs also from Heifetz, Meier, and Schipper (2013).³ See Meier and Schipper (2012) for further discussions and results comparing the two notions of optimization.

³In the working paper version of Heifetz, Meier, and Schipper (2013), we also used the notion in Definition 2 featuring local optimization over actions but following the recommendation of a reviewer in the publication we

Definition 2 (Extensive-form Rationalizable Strategies) Define, inductively, the following sequence of belief systems and strategies of player $i \in I$:

$$\begin{aligned}
B_i^1 &= B_i \\
S_i^1 &= \left\{ s_i \in S_i : \begin{array}{l} \text{there exists a belief system } b_i \in B_i^1 \text{ with which for every} \\ \text{information set } h_i \in H_i \text{ player } i \text{ with } s_i \text{ is rational at } h_i \end{array} \right\} \\
&\vdots \\
B_i^k &= \left\{ b_i \in B_i^{k-1} : \begin{array}{l} \text{for every information set } h_i, \text{ if there exists some profile of the other} \\ \text{players' strategies } s_{-i} \in S_{-i}^{k-1} = \prod_{j \neq i} S_j^{k-1} \text{ such that } s_{-i} \text{ reaches } h_i \\ \text{in the tree } T_{h_i}, \text{ then } b_i(h_i) \text{ assigns probability 1 to } S_{-i}^{k-1, T_{h_i}} \end{array} \right\} \\
S_i^k &= \left\{ s_i \in S_i : \begin{array}{l} \text{there exists a belief system } b_i \in B_i^k \text{ with which for every} \\ \text{information set } h_i \in H_i \text{ player } i \text{ with } s_i \text{ is rational at } h_i \end{array} \right\}
\end{aligned}$$

The set of player i 's extensive-form rationalizable strategies is

$$S_i^\infty = \bigcap_{k=1}^{\infty} S_i^k.$$

For finite generalized extensive-form games with unawareness (and perfect recall), the set of extensive-form rationalizable strategies is nonempty.

Theorem 2 *The set of player i 's extensive-form rationalizable strategies is non-empty.*

Heifetz, Meier, and Schipper (2013) proved this result using optimization by replacements of strategies. The proof of Theorem 2 follows verbatim by replacing it with the local notion of optimization over actions used in this paper. Earlier, Pearce (1984) and Battigalli (1997) showed non-emptiness of extensive-form rationalizable strategies in standard extensive-form games.

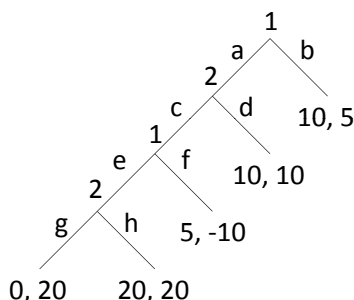
The following subsections we discuss differences between prudent rationalizability and extensive-form rationalizability. The examples should be of general interest as they involve standard extensive-form games without unawareness.

changed it to optimization over strategy-replacements. After the paper had been published, Pierpaolo Battigalli strongly encouraged us in repeated private communication to stick to the local notion of optimization over actions, which we agree is more intuitive and conceptually proper.

4.1 Subgame Perfection, Extensive-form Rationalizability, and Prudent Rationalizability

Some normal-form games have Nash equilibria in weakly dominated strategies. Similarly, the game in Figure 2 is an example⁴ of a standard perfect-information extensive-form game with a subgame-perfect equilibrium involving a strategy which is extensive-form rationalizable but not prudent rationalizable. To wit, this is the subgame-perfect equilibrium (af, dg) . The strategy af is extensive-form rationalizable for player 1 – it is supported by the belief system which starts at the root with the belief that player 2 will play dg , revised completely at 1’s second decision node by the belief that 2 is playing cg ; the strategy dg is extensive-form rationalizable for player 2 with the belief system with which player 2 is initially certain that 1 is playing af , revised completely at 2’s second decision node to the belief that 1 played ac . However, af is rational for player 1 at the root for no full-support belief on 2’s entire strategy set $\{dg, dh, cg, ch\}$, and hence af is not prudent rationalizable.

Figure 2:



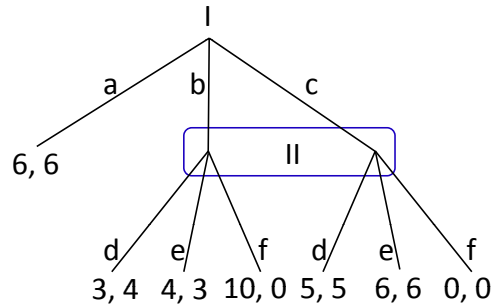
4.2 Prudent Rationalizability versus Extensive-form Rationalizability

In normal-form games, iterated admissibility is a refinement of rationalizability. Yet, in extensive-form games prudent rationalizability is *not* a refinement of extensive-form rationalizability, as the following example (Figure 3) demonstrates.

In this example, player 1 can guarantee herself the payoff 6 by choosing a and ending the game. If player 2 is called to play, should he believe that player 1 chose b or c ? If player 1 is certain that player 2 is rational, she is certain that player 2 will not choose f . Hence, if player 2 is certain that player 1 is certain that he (player 2) is rational, then at his information set player 2 is certain that player 1 chose c . The reason is that among player 1’s actions leading to 2’s information set, c is the only action which, assuming 2 believes c was chosen and that 2 is

⁴This example arose from discussions by one of the authors with Ronen Gradwohl.

Figure 3:



rational and will hence choose e , yields player 1 the payoff 6, which is just as high as the payoff she could guarantee herself with the outside option a . Hence (a, e) and (c, e) are the profiles of extensive-form (correlated) rationalizable strategies (as well as extensive-form rationalizable strategies) in this game.

The notion of prudence, in contrast, embodies the idea that being prudently rational, player 1 shouldn't rule out completely any of 2's possible choices, and hence that c is strictly inferior for player 1 relative to her outside option a . Hence, if 2's information set is ever reached, the only way for 2 to rationalize this is to believe that 1 chose b , based on a belief ascribing a high probability to the event that 2 will foolishly choose f . Player 2's best reply to b is d ; and player 1's best reply to d is a . Thus, the only profile of prudent rationalizable strategies in this game is (a, d) .

This example demonstrates that in dynamic interactions the notions of rationalization and prudence might involve a *tension*. Extensive-form rationalizability embodies a best-rationalization principle (Battigalli 1996); it is driven by the assumption that in each of his information sets, a player assesses the other players' future behavior by attributing to them the 'highest' level of rationality and mutual certainty of rationality consistent with the fact that the information set has indeed been reached. But, with the additional criterion of 'prudence', what should a player believe about the behavior of his opponent if, as in the example, the opponent's only action which is compatible with common certainty of rationality is imprudent on the part of the opponent?

The definition of prudent rationalizability resolves this tension unequivocally in favor of the prudence consideration. It remains open whether and how a more balanced and elaborate definition could resolve the tension in less an extreme fashion. We plan to address this challenge in future work. However, any definition would have to cut the Gordian knot in the above example in one particular way, choosing either d or e , and indeed both potential resolutions are backed by sensible intuitions. This suggests that for dynamic interactions we need not necessarily expect one ultimate definition of rationalizability taking into account both rationalization and

prudence.

Remark 1 *For standard extensive-form games, Brandenburger and Friedenberg (2007) studied the connection between iterative elimination of conditionally weakly dominated strategies and iterative elimination of conditionally dominated strategies. They showed that under a “no relevant convexities” condition, iterative elimination of conditionally dominated strategies coincides with iterative elimination of conditionally weakly dominated strategies. Since iterative elimination of conditionally (resp. weakly) dominated strategies is equivalent to extensive-form (resp. prudent) rationalizability (using the optimization over strategy-replacements rather than locally over actions), we have that in standard games with “no relevant convexities” extensive-form rationalizability coincides with prudent rationalizability. However, the example in Figure 3 does not satisfy this condition, and hence demonstrates that in general, prudent rationalizability is not a refinement of extensive-form rationalizability.*

Nevertheless, as far as paths of play are concerned, in the above example the set of paths induced by prudent rationalizability (the path a) is a subset of the paths induced by extensive-form rationalizability (the paths a and (c, e)).

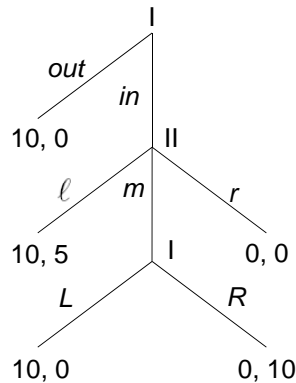
So far we demonstrated that prudent rationalizability is *not a refinement of extensive-form rationalizable strategies*. Nevertheless we conjecture that generally prudent rationalizability refines the set of extensive-form rationalizable paths. In an earlier working paper version of this paper we claimed this conjecture as a result and presented what we considered is a proof of the result. Yet, the work by Chen and Micali (2013) made us realize further subtleties. While we still believe that the conjecture true, a fully satisfactory proof has been alluding us so far.

4.3 Forward Induction: The Tension between Extensive-form Rationalizability and Prudent Rationalizability

In Figure 3 we demonstrated the tension between the considerations of rationalization and prudence when a player tries to divine his opponent’s past actions. A related but distinct tension arises when a player tries to deduce the opponent’s future behavior from past actions of that opponent. Consider the following example in Figure 4.

In this example, in is imprudent for player 1 (since by going out she can guarantee a payoff of 10, while by moving in she risks getting 0 if player 2 would rather foolishly choose r). This means that if player 1 does move in and player 2 gets to play, no prudent strategy in \bar{S}_1^1 reaches 2’s information set. Hence, the beliefs \bar{B}_2^2 of player 2 about player 1’s future actions are not restricted. In particular, it contains beliefs by which if player 2 chooses m , player 1 will foolishly choose R (with a high probability). That’s why both m and ℓ are prudent rationalizable for player 2.

Figure 4:



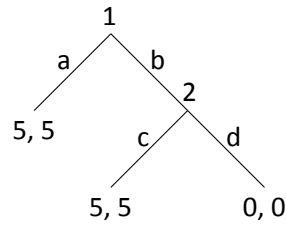
However, it is not very sensible on the part of player 2 to believe that following m player 1 may choose R . After all, when player 2 has to move, player 1 has already proved to be imprudent, but not irrational. Indeed, player 1's rationalizable (though imprudent) strategy (in, L) yields her the payoff 10 in conjunction with 2's only (extensive-form) rationalizable strategy ℓ , as well as in conjunction with 2's prudent rationalizable strategy m ; and this payoff is the same as the payoff player 1 gets from her only prudent rationalizable strategy (out, L) .

Thus, as long as player 1 has been rational (even if imprudent) thus far, it makes more sense for player 2 to believe that player 1 will continue to be rational (though possibly imprudent) in the future. Restricting player 2's beliefs according to this logic would cross out the nonsensical choice m .

Already Pearce (1984) was well aware of this tension, which motivated his definition of *cautious extensive-form rationalizability*. That definition involves refining the set of rationalizable strategies by another round of strategy elimination with full support beliefs about the other players' surviving strategies; and then repeating this entire procedure – the standard iterative elimination process as in the definition of rationalizability, followed by one round assuming full-support beliefs – ad infinitum. In the above example, cautious extensive-form rationalizability does indeed rule out the strategy m for player 2.

However, as Pearce (1984) himself admits, the definition of cautious extensive-form rationalizability is not really satisfactory, as the following simple example of his shows (Figure 5). In this example, the strategy d is irrational for player 2. Once d is crossed out, both a and b are extensive-form rationalizable for player 1, and are actually also cautious extensive-form rationalizable. Notice that in contrast, b *does* get crossed out by *prudent rationalizability*, and the only prudent rationalizable strategy for player 1 is a .

Figure 5:



5 Application: Disclosure of Verifiable Information

In this section we provide an application of prudent rationalizability to the problem of disclosure of verifiable information by interested parties. Disclosure of verifiable information arises in many contexts such as buyer-seller relationships, financial markets, legal proceedings, electoral campaigns etc.; see Milgrom (2008) for a survey. The application also allows us to demonstrate how unawareness affects outcomes of in this important class of strategic problems.

5.1 Disclosure of Verifiable Information with Full Awareness

Consider a merchandise whose quality $q_i \in \{q_1, \dots, q_n\}$ is known to its seller, while a buyer knows only the prior probability distribution (p_1, \dots, p_n) of the qualities, where $p_i > 0$ for all $i = 1, \dots, n$. For each quality level q_i the seller is better off the larger the quantity that she sells, while the utility of the buyer from the merchandise is strictly concave in the quantity purchases with a single peak at $\beta(q_i)$. Furthermore,

$$\beta(q_1) < \dots < \beta(q_n).$$

Before sale takes place, the seller has the option of providing the buyer with a certified signal about the quality of her merchandise, proving to the seller that the quality is within some range $\{q_{\min}, \dots, q_{\max}\}$ containing the actual quality q_i .

Milgrom and Roberts (1986) proved that if the buyer's utility is strictly concave then there is a unique sequential equilibrium, in which when the quality is q_i the seller certifies to the buyer a range (possibly a singleton) $\{q_{\min}, \dots, q_{\max}\}$ in which $q_{\min} = q_i$, while the buyer is skeptical and always buys $\beta(q_{\min})$. Thus, in this unique sequential equilibrium the quality q_i is fully revealed to the seller, who buys the optimal quantity $\beta(q_i)$ for him.

We proceed with the caveat that the quantities which can be demanded by the buyer belong to a finite, fine grid (recall that, formally, in our formulation each player has finitely many available actions in each information set). For simplicity, we assume further that the quantities

$\beta(q_i)$, $i = 1, \dots, n$ belong to this grid. For $1 \leq m < n$ we denote by $[\beta(q_m), \beta(q_n)]$ the set of quantities in this grid at least as large as $\beta(q_m)$ and no larger than $\beta(q_n)$.

Any profile of prudent rationalizable strategies in this game yields the full revelation outcome shown by Milgrom and Roberts (1986) using sequential equilibrium.

Proposition 2 *The strategy to buy $\beta(q_{\min})$ when confronted with the certification that the quality is in the range $\{q_{\min}, \dots, q_{\max}\}$ is also the unique prudent rationalizable strategy for the buyer, and certifying some range $\{q_{\min}, \dots, q_{\max}\}$ in which $q_{\min} = q_i$ constitute the prudent rationalizable strategies of the seller.*

The proof is contained in the appendix.

It is not difficult to see that the above argument does not depend on the assumption that the available certificates consist of ranges of qualities (containing the true quality). For the argument to hold it is enough to assume that for each quality level q_i one of the available certificates is the fully revealing certificate $\{q_i\}$.

Note that the result would not obtain when we employ extensive-form rationalizability rather than prudent rationalizability. The reason is that when the buyer is presented with a certificate $\{q_m, \dots, q_n\}$, then the buyer could optimistically believe that the seller's quality is q_n and buy a larger quantity than with a prudent full support belief. Battigalli (2006) already realized that in disclosure games some restriction on beliefs that imply a degree of scepticism is required for rationalizability to yield unraveling. In his case, the restriction is tailored to the disclosure game in that he assumes that the receiver puts some strict positive probability on the lowest quality consistent with the message. In contrast, we use full support beliefs consistent with the message because we strive for a solution concept that is not tailored to the particular game but generally applicable to games.

5.2 Disclosure of Verifiable Information under Unawareness

Assume now that there are several dimensions of quality along which such certifications could be provided. To fix ideas, consider two dimensions L, H and $^0, ^*$. The four combinations are

$$L^0, H^0, L^*, H^*.$$

So, for instance, in the state L^0 the available certificates are $\{L, H\} \times \{^0, ^*\}$, $\{L\} \times \{^0, ^*\}$, $\{L, H\} \times \{^0\}$ and $\{L\} \times \{^0\}$.

Assume further that

$$\beta(L^*) < \beta(L^0) < \beta(H^0) < \beta(H^*)$$

Since the singleton certificates

$$\{L\} \times \{*\}, \{L\} \times \{^0\}, \{H\} \times \{^0\}, \{H\} \times \{*\}$$

are available, the above argument obtains and full revelation takes place in any profile of prudent rationalizable strategies of the players.

Assume, however, that the buyer is initially aware only of the $\{L, H\}$ dimension and is unaware of the $\{^0, *\}$ dimension; he evaluates the merchandize as having the default quality L^0 when confronted with the certificate $\{L\}$, and similarly, with the certificate $\{H\}$ he evaluates the merchandize as having the default quality H^0 . Assume further that the seller knows this, and that by presenting the certificates $\{*\}$, $\{^0\}$ or $\{^0, *\}$ the seller inter alia makes the buyer aware of the $\{^0, *\}$ dimension.

Intuitively, it is clear that the seller will want to make the buyer aware of this extra dimension when the quality is H^* , because this will lead the buyer to demand the high quantity $\beta(H^*)$. In contrast, when the actual quality is L^* , the seller will prefer not to present any certificate at all along the dimension $\{^0, *\}$: This way the buyer will remain unaware of this extra dimension, and will demand the quantity $\beta(L^0)$ (because unraveling and full revelation will occur only along the $\{L, H\}$ dimension); if the seller were to make the buyer aware of this extra dimension, the buyer would have demanded only $\beta(L^*) < \beta(L^0)$.

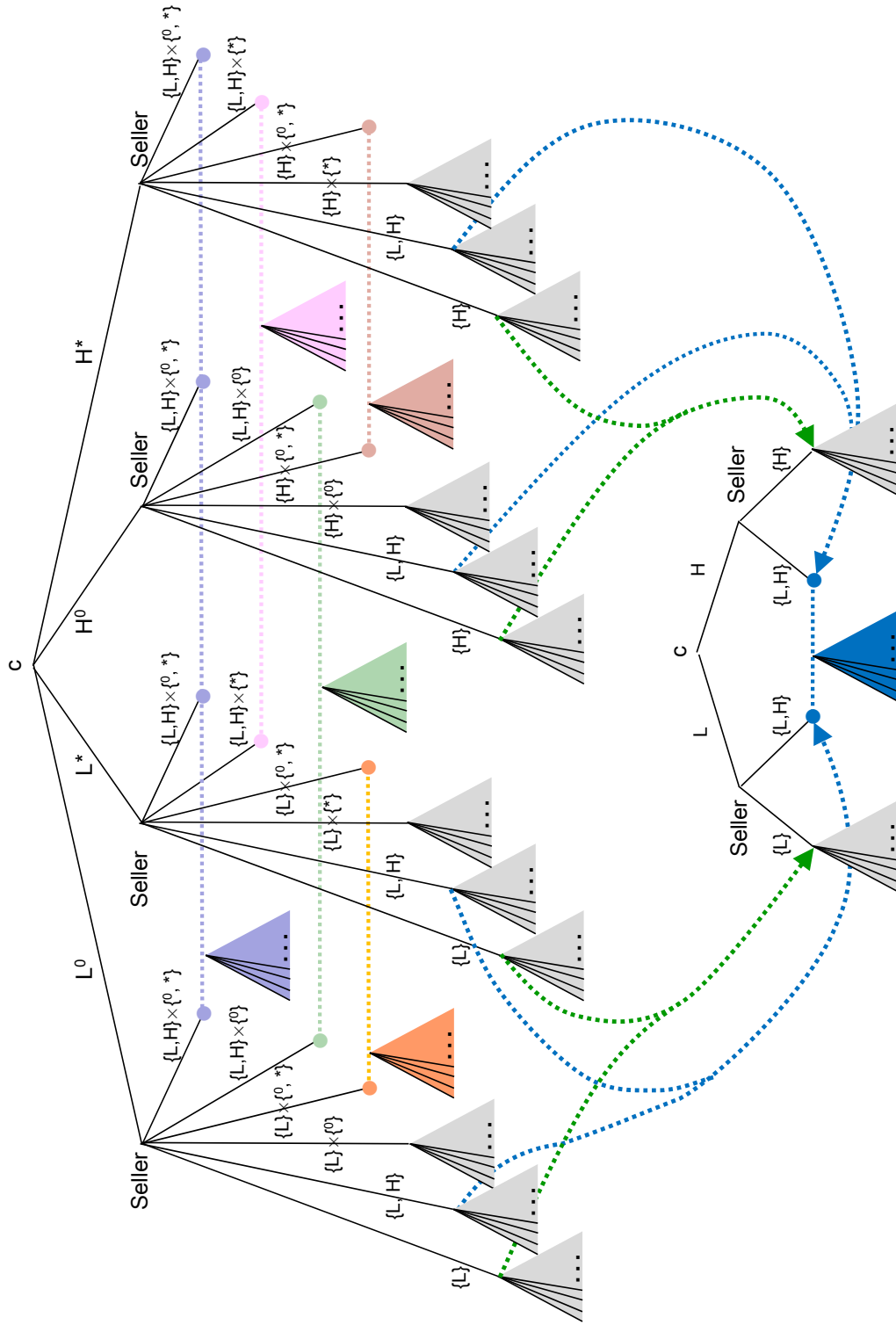
This strategic interaction is represented in the following generalized game form (Figure 6). Initially, nature, c , selects a state out of $\{L^0, L^*, H^0, H^*\}$. The seller observes the state of nature and chooses a certificate. Unless the seller presents a certificate involving the dimension $\{^0, *\}$, the buyer remains unaware of it. This is indicated by the intermitted arrows from nodes in the upper tree to nodes in the lower tree. E.g., if the seller selects the certificate $\{L\}$, then the buyer remains unaware of the $\{^0, *\}$ dimension and views the game as represented by the lower tree. In particular, his information set is a singleton containing the node after nature selects L and the seller reports $\{L\}$ in the lower tree. If the seller presents a certificate involving the $\{^0, *\}$ -dimension, then the buyer becomes aware of it and he conceives of the entire generalized game. For instance, if the seller selects the certificate $\{L, H\} \times \{^0, *\}$, then the buyer's information set is given by the upmost information set drawn as an intermitted line connecting four nodes.

We summarize the discussion in the following proposition.

Proposition 3 *In the verifiable information model in which the buyer is unaware of some dimension of the good's quality, the seller may not fully reveal the quality in any prudent rationalizable outcome.*

This is in sharp contrast to the case with full awareness discussed in the previous subsection.

Figure 6:



6 Further Discussion of Prudent Rationalizability

Definition 2 of extensive-form rationalizable strategies involves, as in Battigalli (1997), an iterative reduction procedure of belief systems (that is, by definition $B_i^k \subseteq B_i^{k-1}$), and this definition implies that strategies get iteratively eliminated ($S_i^k \subseteq S_i^{k-1}$). In contrast, the inductive definition of *prudent* rationalizable strategies involves an iterative elimination of strategies (that is, by definition $\bar{S}_i^k \subseteq \bar{S}_i^{k-1}$, in analogy with the original formulation of Pearce (1984) for extensive-form rationalizability by an iterative elimination procedure), but in the case of prudence it is *not* generally the case that $\bar{B}_i^k \subseteq \bar{B}_i^{k-1}$. Indeed, when $\bar{S}_{-i}^k \subsetneq \bar{S}_{-i}^{k-1}$:

- If the set of strategy profiles in \bar{S}_{-i}^k reaching some information set $h_i \in H_i$ is a proper, non-empty subset of the strategy profiles in \bar{S}_{-i}^{k-1} that reach h_i , then the support of each belief $\bar{b}_i^{k-1}(h_i)$ in each belief system $\bar{b}_i^{k-1} \in \bar{B}_i^{k-1}$ is strictly larger than the support of any belief $\bar{b}_i^k(h_i)$ for $\bar{b}_i^k \in \bar{B}_i^k$.
- For information sets h_i *not* reached by \bar{S}_{-i}^k , there is no restriction (beyond conditional probabilities) on $\bar{b}_i^k(h_i)$ for $\bar{b}_i^k \in \bar{B}_i^k$. No such restriction is needed, because if we define

$$m_{h_i}^k = \max \{ m < k : \text{there exists } s_{-i} \in \bar{S}_{-i}^m \text{ that reaches } h_i \}$$

then for $s_i^k \in \bar{S}_i^k$ the restrictions on i 's actions $s_i^k(h_i)$ at h_i were already determined at stage $m_{h_i}^k$, since by definition $s_i^k \in \bar{S}_i^k \subseteq \bar{S}_i^{m_{h_i}^k}$.

Is it nevertheless feasible to define prudent rationalizability via a reduction process of belief systems? Asheim and Perea (2005) proposed to look at *systems of conditional lexicographic probabilities* – belief systems in which each belief at an information set is itself a *lexicographic probability system* (Blume, Brandenburger and Dekel 1991) about the other players' strategy profiles. Using belief systems which are conditional lexicographic probabilities we could, in the spirit of Stahl (1995), put forward an equivalent definition of prudent rationalizable strategies involving an iterative reduction procedure of belief systems rather than an iterative elimination procedure of strategies. In each round of the procedure, the surviving belief systems would be those in which at each information set, ruled-out strategy profiles of the other players (i.e., strategy profiles outside $\bar{S}_{-i}^{m_{h_i}^k}$) would be deemed infinitely less likely than the surviving strategy profiles, but infinitely more likely than strategy profiles which had already been eliminated in previous rounds. We leave the precise formulation of such an equivalent definition to future work.

In their paper, Asheim and Perea (2005) proposed the notion of *quasi-perfect rationalizability*, which also involves the idea of cautious beliefs. Quasi-perfect rationalizability is distinct from our notion of prudent rationalizability. The difference is that with prudent rationalizabil-

ity (as with extensive-form rationalizability), a player need not believe that another player's future behavior must be rationalizable to a higher order than that exhibited by that other player in the past; in contrast, with the quasi-perfect rationalizable strategies of Asheim and Perea (2005), a player should ascribe to her opponent the highest possible level of rationality in the future even if this opponent has already proved to be less rational in the past. That's why quasi-perfect rationalizability implies backward induction in generic perfect information games, while our prudent rationalizable strategies need not coincide with the backward induction strategies in such games (though they do generically lead to the backward induction path – the argument is the same as in Reny 1992 and Battigalli 1997, since in generic perfect information games prudent rationalizability coincides with extensive-form rationalizability in terms of realized paths).

A Proofs

Proof of Theorem 1

First, observe that $\bar{B}_i^k \neq \emptyset$ for every $k \geq 1$, because if an information set $h_i \in H_i$ is reached by some $s_{-i} \in \bar{S}_{-i}^{k-1}$, then s_{-i} reaches also all of i 's information sets that precede h_i in the tree T_{h_i} .

We proceed by induction. $\bar{S}_i^0 = S_i$ and hence non-empty. Notice also that for every $b_i \in \bar{B}_i^1$, a standard backward induction procedure on the arborescence of information sets H_i yields a strategy $s_i \in \bar{S}_i^1$ with which player i would be rational for all $h_i \in H_i$ given b_i .

Suppose, inductively, we have already shown that $\forall i \in I \quad \bar{S}_i^{k-1} \neq \emptyset$ (and hence that $\bar{S}_{-i}^{k-1} \neq \emptyset$), and also that for every $b_i \in \bar{B}_i^{k-1}$ there exists a strategy $s_i \in \bar{S}_i^{k-1}$ with which player i would be rational for all $h_i \in H_i$ given b_i .

Let $b_i \in \bar{B}_i^k$. Let $\dot{H}_i \subseteq H_i$ be the set of i 's information sets not reached by any profile $s_{-i} \in \bar{S}_{-i}^{k-1}$ but reached by some profile $s_{-i} \in \bar{S}_{-i}^{k-2}$. If $\dot{H}_i \neq \emptyset$, for every $h_i \in \dot{H}_i$ with no predecessor in \dot{H}_i , modify (if necessary) $b_i(h_i)$ so as to have full support on the profiles in \bar{S}_{-i}^{k-2} that reach h_i , and in succeeding information sets modify b_i by Bayes rule whenever possible. Denote the modified belief system by \dot{b}_i . Then by construction also $\dot{b}_i \in \bar{B}_i^k$.

Consider a sequence of belief systems $b_{i,n} \in \bar{B}_i^{k-1}$ such that

$$\dot{b}_i = \left(\dot{b}_i(h'_i) \right)_{h'_i \in H_i} \equiv \left(\lim_{n \rightarrow \infty} b_{i,n}(h'_i) \right)_{h'_i \in H_i} .$$

To construct such a sequence $b_{i,n} \in \bar{B}_i^{k-1}$, for every information set $h'_i \in H_i$ not reached by any $s_{-i} \in \bar{S}_{-i}^{k-1}$ define $b_{i,n}(h'_i) = \dot{b}_i(h'_i)$ for every $n \geq 1$; and for every $h'_i \in H_i$ with no predecessors but reached by some profile $s_{-i} \in \bar{S}_{-i}^{k-1}$ define $b_{i,n}(h'_i) \in \Delta(\bar{S}_{-i}^{k-1})$ to be any

converging sequence of beliefs such that for every $n \geq 1$ the support of $b_{i,n}(h'_i)$ is the subset of profiles in \bar{S}_i^{k-2} that reach h'_i , while $\lim_{n \rightarrow \infty} b_{i,n}(h'_i) = \dot{b}_i(h'_i)$. In succeeding information sets reached by some $s_i \in \bar{S}_i^{k-1}$ define $b_{i,n}(h'_i)$ by conditioning whenever possible.

Given such a sequence of belief systems $b_{i,n} \in \bar{B}_i^{k-1}$ let $s_{i,n} \in \bar{S}_i^{k-1}$ be a corresponding sequence of strategies with the property that given $b_{i,n}$, it is the case that with the strategy $s_{i,n}$ player i would be rational at every $h_i \in H_i$. Since player i has finitely many strategies, some strategy s_i appears infinitely often in the sequence $s_{i,n}$. Since expected utility is linear in beliefs and hence continuous, also given \dot{b}_i it is the case that with the strategy s_i player i would be rational at every $h_i \in H_i$. Hence $s_i \in \bar{S}_i^k$ as well.

Now, since player i 's set of strategies S_i is finite and by definition $\bar{S}_i^{k+1} \subseteq \bar{S}_i^k$ for every $k \geq 1$, for some ℓ we eventually get $\bar{S}_i^\ell = \bar{S}_i^{\ell+1} \forall i \in I$ and hence $\bar{B}_i^{\ell+1} = \bar{B}_i^{\ell+2}$ for all $i \in I$. Inductively,

$$\emptyset \neq \bar{S}_i^\ell = \bar{S}_i^{\ell+1} = \bar{S}_i^{\ell+2} = \dots$$

and therefore

$$\bar{S}_i^\infty = \bigcap_{k=1}^{\infty} \bar{S}_i^k = \bar{S}_i^\ell \neq \emptyset$$

as required. \square

Proof of Proposition 1

\bar{B}_{DM}^1 contains belief systems in which in the information set $\{\delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}}\}$ (which follows the announcement of only ω_2 by the Announcer) the DM's belief assigns high probabilities to $\delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}}$. The strategies in \bar{S}_{DM}^1 corresponding to these belief systems prescribe 'middle' to the DM in the information set $\{\delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}}\}$. The crucial point is that $\bar{B}_{Announcer}^2$ contains *only* belief systems that assign strictly positive probabilities to these strategies of the DM. Thus, with *any* belief system in $\bar{B}_{Announcer}^2$, it is sub-optimal for the Announcer to announce ω_2 in the announcer's information set $\{\delta_1\}$, in which the Announcer is certain of ω_1 .⁵ Hence, $\bar{S}_{Announcer}^2$ does not contain strategies in which the Announcer announces just ω_2 when the announcer's information set is $\{\delta_1\}$. We conclude that \bar{B}_{DM}^3 contains only belief systems in which the belief at the information set $\{\delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}}\}$ assigns probability zero to $\delta_{1|\{\omega_1, \omega_2\}}$. Hence, \bar{S}_{DM}^3 contains only strategies with which the DM chooses 'middle' at the information set $\{\delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}}\}$.

Furthermore, already \bar{S}_{DM}^1 contains only strategies with which the DM chooses 'left' at the information set $\{\delta_{1|\{\omega_1, \omega_3\}}, \delta_{3|\{\omega_1, \omega_3\}}, \delta_{4|\{\omega_1, \omega_3\}}\}$ (i.e. when the Announcer announces just the new state ω_3). This is because prudent rationalizability implies that all the belief systems in

⁵Because according to every belief system in $\bar{B}_{Announcer}^2$, announcing just ω_2 will lead the DM with a positive probability to choose 'middle'.

\bar{B}_{DM}^1 assign a positive probability to strategies of the Announcer with which the Announcer announces the new state ω_3 even when the Announcer's information set (from the point of view of the DM!) is $\{\delta_{1|\{\omega_1, \omega_3\}}\}$ or $\{\delta_{4|\{\omega_1, \omega_3\}}\}$.

Also, \bar{B}_{DM}^1 contains belief systems in which the DM's belief in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ (when the Announcer announces both new states ω_2, ω_3) assigns high probability to δ_2 . The strategies in \bar{S}_{DM}^1 corresponding to these belief systems prescribe 'middle' to the DM in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$. Hence, $\bar{B}_{Announcer}^2$ contains *only* belief systems that assign strictly positive probabilities to these strategies of the DM. Thus, with *any* belief system in $\bar{B}_{Announcer}^2$, it is sub-optimal for the Announcer to announce both ω_2 and ω_3 in the announcer's information sets $\{\delta_1\}$ and $\{\delta_3\}$. Similarly, \bar{B}_{DM}^1 contains belief systems in which the DM's belief in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ assigns high probability to δ_1 . The strategies in \bar{S}_{DM}^1 corresponding to these belief systems prescribe 'left' to the DM in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$. Hence, $\bar{B}_{Announcer}^2$ contains *only* belief systems that assign strictly positive probabilities to these strategies of the DM. Thus, with *any* belief system in $\bar{B}_{Announcer}^2$, it is sub-optimal for the Announcer to announce both ω_2 and ω_3 in the Announcer's information sets $\{\delta_1\}, \{\delta_2\}$ or $\{\delta_3\}$. We conclude that \bar{B}_{DM}^3 contains only belief systems in which the belief at the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ assigns probability zero to $\delta_1, \delta_2, \delta_3$. That is, \bar{B}_{DM}^3 contains only belief systems that assign probability 1 to δ_4 at the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$. Hence, \bar{S}_{DM}^3 contains only strategies with which the DM chooses 'right' at the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$.

We thus conclude that \bar{S}_{DM}^3 contains a unique strategy s_{DM}^* . This strategy prescribes the DM to choose 'left' in the information set $\{\delta_{1|\{\omega_1\}}\}$ (i.e. when the Announcer does not announce any new state), to choose 'middle' in the information set $\{\delta_{1|\{\omega_1, \omega_2\}}, \delta_{2|\{\omega_1, \omega_2\}}, \delta_{4|\{\omega_1, \omega_2\}}\}$ (i.e. when the Announcer announces just the new state ω_2), to choose 'left' in the information set $\{\delta_{1|\{\omega_1, \omega_3\}}, \delta_{3|\{\omega_1, \omega_3\}}, \delta_{4|\{\omega_1, \omega_3\}}\}$ (i.e. when the Announcer announces just the new state ω_3) and to choose 'right' in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ (i.e. when the Announcer announces both new states ω_2, ω_3).⁶ \square

Proof of Proposition 2

When the buyer is confronted with the certificate $\{q_n\}$, his unique level-1 (prudent) rationalizable action is to buy $\beta(q_n)$, while when he is confronted with some range $\{q_m, \dots, q_n\}$ all the quantities in the interval $[\beta(q_m), \beta(q_n)]$ are level-1 (prudent) rationalizable (because any posterior belief of the buyer about the qualities with support $\{q_m, \dots, q_n\}$ can be derived from a belief of the buyer that the seller provides the certificate $\{q_m, \dots, q_n\}$ with an appropriate probability r_i when the seller knows that the quality is $q_i \in \{q_m, \dots, q_n\}$.) Consequently, the only level-2 prudent rationalizable strategies of the seller are those in which she provides

⁶This is also the unique strategy of the DM which is part of an awareness equilibrium satisfying reasoning refinement in Ozbay (2007).

the certificate $\{q_n\}$ when the quality is q_n (because any other certificate that she can provide $\{q_m, \dots, q_n\}$ will yield an expected sale strictly smaller than $\beta(q_n)$ with a full support belief about the level-1 prudent rationalizable strategies of the buyer, that have actions in the range $[\beta(q_m), \beta(q_n)]$).

Assume, inductively, that we have already proved that in all the level- $(2k - 1)$ prudent rationalizable strategies of the buyer, for every $i = 0, \dots, k - 1$ he buys the quantity $\beta(q_{n-i})$ when confronted with a certificate of the form $\{q_{n-i}, \dots, q_\ell\}$, and that in all the level- $2k$ prudent rationalizable strategies of the seller she indeed provides such a certificate when the quality is q_{n-i} . Then in all the level- $(2k + 1)$ (prudent) rationalizable strategies of the buyer he buys the quantity $\beta(q_{n-k})$ when confronted with a certificate of the form $\{q_{n-k}, \dots, q_\ell\}$ (because he believes that such a certificate could only be presented to him with the quality q_{n-k} , as by the induction hypothesis with each higher quality all the level- $2k$ prudent rationalizable strategies of the seller present a certificate where that higher value is the minimal value). Furthermore, when confronted with some range $\{q_m, \dots, q_{n-k}, \dots, q_\ell\}$ all the quantities in the interval $[\beta(q_m), \beta(q_{n-k})]$ are level- $(2k + 1)$ (prudent) rationalizable (because any posterior belief of the buyer about the qualities with support $\{q_m, \dots, q_{n-k}\}$ can be derived from a belief of the buyer on the level- $2k$ prudent rationalizable strategies of the seller in which the seller provides the certificate $\{q_m, \dots, q_\ell\}$ with an appropriate probability r_i when the seller knows that the quality is $q_i \in \{q_m, \dots, q_{n-k}\}$.)

Consequently, in all the level- $(2k + 2)$ prudent rationalizable strategies of the seller she provides the certificate $\{q_{n-k}, \dots, q_\ell\}$ when the quality is q_{n-k} (because any other certificate that she can provide $\{q_m, \dots, q_{n-k}, \dots, q_\ell\}$ will yield an expected sale strictly smaller than $\beta(q_{n-k})$ with a full support belief about the level- $(2k + 1)$ prudent rationalizable strategies of the buyer, that have actions in the range $[\beta(q_m), \beta(q_{n-k})]$).

Hence, the inductive claim obtains in particular for $k = n - 1$, concluding what we wanted to prove. \square

References

- [1] Asheim, G.B. and A. Perea (2005). Sequential and quasi-perfect rationalizability in extensive games, *Games and Economic Behavior* 53, 15–42.
- [2] Battigalli, P. (2006). Rationalization in signaling games: Theory and applications, *International Game Theory Review* 8, 67–93.
- [3] Battigalli, P. (1997). On rationalizability in extensive games, *Journal of Economic Theory* 74, 40–61.

- [4] Battigalli, P. (1996). Strategic rationality orderings and the best rationalization principle, *Games and Economic Behavior* 13, 178–200.
- [5] Battigalli, P. and M. Siniscalchi (2002). Strong belief and forward induction reasoning, *Journal of Economic Theory* 106, 356–391.
- [6] Blume, L., Brandenburger, A., and E. Dekel (1991). Lexicographic probabilities and choice under uncertainty, *Econometrica* 69, 61–79.
- [7] Brandenburger, A. and A. Friedenberg (2007). The relationship between rationality on the matrix and the tree, mimeo.
- [8] Chen J. and S. Micali (2013). The order independence of iterated dominance in extensive games, *Theoretical Economics* 8, 125–163.
- [9] Feinberg, Y. (2012). Games with unawareness, mimeo.
- [10] Filiz-Ozbay, E. (2012). Incorporating unawareness into contract theory, *Games and Economic Behavior* 76, 181–194.
- [11] Grant, S. and J. Quiggin (2013). Inductive reasoning about unawareness, *Economic Theory* 54, 717–755.
- [12] Grossman, S.J. (1981). The informational role of warranties and private disclosure of product quality, *Journal of Law and Economics* 24, 461–483.
- [13] Grossman, S.J. and O. Hart (1980). Disclosure laws and takeover bids, *Journal of Finance* 35, 323–334.
- [14] Guarino, P. (2017). An epistemic characterization of extensive-form rationalizability in dynamic games with unawareness, Maastricht University.
- [15] Halpern, J. and L.C. Rêgo (2014). Extensive games with possibly unaware players, *Mathematical Social Sciences* 70, 42–58.
- [16] Heifetz, A., Meier, M., and B.C. Schipper (2013). Dynamic unawareness and rationalizable behavior, *Games and Economic Behavior* 81, 50–68.
- [17] Heifetz, A. and A. Perea (2015). On the outcome equivalence of backward induction and extensive form rationalizability, *International Journal of Game Theory* 44, 37–59.
- [18] Li, Y.X. and B.C. Schipper (2018a). Strategic reasoning in persuasion games: An Experiment, University of California, Davis.
- [19] Li, Y.X. and B.C. Schipper (2018b). Disclosure under unawareness: An experiment, University of California, Davis.

- [20] Meier, M. and B.C. Schipper (2014). Bayesian games with unawareness and unawareness perfection, *Economic Theory* 56, 219–249.
- [21] Meier, M. and B.C. Schipper (2012). Conditional dominance in games with unawareness, University of California, Davis.
- [22] Milgrom, P. (2008). What the seller won't tell you: persuasion and disclosure in markets, *Journal of Economic Perspectives* 22, 115–131.
- [23] Milgrom, P. (1981). Good news and bad news: representation theorems and applications. *The Bell Journal of Economics* 12, 380–391.
- [24] Milgrom, R. and J. Roberts (1986). Relying on the information of interested parties, *Rand Journal of Economics* 17, 18–32.
- [25] Ozbay, E. (2007). Unawareness and strategic announcements in games with uncertainty, in: Samet, D. (ed.), *Proceedings of the 11th conference on Theoretical Aspects of Rationality and Knowledge*, Presses Universitaires de Louvain, pp. 231-238.
- [26] Pearce, D.G. (1984). Rationalizable strategic behavior and the problem of perfection, *Econometrica* 52, 1029–1050.
- [27] Perea, A. (2018a). Common belief in rationality in games with unawareness, Maastricht University.
- [28] Perea, A. (2018b). Why forward induction leads to the backward induction outcome: A new proof for Battigall's theorem, *Games and Economic Behavior* 110, 120–138.
- [29] Rêgo, L.C. and J. Halpern (2012). Generalized solution concepts in games with possibly unaware players, *International Journal of Game Theory* 41, 131–155.
- [30] Reny, P. (1992). Backward induction, normal form perfection and explicable equilibria, *Econometrica* 60, 627–649.
- [31] Schipper, B.C. (2018a). Kuhn's theorem for extensive-form games with unawareness, University of California, Davis.
- [32] Schipper, B.C. (2018b). Discovery and equilibrium in games with unawareness, University of California, Davis.
- [33] Schipper, B.C. (2016). Network formation in a society of fragmented knowledge and awareness, University of California, Davis.
- [34] Schipper, B.C. (2014). Unawareness - A gentle introduction to both the special issue and the literature, *Mathematical Social Sciences* 70, 1–9.

- [35] Schipper, B.C. and H.Y. Woo (2019). Political awareness, microtargeting of voters, and negative electoral campaigning, *Quarterly Journal of Political Science* 14, 41–88.
- [36] Shimoji, M. and J. Watson (1998). Conditional dominance, rationalizability, and game forms, *Journal of Economic Theory* 83, 161–195.
- [37] Stahl, D. (1995). Lexicographic rationalizability and iterated admissibility, *Economics Letters* 47, 155–159.