Essays on Dynamic Consistency in Games with Ambiguous Information

Dissertation zur Erlangung des Doktorgrades der Wirtschaftswissenschaften (Dr. rer. pol.) an der Fakultät für Wirtschaftswissenschaften der Universität Bielefeld

vorgelegt von

Marieke Pahlke

April 2021

Gutachter

Erster Gutachter:	Prof. Dr. Frank Riedel
	Universität Bielefeld
Zweiter Gutachter:	Prof. Peter Klibanoff, Ph.D.
	Northwestern University

Adresse

Universität Bielefeld
Fakultät für Wirtschaftswissenschaften
Universitätsstraße 25
33615 Bielefeld
Germany

Printed on non-ageing paper – ISO 9706

Acknowledgements

This thesis was a long journey and would not have been possible without the support and help of several people:

First, I would like to express my deepest gratitude to my supervisor Frank Riedel for his constant support, critical questions, and guidance. I sincerely appreciate that he always found time to discuss my ideas and problems and motivate me to go deeper and further. Thanks to him, I have learned a lot about research and academic life in the last few years.

I also want to thank my second supervisor Peter Klibanoff whose support and interest were very beneficial, especially during my stay in Evanston. The experiences and especially the exchange with the researchers there were very enriching.

Furthermore, I would like to thank the members of the ORA research group. In particular, I would like to mention the members of our reading group in Bielefeld as well as Jean-Marc Tallon and Frédéric Koessler for the possibility to visit and present my work in Paris and many exciting and motivating discussions.

My special gratitude also goes to all members of the IMW. There was always someone with an open door for helpful discussions. Especially I would like to thank Patrick (my living math lexicon) and Annika for proofreading this thesis and organizing so many joyful social events, Christmas parties, and hat-craft sessions. Also, the administrative assistance provided by the chairman of our graduate school Herbert Dawid, and our secretaries Bettina, Karin, and Diana has been a great support for me during my studies.

Moreover, I acknowledge the financial support of the DFG within BiGSEM and the ORA project (Ri 1128-9-1). I also thank the DAAD and the BGTS for the financial support of my research stay in Evanston.

This thesis would not have been possible without my family and friends. I want to thank Dagmar, Julia, Nina, Sabine, and Sabine for our motivating, inspiring, and encouraging meetings. I also thank my mother and my grandparents for their tireless support for my entire life. Finally, I want to thank Philipp for his encouragement, patience, and always finding a way to cheer me up.

Contents

General Introduction **1**

1 Dynamic Consistency in Incomplete Information Games with Multiple Priors **8**

1.1 Introduction . . . 8
1.2 Model . . . 10
1.2.1 Dynamic Inconsistency . . . 12
1.2.2 Beliefs . . . 14
1.3 Ex-Ante and Interim Equilibria . . . 29
1.3.1 Definition and Existence . . . 29
1.3.2 Relation of Ex-Ante and Interim Equilibria . . . 31
1.4 Sequential Equilibria . . . 34
1.4.1 Sequential Rationality and Rectangularity . . . 37
1.4.2 Properties of Sequential Equilibria . . . 38
1.5 Conclusion and Discussion . . . 41
1.6 Appendix . . . 44
1.6.1 Proofs . . . 44
1.6.2 Further Results . . . 52
1.6.3 Examples . . . 54

2 Dynamic Consistency in Ambiguous Persuasion **61**

2.1 Introduction . . . 61
2.2 Model . . . 63
2.2.1 Model of BLL . . . 63
2.2.2 Dynamically Consistent Belief Formation Process . . . 66
2.3 Dynamic Consistency and PBE . . . 70
2.4 Value of Information . . . 74
2.5 Conclusion and Discussion . . . 76

3 Dynamic Consistency in Ambiguous Dutch Auctions **79**
3.1 Introduction 79
3.2 Ambiguous Dutch Auction 82
3.2.1 Basic Setting 82
3.2.2 The Modified Dutch Mechanism (MDM) 83
3.2.3 Information and Strategies of Buyers 84
3.2.4 Updating Beliefs and Worst-Case Belief 85
3.2.5 Worst-Case Belief of Bose and Daripa (2009) 86
3.3 Dynamically Consistent Buyers 88
3.3.1 Rectangularity 88
3.3.2 Rectangular Worst-Case Beliefs 89
3.3.3 Surplus Extraction under Dynamic Consistency 93
3.4 Numerical Example 98
3.5 Rectangularity versus Consistent Planning 100
3.5.1 Consistent Planning in the Ambiguous Dutch Auction 100
3.5.2 Consistent Planning and Rectangularity 101
3.6 Conclusion and Discussion 102
3.7 Appendix 104
3.7.1 Derivation Worst-Case Belief 104
3.7.2 Proofs of Section 3.3 105
3.7.3 Correction of the Results of Bose and Daripa (2009) 112

Bibliography **125**

General Introduction

Strategic interactions play a fundamental role in most economic settings. Game theory is the mathematical language we use to describe the underlying logic and implications of strategic interactions. In game theory, we call a situation in which every participant (called player) knows everything that is payoff-relevant a game with complete information. Apart from what every player does (i.e., their actions), every player knows everything that happens in the future and even the other players' preferences.

However, in reality, a bidder who participates in an auction usually does not know the valuation of the opposing bidders, and the outcome of a peace negotiation may depend on the result of an upcoming election.

Harsanyi (1967, 1968a,b) proposes an approach to deal with these, so called, games with incomplete information: At the beginning of a game, all payoff-relevant parameters are determined by the realization of a random variable. The realization of this random variable is often called a type or state.

The approach of Harsanyi and most of the subsequent research assumes that players know the probability distribution of these states or types. However, a bidder may only have a vague idea about the valuation of the opposing bidders and an event in the future, e.g., the result of the United States presidential election in 2024, is still hard to predict. Thus, motivated by the distinction between risk and uncertainty of Knight (1921) and the famous paradox of Ellsberg (1961) such immeasurable uncertainty (or ambiguity) has been introduced in decision and game theory. However, as one can imagine, ambiguity leads to difficulties which can not occur under risk. One example of this difficulties is dynamic inconsistency.

In this doctoral thesis, I analyze different dynamic games with ambiguity and possible problems and applications.

Risk and Ambiguity

Exactly one hundred years ago Knight (1921) distinguished distinct kinds of uncertainty. Whereas it is simple to derive the probability of a coin toss or a dice roll, it is much harder to predict precise probabilities for more complex events, e.g., stock prices (especially during a financial crisis) or the death rate of a new pandemic. Knight (1921) distinguishes between *risk* and *uncertainty*. He uses the term risk for randomness or uncertainty that can be completely captured by one probability distribution, while uncertainty describes situations where the probability distribution of outcomes is unknown. Thereafter, the latter was named *Knightian uncertainty* or *ambiguity*.

Ellsberg (1961) captured the implications of risk and ambiguity in his famous thought experiment called the Ellsberg paradox. The Ellsberg paradox shows that agents prefer to bet on a risky urn instead of an ambiguous urn. This behavior cannot be explained by the theory of subjective expected utility introduced by Savage (1954). The most prominent decision models which capture the difference of risk and ambiguity are: the maxmin expected utility (MEU) by Gilboa and Schmeidler (1989), the Choquet expected utility by Schmeidler (1989), the incomplete preference model by Bewley (2002) and the smooth ambiguity model by Klibanoff et al. (2005).[1]

Since then, ambiguity has been introduced into many decision and game-theoretic models and has been applied, e.g., in auction (Bose et al. (2006)), mechanism design (Di Tillio et al. (2016)), asset pricing (Ju and Miao (2012)), optimal stopping (Riedel (2009)) and cheap talk (Kellner and Le Quement (2018)). However, ambiguity may induce dynamically inconsistent behavior, which complicates the analysis of dynamic settings.

Dynamic Inconsistency

To analyze agents' behavior in dynamic settings with ambiguous beliefs, one must first specify how agents update beliefs. In this thesis, I assume prior-by-prior Bayesian updating (or full Bayesian updating, Pires (2002)), i.e., the set of interim beliefs consists of the Bayesian update of each ex-ante belief. Additional to prior-by-prior Bayesian updating, different updating rules are defined in the literature, e.g, maximum likelihood updating (Gilboa and Schmeidler (1993)), relative maximum likelihood updating (Cheng (2021)) and the updating rules of Hanany and Klibanoff (2007, 2009). Except for the updating rules of Hanany and Klibanoff (2007, 2009), all of them can lead to dynamically inconsistent behavior in combination with maxmin preferences.

[1] For recent surveys of the literature on ambiguity and axiomatic foundation see Gilboa (2009), Gilboa and Marinacci (2016) and Etner et al. (2012).

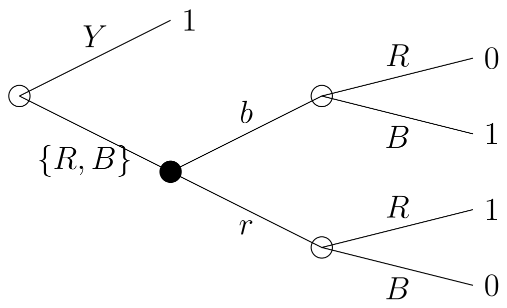

Figure 1: Dynamic Three Colors Ellsberg Experiment

Roughly speaking, new information can lead to a change in the worst-case belief, which induces a different optimal strategy and, therefore, dynamically inconsistent behavior. To illustrate the problem of dynamic inconsistency, consider the following dynamic version of the three-color Ellsberg experiment. An urn contains 30 red (R) balls and 60 blue (B) or yellow (Y) balls. The exact distribution of blue and yellow balls is unknown. Ex-ante the agent only knows that the probability of a red ball being drawn is $\mathbb{P}(R) = \frac{1}{3}$, whereas the probability of a blue (or yellow) ball is $\mathbb{P}(B), \mathbb{P}(Y) \in \left[0, \frac{2}{3}\right]$. The agent first observes if the drawn ball is yellow or not. Then, he can choose between betting on the event "a blue ball is drawn" or the event "a red ball is drawn". The decision problem is depicted in Figure 1. Empty circles represent nature moves and the solid circle the decision node of the agent. Further, r and b denotes the choice of the agent, i.e., betting on red or blue, respectively. Let us assume that the agent's preferences can be modeled by the maxmin expected utility model of Gilboa and Schmeidler (1989). Ex-ante, before the agent learns if the ball is yellow, the worst expected utilities of r and b are

$$\min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \mathbb{E}(u(b)) = \min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \frac{1}{3}\cdot 0 + \mathbb{P}(B) + \frac{2}{3} - \mathbb{P}(B) = \frac{2}{3},$$

$$\min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \mathbb{E}(u(r)) = \min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \frac{1}{3}\cdot 1 + \frac{2}{3} - \mathbb{P}(B) = \frac{1}{3},$$

and the agent prefers betting on blue. Now, suppose the ball is not yellow and the agent updates his belief set with prior-by-prior Bayesian updating. Then, the interim worst-case expected utilities of r and b are

$$\min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \mathbb{E}(u(b)) = \min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \frac{\frac{1}{3}}{\frac{1}{3}+\mathbb{P}(B)}\cdot 0 + \frac{\mathbb{P}(B)}{\frac{1}{3}+\mathbb{P}(B)} = 0,$$

$$\min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \mathbb{E}(u(r)) = \min_{\mathbb{P}(B)\in\left[0,\frac{2}{3}\right]} \frac{\frac{1}{3}}{\frac{1}{3}+\mathbb{P}(B)} + \frac{\mathbb{P}(B)}{\frac{1}{3}+\mathbb{P}(B)}\cdot 0 = \frac{1}{3}.$$

Now, the agent prefers betting on red. Hence, learning that the color of the ball is not yellow changes the optimal action of the agent. He does not follow the optimal ex-ante plan at the interim stage and behaves dynamically inconsistent.

Solving the problem of dynamically inconsistent behavior is not straightforward. There is a well-known conflict between dynamic consistency and consequentialism in the literature on ambiguous beliefs in dynamic settings. Intuitively, consequentialism states that interim preferences do not depend on past discarded actions or events that are not consistent with the given information set. Among others, Ellis (2018) and Aryal and Stauber (2014) show that dynamic consistency, consequentialism, and a common prior assumption are only fulfilled simultaneously if players behave as expected utility maximizers. Hence, ambiguity, dynamic consistency, and consequentialism cannot occur simultaneously. The literature proposes different approaches to overcome this impossibility result. The three main approaches are the following:

- Siniscalchi (2011) axiomatize and generalize the *consistent planning* approach of Strotz (1955) to an ambiguous decision-theoretical setting. He states that

  > "...consistent planning (CP) is a refinement of backward induction. If there are unique optimal actions at any point in the tree, the two concepts coincide. Otherwise, CP complements backward induction with a specific tie-breaking rule: indifferences at a history h are resolved by considering preferences at the history that immediately precedes h."

  Siniscalchi (2011) shows that consistent planning satisfies consequentialism but not dynamic consistency.

  In our Ellsberg urn example of Figure 1, the optimal interim action of an agent who uses consistent planning is betting on red. At the ex-ante stage, the optimal plan that is consistent with the optimal interim choice is betting on red.

- Sarin and Wakker (1998) and Epstein and Schneider (2003) define *rectangularity* for different decision-theoretic settings.[2] They show that rectangularity implies dynamically consistent behavior in their setting. Furthermore, Riedel et al. (2018) explore a dynamic decision-theoretic setting where preference relations are defined on pairs of imprecise probabilistic information and acts. Roughly speaking, they show that an ambiguity-averse agent that behaves dynamically consistently chooses a rectangular subjective ex-ante belief set and evaluates acts according to the worst-case belief given his subjective ex-ante belief set.

  Rectangularity is a condition on belief sets that leads to a generalized version of the law of iterated expectation. Intuitively, the agent takes possible future worst-case beliefs into account. Therefore, rectangularity depends on the information structure of the decision problem. Epstein and Schneider (2003) and Riedel

[2] Sarin and Wakker (1998) do not use the term rectangularity. Instead, they use the term *reduced family of probability measures* for the rectangular hull as defined in Epstein and Schneider (2003).

et al. (2018) show that an agent with a rectangular belief set satisfies dynamic consistency for a given information structure and consequentialism.

A different but related approach is the one of Hill (2020). He solves the conflict of dynamic consistency and consequentialism by reformulating the dynamic consistency axiom on subjective trees. Further, if subjective trees are represented by a partition of the state space, dynamic consistency implies rectangularity.

In our Ellsberg urn example of Figure 1, the agent's optimal ex-ante and interim action with rectangular beliefs is betting on red. The optimal interim action becomes ex-ante optimal.

- The *updating rules of Hanany and Klibanoff (2007, 2009)* satisfy dynamic consistency but violate consequentialism. Roughly speaking, a decision maker with maxmin preferences and a set of priors $\mathcal{P}$ only updates a subset of $\mathcal{P}$. This subset depends on his ex-ante optimal choice and ensures that the ex-ante optimal plan over contingencies is interim optimal. For smooth-ambiguity preferences, Hanany and Klibanoff (2009) derive the so-called *smooth-rule*. The smooth-rule updates the second order belief such that the ex-ante worst case beliefs receives more weight. These updating rules predict a different optimal strategy in our Ellsberg urn example of Figure 1. Now, betting on blue is ex-ante and interim optimal.

Besides, Klibanoff et al. (2009) axiomatize a recursive smooth ambiguity model.

So far, it is still an open question if there exist settings or conditions where these three approaches lead to similar or different optimal behavior. Consistent planning and rectangularity are both related to backward induction. Thus, in some settings, it is argued that they induce equivalent optimal behavior. This is the case in our example of Figure 1. However, we will see in Chapter 3 that settings exist where the equilibrium strategies differ. The updating rules of Hanany and Klibanoff (2007, 2009) focus on the ex-ante optimal choice. Therefore, as in our Ellsberg urn example, they will usually predict different behavior as consistent planning or rectangularity.

Ambiguity in Dynamic Games

The formation of beliefs plays a fundamental role in games with incomplete information and motivated the definition of different equilibrium concepts as, e.g., perfect Bayesian equilibrium (Fudenberg and Tirole (1991a,b)) or sequential equilibrium (Kreps and Wilson (1982)). One essential assumption for these equilibrium concepts is that rational players use (whenever possible) Bayes' rule to update their beliefs after observing new information. Further, in the canonical model with expected utility maximizers, Bayes' rule leads to dynamically consistent behavior. The best response at the ex-ante stage is also optimal at the interim or ex-post stage. This result breaks down if players are ambiguity-averse and maximize their worst-case expected utility. Dynamically inconsistent behavior makes it impossible to use the standard concepts of sequential equilibrium and perfect Bayesian equilibria and complicates the analysis of dynamic games with ambiguity.

Most of the literature on ambiguity in dynamic games focuses on the optimal decision at one stage, e.g., Lo (1998), Kajii and Ui (2005), and Eichberger and Kelsey (1999) focus on the interim optimization. Lo (1999) explores extensive-form games with maxmin expected utility. Instead of imposing conditions to ensure dynamic consistency, he introduces an equilibrium concept that explicitly requires that each player chooses a strategy that is interim optimal in an equilibrium. Kajii and Ui (2005) also formulate an incomplete information game with multiple priors. Their setting only consists of two stages. First, new information arises due to signals which are independent of the strategies. Then, players play a simultaneous-move game. Therefore, the information structure is very close to decision-theoretic settings and does not capture the strategic aspects that, e.g., occur in signaling games. Further, Eichberger and Kelsey (1999) investigate signaling games in which beliefs are represented by capacities. They formulate an equilibrium concept similar to perfect Bayesian equilibrium, but they focus on interim utility maximization.

Hanany et al. (2020) define sequential equilibria for multistage games with incomplete ambiguous information. In their setting, players have smooth-ambiguity preferences and use the smooth-rule of Hanany and Klibanoff (2009) to update beliefs. Battigalli et al. (2019) analyze self-confirming equilibria for players with smooth-ambiguity preferences. But instead of the smooth-rule, they use prior-by-prior Bayesian updating and the consistent planning approach of Siniscalchi (2011). We discuss these papers in more detail in Chapter 1.

So far, the literature on rectangularity in games is small. Liu and Xiong (2016) define rectangularity in a game-theoretical setting. However, they use a similar model as Kajii and Ui (2005) and the information structure is very similar to decision theory and cannot capture strategic aspects. Muraviev et al. (2017) use rectangularity to show outcome equivalence between mixed and behavioral strategies in games with Ellsberg strategies.

Generalizing rectangularity to games is not straightforward. In games, players receive new information by observing actions played by their opponents. If players have heterogeneous information, observing opponents' actions reveals information about the states. These strategic effects cannot occur in decision-theoretical models. Further, as already mentioned Ellis (2018) and Aryal and Stauber (2014) show that dynamic consistency, consequentialism, and a common prior assumption can only be satisfied if the players behave as expected utility maximizer.

Contribution

In this doctoral thesis, I generalize rectangularity to different settings. Chapter 1 studies multistage games with ambiguous incomplete information about states of the world or types of opponents. Here, ambiguity arises due to ambiguity about the choice of nature. There is no ambiguity about the strategies of the opponents. First, I generalize rectangularity to multistage games. Then, I define and show the existence of sequential equilibria with rectangular beliefs. To overcome the impossibility result of Ellis (2018) and Aryal and Stauber (2014) I weaken the dynamic consistency and common prior assumption slightly. Similar to Epstein and Schneider (2003) and Riedel et al. (2018), dynamic consistency is only required for the information structure induced by the game. Furthermore, players may have heterogeneous rectangular belief sets if they receive heterogeneous information during the game.[3]

In Chapter 2, I analyze ambiguous persuasion with dynamically consistent players. In the ambiguous persuasion setting, the Sender can design an ambiguous communication device by choosing a set of communication devices. Here, ambiguity arises due to an ambiguous strategy that cannot be modeled by the setting of Chapter 1. I show that the Sender can gain from ambiguous persuasion even if the Receiver behaves dynamically consistently. Furthermore, I discuss the relation to the (negative) value of information of ambiguous communication for the Receiver.

Chapter 3 investigates a dynamic decreasing price auction with two buyers. The buyers have ambiguous beliefs about the valuation of the opponent buyer. The timing of the auction is discrete. At the beginning of each period, one player is chosen randomly and secretly and gets a price offer p_k. If he rejects the price offer, the other buyer gets the same price offer. If one of the buyers accepts, the auction ends immediately. If no one accepts the current price, the auction proceeds to the next period with a new price offer $p_{k+1} < p_k$. The timing becomes complex since the buyers never learn who gets the price offer first. Therefore, the auction cannot be modeled as a multistage game of Chapter 1. I show that even if buyers behave dynamically consistently, the seller can extract almost all surplus. Furthermore, in this setting, consistent planning, and rectangularity lead to different equilibrium strategies.

[3] See Section 1.2.2.2 for a more detailed discussion.

Chapter 1

Dynamic Consistency in Incomplete Information Games with Multiple Priors

1.1 Introduction

In this chapter, we model multistage incomplete information games with uncertainty about types or states, which include risk and ambiguity. The uncertainty is given by a common set of probability distributions $\mathcal{P}$ over states or types, called imprecise probabilistic information. If $\mathcal{P}$ is a singleton, the uncertainty reduces to risk, and players face a usual incomplete information game. We assume that players have maxmin expected preferences (MEU) as introduced by Gilboa and Schmeidler (1989) and maximize their worst-case expected utility.

We contribute to the literature of ambiguity in dynamic games in two ways. First, we characterize a belief formation process that ensures dynamically consistent behavior. We assume that players update their belief sets prior-by-prior using Bayes' rule whenever possible. Players know which information they could potentially get in the future, i.e., they know the structure of the game. When forming a set of ex-ante beliefs, players combine their knowledge about the information structure and the information given by the common set of imprecise probabilistic information. Formally, this leads to a belief set for each player that satisfies a rectangularity condition and ensures dynamically consistent behavior.

Second, we generalize the concept of sequential equilibrium to incomplete information games with ambiguity. Sequential equilibria require sequential rationality, which cannot be satisfied if players are dynamically inconsistent. Thus, using our belief formation process, we ensure dynamically consistent behavior and, therefore, the existence of a sequential equilibrium. Furthermore, we show that ambiguity can induce sequential equilibria that cannot exist without ambiguity.

Using our belief formation process makes it possible to analyze dynamically consistent behavior of ambiguity-averse players in dynamic games. This allows for an analysis of ex-ante and ex-post stages, which, e.g., facilitates a consistent welfare analysis.

Ellis (2018) and Aryal and Stauber (2014) fix a common set of priors in a game with ambiguity. They argue that generally, a common ex-ante belief set can only be rectangular for all players if ambiguity about other players' types reduces to risk. In our setting, if players receive heterogeneous information during the game, our belief formation process leads to heterogeneous ex-ante belief sets. Intuitively, players interpret the common imprecise probabilistic information differently since they take their own information structure into account. Considering this heterogeneity allows for rectangular ex-ante belief sets of all players despite the common set of imprecise probabilistic information.[1] Therefore, the critique of Ellis (2018) and Aryal and Stauber (2014) does not apply to our setting.

The structure of this chapter is as follows. First, we summarize the related literature. In Section 1.2, we formulate the extensive-form game with ambiguity and define belief sets that satisfy rectangularity. Section 1.3 shows the existence and the relation of ex-ante and interim equilibria. In Section 1.4, we prove the existence of sequential equilibria with rectangular beliefs and discuss the relationship between sequential rationality and rectangularity. Furthermore, we give an example that shows that ambiguity might induce new sequential equilibria. Finally, Section 1.5 concludes.

Related Literature As already mentioned in the general introduction, Epstein and Schneider (2003) and Riedel et al. (2018) axiomatize rectangularity in decision-theoretical settings. However, due to the different information structures in games and decision-theoretical settings and strategic effects, we can not directly apply their definition of rectangularity to games.

Hanany et al. (2020) and Battigalli et al. (2019) apply the approaches of Klibanoff et al. (2009) and Siniscalchi (2011) to games. Hanany et al. (2020) explore a finite extensive-form multistage game with incomplete information but use smooth ambiguity aversion instead of multiple priors. They show that the smooth-rule of Hanany and Klibanoff (2009) is equivalent to sequential optimality and induces the existence of sequential

[1] Our setting does not satisfy Common Ex-Ante Behavior (Axiom 4) of Ellis (2018), since we allow for heterogeneous rectangular ex-ante belief sets. Therefore, his impossibility result does not apply in our setting.

equilibria. Battigalli et al. (2019) explore (rationalizable) self-confirming equilibria in dynamic games with smooth ambiguity-averse players. They use the consistent planning approach of Siniscalchi (2011). We discuss the relation to these papers in Section 1.5.

Rectangularity has been rarely used in games. Liu and Xiong (2016) define rectangularity in a game-theoretical setting similar to Kajii and Ui (2005). Their game only consists of two stages. First, players observe a signal which reveals information about the ambiguous state. Then, they play a simultaneous-move game. The signal is independent of the strategies of the players. Therefore, their setting cannot capture strategic aspects that arise, for example, in signaling games.

Muraviev et al. (2017) explore extensive-form games where players can use Ellsberg strategies, introduced by Riedel and Sass (2014). Ellsberg strategies extend mixed strategies to ambiguous strategies, i.e., instead of playing a probability distribution over the pure strategies, a player chooses a set of probability distributions. They show that a rectangularity condition ensures outcome-equivalence between mixed and behavioral strategies, but they do not formulate a general equilibrium concept for such games. However, similar to this chapter, Muraviev et al. (2017) have to construct a filtration to define rectangularity. The difference lies in the source of ambiguity. In Ellsberg games, ambiguity arises due to ambiguous strategies. In our setting, ambiguity arises due to incomplete ambiguous information about states or types.

1.2 Model

This section defines a finite extensive-form multistage game with incomplete information, multiple priors, and perfect recall. The definition is similar to Hanany et al. (2020), but instead of smooth ambiguity aversion, players face imprecise probabilistic information and maxmin preferences. We will show later that given this imprecise probabilistic information and the information structure of the game, each player constructs a subjective set of ex-ante beliefs. Given these beliefs, each player evaluates a strategy by using maxmin expected utility (MEU).

Definition 1.1. *A tuple* $\Gamma = (N, H, (\mathcal{I}_i)_{i\in N}, (u_i)_{i\in N}, \mathcal{P})$ *is a* ***finite extensive-form multistage game with incomplete information, perfect recall and multiple priors*** *with:*

- N *is a finite set of players.*
- H *is a finite set of (terminal) histories, where each history* h *is of the form* $h = (h_{-1}, (h_{0,i})_{i\in N}, \ldots, (h_{T,i})_{i\in N})$.

For $0 \leq t \leq T+1$, *let* $H^t := \{h^t := (h_{-1}, (h_{0,i})_{i \in N}, \dots, (h_{t-1,i})_{i \in N}) : h \in H\}$ *be the set of partial histories up to but not including stage* t*. For each player* $i \in N$, $0 \leq t \leq T+1$ *and* $h^t \in H^t$, $A_i(h^t) := \{\hat{h}_{t,i} : \hat{h} \in H, \hat{h}^t = h^t\}$ *is the set of actions available to player* i *at* h^t*. The set of uncertain types or states is* H^0.

- $\mathcal{I}_i := \bigcup_{0 \leq t \leq T} \mathcal{I}_i^t$ *are the information sets for player* i*, where each* $\mathcal{I}_i^t$ *is a partition of* H^t *such that for all* h^t, $\hat{h}^t \in H^t, \hat{h}^t \in I_i(h^t)$ *implies* $A_i(h^t) = A_i(\hat{h}^t)$*, where* $I_i(h^t)$ *is the unique element of* $\mathcal{I}_i^t$ *such that* $h^t \in I_i(h^t)$*. Furthermore,* $\mathcal{I} = \bigcup_{i \in N} \mathcal{I}_i$ *denotes the set of all information sets.*
  For $0 \leq t \leq T$ *and* $h^t \in H^t$, $R_i(h^t) := ((I_i(h^s), h^t_{s,i})_{0 \leq s < t}, I_i(h^t))$ *is the ordered list of information sets encountered by player* i *and the action taken by player* i *given the partial history* h^t*. The game satisfies perfect recall in that for each player* i*, each stage* $0 \leq t \leq T$*, and each partial history* $h^t, \hat{h}^t \in H^t$, $I_i(h^t) = I_i(\hat{h}^t)$ *implies* $R_i(h^t) = R_i(\hat{h}^t)$.
- $u_i \colon H \to \mathbb{R}$ *is the (utility) payoff of player* i.
- $\mathcal{P} \subset \Delta(H^0)$ *is the set of imprecise probabilistic information over states or types which is homogeneous across all players. We assume that* $\mathcal{P}$ *is compact and all* $\pi \in \mathcal{P}$ *have full support, i.e.,* $\pi(h^0) > 0$ *for all* $h^0 \in H^0$ *and all* $\pi \in \mathcal{P}$.

The definition above allows for imperfectly observed actions as well as for private information about types or states. The multistage structure assumes that each player chooses an action at each stage. Since $A_i(h^t)$ can be a singleton, this assumption is not restrictive and sequential play can be modeled as well.

The only difference compared to the standard setting without ambiguity is the last bullet point. Instead of having an exact distribution over types, players have imprecise probabilistic information given by a set of possible distributions $\mathcal{P}$. If $\mathcal{P}$ is a singleton, there is no ambiguity, and the game reduces to the standard version without ambiguity. The compactness assumption on $\mathcal{P}$ ensures the existence of a worst-case belief. Full support ensures that an out-of-equilibrium path only occurs because of non-completely mixed strategies. Therefore, for completely mixed strategies, Bayes' rule is always well-defined.

At each stage, conditional on their information set I_i^t, players choose a distribution over their actions which are available at I_i^t. A strategy profile consists of these distributions for each player and information set.

Definition 1.2. *A* ***(behavioral) strategy*** *for player* i *in a game* Γ *is a function* σ_i *such that* $\sigma_i(I_i^t) \in \Delta(A_i(I_i^t))$ *for each* $I_i^t \in \mathcal{I}_i^t$*, where* $\Delta(A_i(I_i^t))$ *denotes the set of all probability vectors over* $A_i(I_i^t)$.
Furthermore, let Σ_i *denote the set of all strategies for player* i, $\sigma := (\sigma_i)_{i \in N}$ *be a strategy*

profile, and $\sigma_{-i} := (\sigma_j)_{j \neq i}$ *be the strategies of all opponents of* i.

A strategy profile induces a transition probability with which a particular (partial) history occurs. For a given strategy profile σ, a history h, and $0 \leq r \leq t \leq T+1$, the probability of reaching h^t starting from h^r is defined by

$$p_\sigma(h^t|h^r) := \prod_{j \in N} \prod_{r \leq s < t} \sigma_j(I_j(h^s))(h_{s,j}).$$

It will be useful to split $p_\sigma(h^t|h^r)$ in one part that only depends on the player himself and another part that represents the actions of all opponents. We define

$$p_{\sigma_i}(h^t|h^r) := \prod_{r \leq s < t} \sigma_i(I_i(h^s))(h_{s,i}),$$

and

$$p_{\sigma_{-i}}(h^t|h^r) := \prod_{j \neq i} \prod_{r \leq s < t} \sigma_j(I_j(h^s))(h_{s,j}).$$

Then, $p_{\sigma_i}(h^t|h^r) p_{\sigma_{-i}}(h^t|h^r) = p_\sigma(h^t|h^r)$.

1.2.1 Dynamic Inconsistency

Multiple priors can lead to dynamically inconsistent behavior. To illustrate dynamic inconsistency, we repeat the three-player example from Aryal and Stauber (2014).[2] We will use this example as a running example in the following sections to illustrate notation and results.

Running Example. *The game, depicted in Figure 1.1, shows that ambiguity and multiple priors can lead to dynamically inconsistent behavior. There are two players, player* 1 *and player* 2*. First, nature chooses the state* L*,* R*, or* O*. Let* l*,* r *and* o *be the probability of* L*,* R *and* O*, respectively. The imprecise probabilistic information is given by an* ϵ*-contamination of the distribution that assigns probability one to* R*, i.e.,* $(l, r, o) = (0, 1, 0)$*. We denote the set of all probability distributions over* $\{L, R, O\}$ *by* Δ*. Then, the imprecise probabilistic information is* [3]

$$\mathcal{P} = \{(1-\epsilon)(0,1,0) + \epsilon(l,r,o) \, : \, (l,r,o) \in \Delta\}.$$

[2] To fit our definition of multistage games, one would have to include a constant action for player 2 at the information set of player 1 and a constant action for player 1 at the information set of player 2. Since this does not change the results of the example, we skip these constant actions due to notational convenience.

[3] $\mathcal{P}$ does not satisfy the full support assumption stated in Definition 1.1. Formally, the full support assumption is needed to guarantee that the probability of reaching an information set is zero if and only if all partial histories leading to this information set have probability zero because of the played strategy profile. Hence, if the probability of reaching an information set is zero for one ex-ante belief $\pi \in \mathcal{P}$, then it is zero for all ex-ante beliefs. Due to the ϵ-contamination structure of $\mathcal{P}$, the probability

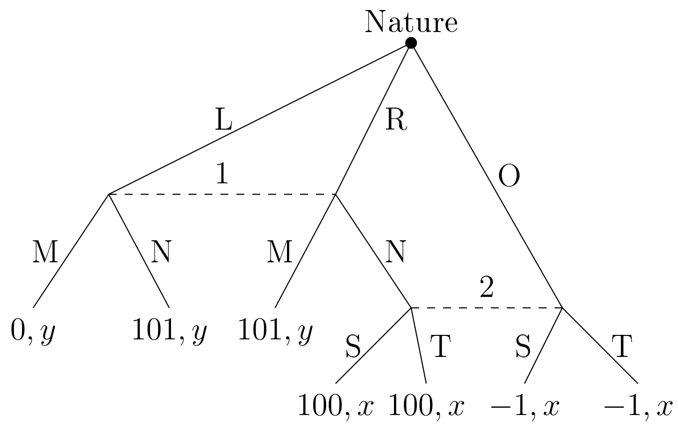

Figure 1.1: Three-Player Game of Aryal and Stauber (2014)

To illustrate the problem of dynamic inconsistency, assume that $\mathcal{P}$ represents the ex-ante beliefs of player 1*.*

After the choice of nature, player 1 *can observe if the state is O or not. If the state is not O, player* 1 *can choose between N and M. If the state is O or R and player* 1 *played N, player* 2 *can choose an action without knowing which of the two cases is true. Dashed lines depict the information sets of both players. For the moment, let us concentrate on player* 1*. His payoffs are independent of the strategy of player* 2*. He gets his lowest payoff, -1, if the state is O. Therefore, his ex-ante worst-case belief gives the highest possible probability to O, i.e., the ex-ante worst-case belief is* $(0, 1-\epsilon, \epsilon)$*. Since the probability of L is zero, his optimal ex-ante strategy is playing M with probability one.*

Now, we check if player 1 *has an incentive to deviate from his optimal ex-ante strategy after observing that the state is not O. Updating $\mathcal{P}$ prior-by-prior using Bayes' rule and conditioning on the event $\{L, R\}$, leads to the following set of updated beliefs*

$$Bay\big(\mathcal{P}|\{L,R\}\big) = \{(l, r, 0) = (1-r, r, 0) \,:\, r \in [1-\epsilon, 1]\}.$$

His interim worst-case belief depends on his strategy. Playing M with probability one would lead to a payoff of zero if the state is L and a payoff of 101 if it is R. Given this strategy, his worst-case belief would be $(\epsilon, 1-\epsilon, 0)$*. But, given this belief, playing M with*

of R is at least $1-\epsilon$ for all ex-ante beliefs. Therefore, the information set of player 1 always has a positive probability. If player 1 plays N with probability zero, the information set of player 2 will not be reached if O has probability zero. In this case, it depends on the ex-ante belief if the probability of reaching the information set of player 2 is strictly positive. Assuming an ex-ante belief set $\mathcal{P}$ satisfying the full support assumption would lead to the same results as long as the minimum probability of L is small enough, i.e., smaller than $\frac{1}{102}$. Since the payoff of player 2 is independent of his actions and since the payoff of player 1 is constant w.r.t. to the action chosen by player 2, we skip the full support assumption due to notational convenience.

probability one is no longer optimal. On the other hand, playing N with probability one leads to a payoff of 101 or 100 if the state is L or R, respectively. Hence, the worst-case belief, if he plays N with probability one, is $(0, 1, 0)$*. But for this belief, playing M with probability one is optimal. One can show that for $\epsilon > \frac{1}{102}$ the optimal interim strategy of player* 1 *is a mixed strategy with probability $\frac{1}{102} < 1$ for M.*

Hence, player 1 *behaves dynamically inconsistently and plays a different strategy after observing that the state is not O. For detailed calculations see Aryal and Stauber (2014).*

The example above shows that new information can change beliefs such that the optimal strategy changes as well. This leads to dynamically inconsistent behavior. However, at the ex-ante stage, player 1 knows that his actions only influence his payoff if the state is not O. Should he not consider his knowledge about the information structure of the game in his ex-ante decision?

In decision-theoretic settings, an essential property of a set of distributions or beliefs to ensure dynamic consistency is *rectangularity*, or sometimes called stability under pasting introduced by Epstein and Schneider (2003) and Sarin and Wakker (1998). Rectangularity can be interpreted as a generalization of the law of iterated expectations. It captures the idea of decomposing any probability measure into its conditionals and marginals. Therefore, at the ex-ante stage, players take their interim worst-case beliefs and the information structure of the game into account. The information structure plays an essential role for rectangularity. The information that a player receives influences his interim beliefs and, therefore, dynamically inconsistent behavior. Since the game structure is known to each player, each player knows the possible information sets for each stage. Knowing the possible sets of updated beliefs, a player constructs his set of ex-ante beliefs in a rational way that is crucial for dynamic consistency. This is given by constructing a set of ex-ante beliefs such that the belief system is rectangular (or stable under pasting). To define rectangularity, we have to consider that each player's information consists of the opponents' observed actions. Therefore, we define beliefs on a more general state space. However, we will see that defining beliefs on the general state space does not change the equilibria of the game. Only ambiguity aversion and dynamic inconsistency lead to new equilibria.

In the next section, we formulate the definition of beliefs and rectangularity.

1.2.2 Beliefs

In standard game-theoretic settings without ambiguity, players have a (common) ex-ante belief over the set of types H^0. Let $\pi \in \Delta H^0$ be such an ex-ante belief. Then, player i faces the following maximization problem given a strategy profile of the opponents σ_{-i}

$$\max_{\sigma_i \in \Delta S_i} \sum_{h \in H} u(h) p_{\sigma_i}(h|h^0) p_{\sigma_{-i}}(h|h^0) \pi(h^0). \tag{1.1}$$

To evaluate a strategy σ_i, player i calculates his expected payoff by multiplying the ex-ante belief and the transition probability induced by his strategy and his conjecture about the strategy profile of his opponents.

Furthermore, in games, new information is influenced by strategic aspects. In decision-theoretic settings, further information usually occurs as an exogenously given signal. In games, the signals are observable actions of the opponents. Therefore, the strategies of the opponents influence the information that a player observes. To take this dependence into account, we define beliefs over the set of (terminal) histories H such that they are consistent with the set of imprecise probabilistic information $\mathcal{P}$ and the strategy profile of the opponents σ_{-i}. We will see that our definition leads to an equivalent maximization problem of player i as Equation (1.1).

Before we start with the definition of beliefs, we need the following definition of a sequence of partitions, which represents the information flow of the game.

Definition 1.3. *Given the set of histories $\mathcal{H}$ and the information sets $\mathcal{I}$, we denote with $(\mathcal{F}_i^t)_{t=0,\dots,T+1}$ the* ***sequence of information partitions*** *of player i, where*

$$\mathcal{F}_i^0 := H,$$
$$\mathcal{F}_i^t := \Big\{ \{h \in H \,:\, h^t \in I_i^t\}_{I_i^t \in \mathcal{I}_i^t} \Big\}.$$

Since there is a one-to-one relation between the elements F_i^t of $\mathcal{F}_i^t$ and the information sets $I_i^t \in \mathcal{I}_i^t$, we sometimes call F_i^t an information set.

Fix a player i and a strategy profile σ_{-i}. First, we define a system of beliefs of player i induced by the partition $(\mathcal{F}_i^t)_{t=0,\dots,T+1}$, the imprecise probabilistic information $\mathcal{P}$, and σ_{-i}. Then, we discuss which properties are needed to have a maximization problem, which is equivalent to Equation (1.1). Finally, we show the existence of a belief system satisfying all these properties.

Definition 1.4. *Given $(\mathcal{F}_i^t)_{t=0,\dots,T+1}$, the set of imprecise probabilistic information $\mathcal{P}$ and a strategy profile σ_{-i}, we call $\Psi_{\sigma_{-i}} = ((\Psi_{\sigma_{-i}}^t(F_i^t))_{F_i^t \in \mathcal{F}_i^t})_{t=0,\dots,T}$ the* ***belief system*** *of player i if*

$$\Psi_{\sigma_{-i}}^t(F_i^t) \subset \Delta(H),$$

with support on F_i^t for all $F_i^t \in \mathcal{F}_i^t$, $t =, 0, \dots T$ and $i \in N$.
Furthermore, we call $\Psi_{\sigma_{-i}}^0 := \Psi_{\sigma_{-i}}^0(F_i^0)$ the ex-ante belief set and $\Psi_{\sigma_{-i}}^t(F_i^t)$ the interim belief set at information set F_i^t.

To be consistent with the set of imprecise probabilistic information and the opponents' strategy profile, a belief system should satisfy the following properties:

1) The ex-ante belief set $\Psi^0_{\sigma_{-i}}$ is consistent with the set of imprecise probabilistic information $\mathcal{P}$ and the strategy profile of the opponent σ_{-i}. Formally, $\phi \in \Psi^0_{\sigma_{-i}}(F^0_i)$ if and only if there exist $\pi \in \mathcal{P}$ and a normalization constant $c_i \in \mathbb{R}$ such that[4]

$$\phi(h) = \frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi(h^0).$$

2) The interim belief sets $\Psi^t_{\sigma_{-i}}(F^t_i)$ are generated using prior-by-prior Bayesian updating whenever possible. Formally,

$$\Psi^t_{\sigma_{-i}}(F^t_i) = \left\{ \text{Bay}(\phi|F^t_i) \,:\, \phi \in \Psi^0_{\sigma_{-i}}(F^0_i) \right\},$$

with

$$\text{Bay}(\phi|F^t_i)(h) = \begin{cases} \frac{\phi(h)}{\phi(F^t_i)} & \text{if } h \in F^t_i, \\ 0 & \text{otherwise} \end{cases}$$

for all F^t_i with $\phi(F^t_i) > 0$, for some $\phi \in \Psi^0_{\sigma_{-i}}(F^0_i)$.[5]

3) For all information sets with positive probability, the interim belief sets are consistent with the set of imprecise probabilistic information $\mathcal{P}$ and the strategy profile of the opponents σ_{-i}. Let $\text{Bay}(\mathcal{P}|F^t_i)$ be the prior-by-prior Bayesian update of $\mathcal{P}$ at the information set F^t_i. Then, similar to 2), $\tilde{\phi} \in \Psi^t_{\sigma_{-i}}(F^t_i)$ if and only if there exist $\tilde{\pi} \in \text{Bay}(\mathcal{P}|F^t_i)$ and a normalization constant $c^t_i \in \mathbb{R}$ such that

$$\tilde{\phi}(h) = \frac{p_{\sigma_{-i}}(h|h^t)}{c^t_i}\tilde{\pi}(h^t)$$

for all F^t_i with $\phi(F^t_i) > 0$.

Property 1) ensures that the ex-ante beliefs are consistent with the information given by the game structure and the set of imprecise probabilistic information $\mathcal{P}$. The second property is an extension of the usual assumption that players update their beliefs using Bayes' rule whenever possible. Property 3) ensures the same relation between the interim belief sets and the Bayesian update of $\mathcal{P}$. Furthermore, if $\Psi^0_{\sigma_{-i}} = \{\phi\}$ is singleton and satisfies all properties, the ex-ante maximization problem of player i is

$$\max_{\sigma_i \in \Delta S_i} \sum_{h \in H} u(h) p_{\sigma_i}(h|h^0)\phi(h) = \max_{\sigma_i \in \Delta S_i} \frac{1}{c_i} \sum_{h \in H} u(h) p_{\sigma_i}(h|h^0) p_{\sigma_{-i}}(h|h^0)\pi(h^0),$$

[4]Please note that the normalization constant is needed to guarantee, that $\phi(\cdot)$ is a probability measure.

[5]The full support assumption on $\mathcal{P}$ implies that $\phi(F^t_i) > 0$ if and only if $\phi'(F^t_i) > 0$ for any $\phi, \phi' \in \Psi^0_{\sigma_{-i}}(F^0_i)$.

which is equivalent to the maximization problem of Equation (1.1). Therefore, using our definition of beliefs over (terminal) histories does not influence the set of Nash Equilibria in an unambiguous game. In Section 1.3, we will see that the equivalence of the maximization problems extends to games with ambiguity if there is no dynamically inconsistent behavior given the set of imprecise probabilistic information.

Before proceeding with the equilibrium analysis, we have to ensure that there exists a belief system over histories satisfying the above properties. The following assumption will ensure the existence.

Assumption 1.1. *We assume that the number of actions is constant across different information sets at the same stage, i.e.,* $|A_i(I_i^t)| = |A_i(\hat{I}_i^t)|$ *for all* $I_i^t, \hat{I}_i^t \in \mathcal{I}^t$ *and* $i \in N$, *where* $A_i(I_i^t)$ *denotes the actions set of player* i *at information set* I_i^t.

Assumption 1.1 may seem restrictive. However, any finitely repeated game with incomplete information does satisfy Assumption 1.1. Furthermore, for any game Γ as defined in Definition 1.1, we can find a game Γ' satisfying Assumption 1.1 such that the equilibria of Γ and Γ' are payoff-equivalent. One can easily construct Γ' by including copies of partial histories of Γ. To be more precise, let t be a stage with two information sets I_i^t and $\hat{I}_i^t$ and assume that $|A_i(I_i^t)| = c_1 \neq c_2 = |A_i(\hat{I}_i^t)|$. Then, we can copy the partial histories starting at I_i^t c_2-times and the partial histories starting at $\hat{I}_i^t$ c_1-times. The new action sets, including the copies, have both a cardinality equal to c_1 times c_2. Since a player is indifferent between any copy of a partial history and the partial history itself, including these copies does not change the equilibrium payoffs. For the rest of the chapter, we assume that Assumption 1.1 is satisfied.

Due to Assumption 1.1, the number of actions player i can choose from at stage t is the same for all information sets. We denote this number with $|A_i^t|$.

Lemma 1.1. *Let* $\bar{c}_i := \prod_{t=1}^{T} |A_i^t|$. *The following sets form a belief system that satisfies Properties 1) to 3).*

$$\Phi^0_{\sigma_{-i}}(F_i^0) := \left\{ \frac{p_{\sigma_{-i}}(h|h^0)}{\bar{c}_i} \pi(h^0) \; : \; \pi \in \mathcal{P} \right\},$$
$$\Phi^t_{\sigma_{-i}}(F_i^t) := \left\{ Bay(\phi|F_i^t) \; : \; \phi \in \Phi^0_{\sigma_{-i}}(F_i^0) \right\}$$

for all information sets with $\phi(F_i^t) > 0$. *Furthermore, any belief system satisfying Properties 1) to 3) has to be equal to* $\Phi_{\sigma_{-i}}$ *at all information sets with* $\phi(F_i^t) > 0$.

Simple calculations show that Assumption 1.1, Property 1), and the normalization of beliefs imply that the normalization constant equals $\bar{c}_i$. Then, Property 1) and 2) follow immediately from the definition of $\Phi_{\sigma_{-i}}$. Bayesian updating implies that $c_i^t = \prod_{s=t}^{T} |A_i^s|$ and Property 3) follows similar to Property 1). The formal proof can be found in Section 1.6.1 in the Appendix.

We come back to our running example to illustrate the definitions above:[6]

Running Example (cont.). *We denote with LM the history, where nature chooses type L and player* 1 *plays* M. *All histories are denoted similarly. Furthermore, denote the probability with which player* 1 *plays* N *with* n *and similarly all probabilities of an action with the corresponding lower case. The set of all histories* H *is then given by*

$$H = \{LM, LN, RM, RNS, RNT, OS, OT\}.$$

At the ex-ante stage, player 1 *and* 2 *have no information about the states. Therefore, their information partitions at the ex-ante stage consist only of one element, which is the set of all histories*

$$\mathcal{F}_i^0 = H.$$

At the interim stage, player 1 *can observe if the state is O or not. His information set consists of three elements. The first set contains all histories starting at L or R. The second and third set represent the case where player* 1 *learns that the state is O:*

$$\mathcal{F}_1^1 = \{F_{1,1}^1, F_{1,2}^1, F_{1,3}^1\} = \big\{\{LM, LN, RM, RNS, RNT\}, \{OS\}, \{OT\}\big\}.$$

Similarly, player 2*'s interim information partition consists of the set that contains all histories starting from O and histories where the state is R and player* 1 *plays* N *and the sets where he learns the exact history:*

$$\mathcal{F}_2^1 = \{F_{2,1}^1, F_{2,2}^1, F_{2,3}^1, F_{2,4}^1\} = \big\{\{RNS, RNT, OS, OT\}, \{LM\}, \{LN\}, \{RM\}\big\}.$$

A strategy $\sigma_{-1} = (s,t)$ *of player* 2 *induces a transition probability* $p_{\sigma_{-1}}(\cdot|\cdot)$ *for each history* $h \in H$ *which is independent of player* 1*'s strategy. Multiplying the imprecise probabilistic information with the transition probability,* $p_{\sigma_{-1}}(\cdot|\cdot)$, *induced by the strategy* $\sigma_{-1} = (s,t)$, *leads to the following set of ex-ante beliefs for player* 1*:*

$$\Phi_{(s,t)}^0 = \left\{\left(\frac{l}{2}, \frac{l}{2}, \frac{r}{2}, \frac{rs}{2}, \frac{rt}{2}, os, ot\right) \,:\, (l,r,o) \in \mathcal{P}\right\}.$$

Similarly, the ex-ante belief set of player 2 *given strategy* $\sigma_{-2} = (m,n)$ *of player* 1 *is*

$$\Phi_{(m,n)}^0 = \left\{\left(lm, ln, rm, \frac{rn}{2}, \frac{rn}{2}, \frac{o}{2}, \frac{o}{2}\right) \,:\, (l,r,o) \in \mathcal{P}\right\}.$$

[6] Please note that our running example does not satisfy the multistage structure and Assumption 1.1. The game can be easily translated into a game that satisfies both assumptions by including trivial moves. However, due to the simple structure and the fact that each player has only one non-trivial move, Assumption 1.1 is not needed for the following constructions and calculations. Therefore, due to simplicity, we use this simpler version of the game.

1.2.2.1 Rectangularity

Given the above notation and definitions, we formulate the formal definition of rectangularity. As already mentioned, rectangularity is a generalization of the law of iterative expectation.

Let us first look at the case without ambiguity, i.e., there exists only one ex-ante belief ϕ, and assume that Bayes' rule is always well defined. For each information set F_i^1 in the next stage, Bayesian updating leads to an updated belief $\tilde{\phi}_{F_i^1}$. Roughly speaking, Bayes' rule is defined such that the denominator of Bayes' rule equals the marginal belief of reaching the information set on which we update the ex-ante belief. Hence, multiplying (or pasting) the updated belief given an information set F_i^1 with the marginal belief of reaching this information set leads to the ex-ante belief restricted to F_i^1. This holds for any information set and, therefore, summation over all information sets leads to the ex-ante belief on H, i.e.,

$$\phi(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \phi(F_i^1)\tilde{\phi}_{F_i^1}(h).$$

Now, we generalize this property to an ambiguous setting. With ambiguity, players have a set of ex-ante beliefs. Rectangularity states that we can take any updated and marginal belief (even if they are not derived from the same ex-ante belief), and the pasting is still an element of the ex-ante belief set.

Definition 1.5. *For beliefs $\phi \in \Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ and $\tilde{\phi} = (\tilde{\phi}_{F_i^t})_{F_i^t \in \mathcal{F}_i^t}$ with $\tilde{\phi}_{F_i^t} \in \Psi_{\sigma_{-i}}^{t}(F_i^t)$, the pasting of marginal and updated belief, $\phi \circ \tilde{\phi}$, is defined as*

$$\phi \circ \tilde{\phi}(\cdot) := \sum_{F_i^t \in \mathcal{F}_i^t} \phi(F_i^t)\tilde{\phi}_{F_i^t}(\cdot).$$

The pasting of $\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ and $(\Psi_{\sigma_{-i}}^{t}(F_i^t))_{F_i^t \in \mathcal{F}_i^t}$ is defined as the set consisting of pasting each element of $\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ with each element of $(\Psi_{\sigma_{-i}}^{t}(F_i^t))_{F_i^t \in \mathcal{F}_i^t}$, i.e.,

$$\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1}) \circ \left(\Psi_{\sigma_{-i}}^{t}(F_i^t)\right)_{F_i^t \in \mathcal{F}_i^t} = \left\{\phi \circ \tilde{\phi} : \phi \in \Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1}) \text{ and } \tilde{\phi}_{F_i^t} \in \Psi_{\sigma_{-i}}^{t}(F_i^t)\right\}.$$

A set of beliefs $\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ is called ***rectangular*** *(stable under pasting) if*

$$\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1}) \circ \left(\Psi_{\sigma_{-i}}^{t}(F_i^t)\right)_{F_i^t \in \mathcal{F}_i^t} = \Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1}).$$

A belief system Ψ is called stable under pasting (rectangular) if $\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ is stable under pasting for all $F_i^t \in \mathcal{F}_i^t$, $i \in N$ and $t = 1, \ldots T+1$.

By the intuition given above, without ambiguity, rectangularity should follow from Bayesian updating. The following remark shows that this is indeed true.

Remark 1.1. *Let Ψ_i^0 be singleton and $\phi(F_i^t) > 0$ for all F_i^t. Then, Bayes' rule is always well defined and rectangularity is equivalent to Bayes' rule. To see this, take into account that since $\Psi_{\sigma_{-i}}^0$ is singleton, $\Psi_{\sigma_{-i}}^t(F_i^t)$ are singleton as well. Denote by $\bar{F}_i^t$ the element of the partition $\mathcal{F}_i^t$ which contains h. First we show that Bayes' rule implies rectangularity:*

$$\phi \circ \tilde{\phi}(h) = \sum_{F_i^t \in \mathcal{F}_i^t} \phi(F_i^t)\tilde{\phi}_{F_i^t}(h) = \phi(\bar{F}_i^t)\frac{\phi(h)}{\phi(\bar{F}_i^t)} = \phi(h).$$

The other direction follows by similar calculations since $\phi(\bar{F}_i^t) > 0$:

$$\phi(h) = \phi \circ \tilde{\phi}(h) = \phi(\bar{F}_i^t)\tilde{\phi}_{\bar{F}_i^t}(h) \quad \Leftrightarrow \quad \tilde{\phi}_{\bar{F}_i^t}(h) = \frac{\phi(h)}{\phi(\bar{F}_i^t)}.$$

Furthermore, rectangularity preserves some nice properties. Remark 1.2 shows that the Bayesian update of $\phi \circ \tilde{\phi}$ equals $\tilde{\phi}$ and the marginal probability of a pasting $\phi \circ \tilde{\phi}$ equals the marginal probability of ϕ. We will see that these properties are beneficial when we explain the construction of a rectangular belief system.

Remark 1.2. *Let $\Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ be rectangular. Then, for any $\phi \in \Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ there exist some $\phi' \in \Psi_{\sigma_{-i}}^{t-1}(F_i^{t-1})$ and $\tilde{\phi} = (\tilde{\phi}_{F_i^t})_{F_i^t \in \mathcal{F}_i^t} \in (\Psi_{\sigma_{-i}}^t(F_i^t))_{F_i^t \in \mathcal{F}_i^t}$ such that*

$$\phi(\bar{h}) = \phi' \circ \tilde{\phi}(\bar{h}) = \sum_{F_i^t \in \mathcal{F}_i^t} \phi'(F_i^t)\tilde{\phi}_{F_i^t}(\bar{h}).$$

Let $\bar{F}_i^t$ denote the element of the partition which contains $\bar{h}$. Then, $\phi(\bar{h}) = \phi'(\bar{F}_i^t)\tilde{\phi}_{\bar{F}_i^t}(\bar{h})$.

- *The Bayesian update of ϕ given $\bar{F}_i^t$ equals*

$$\begin{aligned} Bay(\phi|\bar{F}_i^t)(\bar{h}) =& \frac{\phi(\bar{h})}{\sum_{h \in \bar{F}_i^t} \phi(h)} = \frac{\sum_{F_i^t} \phi'(F_i^t)\tilde{\phi}_{F_i^t}(\bar{h})}{\sum_{h \in \bar{F}_i^t} \sum_{F_i^t} \phi'(F_i^t)\tilde{\phi}_{F_i^t}(h)} \\ =& \frac{\phi'(\bar{F}_i^t)\tilde{\phi}_{\bar{F}_i^t}(\bar{h})}{\phi'(\bar{F}_i^t)\sum_{h \in \bar{F}_i^t} \tilde{\phi}_{\bar{F}_i^t}(h)} = \tilde{\phi}_{\bar{F}_i^t}(\bar{h}), \end{aligned}$$

  where the last equality follows since $\sum_{h \in \bar{F}_i^t} \tilde{\phi}_{\bar{F}_i^t}(h) = 1$ for all $t \geq 0$.

- *The marginal distribution is given by*

$$\sum_{h \in \bar{F}_i^t} \phi(h) = \sum_{h \in \bar{F}_i^t} \sum_{F_i^t} \phi'(F_i^t)\tilde{\phi}_{F_i^t}(h) = \phi'(\bar{F}_i^t) \underbrace{\sum_{h \in \bar{F}_i^t} \tilde{\phi}_{\bar{F}_i^t}(h)}_{=1} = \phi'(\bar{F}_i^t).$$

Hence, marginal and updated distributions of a distribution of a belief set that is rectangular coincide with the marginal and updated distribution from which it is constructed.

The literature on decision theory using rectangularity and cited above shows that a rectangular belief system can always be constructed in the following way. First, given the ex-ante belief set and the information structure, one can calculate the prior-by-prior Bayesian updates for all information sets. Then, one proceeds by backward induction and constructs a rectangular belief set by pasting marginal and updated beliefs. The belief sets constructed like this are the smallest rectangular sets that contain the original belief sets. Therefore, they are called the rectangular hulls of the original belief sets. The same method can be used here to get a rectangular belief system for completely mixed σ_{-i}. When σ_{-i} is completely mixed, Bayes' rule is always well defined, and we can derive the prior-by-prior Bayesian update of $\Phi^0_{\sigma_{-i}}$ for each stage and each information set. Denote with $\Phi^{t-1}_{\sigma_{-i}}(F^{t-1}_i)$ the Bayesian update of $\Phi^0_{\sigma_{-i}}$ at $t-1$ given the information set F^{t-1}_i and, similarly, with $\Phi^t_{\sigma_{-i}}(F^t_i)$ the Bayesian update at t given F^t_i.[7] The rectangular hull $\text{rect}(\Phi^{t-1}_{\sigma_{-i}}(F^{t-1}_i))$ is given by the pasting of $\Phi^{t-1}_{\sigma_{-i}}(F^{t-1}_i)$ and $\Phi^t_{\sigma_{-i}}$, i.e.,

$$\text{rect}(\Phi^{t-1}_{\sigma_{-i}}) = \left\{\phi \circ \tilde{\phi} \,:\, \phi \in \Phi^{t-1}_{\sigma_{-i}}(F^{t-1}_i) \text{ and } \tilde{\phi}_{F^t_i} \in \Phi^t_{\sigma_{-i}}(F^t_i),\ \forall F^t_i \in \mathcal{F}^t_i\right\}.$$

Remark 1.2 shows that the set of Bayesian updates of $\text{rect}(\Phi^{t-1}_{\sigma_{-i}})$ coincides with the set of Bayesian updates of $\Phi^{t-1}_{\sigma_{-i}}$. Therefore, $\text{rect}(\Phi^{t-1}_{\sigma_{-i}})$ is rectangular by construction. This method holds for any arbitrary t. Hence, starting with the last two periods, $T-1$ and T, and proceeding by backward induction we can always close a prior set $\Phi^0_{\sigma_{-i}}$ under pasting and $\text{rect}(\Phi^0_{\sigma_{-i}})$ is the smallest set containing $\Phi^0_{\sigma_{-i}}$ that is rectangular. Furthermore, the construction induces that the Bayesian updates of $\text{rect}(\Phi^0_{\sigma_{-i}})$ are rectangular for any information set.

If σ_{-i} is not completely mixed, there can exist information sets such that the marginal probability of reaching these information sets is zero. Let $\bar{F}^t_i$ be an information set such that there exists $\phi \in \Phi^0_{\sigma_{-i}}$ with $\phi(\bar{F}^t_i) = 0$. The full support assumption of $\mathcal{P}$ implies that $\phi'(\bar{F}^t_i) = 0$ for all $\phi' \in \Phi^0_{\sigma_{-i}}$. Let $\bar{F}^{t-1}_i$ be information set that precedes $\bar{F}^t_i$, i.e., $\bar{F}^t_i \subseteq \bar{F}^{t-1}_i$ and, without loss of generality, let $\phi(\bar{F}^{t-1}_i) > 0$.[8] Furthermore, perfect recall implies that all information sets that are reachable from $\bar{F}^t_i$ have probability zero as well. For information sets with probability zero, Bayes' rule is not well defined. The construction of the rectangular hull as described above can be generalized as follows. For all information sets with positive probability, the set of updated beliefs is derived prior-by-prior Bayesian updating. For information sets with probability zero, players can choose an arbitrary compact set of updated beliefs. Then, the rectangular hull is

[7] We assume that the Bayesian update is a probability distribution over the whole set of full histories H such that histories that are not an element of the observed information set have probability zero.

[8] If $\phi(\bar{F}^{t-1}_i) = 0$ we can replace $\bar{F}^t_i$ by $\bar{F}^{t-1}_i$ and check if the probability of the information set preceding $\bar{F}^{t-1}_i$ has positive probability. Repeating this leads to an information set with probability zero such that the preceding information set has a strictly positive probability.

constructed by backward induction, as described above. The construction may change the set of beliefs at information sets with zero probability. But the construction of the rectangular hull of the belief set at $\bar{F}_i^{t-1}$ is not influenced by the belief set at $\bar{F}_i^t$, since the marginal probability of $\bar{F}_i^t$ is zero. Therefore, the arbitrary choice of updated belief sets at information sets with probability zero does not influence the construction of the rectangular hull, and we use the notation $\text{rect}(\Phi^t_{\sigma_{-i}})$ for any strategy $\sigma_{-i} \in \Sigma_{-i}$.

To illustrate the construction of a rectangular prior set, we come back to our running example.

Running Example (cont.). *We have already shown that*

$$\begin{aligned}
H &= \{LM, LN, RM, RNS, RNT, OS, OT\},\\
\mathcal{F}_i^0 &= H,\\
\mathcal{F}_1^1 &= \{F_{1,1}^1, F_{1,2}^1, F_{1,3}^1\} = \big\{\{LM, LN, RM, RNS, RNT\}, \{OS\}, \{OT\}\big\},\\
\mathcal{F}_2^1 &= \{F_{2,1}^1, F_{2,2}^1, F_{2,3}^1, F_{2,4}^1\} = \big\{\{RNS, RNT, OS, OT\}, \{LM\}, \{LN\}, \{RM\}\big\},\\
\Phi^0_{(s,t)} &= \left\{ \left(\frac{l}{2}, \frac{l}{2}, \frac{r}{2}, \frac{rs}{2}, \frac{rt}{2}, os, ot \right) \,:\, (l, r, o) \in \mathcal{P} \right\},\\
\Phi^0_{(m,n)} &= \left\{ \left(lm, ln, rm, \frac{rn}{2}, \frac{rn}{2}, \frac{o}{2}, \frac{o}{2} \right) \,:\, (l, r, o) \in \mathcal{P} \right\}.
\end{aligned}$$

To construct the rectangular hull of $\Phi^0_{(s,t)}$, we need the marginal and updated beliefs of player 1. The marginal beliefs for an arbitrary $\phi \in \Phi^0_{(s,t)}$ of the information sets of player 1 are

$$\begin{aligned}
\phi(F_{1,1}^1) &= l + \frac{r}{2} + \frac{rs}{2} + \frac{rt}{2} = l + r,\\
\phi(F_{1,2}^1) &= os,\\
\phi(F_{1,3}^1) &= ot.
\end{aligned}$$

For an arbitrary ex-ante belief $\phi \in \Phi^0_{(s,t)}$, the Bayesian update given $F_{1,1}^1$ is

$$Bay(\phi | F_{1,1}^1) = \Big(\frac{l}{2(l+r)}, \frac{l}{2(l+r)}, \frac{r}{2(l+r)}, \frac{rs}{2(l+r)}, \frac{rt}{2(l+r)}, 0, 0 \Big).$$

Hence, the prior-by-prior Bayesian updates of $\Phi^0_{(s,t)}$ given the information sets $F_{1,1}^1$, $F_{1,2}^1$, and $F_{1,3}^1$ are

$$\begin{aligned}
Bay(\Phi^0_{(s,t)} | F_{1,1}^1) =& \Big\{ \Big(\frac{l}{2(l+r)}, \frac{l}{2(l+r)}, \frac{r}{2(l+r)}, \frac{rs}{2(l+r)}, \frac{rt}{2(l+r)}, 0, 0 \Big) : (l, r, o) \in \mathcal{P} \Big\}\\
=& \left\{ \left(\tilde{l}, \tilde{l}, \tilde{r}, \tilde{r}s, \tilde{r}t, 0, 0 \right) \,:\, \tilde{l} \in \left[0, \frac{\epsilon}{2}\right], \tilde{r} \in \left[\frac{1-\epsilon}{2}, \frac{1}{2}\right], \tilde{l} + \tilde{r} = \frac{1}{2} \right\},\\
Bay(\Phi^0_{(s,t)} | F_{1,2}^1) =& \Big\{ \big(0, 0, 0, 0, 0, 1, 0\big) \Big\},\\
Bay(\Phi^0_{(s,t)} | F_{1,3}^1) =& \Big\{ \big(0, 0, 0, 0, 0, 0, 1\big) \Big\}.
\end{aligned}$$

The rectangular hull $\mathrm{rect}(\Phi^0_{(s,t)})$ *consists of all possible combination of marginal and updated beliefs. For the histories of the information sets* $F^1_{1,2}$ *and* $F^1_{1,3}$ *,the updated belief is either zero or one. Therefore, we concentrate on the information set* $F^1_{1,1}$*. Since* $\tilde{l}$*,* $\tilde{r}$*,* l*, and* r *are elements of closed intervals, we can focus on the all possible combination of the lowest and highest values for* $\tilde{l}$*,* $\tilde{r}$*,* l*, and* r*. Then, the convex hull of the pasting of these distributions forms the rectangular hull.*

Let ϕ *be such that* $r = 1$ *and* $l = 0$*. Given this ex-ante belief, the marginal probability of reaching* $F^1_{1,1}$ *is* $\phi(F^1_{1,1}) = 1$*. Let* ϕ' *denote the pasting of the marginal* $\phi(F^1_{1,1})$ *and the update* $\tilde{\phi}$*. Considering lowest and highest values for* $\tilde{r}$ *and* $\tilde{l}$*, there are two updated beliefs* $\tilde{\phi}$ *that can be pasted with this marginal belief:*

- $\tilde{\phi}$ *such that* $\tilde{l} = 0 = \frac{1}{2} - \tilde{r}$*:*
  The pasting is then given by

$$\begin{aligned}\phi' = \phi \circ \tilde{\phi} &= \left(\phi(F^1_{1,1})\tilde{l}, \phi(F^1_{1,1})\tilde{l}, \phi(F^1_{1,1})\tilde{r}, \phi(F^1_{1,1})\tilde{r}s, \phi(F^1_{1,1})\tilde{r}t, 0, 0\right)\\ &= \left(1\cdot 0, 1\cdot 0, 1\cdot\frac{1}{2}, 1\cdot\frac{s}{2}, 1\cdot\frac{t}{2}, 0, 0\right) = \left(0, 0, \frac{1}{2}, \frac{s}{2}, \frac{t}{2}, 0, 0\right).\end{aligned}$$

- $\tilde{\phi}$ *such that* $\tilde{l} = \frac{\epsilon}{2} = \frac{1}{2} - \tilde{r}$*:*
  The pasting is then given by

$$\begin{aligned}\phi' = \phi \circ \tilde{\phi} &= \left(\phi(F^1_{1,1})\tilde{l}, \phi(F^1_{1,1})\tilde{l}, \phi(F^1_{1,1})\tilde{r}, \phi(F^1_{1,1})\tilde{r}s, \phi(F^1_{1,1})\tilde{r}t, 0, 0\right)\\ &= \left(1\cdot\frac{\epsilon}{2}, 1\cdot\frac{\epsilon}{2}, 1\cdot\left(\frac{1}{2} - \frac{\epsilon}{2}\right), 1\cdot\left(\frac{1}{2} - \frac{\epsilon}{2}\right)s, 1\cdot\left(\frac{1}{2} - \frac{\epsilon}{2}\right)t, 0, 0\right)\\ &= \left(\frac{\epsilon}{2}, \frac{\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, 0, 0\right).\end{aligned}$$

Combining any possible combination of (l, r, o) *and* $(\tilde{l}, \tilde{r})$ *in such a way leads to the pastings given in Table 1.1. The probability of OS and OT follows from the pasting*

Marginal	Update	Pasting
$r = 1,\ l = 0$	$\tilde{l} = 0 = \frac{1}{2} - \tilde{r}$	$\left(0, 0, \frac{1}{2}, \frac{s}{2}, \frac{t}{2}, 0, 0\right)$
$r = 1,\ l = 0$	$\tilde{l} = \frac{\epsilon}{2} = \frac{1}{2} - \tilde{r}$	$\left(\frac{\epsilon}{2}, \frac{\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, 0, 0\right)$
$r = 1-\epsilon,\ l = \epsilon$	$\tilde{l} = 0 = \frac{1}{2} - \tilde{r}$	$\left(0, 0, \frac{1}{2}, \frac{s}{2}, \frac{t}{2}, 0, 0\right)$
$r = 1-\epsilon,\ l = \epsilon$	$\tilde{l} = \frac{\epsilon}{2} = \frac{1}{2} - \tilde{r}$	$\left(\frac{\epsilon}{2}, \frac{\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, 0, 0\right)$
$r = 1-\epsilon,\ l = 0$	$\tilde{l} = 0 = \frac{1}{2} - \tilde{r}$	$\left(0, 0, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, \epsilon s, \epsilon t\right)$
$r = 1-\epsilon,\ l = 0$	$\tilde{l} = \frac{\epsilon}{2} = \frac{1}{2} - \tilde{r}$	$\left(\frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)^2}{2}, \frac{(1-\epsilon)^2 s}{2}, \frac{(1-\epsilon)^2 t}{2}, \epsilon s, \epsilon t\right)$

Table 1.1: Pasting for Rectangular Hull of Player 1

with the updated belief given the information sets $F^1_{1,2}$ *and* $F^1_{1,3}$. *For the first four rows the marginal probability of reaching* $F^1_{1,2}$ *or* $F^1_{1,3}$ *is zero since* $o = 1 - r - l = 0$. *For the last two rows,* $l = 0$ *and* $r = 1 - \epsilon$, *imply* $o = \epsilon$. *Since the prior-by-prior Bayesian update given* $F^1_{1,2}$ *or* $F^1_{1,3}$ *consists of just one belief, which gives probability one to* OS *or* OT, *respectively, the pasting of marginal and update for* OS *and* OT *equals the values given above.*

The rectangular hull $\mathrm{rect}(\Phi^0_{(s,t)})$ *of player 1 is then given by the convex hull of the pastings given in Table 1.1:*

$$\mathrm{rect}(\Phi^0_{(s,t)}) = \mathrm{conv}\Bigg\{ \left(0, 0, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, \epsilon s, \epsilon t\right), \left(\frac{\epsilon}{2}, \frac{\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, 0, 0\right), \left(0, 0, \frac{1}{2}, \frac{s}{2}, \frac{t}{2}, 0, 0\right), \left(\frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)^2}{2}, \frac{(1-\epsilon)^2 s}{2}, \frac{(1-\epsilon)^2 t}{2}, \epsilon s, \epsilon t\right) \Bigg\}.$$

To see the difference between the rectangular hull and $\Phi^0_{(s,t)}$, *remember that* $\Phi^0_{(s,t)}$ *is given by*

$$\begin{aligned}\Phi^0_{(s,t)} &= \left\{ \left(\frac{l}{2}, \frac{l}{2}, \frac{r}{2}, \frac{rs}{2}, \frac{rt}{2}, os, ot\right) : (l, r, o) \in \mathcal{P} \right\} \\ &= \mathrm{conv}\Bigg\{ \left(0, 0, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)r}{2}, \frac{\epsilon s}{2}, \frac{\epsilon r}{2}\right), \left(\frac{\epsilon}{2}, \frac{\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)r}{2}, 0, 0\right), \left(0, 0, \frac{1}{2}, \frac{s}{2}, \frac{t}{2}, 0, 0\right) \Bigg\}.\end{aligned}$$

Since the belief $\left(\frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)\epsilon}{,} \frac{(1-\epsilon)^2}{2}, \frac{(1-\epsilon)^2 s}{2}, \frac{(1-\epsilon)^2 t}{2}, \epsilon s, \epsilon t\right)$ *is not an element of* $\Phi^0_{(s,t)}$, *it follows that* $\Phi^0_{(s,t)} \subsetneq \mathrm{rect}(\Phi^0_{(s,t)})$. *The last row in Table 1.1 shows that this belief is constructed by pasting the marginal probability of the ex-ante worst-case belief with the interim worst-case belief. We will see later, that this belief changes the ex-ante optimal behavior such that player 1 plays dynamically consistently.*

Similar calculations as above show that the Bayesian update of $\Phi^0_{(m,n)}$ *given* $F^1_{2,1}$ *and the rectangular hull* $\mathrm{rect}(\Phi^0_{(m,n)})$ *of player 2 are given by*

$$Bay(\Phi^0_{(m,n)}|F^1_{2,1})) = \Bigg\{ (0, 0, 0, \tilde{rn}, \tilde{rn}, \tilde{o}, \tilde{o}) : \tilde{rn} \in \left[\frac{(1-\epsilon)n}{2(1-\epsilon)n + 2\epsilon}, \frac{1}{2}\right], \tilde{o} \in \left[0, \frac{\epsilon}{2n(1-\epsilon) + 2\epsilon}\right], \tilde{o} + \tilde{rn} = \frac{1}{2} \Bigg\},$$

and

$$\text{rect}(\Phi^0_{(m,n)}) = \text{conv}\Bigg\{\left(0,0,(1-\epsilon)m,\frac{(1-\epsilon)n+\epsilon}{2},\frac{(1-\epsilon)n+\epsilon}{2},0,0\right),$$
$$\left(0,0,m,\frac{(1-\epsilon)n^2}{a},\frac{(1-\epsilon)n^2}{a},\frac{\epsilon n}{a},\frac{\epsilon n}{a}\right),$$
$$\left(0,0,(1-\epsilon)m,\frac{(1-\epsilon)n}{2},\frac{(1-\epsilon)n}{2},\frac{\epsilon}{2},\frac{\epsilon}{2}\right),\left(0,0,m,\frac{n}{2},\frac{n}{2},0,0\right),$$
$$\left(\epsilon m,\epsilon n,(1-\epsilon)m,\frac{(1-\epsilon)n}{2},\frac{(1-\epsilon)n}{2},0,0\right),$$
$$\left(\epsilon m,\epsilon n,(1-\epsilon)m,\frac{(1-\epsilon)^2n^2}{a},\frac{(1-\epsilon)^2n^2}{a},\frac{(1-\epsilon)\epsilon n}{a},\frac{(1-\epsilon)\epsilon n}{a}\right)\Bigg\},$$

with $a = 2((1-\epsilon)n+\epsilon)$.

The rectangular belief set of player 1 shows the main differences of $\Phi_{\sigma_{-i}}$ and $\text{rect}(\Phi_{\sigma_{-i}})$. The rectangular hull contains the belief, which is the pasting of the prior and interim worst-case belief. We will see later that due to rectangularity, the ex-ante worst-case belief given the rectangular hull will be the pasting of the marginal belief derived from the ex-ante worst-case belief and the interim worst-case belief. Therefore, updating leads to the interim worst-case belief, and dynamically inconsistent behavior cannot occur.

Given the complex structure of beliefs described in the last section, one might wonder why we are not proceeding by constructing rectangular ex-ante belief sets over the set of types H^0 and updating these sets prior-by-prior using Bayes' rule. For our analysis, it is essential that first, the information structure of the game is given by a sequence of partitions of a fixed set, and second, the rectangular belief set of player i is independent of his strategy σ_i but depends on the strategy of the opponents. The information partition at a stage t is a partition of the set of partial histories H^t up to this stage. Then, the information partition $\mathcal{I}_i^{t+1}$ at stage $t+1$ is a partition of H^{t+1} but not of H^t. Therefore, the first part is not satisfied. Furthermore, there can exist paths that start from the same type or state but lead to different information sets depending on the action of the opponent. When constructing rectangular beliefs for all players, we have to consider this dependence on information sets and actions of opponents.[9] The following example illustrates that the opponents' strategies and their influence on the information that a player receives play an essential role for rectangular beliefs.

[9] Aryal and Stauber (2014) construct a rectangular belief sets of beliefs over $\{L, R, O\}$ of player 1 in our running example. Then, they transfer this belief set to a state space that considers the difference between the partial histories RN and RM and show that this transferred belief set is not rectangular for player 2. We are proceeding the other way around. We first transfer the set of imprecise probabilistic information to heterogeneous beliefs sets on H and then construct the rectangular hull for each player.

Even in a simple two-player signaling game, the rectangular hull may depend on the opponent's strategy.

Example 1.1. *We consider a signaling game with three states L, R, or O. Player* 1 *learns the state and can play either A or B in each state. Player* 2 *only observes the action chosen by player* 1*. We compare the rectangular hull of player* 2 *for two different strategies of player* 1*:* $\sigma_1 = (AAA)$ *denotes the strategy of always playing A and* $\sigma_1' = (AAB)$ *denotes the strategy of playing A at state L and R and B at state O. The game is depicted in Figure 1.2. We highlight the histories which are played with positive probability given* σ_1 *or* σ_1'*. Since we focus on player* 2*, we only specify the payoffs of player* 2*. The imprecise probabilistic information is the same as in our running example*

$$\mathcal{P} = \{(1-\epsilon)(0,1,0) + \epsilon(l,r,o) : (l,r,o) \in \Delta\},$$

where l, r, and o denote the probability of L, R, and O, respectively, and $\epsilon > 0$*. Given the strategy* σ_1*, i.e., player* 1 *always plays A, player* 2 *does not learn anything*

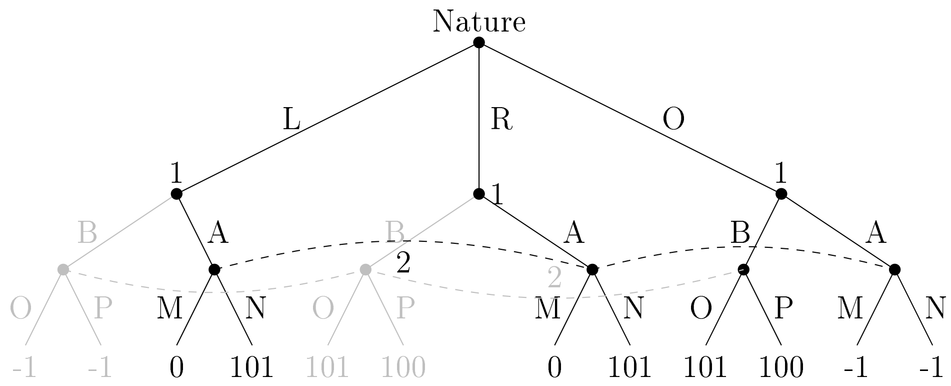

Figure 1.2: Example 1.1.

about nature's choice. His ex-ante and interim decision problems are similar to the ex-ante decision problem of player 1 *in our running example. One can easily show that* $\Phi^0_{\sigma_1}$ *is rectangular. Given the strategy* σ_1'*, player* 2 *learns if the state is O or* $\{L, R\}$*. His ex-ante and interim decision problems are similar to the ex-ante and interim decision problem of player* 1 *in our running example. Similar to our running example, one can show that* $\Phi^0_{\sigma_1'}$ *is not rectangular. The detailed calculation may be found in Section 1.6.3.1 in the Appendix.*

1.2.2.2 Common prior assumption

There is a well-known conflict between dynamic consistency and consequentialism in the literature on ambiguous beliefs in dynamic settings. Intuitively, consequentialism states that at any information set, the preferences do not depend on past discarded actions or events that are not consistent with the given information set.

Aryal and Stauber (2014) and Ellis (2018) show that dynamic consistency, consequentialism, and a common prior assumption are only fulfilled simultaneously if the players behave as expected utility maximizers. Rectangular belief sets satisfy consequentialism and dynamic consistency for a fixed information structure. We allow for heterogeneous belief sets by assuming that each player may interpret the common set of imprecise probabilistic information $\mathcal{P}$ differently by taking his own information structure into account. Therefore, we overcome the impossibility result of Aryal and Stauber (2014) and Ellis (2018) by first, requiring for each player dynamic consistency only for his information structure and second, allowing for heterogeneous belief sets.

However, even if we allow for heterogeneous belief sets across players, the heterogeneity is restricted by rectangularity and the common prior assumption on $\mathcal{P}$. First, note that a belief $\phi \in \text{rect}\Phi^0_{\sigma_{-i}}$ depends on the the strategy of the opponents. Therefore, we say that rectangular ex-ante beliefs satisfy a common prior assumption if the set of marginal beliefs over states or types H^0

$$\left\{ \phi(\tilde{h}^0) = \sum_{\substack{\bar{h} \in H \\ : \bar{h}^0 = \tilde{h}^0}} \phi(\bar{h}) : \phi \in \text{rect}\Phi^0_{\sigma_{-i}} \right\}$$

is the same for all players. In the next remark we discuss two special cases in which the rectangular ex-ante beliefs satisfy a common prior assumption. In the first case, players receive homogeneous information. The second case allows heterogeneous information but requires $\Phi^0_{\sigma_{-i}} = \text{rect}(\Phi^0_{\sigma_{-i}})$. Thus, heterogeneous ex-ante belief sets only occur if players receive heterogeneous information and behave dynamically inconsistently given non-rectangular belief sets $\Phi^0_{\sigma_{-i}}$.

Remark 1.3. Case 1) *Homogeneous information across all players implies that all actions are observable and all players have the same information structure about the set of states or types H^0, i.e., $\mathcal{I}^1_i = \mathcal{I}^1_j$, for all players i and j.*[10] *Since the information sets are the same for all players, we omit the subscript i. Let $H^0(F^1_i) = \{h^0 \in H^0 : \exists\, \hat{h} \in F^1_i \text{ with } h^0 = \hat{h}^0\}$ denote the set of all states or types leading to F^1_i. First observe*

$$\sum_{h \in F^1_i} \frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{c_i} = \sum_{\hat{h}^0 \in H^0(F^1_i)} \pi(\hat{h}^0) \sum_{\substack{h \in F^1_i \\ : h^0 = \hat{h}^0}} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i}$$
$$= \sum_{\hat{h}^0 \in H^0(F^1_i)} \pi(h^0), \tag{1.2}$$

[10] Due to perfect recall, a player always remembers his own action. Therefore, to ensure homogeneous information, it is necessary that all actions are observable.

where the second step follows since $\sum_{\substack{h\in F_i^1\\ :h^0=\tilde{h}^0}} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} = 1$.[11] *Any arbitrary belief* $\phi_i \in \text{rect}(\Phi^0_{\sigma_{-i}})$ *of a fixed player* i *can be represented by the pasting of a marginal and updated belief. Further, fix an arbitrary state or type* $\tilde{h}^0$. *There exists exactly one information set at stage one,* $\tilde{F}_i^1$, *such that all histories starting at* $\tilde{h}^0$ *are an element of* $\tilde{F}_i^1$. *The homogeneous information structure implies that* $\tilde{F}_j^1 = \tilde{F}_i^1$ *for all* $i, j \in N$ *and we omit the subscript* i. *Then, Equation* (1.2) *implies* [12]

$$\begin{aligned}
\sum_{\substack{\bar{h}\in H\\ :\bar{h}^0=\tilde{h}^0}} \phi_i(\bar{h}) &= \sum_{\substack{\bar{h}\in \tilde{F}^1\\ :\bar{h}^0=\tilde{h}^0}} \phi_i(\bar{h})\\
&= \sum_{\substack{\bar{h}\in \tilde{F}^1\\ :\bar{h}^0=\tilde{h}^0}} \left(\sum_{h\in\tilde{F}^1} \frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{c_i}\right) \frac{\frac{p_{\sigma_{-i}}(\bar{h}|\bar{h}^0)\pi'(\bar{h}^0)}{c_i}}{\sum_{h\in\tilde{F}^1}\frac{p_{\sigma_{-i}}(h|h^0)\pi'(h^0)}{c_i}}\\
&= \frac{\sum_{\hat{h}^0\in H^0(\tilde{F}^1)}\pi(h^0)}{\sum_{\hat{h}^0\in H^0(\tilde{F}^1)}\pi'(h^0)}\pi'(\tilde{h}^0) \sum_{\substack{\bar{h}\in \tilde{F}^1\\ :\bar{h}^0=\tilde{h}^0}} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\\
&= \frac{\sum_{\hat{h}^0\in H^0(\tilde{F}^1)}\pi(h^0)}{\sum_{\hat{h}^0\in H^0(\tilde{F}^1)}\pi'(h^0)}\pi'(\tilde{h}^0),
\end{aligned}$$

where the last step uses again $\sum_{\substack{\bar{h}\in\tilde{F}^1\\ :\bar{h}^0=\tilde{h}^0}} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} = 1$. *Now, the last part of the equation does not depend on player* i. *Hence, the marginal beliefs of a state or type* $\tilde{h}^0$ *are the same for all* $i, j \in N$, *that is*

$$\sum_{\substack{\bar{h}\in H\\ s.t.\ \bar{h}^0=\tilde{h}^0}} \phi_i(\bar{h}) = \sum_{\substack{\bar{h}\in H\\ s.t.\ \bar{h}^0=\tilde{h}^0}} \phi_j(\bar{h}).$$

Case 2) *If* $\Phi^0_{\sigma_{-i}} = \text{rect}(\Phi^0_{\sigma_{-i}})$ *for all* $i \in N$, *any* $\phi \in \text{rect}(\Phi^0_{\sigma_{-i}})$ *can be represented as*

$$\phi(h) = \frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{c_i}.$$

Then, similar to Equation (1.2), *one can show that even without homogeneous information, there exists* $\pi \in \mathcal{P}$ *such that*

$$\sum_{\substack{\bar{h}\in H\\ s.t.\ \bar{h}^0=\tilde{h}^0}} \phi(h) = \pi(\tilde{h}^0)$$

for any type or state $\tilde{h}^0$. *Hence, the set of marginal beliefs over states or types equals the common set of imprecise probabilistic information* $\mathcal{P}$.

[11] To see this, remember that by $c_i = \bar{c}_i = \prod_{t=1}^{T}|A_i^t|$. We could interpret $\frac{1}{\bar{c}_i}$ as the transition probability of the strategy that chooses at each information set a uniform distribution over the set of acts A_i^t. Therefore, there exists a strategy of player i, σ_i such that $p_{\sigma_i}(h|h^0) = \frac{1}{\bar{c}_i} = \frac{1}{c_i}$.

[12] Note, that assuming homogeneous information leads to a similar setting as in Liu and Xiong (2016).

1.3 Ex-Ante and Interim Equilibria

In this section, we define ex-ante and interim expected utility, and equilibria. We prove the existence of ex-ante and interim expected equilibria with rectangular beliefs. Then, we show that rectangularity leads to dynamically consistent behavior.

1.3.1 Definition and Existence

Given the ex-ante or interim set of beliefs, we assume that players evaluate a strategy profile by maximizing their worst-case expected utility.

Definition 1.6. *The* ***ex-ante expected utility*** *of a strategy profile $\sigma = (\sigma_i, \sigma_{-i})$ for player i is given by*

$$U^e((\sigma_i, \sigma_{-i})) := \min_{\phi \in \Psi^0_{\sigma_{-i}}} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^0)\phi(h).$$

Similarly, the ***interim expected utility*** *at $F_i^t \in \mathcal{F}_i^t$ at stage $t = 0, \ldots, T$ given a belief set $\Psi^t_{\sigma_{-i}}(F_i^t)$ is*

$$U_i^i(\sigma, \Psi^t_{\sigma_{-i}}(F_i^t)) := \min_{\phi \in \Psi^t_{\sigma_{-i}}(F_i^t)} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^t)\phi(h).$$

Given the ex-ante and interim expected utility of the players, the definitions of an ex-ante and interim equilibrium are straightforward and follow the standard idea of Nash Equilibrium.

Definition 1.7. *A strategy profile σ^* is an* ***ex-ante equilibrium with rectangular beliefs*** *if and only if $\Psi^0_{\sigma_{-i}} = \mathrm{rect}(\Phi^0_{\sigma^*_{-i}})$ and*

$$U_i^e(\sigma^*) \geq U_i^e((\sigma_i', \sigma^*_{-i}))$$

for all $\sigma_i' \in \Sigma_i$ and $i \in N$.

Definition 1.8. *A tuple $(\sigma^*, (\Psi^t_{\sigma^*_{-i}})_{i \in N})$, consisting of a strategy profile σ^* and collection of beliefs $\Psi^t_{\sigma^*_{-i}}$ for each information set at stage t and each player, is an* ***interim equilibrium with rectangular beliefs*** *at stage t if and only if players have rectangular beliefs $\Psi^t_{\sigma_{-i}}(F_i^t) = \mathrm{rect}(\Phi^t_{\sigma^*_{-i}}(F_i^t))$ for all $F_i^t \in \mathcal{F}_i^t$ and*

$$U_i^i(\sigma^*, \Psi^t_{\sigma_{-i}}(F_i^t)) \geq U_i^i((\sigma_i', \sigma^*_{-i}), \Psi^t_{\sigma_{-i}}(F_i^t))$$

for all $\sigma_i' \in \Sigma_i$, $F_i^t \in \mathcal{F}_i^t$ and all $i \in N$.

Since the normalization constant c_i of an ex-ante belief over histories is constant across all histories, it does not influence the maximization problem of player i. Formally, if

$\Psi_{\sigma_{-i}} = \Phi_{\sigma_{-i}}$:

$$U^e((\sigma_i, \sigma_{-i})) = \frac{1}{c_i} \min_{\pi \in \mathcal{P}} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^t) p_{\sigma_{-i}}(h|h^t) \pi(h^0),$$

and similar for the interim stages with the constant c_i^t. Therefore, defining belief over histories instead of partial histories as in classical approaches does not influence the set of equilibria.

Without assuming rectangularity existence of an ex-ante equilibrium follows from standard arguments using that $\phi(h) = \frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi(h^0)$ for all $\phi \in \Phi^0_{\sigma_{-i}}$. With rectangularity, this simple characterization of the beliefs does not hold in general.

Theorem 1.1. *There exists an ex-ante equilibrium with rectangular beliefs.*

The proof of Theorem 1.1 follows the usual idea using Kakutani's fixed point theorem and can be found in the Appendix. For Kakutani's fixed point theorem, it is essentially that $U_i^e((\sigma_i, \sigma_{-i}))$ is jointly continuous in (σ_i, σ_{-i}). Due to our definition of beliefs, the opponents' strategies influence the set of beliefs over which a player minimizes. Therefore, continuity is not trivial, and we need the following lemma for the proof of Theorem 1.1.

Lemma 1.2. *$U_i^e((\sigma_i, \sigma_{-i}))$ with $\Psi^0_{\sigma_{-i}} = \text{rect}(\Phi^0_{\sigma_{-i}})$ is jointly continuous in (σ_i, σ_{-i}).*

To prove continuity we use that any belief in $\text{rect}(\Phi^0_{\sigma_{-i}})$ can be represented by the pasting of marginal and updated belief. Due to Remark 1.2 marginal and updated beliefs can be represented by multiplying $p_{\sigma_{-i}}(h|\cdot)$ and $\pi \in \mathcal{P}$. This leads to the following representation of an arbitrary element $\phi \in \text{rect}(\Phi^0_{\sigma_{-i}})$.

$$\phi(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \left(\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \pi(h^0) \right) \frac{\frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \pi'_{F_i^1}(h^0)}{\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \pi'_{F_i^1}(h^0)}.$$

Using this representation, we transform $U_i^e((\sigma_i, \sigma_{-i}))$ to a minimization problem over the set $(\pi, (\pi')_{F_i^1}) \in \mathcal{P} \times \mathcal{P}^{|\mathcal{F}_i^1|}$ which is independent of σ_{-i}. Then, jointly continuity of the transformed problem follows by standard methods.

Theorem 1.2. *There exists an interim equilibrium with rectangular beliefs at stage t.*

Proof. The compactness of $\mathcal{P}$ implies compactness of $\text{Bay}(\mathcal{P}|F_i^t)$ for any F_i^t. Hence, replacing $\text{rect}(\Phi^0_{\sigma_{-i}})$ in Lemma 1.2 by $\text{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))$ shows jointly continuity of $U_i^i(\sigma_i, \sigma_{-i})$. Then, the proof follows the same line as the proof of Theorem 1.1. □

1.3.2 Relation of Ex-Ante and Interim Equilibria

Now, we come back to the problem of dynamic consistency and show that rectangularity induces dynamically consistent behavior.

Our next theorem shows that under rectangularity, a completely mixed ex-ante equilibrium implies an interim equilibrium. Therefore, as in games without ambiguity, a player would deviate from an ex-ante optimal strategy only at out-of-equilibrium information sets.

Theorem 1.3 (Ex-ante implies Interim)**.** *Let σ^* be a completely mixed ex-ante equilibrium with rectangular beliefs. Then $(\sigma^*, \Psi^t_{\sigma_{-i}})$ with $\Psi^t_{\sigma_{-i}}(F^t_i) = \text{rect}(\Phi^t_{\sigma^*_{-i}}(F^t_i))$ is an interim equilibrium with rectangular beliefs at stage t.*

We will prove this theorem by showing that due to rectangularity, a completely mixed interim equilibrium at $t-1$ implies a completely mixed equilibrium at t. This holds for any arbitrary $t = 0, \dots T+1$ such that Theorem 1.3 follows by iteration. The next corollary follows immediately from the recursive structure of the proof.

Corollary 1.1. *Let $(\sigma^*, \Psi^{t-1}_{\sigma_{-i}})$ be a completely mixed interim equilibrium at $t-1$ with rectangular beliefs, i.e., $\Psi^{t-1}_{\sigma_{-i}}(F^{t-1}_i) = \text{rect}(\Phi^{t-1}_{\sigma^*_{-i}}(F^{t-1}_i))$. Then $(\sigma^*, \Psi^t_{\sigma_{-i}})$ with $\Psi^t_{\sigma_{-i}}(F^t_i) = \text{rect}(\Phi^t_{\sigma^*_{-i}}(F^t_i))$ is an interim equilibrium with rectangular beliefs at stage t.*

The formal proof of Theorem 1.3 can be found in Section 1.6.1.2 in the Appendix. To give an intuition of the result, we need the next lemma. It shows the relation between the worst-case expected utility at different stages and is essential for the relation between interim and ex-ante equilibria.

Lemma 1.3. *Let $\phi^* \in \arg\min_{\phi \in \text{rect}(\Phi^{t-1}_{\sigma_{-i}}(F^{t-1}_i))} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^{t-1})\phi(h)$. Then,*

$$\sum_{h \in F^{t-1}_i} u_i(h) p_{\sigma_i}(h|h^{t-1})\phi^*(h)$$
$$= \sum_{F^t_i \in \mathcal{F}^t_i} \phi^*(F^t_i) p_{\sigma_i}(h^t|h^{t-1}) \min_{\tilde{\phi} \in \text{rect}(\Phi^t_{\sigma_{-i}}(F^t_i))} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^t)\tilde{\phi}(h). \quad (1.3)$$

We give a sketch of the proof since it helps to understand the role of rectangularity. The proof consists of two steps. First, Bayesian updating implies that the left-hand side of Equation (1.3) is greater or equal than the right-hand side. The other direction follows from rectangularity. Due to rectangularity, there exists a $\phi' \in \Psi^{t-1}_{\sigma_{-i}}(F^{t-1}_i)$ such that

$$\phi'(h) = \sum_{F^t_i \in \mathcal{F}^t_i} \phi^*(F^t_i) \tilde{\phi}^*_{F^t_i}(h), \quad (1.4)$$

where $\tilde{\phi}^*_{F_i^t}$ and ϕ^* are the worst-case beliefs at F_i^t or F_i^{t-1}, respectively, i.e. ϕ' is the pasting of the worst-case beliefs at t and $t-1$. We still do not know if ϕ' is the worst-case belief at $t-1$. Therefore, the left-hand side of Equation (1.3) can be smaller or equal than the left-hand side evaluated with the belief ϕ' instead of ϕ^*. Then, using that ϕ' is the pasting of the worst-case beliefs ϕ^* and $\tilde{\phi}^*_{F_i^t}$, we can prove that the left-hand side of Equation (1.3) evaluated with ϕ' equals the right-hand side of Equation (1.3).

As the next remark shows, Lemma 1.3 implies that the worst-case belief at F_i^{t-1} is the pasting of the worst-case belief at F_i^{t-1} and the worst-case beliefs at t. Therefore, the Bayesian update of the worst-case belief at F_i^{t-1} leads to the worst-case belief at all subsequent information sets at t.

Remark 1.4. *The proof of Lemma 1.3 shows the existence of a belief* $\phi' \in \Psi^{t-1}_{\sigma_{-i}}(F_i^{t-1}) = \mathrm{rect}(\Psi^{t-1}_{\sigma_{-i}}(F_i^{t-1}))$ *which satisfies Equation* (1.4)*, i.e.,*

$$\phi'(h) = \sum_{F_i^t \in \mathcal{F}_i^t} \phi^*(F_i^t)\tilde{\phi}^*_{F_i^t}(h),$$

where ϕ^* *and* $\tilde{\phi}^*_{F_i^t}$ *are the worst-case beliefs at* F_i^{t-1} *and* F_i^t*, respectively. Furthermore, the proof states*

$$\begin{aligned}
&\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i}(h|h^{t-1})\phi^*(h) \\
&\qquad\leq \sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i}(h|h^{t-1})\phi'(h) \\
&\qquad= \sum_{F_i^t \in \mathcal{F}_i^t} \phi^*(F_i^t) p_{\sigma_i}(h^t|h^{t-1}) \min_{\tilde{\phi} \in \Phi^t_{\sigma_{-i}}(F_i^t)} \sum_{h \in F_i^t} u_i(h) p_{\sigma_i}(h|h^t)\tilde{\phi}(h).
\end{aligned}$$

But by Lemma 1.3, we know that the inequality is an equality. Therefore, ϕ' *is a worst-case belief at* F_i^{t-1} *and Remark 1.2 implies that the worst-case belief at an information set at stage* t *is the Bayesian update of the worst-case belief of the previous information set at stage* $t-1$.

Remark 1.4 and Lemma 1.3 show how rectangularity leads to dynamically consistent behavior which is necessary for the proof of Theorem 1.3.

The proof of Theorem 1.3 follows the usual idea of contraposition. If there would exist a profitable deviation at t, this deviation would be profitable at $t-1$ as well. Therefore, an equilibrium at $t-1$ implies an equilibrium at t. Then, the theorem follows from iteration. However, one has to consider the worst-case beliefs. The belief set only depends on the strategy of the opponents. Therefore, fixing the strategy of the opponents leads to fixed belief sets. However, the worst-case belief of player i may change if he deviates from the equilibrium strategy. To prove Theorem 1.3, we have to define the pasting of the worst-case belief at t given the equilibrium strategy with the worst-case belief at $t-1$ given the deviation strategy. Due to rectangularity, this pasting is an element of the belief set at $t-1$. Then, we use Lemma 1.3 to show that a profitable deviation at t implies a profitable deviation at $t-1$.

Theorem 1.3 shows the relation between equilibria at different stages. Roughly speaking, due to rectangularity, players update their beliefs such that their worst-case belief at t is the Bayesian update of the worst-case belief at $t-1$. This implies dynamically consistent behavior and leads to the relation between ex-ante and interim equilibria stated in Theorem 1.3.

We come back to our running example and show that rectangularity rules out dynamic inconsistency.

Running Example (cont.). *Remember the results from above. Without rectangularity, player* 1 *behaves dynamically inconsistently. His optimal ex-ante strategy is to play M with probability one. After learning that the state is not O, his optimal interim strategy is to play M with probability $m = \frac{1}{102}$ if $\epsilon > \frac{1}{102}$. Now, we will show that beliefs that are rectangular lead to dynamically consistent behavior. Since player* 2 *is indifferent between S and T, we still focus on player* 1*. We already know the information partitions and rectangular beliefs of player* 1*:*

$$\begin{aligned}
H &= \{LM, LN, RM, RNS, RNT, OS, OT\},\\
\mathcal{F}_i^0 &= H,\\
\mathcal{F}_1^1 &= \{F_{1,1}^1, F_{1,2}^1\} = \big\{\{LM, LN, RM, RNS, RNT\}, \{OS, OT\}\big\},\\
\Psi_1^0 &= \mathrm{rect}(\Phi_{(s,t)}^0)\\
&= \mathrm{conv}\left\{\left(0, 0, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, \epsilon s, \epsilon t\right),\right.\\
&\qquad \left(\frac{\epsilon}{2}, \frac{\epsilon}{2}, \frac{1-\epsilon}{2}, \frac{(1-\epsilon)s}{2}, \frac{(1-\epsilon)t}{2}, 0, 0\right), \left(0, 0, \frac{1}{2}, \frac{s}{2}, \frac{t}{2}, 0, 0\right),\\
&\qquad \left.\left(\frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)\epsilon}{2}, \frac{(1-\epsilon)^2}{2}, \frac{(1-\epsilon)^2 s}{2}, \frac{(1-\epsilon)^2 t}{2}, \epsilon s, \epsilon t\right)\right\}.
\end{aligned}$$

From Remark 1.2, we know that the prior-by-prior Bayesian updates of $\mathrm{rect}(\Phi_{(s,t)}^0)$ *and* $\Phi_{(s,t)}^0$ *are the same. Therefore, the optimal interim strategy with rectangular beliefs is the same as without rectangularity, i.e., $m^* = \frac{1}{102}$ if $\epsilon > \frac{1}{102}$. For the optimal ex-ante strategy with rectangular beliefs we solve the following problem*

$$\begin{aligned}
&\max_{(1-m,m)} U_1^e((1-m, m))\\
&= \max_{(1-m,m)} \min_{\phi \in \mathrm{rect}(\Phi_{(s,t)}^0)} 101(1-m)\phi(LN) + 101m\phi(RM) + 100(1-m)\phi(RNT)
\end{aligned}$$

$$+\,100(1-m)\phi(RNS)-\phi(OS)-\phi(OT)$$

$$=\max_{(1-m,m)}\min_{\substack{(\frac{l}{2},\frac{l}{2},\frac{r}{2},\frac{rs}{2},\frac{rt}{2},os,ot)\\ \in \text{rect}(\Phi^0_{(s,t)})}} 101(1-m)\frac{l}{2}+101m\frac{r}{2}+100(1-m)\frac{r}{2}-(1-r-l)$$

$$=\max_{(1-m,m)}\min_{\substack{(\frac{l}{2},\frac{l}{2},\frac{r}{2},\frac{rs}{2},\frac{rt}{2},os,ot)\\ \in \text{rect}(\Phi^0_{(s,t)})}} l\left(\frac{101(1-m)}{2}-1\right)+r\left(\frac{100+m}{2}-1\right)+1.$$

The worst-case belief depends on m. If $101(1-m)>100+m$, the worst-case belief gives the lowest possible value to l, the highest value to o, and $r=1-l-o$. If $101(1-m)\leq 100+m$, the worst-case belief gives the highest value to o, the lowest value to r, and $l=1-o-r$. Hence, the worst-case belief is

$$\phi^*=\begin{cases}\left(0,0,\frac{1-\epsilon}{2},\frac{(1-\epsilon)s}{2},\frac{(1-\epsilon)t}{2},\epsilon s,\epsilon t\right) & \text{if } m<\frac{1}{102},\\ \left(\frac{\epsilon(1-\epsilon)}{2},\frac{\epsilon(1-\epsilon)}{2},\frac{(1-\epsilon)^2}{2},\frac{(1-\epsilon)^2 s}{2},\frac{(1-\epsilon)^2 t}{2},\epsilon s,\epsilon t\right) & \text{if } m\geq\frac{1}{102}.\end{cases}$$

The worst-case ex-ante utility is

$$U_1^e((1-m,m))=\begin{cases}(1-\epsilon)(\frac{100+m}{2}-1)+1 & \text{if } m<\frac{1}{102},\\ (1-\epsilon)\epsilon(\frac{101(1-m)}{2}-1)+(1-\epsilon)^2(\frac{100+m}{2}-1)+1 & \text{if } m\geq\frac{1}{102},\end{cases}$$

$$=\begin{cases}(1-\epsilon)(49+\frac{m}{2})+1 & \text{if } m<\frac{1}{102},\\ m\frac{(1-\epsilon)}{2}(1-102\epsilon)+(1-\epsilon)(49+\frac{\epsilon}{2})+1 & \text{if } m\geq\frac{1}{102}.\end{cases}$$

Hence, the optimal ex-ante strategy is $m^*=\frac{1}{102}$ if $\epsilon>\frac{1}{102}$ which proves dynamic consistency.

Given the relation between ex-ante and interim equilibria discussed in this section, we can now define and prove the existence of sequential equilibria.

1.4 Sequential Equilibria

Kreps and Wilson (1982) define a sequential equilibrium in a game without ambiguity as a tuple of a strategy profile and a belief system such that the strategy profile is sequentially rational and the belief system is consistent with respect to the strategy profile. Consistency with respect to a strategy profile σ means that there exists a sequence of completely mixed strategy profiles that converges to σ such that the sequence of beliefs constructed by Bayesian updating given the completely mixed strategy profiles converges to the equilibrium belief. We use a similar notion of consistency that includes rectangularity.

Fix a sequence $\epsilon^k=(\epsilon^k_I)_{I\in\cup_{i\in N}\mathcal{I}_i}$ with $0<\epsilon^k_{I_i}\leq\frac{1}{|A_i(I_i)|}$ for all player i and information sets I_i that converges in the sup-norm to zero. For any k, let Γ^k denote the restriction of

Γ such that the set of feasible strategies is the set of all completely mixed σ^k satisfying $\sigma_i^k(I_i)(a_i) \geq \epsilon_{I_i}^k$ for all players, information sets, and actions $a_i \in A_i(I_i)$. Let Σ^k denote the set of strategy profiles satisfying this constraint. For every strategy profile in Σ^k, Bayes' rule is always well defined. Let $(\sigma^k)_k$ with $\sigma^k \in \Sigma^k$ converge to $\sigma \in \Sigma$ as k goes to infinity. For each player i and each σ_{-i}^k, we can construct a ex-ante belief set $\text{rect}(\Phi^0_{\sigma_{-i}^k})$ which is rectangular. Now, we construct an ex-ante belief system given σ which is rectangular and consistent with σ_{-i}^k. First, note that

$$\lim_{k\to\infty} \frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i} = \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \tag{1.5}$$

for all $h \in H$. Take an arbitrary tuple $(\pi, (\pi'_{F_i^1})_{F_i^1 \in \mathcal{F}_i^1}) \in \mathcal{P} \times \mathcal{P}^{|\mathcal{F}_i^1|}$. Then, there exists a sequence of $\phi^k \in \text{rect}(\Phi^0_{\sigma_{-i}^k})$ such that

$$\phi^k(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \phi^k(F_i^1)\bar{\phi}^k_{F_i^1}(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \Big(\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi(h^0)\Big) \frac{\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}.$$

Please note that the latter fraction $\frac{p_{\sigma_{-i}^k}(h|h^0)\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}$ is an element of the Bayesian update of $\mathcal{P}$ given F_i^1 and σ^k for each k and therefore an element of $\Delta(H^1)$, the set of probability distributions over partial histories at stage one. Taking the limit of ϕ^k only influences the path probability induced by the strategy of the opponents, i.e., π and $\pi'_{F_i^1}$ are fixed. Hence, Equation (1.5) implies

$$\lim_{k\to\infty}\phi^k(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \Big(\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi(h^0)\Big) \lim_{k\to\infty} \frac{\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}.$$

Then, the compactness of $\Delta(H^1)$ implies that every sequence of $\frac{\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}^k}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}$ has a convergent subsequence.

The limit of $\text{rect}(\Phi_{\sigma_{-i}^k})$ is then defined as

$$\begin{aligned}&\lim_{k\to\infty}\text{rect}(\Phi^0_{\sigma_{-i}^k})\\ &:= \Big\{\phi \in [0,1]^H \ : \ \exists(\phi^k)_{k=1,2,\dots} \in \big(\text{rect}(\Phi^0_{\sigma_{-i}^k})\big)_{k=1,2,\dots} \text{ with } \phi(h) = \lim_{k\to\infty}\phi^k(h)\Big\}.\end{aligned}$$

Similarly, one can define the limit of rectangular interim belief sets at stages $t > 0$.

By construction, $\lim_k \operatorname{rect}(\Phi^{t-1}_{\sigma^k_{-i}}(F_i^{t-1}))$ is rectangular as the following calculations show. Let $\phi \in \lim_k \operatorname{rect}(\Phi^{t-1}_{\sigma^k_{-i}}(F_i^{t-1}))$ and $(\phi_{F_i^t})_{F_i^t} \in \Big(\mathrm{Bay}\big(\lim_k \operatorname{rect}(\Phi^{t-1}_{\sigma^k_{-i}}(F_i^{t-1}))|F_i^t\big)\Big)_{F_i^t}$. We have to show that the pasting of ϕ and $(\phi_{F_i^t})_{F_i^t}$ is an element of $\lim_k \operatorname{rect}(\Phi^{t-1}_{\sigma^k_{-i}}(F_i^{t-1}))$. The pasting is given by

$$\phi \circ (\phi_{F_i^t})_{F_i^t}(\cdot) = \sum_{F_i^t \in \mathcal{F}_i^t} \lim_k \phi^k(F_i^t) \lim_k \phi_{F_i^t}(\cdot) = \lim_k \sum_{F_i^t \in \mathcal{F}_i^t} \phi^k(F_i^t)\phi_{F_i^t}(\cdot).$$

Then, since $\sum_{F_i^t \in \mathcal{F}_i^t} \phi^k(F_i^t)\phi_{F_i^t}(\cdot) \in \operatorname{rect}(\Phi^{t-1}_{\sigma^k_{-i}})$ we get

$$\phi \circ (\phi_{F_i^t})_{F_i^t}(\cdot) \in \lim_k \operatorname{rect}(\Phi^{t-1}_{\sigma^k_{-i}}(F_i^{t-1}))$$

and rectangularity is maintained under the limit. Now, we can define consistency with respect to a strategy profile σ and sequential rationality for rectangular beliefs.

Definition 1.9. *We say that a belief system Ψ is* ***consistent w.r.t.*** *σ if there exists a sequence $(\sigma^k)_{k=1,\dots}$ such that*

- $\sigma^k \in \Sigma^k$ *for all k,*
- $\sigma = \lim_k \sigma^k$,
- $\Psi^t_{\sigma_{-i}}(F_i^t) = \{\phi : \phi(h) = \lim_k \phi^k(h), \phi^k \in Bay(\operatorname{rect}(\Phi^0_{\sigma^k_{-i}}|F_i^t))\}$ *for all $F_i^t \in \mathcal{F}_i^t$ and $t \geq 0$.*

The definition of consistency w.r.t. σ and the discussion above show that a belief system, which is consistent w.r.t. σ and rectangular for σ^k, is rectangular for σ.

The second property of sequential equilibria is sequential rationality. Roughly speaking, sequential rationality captures the idea that a strategy is optimal at each stage and each information set. Therefore, a strategy is sequentially rational if it is an ex-ante and interim equilibrium at each stage.

Definition 1.10. *A tuple (σ, Ψ), consisting of a strategy profile and a belief system, is* ***sequentially rational*** *if*

- *σ is an ex-ante equilibrium with rectangular beliefs $\Psi^0_{\sigma_{-i}} = \operatorname{rect}(\Phi^0_{\sigma_{-i}})$ for all $i \in N$ and*
- *for all $t > 0$, the tupel (σ, Ψ^t) is an interim equilibrium with rectangular beliefs $\Psi^t_{\sigma_{-i}}(F_i^t) = \operatorname{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))$ for all $i \in N$ at stage t.*

Now, we can define a sequential equilibrium.

Definition 1.11. *The tupel* (σ^*, Ψ) *consisting of a strategy profile and a belief system with* $\Psi^0_{\sigma_{-i}} = \text{rect}(\Phi^0_{\sigma^*_{-i}})$ *is a* ***sequential equilibrium with rectangular beliefs*** *if*

- (σ^*, Ψ) *is sequentially rational and*
- Ψ *is consistent w.r.t.* σ.

Given dynamic consistency, the existence proof follows a similar idea as without ambiguity.

Theorem 1.4. *There exists a sequential equilibrium with rectangular beliefs.*

Proof. Let ϵ^k, Γ^k, and σ^k as above. For each Γ^k, we can construct a belief system Ψ^k that is rectangular, i.e., $\Psi^{t,k}_i(F^t_i) = \text{rect}(\Phi^t_{\sigma^k_{-i}}(F^t_i))$. Furthermore, by Theorem 1.1 there exists an ex-ante equilibrium $\hat{\sigma}^k$ with rectangular beliefs $\Psi^{0,k}$ for each Γ^k. Theorem 1.3 shows that $(\hat{\sigma}^k, \Psi^{t,k})$ is an interim equilibrium with rectangular beliefs at stage t. By compactness of the set of strategy profiles, there exists a sub-sequence of σ^k which converges to $\hat{\sigma}$. For this sub-sequence, we can construct a system of rectangular beliefs Ψ such that, $\Psi^t_{\sigma_{-i}}(F^t_i) = \lim_k \Psi^{t,k}_{\sigma_{-i}}(F^t_i)$. Then by construction Ψ is rectangular and satisfies consistency w.r.t. $\hat{\sigma}$.
By Lemma 1.2 $U^e_i(\sigma)$ and $U^i_i(\sigma, \Phi^t)$ are jointly continuous in σ. Then, since Σ is the closure of $\bigcup_k \Sigma^k$ the strategy profile $\hat{\sigma}$ satisfies sequential rationality. □

1.4.1 Sequential Rationality and Rectangularity

The assumption that $\Psi^0_{\sigma_{-i}} = \text{rect}(\Phi^0_{\sigma^*_{-i}})$ is essential. Epstein and Schneider (2003) and Riedel et al. (2018) show that dynamic consistency implies rectangular belief sets. However, sequential rationality is a weaker condition than the dynamic consistency axiom required by Epstein and Schneider (2003) and Riedel et al. (2018). The following version of our running example shows that sequential rationality does not imply rectangularity.

Running Example (cont.). *In Section 1.2.1, we show that the interim optimal strategy of player* 1 *given* $\mathcal{P}$ *is* $m^* = \frac{1}{102}$ *if* $\epsilon > \frac{1}{102}$. *For* $\epsilon < \frac{1}{102}$, *playing* $m^* = 1$ *is the optimal interim strategy. Hence, given* $\epsilon < \frac{1}{102}$, *playing* M *with probability one is an ex-ante and interim optimal choice. We have already shown, that*

$$\Phi^0_{(s,t)} = \left\{ \left(\frac{l}{2}, \frac{l}{2}, \frac{r}{2}, \frac{rs}{2}, \frac{rt}{2}, os, ot \right) \ : \ (l, r, o) \in \mathcal{P} \right\}$$

is not rectangular for arbitrary ϵ. *However, one can easily show that* $m^* = 1$ *and* $\Phi_{(s,t)}$ *are sequentially rational for* $\epsilon < \frac{1}{102}$. *Therefore, sequential rationality does not imply rectangular belief sets.*

Even if the equivalence of sequential rationality and rectangularity does not hold in general, we can prove for a special case that sequential rationality of (σ^*, Ψ) implies that σ^* is sequentially rational with respect to $\mathrm{rect}(\Psi)$.

Theorem 1.5. *Let $(\sigma^*, \Psi_{\sigma_{-i}})$ with $\Psi_{\sigma_{-i}} = \Phi_{\sigma_{-i}}$ be sequentially rational. Furthermore, let $((\phi^*_{F_i^t})_{F_i^t})_{t=0,\dots T}$ denote the collection of worst-case beliefs given σ^* at different information sets and stages. Assume that for all players $i \in N$, all stages $t = 0, \dots T$, and information sets $F_i^t \in \mathcal{F}_i^t$, there exists no $\phi_{F_i^t} \in \Psi^t_{\sigma_{-i}(F_i^t)}$ such that $\phi_{F_i^t} \neq \phi^*_{F_i^t}$ and σ_i^* is an ex-ante or interim best response for the same game with singleton belief $\phi_{F_i^t}$. Then, $(\sigma^*, \mathrm{rect}(\Phi_{\sigma_{-i}}))$ is sequentially rational.*

Intuitively, the assumptions ensure a unique worst case belief. Then, the only ex-ante belief that ensures sequential rationality is the pasting of all worst-case beliefs. Since this is the worst-case belief given the rectangular hull, σ^* is sequentially rational given the rectangular hull $\mathrm{rect}(\Psi_{\sigma_{-i}})$. The example above does not satisfy the above assumption since $m = 1$ is optimal given any ex-ante belief in $\Psi^0_{(s,t)}$. We guess that similar results do hold for more general cases. However, exploring the relation of sequential rationality and rectangularity for more general cases is left for future research.

1.4.2 Properties of Sequential Equilibria

There are two properties of sequential equilibria with rectangular beliefs that we would like to highlight.

Remark 1.5. *First, due to ambiguity and multiple prior preferences, players may have heterogeneous worst-case beliefs. Since each player maximizes his worst-case utility, ambiguity can induce sequential equilibria that cannot exist without ambiguity and common priors.*

Second, as in games without ambiguity, sequential equilibria are an equilibrium refinement that rules out non-credible threats.

The following example illustrates these properties of sequential equilibria with rectangular beliefs.

Example 1.2. *This example follows the idea of Greenberg (2000) and is similar to the running example of Hanany et al. (2020). There are three countries, two small countries A and B, and one influential country C. Country A and B are involved in peace negotiations, which are successful if both countries agree on peace. At stage one, A decides whether to agree with the peace agreement (peace) or not (war). If A plays peace, country B can choose peace or war. If one of the countries chooses war, the peace negotiation fails. Country C observes if the peace negotiations failed or not, but it cannot distinguish which of the two countries broke up the negotiation. After* *observing that the peace negotiations failed, country C can either punish country A or B, denoted by p_A and p_B, respectively or stay neutral, denoted by n. If the peace agreement is successful, C favors either A or B, denoted with f_A and f_B. To introduce ambiguity, we assume that C can condition his action on a payoff-irrelevant ambiguous state, I, or II. This means that C observes the state, whereas countries A and B do not know the state. With p_A^I and p_A^{II} and, similarly, p_B^I, p_B^{II}, n^I, n^{II}, we denote actions p_A, p_B, n conditioned on the state I or II, respectively. The game is depicted in Figure 1.3.*

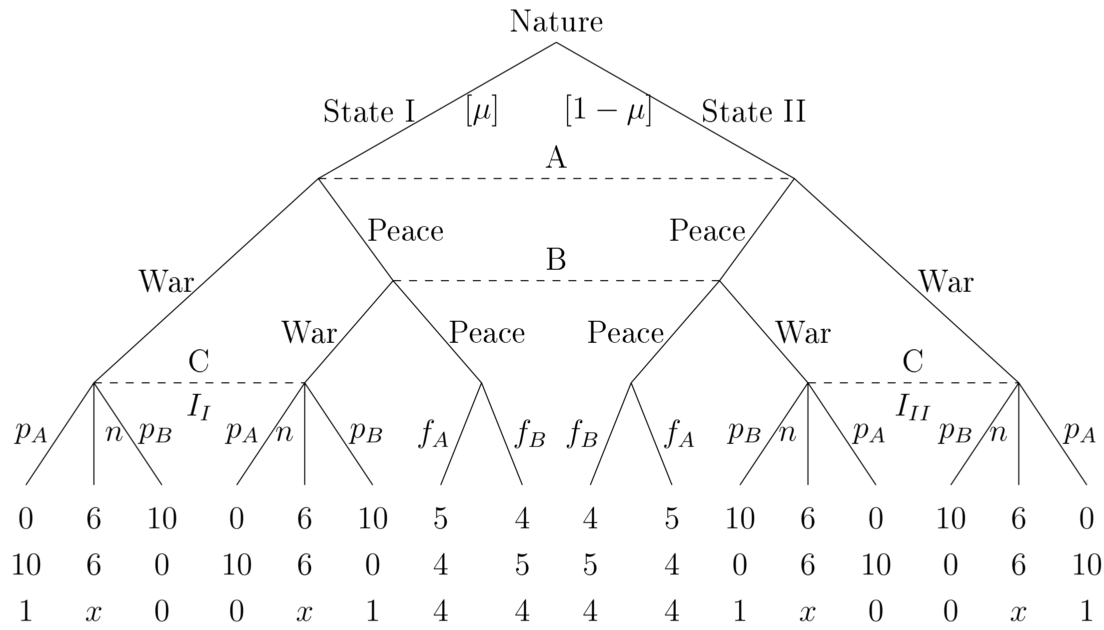

Figure 1.3: Example Peace Negotiation.

First, we show that there exists no ex-ante equilibrium in which the peace agreement takes place without ambiguity. Therefore, without ambiguity there cannot exist a sequential equilibrium with a peace agreement. Let μ represent the ex-ante belief of countries A and B that the state is I in a game without ambiguity. Furthermore, let α and β denote the probability with which countries A and B choose war, respectively. In Section 1.6.3.2 in the Appendix, we calculate all ex-ante equilibria of this game without ambiguity and show that in all of them, at least one of the countries A or B plays war with probability one. Intuitively, A and B cannot distinguish between state I and II and choose the same action at both states. Therefore, the updated belief of C given that the negotiations failed is the same at both information sets I_I and I_{II}. It is then either optimal for C always to punish A or B or, if C is indifferent, to mix and, e.g., punish A if the state is I and B if the state is II, or vice versa. To motivate A and B to play peace, both have to believe that they will be punished with a high probability. Without ambiguity and with common beliefs, the ex-ante belief of A that himself will be punished *is* $\mu p_A^I + (1-\mu)p_A^{II}$ *and, similarly, country B's belief to be punished is* $\mu p_B^I + (1-\mu)p_B^{II}$*. These beliefs sum up to a value smaller or equal to one.*[13] *Hence, both countries cannot believe simultaneously that they will be punished with a sufficiently high probability to motivate them to play peace.*

With multiple priors, ambiguity induces a new equilibrium in which the peace agreement does not fail. Instead of one ex-ante belief μ*, the players are faced with the imprecise probabilistic information* $\mathcal{P} = [\underline{\mu}, \bar{\mu}]$*. In Section 1.6.3.2, we construct the ex-ante belief sets* $\Phi^0_{\sigma_{-A}}$*,* $\Phi^0_{\sigma_{-B}}$*, and* $\Phi^0_{\sigma_{-C}}$ *and show that they are rectangular. With ambiguity, there exists an ex-ante equilibrium where A and B play peace with probability one. Consider the following strategy profile* σ^**:*

$$\alpha^* = 0, \quad \beta^* = 0, \quad f_A^I = f_A^{II} = \frac{1}{2}, \quad p_A^I = p_B^{II} = 1.$$

This strategy profile forms an ex-ante equilibrium as long as $\bar{\mu} > 0.55$ *and* $\underline{\mu} < 0.45$*. Given* $\alpha^* = 0$ *and* $\beta^* = 0$*, C has no incentive to deviate since the information sets* I_I *and* I_{II} *are reached with probability zero. To show that A does not have an incentive to deviate, we compare his worst-case payoff from playing* α^* *or* $\alpha = 1$*. The payoff from playing* α^* *is*

$$\min_{\mu \in [\underline{\mu}, \bar{\mu}]} \mu\big[(1-\alpha^*)(1-\beta^*)(5f_A^I + 4f_B^I)\big] + (1-\mu)\big[(1-\alpha^*)(1-\beta^*)(5f_A^{II} + 4f_B^{II})\big] = 4.5.$$

The payoff from deviating to war, i.e., $\alpha = 1$ *is*

$$\min_{\mu \in [\underline{\mu}, \bar{\mu}]} \mu \cdot 0 + 10(1-\mu) = 10(1-\bar{\mu}).$$

Therefore deviating to war is not profitable as long as $\bar{\mu} > 0.55$*. Similarly, one can show that B does not deviate from* β^* *as long as* $\underline{\mu} < 0.45$*. This proves that under ambiguity with* $[0.45, 0.55] \subset [\underline{\mu}, \bar{\mu}]$ *there exists an ex-ante equilibrium in which both countries play peace with probability one. Intuitively, due to the worst-case beliefs, each small country believes that it will be punished with a high enough probability to deter it from playing war.*

Next, we discuss that the ex-ante equilibrium under ambiguity specified above is a sequential equilibrium if $x \leq 0.5$*. Furthermore, for* $x > 0.5$*, it is not a sequential equilibrium. The formal proof is given in Section 1.6.3.2. If* $x \leq 0.5$*, we can find a sequence of* α^k *and* β^k*, that converge to 0 as k goes to infinity such that C is indifferent between* p_A *and* p_B *at both information sets for all k. Since* $x \leq 0.5$*, playing* $n^I = n^{II} = 0$ *is optimal at the interim stage.*[14] *Therefore, we can find a sequence of strategy profiles with completely mixed strategies* $(\sigma^k)_k$ *that converges to the ex-ante equilibrium* σ^*

[13]It is equal to one if C plays n with probability zero and strictly smaller than one if n is played with a strictly positive probability.

[14]If $x = 0.5$, country C is indifferent between all 3 actions. Therefore, $p_A^I = 1, p_B^{II} = 1, n^I = n^{II} = 0$ is a best response.

specified above and satisfies sequential rationality. Furthermore, the limit of the belief systems constructed by Bayes' rule for each completely mixed strategy profile σ^k satisfies consistency by construction. If $x > 0.5$, country C has an incentive to deviate to $n^I = n^{II} = 1$ at the interim stage. Hence, there does not exist a sequence of completely mixed strategy profiles that converges to σ^* and satisfies sequential rationality.

The analysis above shows that, due to ambiguity, there exist ex-ante equilibria in which peace is played with probability one. Furthermore, these equilibrium strategy profiles are part of a sequential equilibrium as long as x is small enough, i.e., as long as punishing is a credible strategy for country C.

1.5 Conclusion and Discussion

In this chapter, we introduce rectangularity to finite multistage games with ambiguous incomplete information. Players face imprecise probabilistic information about states or types of opponents. Furthermore, they know the information structure of the game. Given the imprecise probabilistic information and the knowledge about the information structure, each player constructs an ex-ante belief set, which is rectangular. We show that rectangularity ensures dynamically consistent behavior in multistage games with multiple priors and, therefore, the existence of sequential equilibria. Furthermore, we show that in multistage games with rectangular beliefs, ambiguity can create new sequential equilibria that do not exist in games without ambiguity. To conclude, we discuss some related issues and corresponding literature.

Singleton Subjective Beliefs One could argue that our results could be obtained by choosing a single subjective ex-ante belief for each player that equals the worst-case ex-ante belief given rectangular beliefs. Due to rectangularity, the worst-case beliefs at the interim stages are the Bayesian updates of the worst-case ex-ante belief. Therefore, updating this single subjective ex-ante belief would lead to the same beliefs at the interim stages as in the setting with a subjective set of beliefs. But the setting with single subjective beliefs would lack an explanation of how players derive their subjective beliefs. In our model, the subjective set of ex-ante beliefs is endogenously derived from the common imprecise probabilistic information and the knowledge about the game's information structure. This combination leads to rectangular belief sets for dynamically consistent players. The choice of the worst-case beliefs arises due to ambiguity aversion and MEU.

Other Approaches to Deal with Dynamic Inconsistency Siniscalchi (2011) characterizes a consistent planning approach for dynamic choices with dynamically inconsistent preferences. He defines an individual's preferences over decision trees. Consistent planning is a refinement of backward induction that assumes that a decision maker can correctly forecast future decisions. In our approach, rectangular beliefs are constructed by a backward induction method on the beliefs. Auster and Kellner (2020) use the consistent planning approach to analyze Dutch auctions in an ambiguous independent private value setting. In their setting, our approach leads to similar equilibrium outcomes as consistent planning.

Battigalli et al. (2019) use the consistent planning approach to define self-confirming equilibria (SCE) for sequential games with players who admit smooth-ambiguity preferences of Klibanoff et al. (2005). They show that the SCE of a sequential game is not equivalent to the SCE of the strategic form of the game. Further, they analyze sufficient conditions to generalize the monotonicity result of Battigalli et al. (2015) to dynamic games, which states that in static games, the set of SCE expands as ambiguity aversion increases.

Another way to rule out dynamic inconsistency in decision-theoretic settings is the updating rules proposed by Hanany and Klibanoff (2007, 2009). They propose updating rules that update only a subset of the ex-ante belief set using Bayes' rule. Which subset is updated depends on the optimal ex-ante choice. Updating this subset of beliefs leads to an interim belief set supporting the ex-ante optimal choice and ensures dynamic consistency. However, comparing this approach to our model shows that the updating rules of Hanany and Klibanoff (2007, 2009) give a higher weight to the ex-ante optimal choice. This approach may generate different equilibria compared to our approach. For example, in our running example, the updating rules of Hanany and Klibanoff (2007, 2009) would lead to an equilibrium satisfying sequential rationality in which $m = 1$ is ex-ante and interim optimal.

An essential assumption to use rectangularity is that all players know the information structure. To use the updating rules of Hanany and Klibanoff (2007, 2009), agents only have to know the information they are getting. But in games, it is often assumed that the players know the game tree and the information structure. If the players know which information they could get in the future, it seems intuitive that they take this information into account when constructing their ex-ante belief sets.

Hanany et al. (2020) define sequential equilibria for similar multistage games as we do but assume smooth ambiguity preferences instead of maxmin preferences. (Hanany and Klibanoff, 2009) define an updating rule, the *smooth-rule*. Hanany et al. (2020) extend the smooth-rule to multistage games. They show that sequential optimality is equivalent to sequential optimality with respect to beliefs updated with the smooth-rule. Given sequential optimality, they can define sequential equilibria. Further, they show that the set of equilibria extends under a common belief assumption if ambiguity aversion increases.

Ellsberg Games As mentioned in the introduction, Ellsberg games introduce ambiguous strategies, i.e., instead of playing a probability distribution over the pure strategies, players can choose a set of probability distributions. In extensive-form games with Ellsberg strategies, ambiguity arises due to the strategy of the players. In our model, ambiguity occurs due to ambiguous information about types or states. This facilitates the definition of sequential equilibria since strategies are not ambiguous. Muraviev et al. (2017) illustrate the implications of their results for equilibrium concepts in extensive-form games with Ellsberg strategies with an example. They show the existence of a dynamically consistent Ellsberg equilibrium for this example. However, a general formulation of equilibrium concepts for extensive-form games with Ellsberg strategies and existence results is left for future research. Our results support their conjecture that rectangularity implies the existence of dynamically consistent Ellsberg equilibria.

Games with Two Types All examples that show the issue of dynamic inconsistency under MEU in games have at least three possible states or types. The reason for this is that dynamically inconsistent behavior cannot occur in games with only two types or states, i.e., $|H^0| = 2$. With two types, there are only two cases that can arise in an interim stage. Either the player learns the correct type or not. If the player knows the correct type, updated beliefs about types are either zero or one. It is easy to show that then rectangularity is always satisfied. If the player does not learn the type, Bayes' rule is always well defined. Further, since there are only two types, the Bayesian update is monotone in the prior probability. This monotonicity implies dynamically consistent behavior for the second case. But as Example 1.2 shows, ambiguity also induces new sequential equilibria in dynamic games with two states or types. Hence, new equilibria may arise even if there is no dynamically inconsistent behavior. Therefore, there are two effects in games with dynamically inconsistent behavior: ambiguous beliefs and dynamically inconsistent behavior.

1.6 Appendix

1.6.1 Proofs

Proof of Lemma 1.1. Property 1) and the normalization of beliefs imply that for any $\phi \in \Psi^0_{\sigma_{-i}}$ we have

$$\begin{aligned}
1 &= \sum_{h\in H} \phi(h) = \sum_{h\in H} \frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{c_i} = \frac{1}{c_i}\sum_{h\in H} p_{\sigma_{-i}}(h|h^0)\pi(h^0) \\
\Leftrightarrow c_i &= \sum_{h\in H} p_{\sigma_{-i}}(h|h^0)\pi(h^0) \\
&= \sum_{h^0\in H^0} \pi(h^0) \sum_{\substack{\bar{h}^{T-1}\in H^{T-1}\\ \text{s.t. } \bar{h}^0=h^0}} p_{\sigma_{-i}}(\bar{h}^{T-1}|h^0) \underbrace{\sum_{\substack{\hat{h}\in H\\ \text{s.t. } \hat{h}^{T-1}=\bar{h}^{T-1}}} p_{\sigma_{-i}}(\hat{h}|\bar{h}^{T-1})}_{=|A_i^T|} \\
&= |A_i^T| \sum_{h^0\in H^0} \pi(h^0) \sum_{\substack{\bar{h}^{T-1}\in H^{T-1}\\ \text{s.t. } \bar{h}^0=h^0}} p_{\sigma_{-i}}(\bar{h}^{T-1}|h^0).
\end{aligned}$$

Proceeding in the same way by backward induction then shows that $c_i = \prod_{s=1}^T |A_i^s|$. Then, Property 1) and 2) follow by definition. Property 3) follows from Bayesian updating. Let F_i^t be an information set with $\phi(F_i^t) > 0$. Then, any $\tilde{\phi} \in \Psi^t_{\sigma_{-i}}(F_i^t)$ has the following form for all $h \in F_i^t$

$$\tilde{\phi}(h) = \frac{\frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{c_i}}{\sum_{\bar{h}\in F_i^t} \frac{p_{\sigma_{-i}}(\bar{h}|\bar{h}^0)\pi(\bar{h}^0)}{c_i}}.$$

We can rewrite the denominator in the following way:

$$\sum_{\bar{h}\in F_i^t} \frac{p_{\sigma_{-i}}(\bar{h}|\bar{h}^0)\pi(\bar{h}^0)}{c_i} = \sum_{\bar{h}^t\in F_i^t} \frac{p_{\sigma_{-i}}(\bar{h}^t|\bar{h}^0)\pi(\bar{h}^0)}{\prod_{s=1}^{t-1}|A_i^s|} \underbrace{\sum_{\substack{h\in F_i^t\\ \text{s.t. } h^t=\bar{h}^t}} \frac{p_{\sigma_{-i}}(h|\bar{h}^t)}{\prod_{s=t}^{T}|A_i^s|}}_{=1}.$$

Then,

$$\begin{aligned}
\tilde{\phi}(h) &= \frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{\frac{c_i}{\prod_{s=1}^{t-1}|A_i^s|}\sum_{\bar{h}^t\in F_i^t} p_{\sigma_{-i}}(\bar{h}^t|\bar{h}^0)\pi(\bar{h}^0)} \\
&= \frac{p_{\sigma_{-i}}(h^t|h^0)\pi(h^0)}{\sum_{\bar{h}^t\in F_i^t} p_{\sigma_{-i}}(\bar{h}^t|\bar{h}^0)\pi(\bar{h}^0)} \frac{p_{\sigma_{-i}}(h|h^t)}{c_i^t} = \tilde{\pi}(h^t)\frac{p_{\sigma_{-i}}(h|h^t)}{c_i^t},
\end{aligned}$$

where $\tilde{\pi} \in \text{Bay}(\mathcal{P}|F_i^t)$ and $c_i^t = \prod_{s=t}^T |A_i^s|$. $\square$

1.6.1.1 Proofs: Existence of Ex-Ante Equilibria

Proof of Lemma 1.2. First, consider the case where σ_{-i} is completely mixed and Bayes' rule is always well defined. Then, any $\phi \in \operatorname{rect}(\Phi^0_{\sigma_{-i}})$ has the following form

$$\phi(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \phi'(F_i^1)\tilde{\phi}_{F_i^1}(h).$$

By Remark 1.2, we can assume without loss of generality that $\phi' \in \Phi^0_{\sigma_{-i}}$ and $\tilde{\phi}_{F_i^1} \in \Phi^1_{\sigma_{-i}}(F_i^1)$. Furthermore, Bayes' rule is always well defined. Therefore, there exist $\pi \in \mathcal{P}$ and $\pi'_{F_i^1} \in \mathcal{P}$ for each F_i^1 such that

$$\phi(h) = \sum_{F_i^1 \in \mathcal{F}_i^1} \Big(\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi(h^0)\Big) \frac{\frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}.$$

With this formulation of ϕ and since $p_{\sigma_{-i}}(\cdot|\cdot)$ is completely characterized by σ_{-i}, we can now write U_i^e as follows

$$\begin{aligned}
U_i^e(\sigma) &= \min_{\phi \in \operatorname{rect}(\Phi^0_{\sigma_{-i}})} \sum_{h\in H} u_i(h) p_{\sigma_i}(h|h^0)\phi(h) \\
&= \min_{\substack{(\pi,(\pi')_{F_i^1}) \\ \in \mathcal{P}\times\mathcal{P}^{|\mathcal{F}_i^1|}}} \sum_{h\in H} u_i(h) p_{\sigma_i}(h|h^0)\left(\sum_{F_i^1 \in \mathcal{F}_i^1}\Big(\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}}(h|h^0)\pi(h^0)}{c_i}\Big)\frac{\frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}}(h|h^0)\pi'_{F_i^1}(h^0)}{c_i}}\right) \\
&=: \hat{U}_i^e((\sigma_i,\sigma_{-i})).
\end{aligned}$$

Now, we show that $\hat{U}_i^e((\sigma_i,\sigma_{-i}))$, and therefore $U_i^e((\sigma_i,\sigma_{-i}))$, are jointly continuous in (σ_i,σ_{-i}). First note, that

$$\sum_{h\in H} u_i(h) p_{\sigma_i}(h|h^0)\Big(\sum_{F_i^1 \in \mathcal{F}_i^1}\Big(\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi(h^0)\Big)\frac{\frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{\sigma_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}\Big) \tag{1.6}$$

is continuous in (σ_i,σ_{-i}). Then, for all $\epsilon > 0$ exists a $\delta > 0$ such that $|a-b| < \delta$ implies

$$\begin{aligned}
&\Bigg|\sum_{h\in H} u_i(h) p_{a_i}(h|h^0)\left(\sum_{F_i^1 \in \mathcal{F}_i^1}\left(\sum_{h\in F_i^1}\frac{p_{a_{-i}}(h|h^0)}{c_i}\pi(h^0)\right)\frac{\frac{p_{a_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{a_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}\right) \\
&-\sum_{h\in H} u_i(h) p_{b_i}(h|h^0)\left(\sum_{F_i^1 \in \mathcal{F}_i^1}\left(\sum_{h\in F_i^1}\frac{p_{b_{-i}}(h|h^0)}{c_i}\pi(h^0)\right)\frac{\frac{p_{b_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{b_{-i}}(h|h^0)}{c_i}\pi'_{F_i^1}(h^0)}\right)\Bigg| < \epsilon.
\end{aligned}$$

By the compactness of $\mathcal{P}$, there exist $(\pi^a,(\pi'^{,a}_{F_i^1})_{F_i^1})$ and $(\pi^b,(\pi'^{,b}_{F_i^1})_{F_i^1})$ in $\mathcal{P}\times\mathcal{P}^{|\mathcal{F}_i^1|}$ such that

$$\hat{U}_i^e(a) = \sum_{h\in H} u_i(h) p_{a_i}(h|h^0)\left(\sum_{F_i^1 \in \mathcal{F}_i^1}\left(\sum_{h\in F_i^1}\frac{p_{a_{-i}}(h|h^0)}{c_i}\pi^a(h^0)\right)\frac{\frac{p_{a_{-i}}(h|h^0)}{c_i}\pi'^{,a}_{F_i^1}(h^0)}{\sum_{h\in F_i^1}\frac{p_{a_{-i}}(h|h^0)}{c_i}\pi'^{,a}_{F_i^1}(h^0)}\right)$$

and

$$\hat{U}_i^e(b) = \sum_{h \in H} u_i(h) p_{b_i}(h|h^0) \left(\sum_{F_i^1 \in \mathcal{F}_i^1} \left(\sum_{h \in F_i^1} \frac{p_{b_{-i}}(h|h^0)}{c_i} \pi^b(h^0) \right) \frac{\frac{p_{b_{-i}}(h|h^0)}{c_i} \pi'^{,b}_{F_i^1}(h^0)}{\sum_{h \in F_i^1} \frac{p_{b_{-i}}(h|h^0)}{c_i} \pi'^{,b}_{F_i^1}(h^0)} \right).$$

Without loss of generality assume that $\hat{U}_i^e(a) \geq \hat{U}_i^e(b)$. Then for all $\epsilon > 0$ exists $\delta > 0$ such that for $|a-b| < \delta$ it follows

$$\begin{aligned}
&|\hat{U}_i^e(a) - \hat{U}_i^e(b)| \\
\leq &\left| \sum_{h \in H} u_i(h) p_{a_i}(h|h^0) \left(\sum_{F_i^1 \in \mathcal{F}_i^1} \left(\sum_{h \in F_i^1} \frac{p_{a_{-i}}(h|h^0)}{c_i} \pi^b(h^0) \right) \frac{\frac{p_{a_{-i}}(h|h^0)}{c_i} \pi'^{,b}_{F_i^1}(h^0)}{\sum_{h \in F_i^1} \frac{p_{a_{-i}}(h|h^0)}{c_i} \pi'^{,b}_{F_i^1}(h^0)} \right) \right. \\
&\left. - \hat{U}_i^e(b) \right| \\
< &\epsilon.
\end{aligned}$$

The first inequality holds since $(\pi^b, \pi'^{,b})$ is in general not a worst-case belief given strategy a. The second inequality follows by the continuity of Equation (1.6) in (σ_i, σ_{-i}). Hence, $\hat{U}_i^e(\sigma)$ is jointly continuous in $\sigma = (\sigma_i, \sigma_{-i})$ for completely mixed σ_{-i}.

If σ_{-i} is not completely mixed and Bayes' rule is not well defined for some F_i^1, the denominator $\sum_{h \in F_i^1} \phi'_{F_i^1}(h)$ equals zero. However, this is equivalent to $\phi(F_i^1) = 0$.[15] Therefore,

$$\begin{aligned}
&\hat{U}_i^e((\sigma_i, \sigma_{-i})) \\
= &\min_{\substack{(\pi, (\pi')_{F_i^1}) \\ \in \mathcal{P} \times \mathcal{P}^{|\mathcal{F}_i^1|}}} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^0) \\
&\left(\sum_{\substack{F_i^1 \in \mathcal{F}_i^1 \\ \text{s.t. } \phi(F_i^1) \neq 0}} \left(\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \pi(h^0) \right) \frac{\frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \pi'_{F_i^1}(h^0)}{\sum_{h \in F_i^1} \frac{p_{\sigma_{-i}}(h|h^0)}{c_i} \pi'_{F_i^1}(h^0)} \right).
\end{aligned}$$

and we can ignore information sets F_i^1 where Bayes' rule is not well defined. □

Now we can prove Theorem 1.1.

Proof of Theorem 1.1. First remember that the set of histories H, the set of information sets $\mathcal{I}$, the set of actions for each player at each information set $A_i(I_i^t)$, and

[15] This equivalence follows from the full support assumption of $\mathcal{P}$. Since $\pi(h^0) > 0$ for all h^0 and all $\pi \in \mathcal{P}$ an information set has only probability zero if the transition probability of all histories contained in this information set are 0. This implies that $p_{\sigma_{-i}}(h|h^0)\pi(h^0) = 0$ for all $\pi \in \mathcal{P}$.

the set of players N are finite. A behavior strategy of player i was defined such that $\sigma_i(I_i^t) \in \Delta(A_I(I_i^t))$. The set of strategies of player i is then $\Sigma_i = \bigtimes_{I_i^t \in \mathcal{I}_i} \Delta(A_i(I_i^t))$ and the set of strategy profiles $\Sigma = \bigtimes_{i \in N} \Sigma_i$. We define the best response of player i given the strategy of the opponents σ_{-i} as the correspondence $B_i : \Sigma_{-i} \rightarrow \Sigma_i$ with

$$B_i(\sigma_{-i}) = \{\sigma_i \in \Sigma_i \, : \, \sigma_i \in \arg\max_{\sigma_i \in \Sigma_i} U_i^e(\sigma_i, \sigma_{-i})\}.$$

Then, the correspondence $B : \Sigma \rightarrow \Sigma$ with

$$B(\sigma) = \bigtimes_{i \in N} B_i(\sigma_{-i})$$

defines the best response.

We will use Kakutani's fixed point theorem to show that $B(\cdot)$ has a fixed point, and therefore, the existence of an ex-ante equilibrium. To apply Kakutani's fixed point theorem, we need the following conditions:

i) Σ is non-empty, convex, and compact.

ii) $B : \Sigma \rightarrow \Sigma$ is a upper-hemicontinuous correspondence and $B(\sigma)$ is non-empty and closed $\forall \sigma \in \Sigma$.

iii) $B(\sigma)$ is convex $\forall \sigma \in \Sigma$.

We will show this conditions step by step:

i) Since $\mathcal{I}_i$ and $A_i(I_i^t)$ are finite for all information sets I_i^t and all player i, $\Delta(A_i(I_i^t))$ is non-empty, compact, and convex. Therefore, Σ_i and Σ are non-empty, compact, and convex as well.

ii) To show the second point, we use Berge's maximum theorem. Let $C : \Sigma_{-i} \rightarrow \Sigma_i$ be a correspondence such that $C(\sigma_{-i}) = \Sigma_i$ for all σ_{-i}. Then, C is upper and lower hemicontinuous as the following explanation shows and therefore continuous.

The definition of lower hemicontinuity says: C is lower hemicontinuous at a if for all open sets V intersecting $C(a)$ exists a neighbourhood U of a such that $C(x)$ intersects V for all $x \in U$. Since $C(a) = \Sigma_i = C(x)$ for all $x \in \Sigma_{-i}$, the definition is satisfied for each a, U, and V.

For upper hemicontinuity, we use the graph-theoretic characterization: Let

$$Gr(C) := \{(a, b) \in \Sigma_{-i} \times \Sigma_i \, : \, b \in C(a)\}.$$

If Σ_i is compact and $Gr(C)$ closed, $C : \Sigma_{-i} \rightarrow \Sigma_i$ is a upper hemicontinuous correspondence with closed domain and closed values. By the definition of C it follows that $Gr(C) = \Sigma_{-i} \times \Sigma_i$. Σ_i and Σ_{-i} are compact by i) and therefore closed. Hence, C is upper hemicontinuous.

Now, we can apply Berge's maximum theorem: With our notation

$$U_i^e : \Sigma_i \times \Sigma_{-i} \to \mathbb{R},$$
$$C : \Sigma_{-i} \to \Sigma_i \quad \text{s.t } C(\sigma_{-i}) = \Sigma,$$
$$C^*(\sigma_{-i}) \coloneqq \arg\max\{U_i^e(\sigma_i, \sigma_{-i}) \; : \; \sigma_i \in C(\sigma_{-i}) = \Sigma_i\} = B_i(\sigma_{-i}).$$

Berge's maximum theorem states that if U_i^e is jointly continuous in both arguments and C is continuous in σ_{-i}, then C^* is non-empty, convex valued, and upper hemicontinuous in σ_{-i}. Hence, by Lemma 1.2 B_i is a upper-hemicontinuous correspondence and $B_i(\sigma_{-i})$ is non-empty and closed $\forall \sigma_{-i} \in \Sigma_{-i}$. Since $B(\sigma) = \bigtimes_{i\in N} B_i(\sigma_{-i})$ the same holds for $B(\cdot)$.

iii) To show the convexity of $B(\sigma)$, we first show that $U_i^e(\cdot)$ is concave in σ_i. Let $\bar{\sigma}_i$ and $\tilde{\sigma}_i \in \Sigma_i$ and $\alpha \in [0,1]$. Then,

$$U_i^e(\alpha\bar{\sigma}_i + (1-\alpha)\tilde{\sigma}, \sigma_{-i}) = \min_{\phi\in\text{rect}(\Phi^0_{\sigma_{-i}})} \sum_{h\in H} u_i(h) p_{\alpha\bar{\sigma}_i+(1-\alpha)\tilde{\sigma}}(h|h^0)\phi(h)$$
$$= \min_{\phi\in\text{rect}(\Phi^0_{\sigma_{-i}})} \sum_{h\in H} u_i(h) \prod_{t=0}^{T} \Big(\alpha \underbrace{\bar{\sigma}_i(I_i(h^t))(h_{t,i})}_{\geq 0} + (1-\alpha) \underbrace{\tilde{\sigma}_i(I_i(h^t))(h_{t,i})}_{\geq 0}\Big)\phi(h)$$
$$\geq \min_{\phi\in\text{rect}(\Phi^0_{\sigma_{-i}})} \sum_{h\in H} u_i(h) \Big(\prod_{t=0}^{T} \alpha\bar{\sigma}_i(I_i(h^t))(h_{t,i}) + \prod_{t=0}^{T}(1-\alpha)\tilde{\sigma}_i(I_i(h^t))(h_{t,i})\Big)\phi(h)$$
$$= \min_{\phi\in\text{rect}(\Phi^0_{\sigma_{-i}})} \Big(\alpha \sum_{h\in H} u_i(h) p_{\bar{\sigma}_i}(h|h^0)\phi(h) + (1-\alpha)\sum_{h\in H} u_i(h) p_{\tilde{\sigma}_i}(h|h^0)\phi(h)\Big)$$
$$\geq \alpha \min_{\phi\in\text{rect}(\Phi^0_{\sigma_{-i}})} \sum_{h\in H} u_i(h) p_{\bar{\sigma}_i}(h|h^0)\phi(h) + (1-\alpha) \min_{\phi\in\text{rect}(\Phi^0_{\sigma_{-i}})} \sum_{h\in H} u_i(h) p_{\tilde{\sigma}_i}(h|h^0)\phi(h)$$
$$= \alpha U_i^e(\bar{\sigma}_i, \sigma_{-i}) + (1-\alpha) U_i^e(\tilde{\sigma}_i, \sigma_{-i}).$$

With the concavity of $U_i^e(\cdot)$ we can prove that $B_i(\sigma_{-i})$ is convex for all σ_{-i}. Fix some arbitrary σ_{-i}, let $\bar{\sigma}_i$, $\tilde{\sigma}_i \in B_i(\sigma_{-i})$, and $\alpha \in [0,1]$. We have to show, that $\alpha\bar{\sigma}_i + (1-\alpha)\tilde{\sigma}_i \in B_i(\sigma_{-i}) = \arg\max_{\sigma_i\in\Sigma_i} U_i^e(\sigma_i, \sigma_{-i})$. Since $\bar{\sigma}_i$, $\tilde{\sigma}_i \in B_i(\sigma_{-i})$ it follows that

$$\begin{aligned} U_i^e(\bar{\sigma}_i, \sigma_{-i}) = U_i^e(\tilde{\sigma}_i, \sigma_{-i}) &= \max_{\sigma_i\in\Sigma_i} U_i^e(\sigma_i, \sigma_{-i}) \\ &\geq U_i^e(\alpha\bar{\sigma}_i + (1-\alpha)\tilde{\sigma}_i, \sigma_{-i}) \\ &\geq \alpha U_i^e(\bar{\sigma}_i, \sigma_{-i}) + (1-\alpha) U_i^e(\tilde{\sigma}_i, \sigma_{-i}) = \max_{\sigma_i\in\Sigma_i} U_i^e(\sigma_i, \sigma_{-i}), \end{aligned}$$

where the last inequality follows from the concavity of U_i^e. Then,

$$U_i^e(\alpha\bar{\sigma}_i + (1-\alpha)\tilde{\sigma}_i, \sigma_{-i}) = \max_{\sigma_i\in\Sigma_i} U_i^e(\sigma_i, \sigma_{-i}).$$

Thus, we have

$$\alpha\bar{\sigma}_i + (1-\alpha)\tilde{\sigma}_i \in \arg\max_{\sigma_i \in \Sigma_i} U_i^e(\sigma_i, \sigma_{-i}) = B_i(\sigma_{-i}).$$

Hence, $B_i(\sigma_{-i})$ is convex valued for all $\sigma_{-i} \in \Sigma_{-i}$. Since this is true for all $i \in N$, it follows that $B(\sigma)$ is convex valued for all $\sigma \in \Sigma$.

Now, we can apply Kakutani's fixed point theorem, which shows that the best response correspondence B has a fixed point and, therefore, proves the existence of an ex-ante equilibrium with rectangular beliefs. □

1.6.1.2 Proofs: Relation of Ex-Ante and Interim Equilibria

Proof of Lemma 1.3. The proof consists of two steps:

i) First, we show that

$$\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i}(h|h^{t-1}) \phi^*(h) \geq \sum_{F_i^t \in \mathcal{F}_i^t} \phi^*(F_i^t) p_{\sigma_i}(h^t|h^{t-1}) \min_{\tilde{\phi} \in \text{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))} \sum_{h \in H} u_i(h) p_{\sigma_i}(h|h^t) \tilde{\phi}(h)$$

which follows directly from Bayesian updating. Since $\phi^*(F_i^t) = 0$ is equivalent to $\phi^*(h) = 0$ for all $h \in F_i^t$, we get

$$\begin{aligned}
\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i}(h|h^{t-1}) \phi^*(h) &= \sum_{F_i^t \in \mathcal{F}_i^t} \sum_{h \in F_i^t} u_i(h) p_{\sigma_i}(h|h^{t-1}) \phi^*(h) \\
&= \sum_{\substack{F_i^t \in \mathcal{F}_i^t \\ \text{s.t. } \phi^*(F_i^t) > 0}} \frac{\phi^*(F_i^t)}{\phi^*(F_i^t)} \sum_{h \in F_i^t} u_i(h) p_{\sigma_i}(h|h^{t-1}) \phi^*(h) \\
&= \sum_{F_i^t \in \mathcal{F}_i^t} \phi^*(F_i^t) \sum_{h \in F_i^t} u_i(h) p_{\sigma_i}(h|h^{t-1}) \underbrace{\frac{\phi^*(h)}{\phi^*(F_i^t)}}_{\in \text{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))} \\
&\geq \sum_{F_i^t \in \mathcal{F}_i^t} \phi^*(F_i^t) \min_{\tilde{\phi} \in \text{rect}(\Phi^1_{\sigma_{-i}}(F_i^t))} \sum_{h \in F_i^t} u_i(h) p_{\sigma_i}(h|h^{t-1}) \tilde{\phi}(h) \\
&= \sum_{F_i^t \in \mathcal{F}_i^t} \phi^*(F_i^t) p_{\sigma_i}(h^t|h^{t-1}) \min_{\tilde{\phi} \in \text{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))} \sum_{h \in F_i^t} u_i(h) p_{\sigma_i}(h|h^t) \tilde{\phi}(h),
\end{aligned}$$

where the last equality follows from $p_{\sigma_i}(h^t|h^{t-1}) = p_{\sigma_i}(\bar{h}^t|\bar{h}^{t-1})$ for all $h, \bar{h}$ in F_i^t.

ii) For the other direction we show that left-hand side is smaller or equal than the right-hand side, i.e.,

$$\sum_{h\in F_i^{t-1}} u_i(h)p_{\sigma_i}(h|h^{t-1})\phi^*(h) \leq \sum_{F_i^t\in\mathcal{F}_i^t} \phi^*(F_i^t)p_{\sigma_i}(h^t|h^{t-1}) \min_{\tilde{\phi}\in\text{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))} \sum_{h\in F_i^t} u_i(h)p_{\sigma_i}(h|h^t)\tilde{\phi}(h),$$

let $\tilde{\phi}^*_{F_i^t} \in \arg\min_{\tilde{\phi}\in\text{rect}(\Phi^t_{\sigma_{-i}}(F_i^t))}\sum_{h\in H} u_i(h)p_{\sigma_i}(h|h^t)\tilde{\phi}(h)$ for all F_i^t and ϕ^* as above. Rectangularity implies that there exists a $\phi' \in \Phi^{t-1}_{\sigma_{-i}}(F_i^{t-1})$ such that

$$\phi'(h) = \sum_{F_i^t\in\mathcal{F}_i^t} \phi^*(F_i^t)\tilde{\phi}^*_{F_i^t}(h)$$

which in general is not a worst-case belief. Then,

$$\begin{aligned}\sum_{h\in F_i^{t-1}} u_i(h)p_{\sigma_i}(h|h^{t-1})\phi^*(h) &\leq \sum_{h\in F_i^{t-1}} u_i(h)p_{\sigma_i}(h|h^{t-1})\phi'(h)\\ &= \sum_{h\in F_i^{t-1}} u_i(h)p_{\sigma_i}(h|h^{t-1}) \sum_{F_i^t\in\mathcal{F}_i^t}\phi^*(F_i^t)\tilde{\phi}^*(h)\\ &= \sum_{F_i^t\in\mathcal{F}_i^t}\phi^*(F_i^t)\sum_{h\in F_i^t} u_i(h)p_{\sigma_i}(h|h^{t-1})\tilde{\phi}^*(h)\\ &= \sum_{F_i^t\in\mathcal{F}_i^t}\phi^*(F_i^t)p_{\sigma_i}(h^t|h^{t-1}) \min_{\tilde{\phi}\in\Phi^t_{\sigma_{-i}}(F_i^t)} \sum_{h\in F_i^t} u_i(h)p_{\sigma_i}(h|h^t)\tilde{\phi}(h).\end{aligned}$$

Combining Step i) and ii) proves Lemma 1.3. □

Proof of Theorem 1.3 (ex-ante implies interim). We show that for an arbitrary t a completely mixed interim equilibrium with rectangular beliefs at $t-1$ implies an interim equilibrium with rectangular beliefs at t. Then, the theorem follows by iteration.

Let $(\sigma^*, \text{rect}(\Phi^{t-1}_{\sigma_{-i}})$ be a completely mixed interim equilibrium with rectangular beliefs at $t-1$ and assume that $(\sigma^*, \text{rect}(\Phi^t_{\sigma_{-i}}))$ is not an interim equilibrium at t. Hence, there exist a player i, an information set F_i^t, and a strategy profile $(\sigma_i', \sigma^*_{-i})$ such that player i deviates from σ^*. Let F_i^{t-1} be the information set that precedes F_i^t and ϕ^* the worst-case belief at F_i^{t-1} given σ^*_{-i}, i.e.,

$$\phi^* \in \mathop{\arg\min}_{\phi\in\text{rect}(\Phi^{t-1}_{\sigma^*_{-i}}(F_i^{t-1}))} \sum_{h\in F_i^{t-1}} u_i(h)p_{\sigma_i^*}(h|h^{t-1})\phi(h).$$

Similarly, let $\phi^{*,t}$ denote the worst-case belief at F_i^t given σ_i^* and $\bar{\phi}$ denote the worst-case at F_i^{t-1} belief given the strategy σ_i which equals σ_i' for $s \geq t$ and equals σ_i^* for

$s < t$. Furthermore, let ϕ' be the pasting of $\bar{\phi}$ and $\phi^{*,t}$. Then, ϕ' is in general not a worst-case belief. Similar to Step ii) of the proof of Lemma 1.3, it follows that

$$\begin{aligned} U_i^i(\sigma^*, \operatorname{rect}(\Phi^{t-1}(F_i^{t-1}))) &\leq \sum_{h \in H} u_i(h) p_{\sigma_i^*}(h|h^{t-1})\phi'(h) \\ &= \sum_{F_i^t \in \mathcal{F}_i^t} \bar{\phi}(F_i^t) p_{\sigma_i^*}(h^t|h^{t-1}) U_i^i(\sigma^*, \operatorname{rect}(\Phi^t_{\sigma^*_{-i}}(F_i^t))). \end{aligned} \tag{1.7}$$

By our assumption $(\sigma^*, \operatorname{rect}(\Phi^t_{\sigma^*_{-i}}))$ is not an interim equilibrium at t, i.e., there exist F_i^t and σ_i' such that

$$U_i^i(\sigma^*, \operatorname{rect}(\Phi^t_{\sigma^*_{-i}}(F_i^t))) < U_i^i((\sigma_i', \sigma^*_{-i}), \operatorname{rect}(\Phi^t_{\sigma^*_{-i}}(F_i^t))). \tag{1.8}$$

Furthermore, since σ^* is completely mixed and $\mathcal{P}$ has full support, it follows that

$$\bar{\phi}(F_i^t) p_{\sigma_i^*}(h^t|h^{t-1}) > 0 \tag{1.9}$$

for all F_i^t. Combining Equation (1.7), Equation (1.8), and Equation (1.9) leads to

$$\begin{aligned} U_i^i(\sigma^*, \operatorname{rect}(\Phi^{t-1}_{\sigma^*_{-i}}(F_i^{t-1}))) &< \sum_{F_i^t \in \mathcal{F}_i^t} \bar{\phi}(F_i^t) p_{\sigma_i^*}(h^t|h^{t-1}) U_i^i((\sigma_i', \sigma^*_{-i}), \operatorname{rect}(\Phi^t_{\sigma^*_{-i}}(F_i^t))) \\ &= \min_{\phi \in \operatorname{rect}(\Phi^{t-1}_{\sigma^*_{-i}})} \sum_{h \in H} u_i(h) p_{\sigma_i^*}(h^t|h^{t-1}) p_{\sigma_i'}(h|h^t)\phi(h). \end{aligned} \tag{1.10}$$

The last equality in Equation (1.10) follows by construction and Remark 1.4: The pasting of $\bar{\phi}$ and the interim worst-case belief at t given σ_i', is the worst-case belief at $t-1$ given the strategy σ_i which equals σ_i' for $s \geq t$ and equals σ_i^* for $s < t$. Then, Equation (1.10) follows from Lemma 1.3.

The calculation above forms a contradiction since $(\sigma^*, \operatorname{rect}(\Phi^{t-1}_{\sigma^*_{-i}}))$ is an interim equilibrium with rectangular beliefs at stage $t-1$. Hence, $(\sigma^*, \operatorname{rect}(\Phi^t_{\sigma^*_{-i}}))$ is an interim equilibrium with rectangular beliefs at stage t. □

1.6.1.3 Proofs: Sequential Rationality and Rectangularity

Proof of Theorem 1.5. First, we show that $\phi^*_{F_i^0}$ is the pasting of the worst-case beliefs of all interim stages. Assume that this is not the case, then there exists an information set F_i^t such that $\operatorname{Bay}(\phi^*_{F_i^0}|F_i^t) \neq \phi^*_{F_i^t}$. Furthermore, σ^* is sequentially rational with respect to $\Psi_{\sigma_{-i}}$. Hence, σ_i^* is an ex-ante best response given the belief $\phi^*_{F_i^0}$ and therefore an interim best response given the belief $\operatorname{Bay}(\phi^*_{F_i^0}|F_i^t)$ as well. This contradicts the assumption that there does not exist a belief $\phi_{F_i^t} \neq \phi^*_{F_i^t}$ such that σ_i^* is a best response given $\phi_{F_i^t}$. Hence, $\phi^*_{F_i^0}$ is the pasting of the worst-case beliefs of all interim stages and $\phi^*_{F_i^0} \in \Psi^0_{\sigma_{-i}}$. Therefore, σ_i^* is an ex-ante best response given $\operatorname{rect}(\Psi_{\sigma_{-i}})$. Then, rectangularity and the fact, that σ^* is sequentially rational given $\Psi_{\sigma_{-i}}$, implies that $(\sigma^*, \operatorname{rect}(\Psi_{\sigma_{-i}})$ is sequentially rational. □

1.6.2 Further Results

Additional to Theorem 1.3, we can show that an equilibrium at stage t implies an equilibrium at stage $t-1$ if a no-profitable one-stage-deviation property is satisfied.

Definition 1.12. *A tuple (σ^*, Ψ) of an strategy profile and a belief system satisfies the **no-profitable one-stage-deviation property** at stage t if for all F_i^t it holds that*

$$U_i^i(\sigma^*, \Psi^t(F_i^t)) \geq U_i^i((\sigma_i', \sigma_{-i}^*), \Psi^t(F_i^t))$$

for all σ_i' such that σ_i' equals σ_i^* everywhere except at F_i^t.

Theorem 1.6 (Interim implies ex-ante equilibria)**.** *Assume that Ψ is a belief system which is rectangular and that (σ^*, Ψ^s) satisfies the no-profitable one-stage-deviation property for all $s < t$. If (σ^*, Ψ^t) is an interim equilibrium with rectangular beliefs at stage t, then σ^* is an ex-ante equilibrium with rectangular beliefs.*

Similarly to Theorem 1.3, we prove Theorem 1.6 by showing that due to rectangularity an interim equilibrium at stage t implies an interim equilibrium at stage $t-1$. Then, the recursive structure implies the following corollary.

Corollary 1.2. *Assume that Ψ is a belief system which is rectangular and that (σ^*, Ψ^s) satisfies the no-profitable one-stage-deviation property for all $s < t$. If (σ^*, Ψ^t) is an interim equilibrium with rectangular beliefs at stage t, then (σ^*, Ψ^{t-1}) is an interim equilibrium with rectangular beliefs at stage $t-1$.*

Furthermore, Theorem 1.6 shows that rectangularity and the non-profitable one-stage-deviation property for all stages imply sequential rationality.

Corollary 1.3. *Let Ψ be a belief system which is stable under pasting and assume that (σ^*, Ψ^t) satisfies the no-profitable one-stage-deviation property for all $t \geq 0$. Then, (σ^*, Ψ) is sequentially rational.*

Proof. The result follows immediately from Theorem 1.6. The no-profitable one-stage-deviation property of (σ^*, Ψ^T) at the last stage and rectangularity imply that (σ^*, Ψ^T) is an interim equilibrium with rectangular beliefs at the last stage. Then, Theorem 1.6 implies that (σ^*, Ψ^t) is an interim equilibrium with rectangular beliefs at all stages t and σ^* is an ex-ante equilibrium with rectangular beliefs. □

As in the proof of Theorem 1.3 the relation of interim and ex-ante worst-case belief and, therefore, Lemma 1.3 and Remark 1.4 are essential to prove Theorem 1.6.

Proof of Theorem 1.6. We will prove that an interim equilibrium with rectangular beliefs at stage t implies an interim equilibrium with rectangular beliefs at stage $t-1$. Since this holds for arbitrary t, iteration proves the theorem.

Assume that (σ^*, Ψ^t) is an interim equilibrium with rectangular beliefs at stage t. We prove that an arbitrary player i has no incentive to deviate from σ^* at an arbitrary information set F_i^{t-1} if all other players $j \neq i$ play σ_j^*.
Fix some arbitrary F_i^t such that the probability of reaching F_i^t from F_i^{t-1} given σ_{-i}^* is positive.[16] Let $\phi^{*,t}$ denote the worst-case belief at F_i^t given σ_i^*. Furthermore, let $\bar{\sigma}_i$ denote a strategy which is equal to σ_i^* at all stages $s \geq t$ and equal to σ_i' at stage $t-1$. The worst-case belief at F_i^{t-1} given $\bar{\sigma}_i$ is denoted by $\bar{\phi}^{t-1}$. The worst-case belief at F_i^{t-1} given σ_i' is denoted by $\phi^{\prime,t-1}$.

Since σ_{-i}^* is fixed, the belief sets of player i are fixed as well. Furthermore, the worst-case belief at F_i^t depends only on the part of the strategy of player i which is chosen at stages $s \geq t$. Hence, the worst-case beliefs at F_i^t given σ_i^* and $\bar{\sigma}_i$ are the same by the definition of $\bar{\sigma}_i$. By Remark 1.4, we know that the Bayesian update of the worst-case belief at $t-1$ is the worst-case belief at t. This implies

$$\mathrm{Bay}(\bar{\phi}^{t-1}) = \phi^{*,t}. \tag{1.11}$$

The optimality of σ_i^* at F_i^t implies

$$\sum_{h \in F_i^t} u_i(h) p_{\sigma_i^*}(h|h^t) \phi^{*,t} \geq \sum_{h \in F_i^t} u_i(h) p_{\sigma_i'}(h|h^t) \phi^{*,t}. \tag{1.12}$$

Combining Equation (1.11) and Equation (1.12) gives us

$$\sum_{h \in F_i^t} u_i(h) p_{\sigma_i^*}(h|h^t) \frac{\bar{\phi}^{t-1}(h)}{\sum_{h \in F_i^t} \bar{\phi}^{t-1}(h)} \geq \sum_{h \in F_i^t} u_i(h) p_{\sigma_i'}(h|h^t) \frac{\bar{\phi}^{t-1}(h)}{\sum_{h \in F_i^t} \bar{\phi}^{t-1}(h)}.$$

Now, we can cancel the normalization terms of Bayes' rule on both sides and multiply each side with $p_{\sigma_i'}(h^t|h^{t-1})$. Then, replacing $\bar{\phi}^{t-1}$ with the worst-case belief at $t-1$ given σ_i' leads to

$$\begin{aligned}\sum_{h \in F_i^t} u_i(h) p_{\sigma_i^*}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \bar{\phi}^{t-1}(h) &\geq \sum_{h \in F_i^t} u_i(h) p_{\sigma_i'}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \bar{\phi}^{t-1}(h) \\ &\geq \sum_{h \in F_i^t} u_i(h) p_{\sigma_i'}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \phi^{\prime,t-1}(h).\end{aligned}$$

This holds for any F_i^t which is reachable from F_i^{t-1}. Hence, summation over all this F_i^t leads to

$$\begin{aligned}&\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i^*}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \bar{\phi}^{t-1}(h) \\ \geq &\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i'}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \phi^{\prime,t-1}(h).\end{aligned} \tag{1.13}$$

[16]For information sets with zero probability, it follows by the full support assumption on $\mathcal{P}$ that $\phi^{t-1}(h) = 0$ for all $h \in F_i^t$ for any $\phi^{t-1} \in \Phi_i^{t-1}(F_i^{t-1})$. Hence, the histories $h \in F_i^t$ do not influence the expected utility at F_i^{t-1}.

Furthermore, by the no-profitable one-stage-deviation property, it follows that

$$\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i^*}(h|h^{t-1}) \phi^{*,t-1}(h) \geq \sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i^*}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \bar{\phi}^{t-1}(h). \quad (1.14)$$

Combining Equation (1.13) and Equation (1.14) leads to

$$\sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i^*}(h|h^{t-1}) \phi^{*,t-1} \geq \sum_{h \in F_i^{t-1}} u_i(h) p_{\sigma_i'}(h|h^t) p_{\sigma_i'}(h^t|h^{t-1}) \phi'^{,t-1}(h)$$

which proves the optimality of σ_i^* at F_i^{t-1}. This holds for any arbitrary F_i^{t-1} and for any arbitrary player i. Hence, (σ^*, Ψ^{t-1}) is an interim equilibrium with rectangular beliefs at $t-1$. □

1.6.3 Examples

1.6.3.1 Calculations Example 1.1

Let a_L, a_R, and b_L and so on denote the probabilities that player 1 plays A or B at state L, R, and O, respectively. Similar, m, n, p, and o denote the probabilities that player 2 plays M, N, P, or O. The set of histories H, the information partition $\mathcal{F}_2^1$, and the ex-ante belief set $\Phi^0_{\sigma_{-2}}$ of player 2 are given by

$$\begin{aligned} H &= \{LBO, LBP, LAN, LAM, RBO, RBP, RAN, RAM, OBO, OBP, OAM, OAN\}, \\ \mathcal{F}_2^1 &= \big\{\{LAN, LAM, RAN, RAM, OAM, OAN\}, \\ &\qquad \{LBO, LBP, RBO, RBP, OBO, OBP\}\big\}, \\ \Phi^0_{\sigma_{-2}} &= \left\{ \left(\frac{lb_L}{2}, \frac{lb_L}{2}, \frac{la_L}{2}, \frac{la_L}{2,} \frac{rb_R}{2}, \frac{rb_R}{2}, \frac{ra_R}{2}, \frac{ra_R}{2}, \frac{ob_O}{2}, \frac{ob_O}{2}, \frac{oa_O}{2}, \frac{oa_O}{2} \right) : (l, r, o) \in \mathcal{P} \right\}. \end{aligned}$$

Moreover, let $F^1_{2,k}$ for $k = 1, 2$ denote the elements of the partition $\mathcal{F}_2^1$ in the same order as they are denoted above. We first look at the case where player 1 always plays A. Then, we compare it with the case where player 1 plays A if the state is L or R and B if the state is O.

- Player 1 plays σ_1, i.e., $a_L = a_R = a_O = 1$:
  The ex-ante belief set and marginal beliefs are given by

$$\begin{aligned} \Phi^0_{\sigma_1} &= \left\{ \left(0, 0, \frac{l}{2}, \frac{l}{2}, 0, 0, \frac{r}{2}, \frac{r}{2}, 0, 0, \frac{o}{2}, \frac{o}{2}\right) : (l, r, o) \in \mathcal{P} \right\}, \\ \phi(F^1_{2,1}) &= 1, \\ \phi(F^1_{2,2}) &= 0. \end{aligned}$$

  For information sets with positive marginal probability Bayes' rule is well defined and the Bayesian updates are given by

$$\text{Bay}(\Phi^0_{\sigma_1} | F^1_{2,1}) = \left\{ \left(0, 0, \frac{l}{2}, \frac{l}{2}, 0, 0, \frac{r}{2}, \frac{r}{2}, 0, 0, \frac{o}{2}, \frac{o}{2}\right) : (l, r, o) \in \mathcal{P} \right\}.$$

Then, $\text{rect}(\Phi^0_{\sigma_1}) = \Phi^0_{\sigma_1}$.

- Player 1 plays σ_1', i.e., $a_L = a_R = b_O = 1$:
The ex-ante belief set and the marginal beliefs are given by

$$\Phi^0_{\sigma_1'} = \left\{ \left(0, 0, \frac{l}{2}, \frac{l}{2}, 0, 0, \frac{r}{2}, \frac{r}{2}, \frac{o}{2}, \frac{o}{2}, 0, 0\right) : (l, r, o) \in \mathcal{P} \right\},$$
$$\phi(F^1_{2,1}) = l + r,$$
$$\phi(F^1_{2,2}) = o.$$

The Bayesian updates for information sets with positive marginal probability are

$$\text{Bay}(\Phi^0_{\sigma_1'} | F^1_{2,1})$$
$$= \left\{ \left(0, 0, \frac{l}{2(l+r)}, \frac{l}{2(l+r)}, 0, 0, \frac{r}{2(l+r)}, \frac{r}{2(l+r)}, 0, 0, 0, 0\right) : (l, r, o) \in \mathcal{P} \right\},$$
$$\text{Bay}(\Phi^0_A | F^1_{2,3}) = \left(0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0\right).$$

The rectangular hull $\text{rect}(\Phi^0_{\sigma_1'})$ is

$$\text{rect}(\Phi^0_{\sigma_1'})$$
$$= \text{conv}\left\{ \left(0, 0, \frac{\epsilon}{2}, \frac{\epsilon}{2}, 0, 0, \frac{(1-\epsilon)}{2}, \frac{(1-\epsilon)}{2}, 0, 0, 0, 0\right), \right.$$
$$\left(0, 0, 0, 0, 0, 0, \frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0\right), \left(0, 0, 0, 0, 0, 0, \frac{(1-\epsilon)}{2}, \frac{(1-\epsilon)}{2}, \frac{\epsilon}{2}, \frac{\epsilon}{2}, 0, 0\right),$$
$$\left. \left(0, 0, \frac{\epsilon(1-\epsilon)}{2}, \frac{\epsilon(1-\epsilon)}{2}, 0, 0, \frac{(1-\epsilon)^2}{2}, \frac{(1-\epsilon)^2}{2}, \frac{\epsilon}{2}, \frac{\epsilon}{2}, 0, 0\right) \right\}.$$

1.6.3.2 Calculations Example 1.2

Denote the history that the state is I, country A chooses war, and country C punishes A by IWP_A. The other histories are denoted similarly. First, we construct the belief sets $\Phi^0_{\sigma_{-i}}$ and show that they are rectangular. Then, we show that for the game with the restricted strategy set Σ^k, there exists an ex-ante and sequential equilibrium such that no player chooses war.

Belief Sets and Rectangularity The information filtration $\mathcal{F}^t_i$ and the ex-ante beliefs $\Phi^0_{\sigma_{-i}}$ are given by

$$H = \{IWP_A, IWN, IWP_B, IPWP_A, IPWN, IPWP_B, IPPF_A, IPPF_B,$$
$$IIPPF_B, IIPPF_A, IIPWP_B, IIPWN, IIPWP_A, IIWP_B, IIWN, IIWP_A\},$$
$$\mathcal{F}^0 = H,$$

$$
\begin{aligned}
\mathcal{F}_A^1 &= H,\\
\mathcal{F}_B^1 &= \big\{\{IWP_A, IWM, IWP_B, IIWP_B, IIWN, IIWP_A\},\\
&\quad \{IPWP_A, IPWN, IPWP_B, IPPF_A, IPPF_B, IIPPF_B,\\
&\quad IIPPF_A, IIPWP_B, IIPWN, IIPWP_A\}\big\},\\
\mathcal{F}_C^1 &= \big\{\{IWP_A, IWN, IWP_B, IPWP_A, IPWN, IPWP_B\},\\
&\quad \{IPPF_A, IPPF_B\}, \{IIPPF_B, IIPPF_A\},\\
&\quad \{IIPWP_B, IIPWN, IIPWP_A, IIWP_B, IIWN, IIWP_A\},
\end{aligned}
$$

and

$$
\begin{aligned}
\Phi^0_{\sigma_{-A}} = \Bigg\{ \Bigg(& \frac{\mu p_A^I}{2}, \frac{\mu n^I}{2}, \frac{\mu p_B^I}{2}, \frac{\mu\beta p_A^I}{2}, \frac{\mu\beta n^I}{2}, \frac{\mu\beta p_B^I}{2}, \frac{\mu(1-\beta) f_A^I}{2}, \frac{\mu(1-\beta) f_B^I}{2},\\
& \frac{(1-\mu)(1-\beta) f_B^{II}}{2}, \frac{(1-\mu)(1-\beta) f_A^{II}}{2}, \frac{(1-\mu)\beta p_B^{II}}{2}, \frac{(1-\mu)\beta n^{II}}{2},\\
& \frac{(1-\mu)\beta p_A^{II}}{2}, \frac{(1-\mu) p_B^{II}}{2}, \frac{(1-\mu) n^{II}}{2}, \frac{(1-\mu) p_A^{II}}{2} \Bigg) \; : \; \mu \in [\underline{\mu}, \bar{\mu}] \Bigg\},\\
\Phi^0_{\sigma_{-B}} = \Bigg\{ \Bigg(& \mu\alpha p_A^I, \mu\alpha n^I, \mu\alpha p_B^I, \frac{\mu(1-\alpha) p_A^I}{2}, \frac{\mu(1-\alpha) n^I}{2}, \frac{\mu(1-\alpha) p_B^I}{2}, \frac{\mu(1-\alpha) f_A^I}{2},\\
& \frac{\mu(1-\alpha) f_B^I}{2}, \frac{(1-\mu)(1-\alpha) f_B^{II}}{2}, \frac{(1-\mu)(1-\alpha) f_A^{II}}{2}, \frac{(1-\mu)(1-\alpha) p_B^{II}}{2},\\
& \frac{(1-\mu)(1-\alpha) n^{II}}{2}, \frac{(1-\mu)(1-\alpha) p_A^{II}}{2}, (1-\mu)\alpha p_B^{II}, (1-\mu)\alpha n^{II},\\
& (1-\mu)\alpha p_A^{II} \Bigg) \; : \; \mu \in [\underline{\mu}, \bar{\mu}] \Bigg\},\\
\Phi^0_{\sigma_{-C}} = \Bigg\{ \Bigg(& \frac{\mu\alpha}{3}, \frac{\mu\alpha}{3}, \frac{\mu\alpha}{3}, \frac{\mu(1-\alpha)\beta}{3}, \frac{\mu(1-\alpha)\beta}{3}, \frac{\mu(1-\alpha)\beta}{3}, \frac{\mu(1-\alpha)(1-\beta)}{2}\\
& \frac{\mu(1-\alpha)(1-\beta)}{2}, \frac{(1-\mu)(1-\alpha)(1-\beta)}{2}, \frac{(1-\mu)(1-\alpha)(1-\beta)}{2}\\
& \frac{(1-\mu)(1-\alpha)\beta}{3}, \frac{(1-\mu)(1-\alpha)\beta}{3}, \frac{(1-\mu)(1-\alpha)\beta}{3}, \frac{(1-\mu)\alpha}{3}, \frac{(1-\mu)\alpha}{3}\\
& \frac{(1-\mu)\alpha}{3} \Bigg) \; : \; \mu \in [\underline{\mu}, \bar{\mu}] \Bigg\}.
\end{aligned}
$$

We denote with $F^1_{B,1}$ and $F^1_{B,2}$ the first and second element of $\mathcal{F}^1_B$. Similarly, $F^1_{C,1}$, $F^1_{C,2}$, $F^1_{C,3}$ and $F^1_{C,4}$ denote the elements of $\mathcal{F}^1_C$. It is easy to verify that $\Phi^0_{\sigma_{-i}}$ is rectangular for $i = A, B, C$.

The marginal belief of player A is $\phi(F_A^1) = 1$. Updating $\Phi^0_{\sigma_{-A}}$ prior-by-prior leads to

$$\Phi^1_{\sigma_{-A}} = \Big\{\Big(\frac{\mu p_A^I}{2}, \frac{\mu n^I}{2}, \frac{\mu p_B^I}{2}, \frac{\mu\beta p_A^I}{2}, \frac{\mu\beta n^I}{2}, \frac{\mu\beta p_B^I}{2}, \frac{\mu(1-\beta) f_A^I}{2}, \frac{\mu(1-\beta) f_B^I}{2}, \frac{(1-\mu)(1-\beta) f_A^{II}}{2}, \frac{(1-\mu)(1-\beta) f_B^{II}}{2}, \frac{(1-\mu)\beta p_B^{II}}{2}, \frac{(1-\mu)\beta n^{II}}{2}, \frac{(1-\mu)\beta p_A^{II}}{2}, \frac{(1-\mu) p_B^{II}}{2}, \frac{(1-\mu) n^{II}}{2}, \frac{(1-\mu) p_A^{II}}{2}\Big) \,:\, \mu \in [\underline{\mu}, \bar{\mu}]\Big\}.$$

Then, the pasting of marginal and updated beliefs shows that $\mathrm{rect}(\Phi^0_{\sigma_{-A}}) = \Phi^0_{\sigma_{-A}}$.

For country B, we have to differ between the information sets $F^1_{B,1}$ and $F^1_{B,2}$. Marginals and updated beliefs are given by

$$\phi(F^1_{B,1}) = \alpha,$$
$$\phi(F^1_{B,2}) = 1 - \alpha,$$

$$\Phi^1_{\sigma_{-B}}(F^1_{B,1}) = \Big\{\Big(\mu p_A^I, \mu n^I, \mu p_B^I, 0,0,0,0,0,0,0,0,0,0, (1-\mu) p_A^I, (1-\mu) n^I, (1-\mu) p_B^I\Big) \,:\, \mu \in [\underline{\mu}, \bar{\mu}]\Big\},$$
$$\Phi^1_{\sigma_{-B}}(F^2_{B,2}) = \Big\{\Big(0,0,0, \frac{\mu p_A^I}{2}, \frac{\mu n^I}{2}, \frac{\mu p_B^I}{2}, \frac{\mu f_A^I}{2}, \frac{\mu f_B^I}{2}, \frac{(1-\mu) f_B^{II}}{2}, \frac{(1-\mu) f_A^{II}}{2}, \frac{(1-\mu) p_B^{II}}{2}, \frac{(1-\mu) n^{II}}{2}, \frac{(1-\mu) p_A^{II}}{2}, 0,0,0\Big) \,:\, \mu \in [\underline{\mu}, \bar{\mu}]\Big\}.$$

The pasting of marginal and updated beliefs shows that $\mathrm{rect}(\Phi^0_{\sigma_{-B}}) = \Phi^0_{\sigma_{-B}}$.

Country C has four information sets. The marginal and updated beliefs are

$$\phi(F^1_{C,1}) = \mu(\alpha + (1-\alpha)\beta),$$
$$\phi(F^1_{C,2}) = \mu(1-\alpha)(1-\beta),$$
$$\phi(F^1_{C,3}) = (1-\mu)(1-\alpha)(1-\beta),$$
$$\phi(F^1_{C,4}) = (1-\mu)(\alpha + (1-\alpha)\beta),$$

and

$$\Phi^1_{\sigma_{-C}}(F^1_{C,1}) = \Big\{\Big(\frac{\alpha}{3(\alpha+(1-\alpha)\beta)}, \frac{\alpha}{3(\alpha+(1-\alpha)\beta)}, \frac{\alpha}{3(\alpha+(1-\alpha)\beta)}, \frac{(1-\alpha)\beta}{3(\alpha+(1-\alpha)\beta)}, \frac{(1-\alpha)\beta}{3(\alpha+(1-\alpha)\beta)}, \frac{(1-\alpha)\beta}{3(\alpha+(1-\alpha)\beta)}, 0,0,0,0,0,0,0,0,0,0\Big) \,:\, \mu \in [\underline{\mu}, \bar{\mu}]\Big\},$$
$$\Phi^1_{\sigma_{-C}}(F^1_{C,2}) = \Big\{\Big(0,0,0,0,0,0, \frac{1}{2}, \frac{1}{2}, 0,0,0,0,0,0,0,0\Big)\Big\},$$
$$\Phi^1_{\sigma_{-C}}(F^1_{C,3}) = \Big\{\Big(0,0,0,0,0,0,0,0, \frac{1}{2}, \frac{1}{2}, 0,0,0,0,0,0\Big)\Big\},$$

$$\Phi^1_{\sigma_{-C}}(F^1_{C,3}) = \Big\{\Big(0,0,0,0,0,0,0,0,0,0, \frac{(1-\alpha)\beta}{3(\alpha+(1-\alpha)\beta)}, \frac{(1-\alpha)\beta}{3(\alpha+(1-\alpha)\beta)}, \frac{(1-\alpha)\beta}{3(\alpha+(1-\alpha)\beta)}, \frac{\alpha}{3(\alpha+(1-\alpha)\beta)}, \frac{\alpha}{3(\alpha+(1-\alpha)\beta)}, \frac{\alpha}{3(\alpha+(1-\alpha)\beta)}\Big) : \mu \in [\underline{\mu}, \bar{\mu}]\Big\}.$$

The pasting of marginal and updated beliefs leads to $\text{rect}(\Phi^0_{\sigma_{-C}}) = \Phi^0_{\sigma_{-C}}$. Hence, $\Phi^0_{\sigma_{-A}}$, $\Phi^0_{\sigma_{-B}}$, and $\Phi^0_{\sigma_{-C}}$ are rectangular.

Sequential Equilibria From the analysis in Example 1.2, we know that σ^* with

$$\alpha^* = \beta^* = 0, \quad f_A^{I,*} = f_A^{II,*} = \frac{1}{2}, \quad p_A^{I,*} = p_B^{II,*} = 1$$

is an ex-ante equilibrium if $[0.45, 0.55] \subset [\underline{\mu}, \bar{\mu}]$. To show that σ^* is an sequential equilibrium if $x \leq 0.5$, we have to find a sequence of completely mixed strategy profiles that converges to σ^*, such that sequential rationality and consistency w.r.t σ^* are satisfied. Let $\Sigma_A^k = [\underline{\alpha}^k, 1-\underline{\alpha}^k]$ and $\Sigma_B^k = [\underline{\beta}^k, 1-\underline{\beta}^k]$ be strategy sets of country A and B such that $\underline{\alpha}^k, \underline{\beta}^k \to 0$ if $k \to \infty$. Furthermore, assume that $\underline{\alpha}^k = (1-\underline{\alpha}^k)\underline{\beta}^k$. For country C, we define Σ_C^k such that $p_A^I, p_A^{II} \in [\epsilon^k, 1-\epsilon^k]$ with $\epsilon^k \to 0$ if $k \to \infty$. Now, we show that for all k the strategy profile σ^k with

$$\alpha^k = \underline{\alpha}^k, \quad \beta^k = \underline{\beta}^k, \quad f_A^{I,k} = f_A^{II,k} = \frac{1}{2}, \quad p_A^{I,k} = p_A^{II,k} = 1-\epsilon^k$$

is an ex-ante equilibrium and σ^k together with the belief system Φ^k constructed using Bayes' rule is an interim equilibrium of the game with the restricted strategy sets Σ_i^k specified above. Similarly to above, one can show that A, B, and C have no incentive to deviate from σ^k at the ex-ante stage. At the interim stage, the maximization problem of country A does not change and $\underline{\alpha}^k$ is still optimal. The expected interim payoff of country B from playing war is given by

$$\min_{\mu \in [\underline{\mu}, \bar{\mu}]} 10\frac{\mu p_A^I}{2} + (1-\mu)\cdot 0 = 5\underline{\mu}.$$

The expected interim payoff from playing peace is

$$\min_{\mu \in [\underline{\mu}, \bar{\mu}]} \frac{\mu}{2}(5f_A^I + 4f_B^I) + \frac{1-\mu}{2}(5f_A^{II} + 4f_B^{II}) = \frac{9}{4}.$$

Since $\underline{\mu} < 0.45$, it is optimal for B to play $\beta^k = \underline{\beta}^k$. Furthermore, since $\underline{\alpha}^k$ and $\underline{\beta}^k$ are such that $\underline{\alpha}^k = (1-\underline{\alpha}^k)\underline{\beta}^k$, the sets of updated beliefs of C after observing war are

$$\Phi^1_{\sigma_{-C}}(F^1_{C,1}) = \left\{\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 0,0,0,0,0,0,0,0,0,0\right)\right\},$$

$$\Phi^1_{\sigma_{-C}}(F^1_{C,4}) = \left\{\left(0,0,0,0,0,0,0,0,0,0,\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\right)\right\}.$$

Therefore, country C is indifferent between punishing A or B if $x < 0.5$ and indifferent between all three actions if $x = 0.5$. By the definition of σ^k, it follows that σ^k converges to σ^*. This shows sequential rationality. Finally, the belief system Φ^k satisfies consistency w.r.t. σ^* by construction. Hence, σ^* is part of a sequential equilibrium.

If $x > 0.5$, country C strictly prefers to play n with probability one. Therefore σ^* is not interim optimal for C and does not form a sequential equilibrium.

Equilibria without Ambiguity In this part, we show that without ambiguity there is no ex-ante equilibrium in which countries A and B play peace.

The expected payoff of country A of playing war with probability α is

$$\begin{aligned}
&\mu\big[6n^I\alpha + 10p^I_B\alpha + 6\beta n^I(1-\alpha) + 10\beta p^I_B(1-\alpha) + (1-\alpha)(1-\beta)(5f^I_A + 4f^I_B)\big]\\
&\quad + (1-\mu)\Big[6n^{II}\alpha + 10p^{II}_B\alpha + 6\beta n^{II}(1-\alpha) + 10\beta p^{II}_B(1-\alpha)\Big]\\
&\quad + (1-\mu)(1-\alpha)(1-\beta)(5f^{II}_A + 4f^{II}_B)\\
=&\alpha(1-\beta)\Big[\mu(6n^I + 10p^I_B) + (1-\mu)(6n^{II} + 10p^{II}_B)\Big]\\
&\quad - \alpha(1-\beta)\Big[\mu(5f^I_A + 4f^I_B) + (1-\mu)(5f^{II}_A + 4f^{II}_B)\Big] + T,
\end{aligned}$$

where T is independent of the strategy of A. Hence, maximizing the expected payoff of country A leads to the following best response α^*:

$$\alpha^* = \begin{cases} 1 & \text{if } \mu(6n^I + 10p^I_B) + (1-\mu)(6n^{II} + 10p^{II}_B) > \mu(5f^I_A + 4f^I_B) + (1-\mu)(5f^{II}_A + 4f^{II}_B),\\ [0,1] & \text{if } \mu(6n^I + 10p^I_B) + (1-\mu)(6n^{II} + 10p^{II}_B) = \mu(5f^I_A + 4f^I_B) + (1-\mu)(5f^{II}_A + 4f^{II}_B),\\ 0 & \text{if } \mu(6n^I + 10p^I_B) + (1-\mu)(6n^{II} + 10p^{II}_B) < \mu(5f^I_A + 4f^I_B) + (1-\mu)(5f^{II}_A + 4f^{II}_B), \end{cases}$$

if $\beta < 1$ and $\alpha^* \in [0,1]$ if $\beta = 1$. Similarly, one can calculate the best response of country B:

$$\beta^* = \begin{cases} 1 & \text{if } \mu(6n^I + 10p^I_A) + (1-\mu)(6n^{II} + 10p^{II}_A) > \mu(5f^I_B + 4f^I_A) + (1-\mu)(5f^{II}_B + 4f^{II}_A),\\ [0,1] & \text{if } \mu(6n^I + 10p^I_A) + (1-\mu)(6n^{II} + 10p^{II}_A) = \mu(5f^I_B + 4f^I_A) + (1-\mu)(5f^{II}_B + 4f^{II}_A),\\ 0 & \text{if } \mu(6n^I + 10p^I_A) + (1-\mu)(6n^{II} + 10p^{II}_A) < \mu(5f^I_B + 4f^I_A) + (1-\mu)(5f^{II}_B + 4f^{II}_A), \end{cases}$$

if $\alpha < 1$ and $\beta^* \in [0,1]$ if $\alpha = 1$. The best response of country C depends on the payoff x. We distinguish between the following three cases:

- If $x > 1$, playing $n^{I,*} = n^{II,*} = 1$ is a dominant strategy.
- If $x < 0.5$, country C will never play n. The probability of playing p_A or p_B

depends on α and β. In particular,

$$p_A^{I,*}, p_A^{II,*} = \begin{cases} 1 & \text{if } \alpha > (1-\alpha)\beta, \\ [0,1] & \text{if } \alpha = (1-\alpha)\beta, \\ 0 & \text{if } \alpha < (1-\alpha)\beta. \end{cases}$$

- If $x \in [0.5, 1]$, the best response depends on the relation between α, β and x. If $\alpha > (1-\alpha)\beta$, the probability of punishing B is zero, i.e., $p_B^{I,*} = p_B^{II,*} = 0$. Furthermore,

$$p_A^{I,*}, p_A^{II,*} = \begin{cases} 0 & \text{if } \alpha < x, \\ [0,1] & \text{if } \alpha = x, \\ 1 & \text{if } \alpha > x, \end{cases}$$

and $n^{I,*} = 1 - p_A^{I,*}, n^{II,*} = 1 - p_A^{II,*}$. Similarly, if $\alpha < (1-\alpha)\beta$ the probability of punishing A is zero, i.e., $p_A^{I,*} = p_A^{II,*} = 0$,

$$p_B^{I,*}, p_B^{II,*} = \begin{cases} 0 & \text{if } (1-\alpha)\beta < x, \\ [0,1] & \text{if } (1-\alpha)\beta = x, \\ 1 & \text{if } (1-\alpha)\beta > x, \end{cases}$$

and $n^{I,*} = 1 - p_B^{I,*}, n^{II,*} = 1 - p_B^{II,*}$.

Using the best responses, one can show that the following equilibria occur depending on the payoff x:

- $x > 1$:
  In any equilibrium, it has to hold that $n^{I,*} = n^{II,*} = 1$ and either $\alpha^* = 1$ and $\beta^* \in [0,1]$ or $\alpha^* \in [0,1]$ and $\beta^* = 1$.

- $x < 0.5$:
  In any equilibrium, the strategies of A and B are $\alpha^* = \frac{1}{2}$ and $\beta^* = 1$. The strategy of C has to satisfy $n^{I,*} = n^{II,*} = 0$ and

$$10(\mu p_A^I + (1-\mu) p_A^{II}) \geq 4 + \mu f_B^I + (1-\mu) f_B^{II}. \tag{1.15}$$

- $x \in [0.5, 1]$:
  There are two types of equilibria: Either $\alpha^* = \frac{1}{2}$, $\beta^* = 1$ and $n^{I,*} = n^{II,*} = 1$ or $\alpha^* = x$, $\beta^* = 1, n^{I,*} = n^{II,*} = 1$.

This shows that in all equilibria, at least one small country plays war with probability one. Hence, there does not exist an equilibrium in which the peace agreement is successful. Please note that we did not specify conditions on f_A^I, f_B^I, f_A^{II} and f_B^{II} since C is always indifferent between f_A^I and f_B^I or f_A^{II} and f_B^{II}.

Chapter 2

Dynamic Consistency in Ambiguous Persuasion

2.1 Introduction

The standard Bayesian persuasion literature analyzes the communication of a Sender and a Receiver. The Sender tries to persuade the Receiver by designing a communication device. Depending on an unknown state, the communication devices generate a signal realization. Given this signal realization, the Receiver chooses an action that influences the payoff of Sender and Receiver. This setting without ambiguity was firstly studied by Kamenica and Gentzkow (2011).

Beauchêne et al. (2019), henceforth abbreviated by BLL, introduce ambiguity in a standard Bayesian persuasion setting and characterize conditions under which the Sender can gain from ambiguous communication. BLL deal with the problem of dynamically inconsistent behavior by restricting their analysis to interim equilibria. They introduce ambiguity in the standard Bayesian persuasion setting of Kamenica and Gentzkow (2011) by allowing the Sender to choose a set of communication devices. Each communication device can generate a signal that reveals information about an unknown (risky) state $\omega \in \Omega$. Sender and Receiver only observe the signal realization without knowing which communication device generated the signal realization. Therefore, ambiguity about the communication device induces ambiguity about the risky state ω. However, they claim that there is no gain of ambiguous persuasion compared to Bayesian persuasion if the players behave dynamically consistently.[1]

In this chapter, we first show that we can restrict without loss of generality to messages that produce recommended actions or synonyms of recommended actions. A synonym

[1]See Proposition 5 of Beauchêne et al. (2019).

m' of a message m is a message that induces the same posterior belief or best response of the Receiver as the message m. This result generalizes the well-known Proposition 1 by Kamenica and Gentzkow (2011), which states that one can restrict without loss of generality to so-called straightforward signals to the ambiguous persuasion setting.

Then, we define beliefs over a more general state space of straightforward messages and states. The more general state space allows for rectangular ambiguous beliefs. These beliefs take the dependence of the ambiguous signal and the ex-ante risky state into account and allow for a non-singleton ex-ante belief set. Given these beliefs, the optimal interim strategy of the Receiver in BLL is ex-ante optimal and, therefore, dynamically consistent. Hence, ambiguous persuasion can generate a higher value for the Sender even under dynamically consistent behavior.

This chapter is organized as follows: First, we discuss the related literature. In Section 2.2, we formulate the ambiguous persuasion model and give an example that illustrates the gain of an ambiguous strategy and the dynamically inconsistent behavior. Section 2.2.2 generalizes Proposition 1 of Kamenica and Gentzkow (2011) and defines rectangular beliefs. Section 2.3 defines perfect Bayesian equilibria under rectangular beliefs and generalizes the results of BLL. Furthermore, in Section 2.4 we discuss the value of information under ambiguous persuasion. Finally, Section 2.5 concludes and discusses related literature in more detail.

Related Literature As in Chapter 1 we follow the approach of Epstein and Schneider (2003) and Riedel et al. (2018) and define rectangularity in the ambiguous persuasion setting. However, the results of Chapter 1 cannot be applied straightforwardly in the ambiguous persuasion setting. In ambiguous persuasion, ambiguity arises due to an ambiguous communication device of the Sender. Hence, the strategy of the Sender is ambiguous. In Chapter 1, we only consider ambiguity about the state space and not about the opponents' strategy. However, restricting to straightforward signals, we can define rectangular beliefs on a more general state space in the setting of BLL. The relation to the consistent planning approach of Siniscalchi (2011) and the updating rules by Hanany and Klibanoff (2007) are discussed in Section 2.5.

The literature on ambiguous communication or information is still relatively small. Kellner and Le Quement (2018) introduce ambiguity in a cheap talk setting by allowing the Sender to commit his signal on an ambiguous payoff-irrelevant state. They show that ambiguity may lead to a pareto improvement compared to non-ambiguous cheap talk. Cheng (2020) analyzes the ambiguous persuasion setting of BLL from an ex-ante perspective. He shows that if Sender and Receiver use the updating rule of Hanany and Klibanoff (2007), the Sender cannot benefit from ambiguous persuasion.

Furthermore, there is increasing literature on the value of information under ambiguity. Hill (2020) defines the value of ambiguous information in a decision-theoretical model which uses so-called subjective trees. Li (2020) studies the relation of ambiguity aversion and an aversion of (partial) information. Kops and Pasichnichenko (2020) experimentally show that for a specific decision problem, a majority of the subjects have a negative value of ambiguous information, which is correlated with ambiguity aversion. However, in a different experiment Ortoleva and Shishkin (2020) find an opposite result. They do not find any evidence for a negative value of information for ambiguity avers agents. In Section 2.5, we discuss the relation to Hill (2020) and Li (2020) in more detail.

2.2 Model

Except for the belief formation process, we follow the model of BLL. Let us first summarize their model.

2.2.1 Model of BLL

The persuasion game consists of two players, a Sender (he) and a Receiver (she). The utility of both players depends on the state of the world $\omega \in \Omega$ and an action $a \in A$ chosen by the Receiver. We denote with $u(a, \omega)$ and $\nu(a, \omega)$ the utilities of the Receiver and the Sender, respectively. Ω and A are compact subsets of the Euclidean space. Ex-ante, the state ω is unknown, and both players have the a common prior state belief $p_0 \in \Delta\Omega$, where $\Delta\Omega$ denotes the set of all distribution functions on Ω. Thus, ex-ante there exists no ambiguity about the state.[2] The Sender tries to persuade the Receiver by choosing a signal that reveals information about the state. A signal consists of a finite set of signal realizations or messages M and a set of communication devices $\Pi = \{\pi_k\}_{k \in K}$.[3] Each communication device is a distribution over the set of messages M for each $\omega \in \Omega$, i.e., $\pi_k(\cdot|\omega) \in \Delta M$ for all $\omega \in \Omega$. Again, ΔM denotes the set of all distribution functions on M. As in BLL, we assume that the π_k's have common support for all $k \in K$. The only difference to the standard Bayesian persuasion setting is that the Sender chooses a set of communication devices instead of one communication device. It is ambiguous to both players, which of the communication devices generates the observed message. After observing a message m, the Receiver updates her prior state belief using Bayes' rule. Since she does not know which communication device

[2] Our definition of belief differs from the one of BLL. To avoid confusion, we use the term state belief whenever we refer to beliefs in the sense of BLL.

[3] Please note that we deviate from the model of BLL by defining Π as the set of communication devices. BLL define Π as the convex hull of the set of communication devices. Since Sender and Receiver have maxmin preferences, the minimization problems over $\{\pi_k\}$ or $\text{co}\big(\{\pi_k\}\big)$ coincide.

generated the message, she updates p_0 with respect to each communication device π_k, which leads to the following set of posterior state beliefs after observing the message $m \in M$:

$$P_m = \left\{ p_m^{\pi_k}(\cdot) \in \Delta\Omega \,:\, p_m^{\pi_k}(\cdot) = \frac{p_0(\cdot)\pi_k(m|\cdot)}{\int_\Omega p_0(\omega)\pi_k(m|\omega)\,\mathrm{d}\omega}, \pi_k \in \Pi \right\}.$$

Sender and Receiver have maxmin preferences á la Gilboa and Schmeidler (1989), i.e., they maximize their worst-case expected utility. BLL assume that the Receiver maximizes her interim worst-case expected utility given that message m was observed. Hence, for all $m \in M$ the expected utility is given by

$$U(a, P_m) = \min_{p_m \in P_m} \mathbb{E}_{p_m}(u(a, \omega)).$$

As usual in the persuasion literature, we assume that the Receiver chooses the sender-preferred action if she has multiple maximizers. We denote with $\hat{a}_m$ the (sender-preferred) best response of the Receiver after observing the message m. The Sender chooses the signal (M, Π) that maximizes his ex-ante worst-case expected utility

$$\sup_{(M,\Pi)} \min_{\pi \in \Pi} \mathbb{E}_{p_0} \Big[\mathbb{E}_\pi \big[\nu(\hat{a}_m, \omega) | \omega \big] \Big].$$

Since the Sender only chooses an action at the ex-ante stage, he can never behave dynamically inconsistently. However, the interim best response of the Receiver is, in general, not ex-ante optimal. Intuitively, ex-ante, the Receiver can hedge against ambiguity by playing any constant strategy. The following example from BLL shows that ambiguity can lead to a higher expected payoff for the Sender. Furthermore, we show that the interim equilibrium strategy of the Receiver is not ex-ante ante optimal.

Example 2.1. *Assume that the Sender is a brand-name drug producer. The Receiver is a physician who can choose between prescribing the brand name drug ($a = a_B$) or a generic competitor ($a = a_G$). The Sender always prefers that the Receiver prescribes the brand name drug. The Receiver's preferences depend on the state, which reflects the effectiveness of the generic drug. If the generic drug is effective ($\omega = \omega_e$), the Receiver prefers the generic drug. If not ($\omega = \omega_i$), she prefers the brand name drug. The payoffs of Sender and Receiver are given in Table 2.1.*

	ω_e	ω_i
a_B	$(1, 2)$	$(1, 2)$
a_G	$(0, 3)$	$(0, -1)$

Table 2.1: Payoffs (S, R)

Sender and Receiver have a common ex-ante state belief $p_0 = \mathbb{P}(\omega = \omega_i) < \frac{1}{4}$.[4] *BLL show that the optimal Bayesian persuasion signal is such that the set of messages* M *consists of two messages* $M = \{i, e\}$ *and the communication device is given by*

$$\pi(e|\omega_e) = \frac{1-4p_0}{1-p_0} = 1 - \pi(i|\omega_e),$$
$$\pi(e|\omega_i) = 0 = 1 - \pi(i|\omega_i).$$

Then, the ex-ante expected payoff of the Sender given the optimal Bayesian persuasion is attained by $\mathbb{P}(m = i) \cdot 1 + \mathbb{P}(m = e) \cdot 0 = 4p_0 < 1$.

Furthermore, BLL construct an ambiguous persuasion signal that leads to a higher expected payoff of the Sender. Let $M = \{e, i\}$ *be as before. The set of communication devices* $\Pi = \{\pi, \pi'\}$ *is given by a communication device that always reveals the true state and a communication device that does the opposite, i.e.,*

$$\pi(i|\omega_i) = 1 = 1 - \pi(e|\omega_i), \qquad \pi(i|\omega_e) = 0 = 1 - \pi(e|\omega_e),$$
$$\pi'(i|\omega_i) = 0 = 1 - \pi'(e|\omega_i), \qquad \pi'(i|\omega_e) = 1 = 1 - \pi'(e|\omega_e).$$

Given this ambiguous communication device, the interim state beliefs are

$$P_m = \{(0,1), (1,0)\}$$

for $m \in \{e, i\}$. *Due to the maxmin preferences, the interim worst-case belief for both messages always gives probability one to the inefficient state* ω_i. *Therefore, the Receiver chooses the brand name drug with probability one. Then, the ex-ante expected payoff of the Sender is one which is greater than the ex-ante expected payoff given the optimal Bayesian persuasion.*

However, the ex-ante expected payoff of the Receiver is given by

$$\min_{\pi \in \Pi} \sum_{m \in \{e,i\}} \big(\pi(m|\omega_e) + \pi(m|\omega_i)\big) \mathbb{E}_{p_m^\pi}\big(u(a_m, \omega)\big),$$

where a_m *denotes her action after observing the messages* m. *If she chooses the brand name drug independently of the signal that she will observe, her ex-ante expected payoff equals*

$$2 \cdot \mathbb{P}(\omega = \omega_e) + 2 \cdot \mathbb{P}(\omega = \omega_i) = 2.$$

Her expected payoff if she always choose the generic drug is

$$3 \cdot \mathbb{P}(\omega = \omega_e) + 1 \cdot \mathbb{P}(\omega = \omega_i) = 3 - 4p_0.$$

Since $p_0 < \frac{1}{4}$, *the optimal interim strategy of always prescribing the brand name drug is not ex-ante optimal, and the Receiver behaves dynamically inconsistently.*

[4] Please note, that for simplicity we deviate from the illustrating example of BLL (page 317) by assuming $u_H = 3$, $u_L = -1$ and $c = 1$, which is consistent with the payoffs in Example 2 of BLL.

2.2.2 Dynamically Consistent Belief Formation Process

In this model, ambiguity arises due to the ambiguous communication device. Ambiguous interim beliefs only occur due to the combination of a risky state and an ambiguous signal. Consider the following two situations at the ex-ante stage:

1) The Receiver does not observe any message. All information about the state $\omega \in \Omega$ is represented by p_0.

2) As in situation 1) the Receiver knows p_0. Additionally, she knows that she will receive an ambiguous message before making her decision.

In the first situation, the Receiver knows that there will be no additional information. She chooses her optimal action, given the expected utility with respect to p_0. In the second situation, the Receiver has ex-ante the same information about the state as in Situation 1). However, she knows that she will receive additional but ambiguous information before making her decision. Further, she knows that this ambiguous information influences her interim beliefs and, thus, her best response. A rational player should consider this knowledge about a game's information structure at the ex-ante stage. Rectangularity takes the interplay of the prior state belief p_0 and the knowledge about the information structure into account. This ensures dynamic consistency.

This section shows that defining beliefs over a general state space allows the definition of non-singleton rectangular belief sets. Then, given rectangular beliefs, the Receiver behaves dynamically consistently, and the equilibrium of BLL is a perfect Bayesian equilibrium.

2.2.2.1 Straightforward Messages

In the ambiguous persuasion setting, the set of messages M is part of the Sender's strategy. In a Bayesian persuasion setting, Kamenica and Gentzkow (2011) call a communication device *straightforward* if $M \subseteq A$. They show that one can restrict without loss of generality to straightforward communication devices in a Bayesian persuasion setting. The next proposition generalizes this result to our ambiguous persuasion setting. It shows that the Sender chooses without loss of generality $M \subseteq A \cup \tilde{A}$, where $\tilde{A}$ is a duplicated set of A such that there exists a bijection $b(\cdot)$ between A and $\tilde{A}$. Given this result, we can define rectangular ex-ante beliefs over $\Omega \times (A \cup \tilde{A})$.

Proposition 2.1. *Let* $(M, \Pi) \in \arg\sup \min_{\pi \in \Pi} \mathbb{E}_{p_0} [\mathbb{E}_\pi [\nu(\hat{a}_m, \omega)|\omega]]$. *Let* $\tilde{A}$ *be such that there exists a bijection* $b(\cdot) : A \to \tilde{A}$ *between* A *and* $\tilde{A}$. *Then, there exist a tuple* (M', Π') *with* $M' \subseteq A \cup \tilde{A}$ *and* $\Pi' = \{\pi'_1, \pi'_2\}$ *such that* (M', Π') *generates the same value for the Sender as* (M, Π).

The intuition of the result is as follows. Kamenica and Gentzkow (2011) show that for Bayesian persuasion, it is without loss of generality that $M \subseteq A$. BLL show that ambiguous persuasion increases the value for the Sender compared to Bayesian persuasion only if the Sender uses a signal with synonyms. Synonyms are messages that copy the meaning of another message, i.e., they induce the same posterior state belief set or best response of the Receiver. Furthermore, they show that for any ambiguous signal, one can find an ambiguous signal which only consists of two communication devices and leads to the same value. Hence, in order to allow synonyms, we have to duplicate the message space. Further, duplication is enough to generate the same value as any ambiguous signal. Therefore, $M \subset A \cup \tilde{A}$.

Proof of Proposition 1. Corollary 1 of BLL shows that there exist π_1 and π_2 such that $(M, \{\pi_1, \pi_2\})$ generates the same value as (M, Π). Hence, we have to show that (M', Π') generates the same value as $(M, \{\pi_1, \pi_2\})$. We first look at the case where the Sender does not use synonyms.

i) The Sender does not use synonyms:

   Since $(M, \{\pi_1, \pi_2\})$ does not use synonyms, there does not exist any $m, m' \in M$ with $m \neq m'$ such that $\hat{a}_m = \hat{a}_{m'}$. Remember, that p_m^π denotes the posterior state belief of the Receiver given the message m and the communication device π. Furthermore, $\hat{a}_m$ denotes the Receivers' best response given message $m \in M$ and the communication devices $\{\pi_1, \pi_2\}$. Since $(M, \{\pi_1, \pi_2\})$ does not use synonyms, there exists at most one $m \in M$ for each $a \in A$ such that $a = \hat{a}_m$. We define $\bar{\pi}_i(\cdot|\omega) \in \Delta M'$ with $M' \subset A$ such that

$$\bar{\pi}_i(a|\omega) = \begin{cases} \pi_i(m|\omega) & \text{if } \exists m \in M \text{ with } a = \hat{a}_m, \\ 0 & \text{otherwise.} \end{cases}$$

   Then, the posterior state belief $p_m^{\pi_i}$ equals the posterior state belief $p_a^{\bar{\pi}_i}$ if $a = \hat{a}_m$. Therefore, $(M, \{\pi_1, \pi_2\})$ and $(M', \{\bar{\pi}_1, \bar{\pi}_2\})$ generate the same set of posterior state beliefs and the same best response of the Receiver. Since the best response does not change, the value of the Sender is the same for both signals.

ii) The Sender uses synonyms:

   If $(M, \{\pi_1, \pi_2\})$ uses synonyms, we can split M in M_1 and M_2 such that there exists a bijection between M_1 and M_2 and $M_1 \cup M_2 = M$. Then, $(M_1, \{\hat{\pi}_1, \hat{\pi}_2\})$ with

$$\hat{\pi}_i(m|\omega) = \frac{\pi_i(m|\omega)}{\sum_{m \in M_1} \pi_i(m|\omega)}$$

defines a signal that does not use synonyms. Hence, as in Case i), there exists $(M_1', \{\bar{\pi}_1, \bar{\pi}_2\})$ with $M_1' \subset A$ that generates the same value as $(M_1, \{\hat{\pi}_1, \hat{\pi}_2\})$. Similarly, one can define the restriction of π_i to M_2 and find $(M_2', \{\tilde{\pi}_1, \tilde{\pi}_1\})$ with $M_2' \subset \tilde{A}$, that generates the same value as M_2 and the restriction of π_i to M_2. Then, $(M', \{\pi_1', \pi_2'\})$ with $M' = M_1' \cup M_2'$ and

$$\pi_i'(a|\omega) = \begin{cases} \bar{\pi}_i(a|\omega) \sum_{m \in M_1} \pi_i(m|\omega) & \text{if } a \in A, \\ \tilde{\pi}_i(a|\omega) \sum_{m \in M_2} \pi_i(m|\omega) & \text{if } a \in \tilde{A}, \end{cases}$$

generates the same value as $(M, \{\pi_1, \pi_2\})$.

□

Proposition 2.1 shows that without loss of generality we can assume that $M \subset A \cup \tilde{A}$. Due to the assumption that all π_k have common full support on M, a strategy of the Sender (M, Π) is completely characterized by Π. For the rest of the chapter, we will use the term strategy of the Sender for such a Π. Furthermore, we denote with $\text{supp}(\Pi) = \text{supp}(\pi_k(\cdot|\omega))$ the support of $\pi_k \in \Pi$ for all $k \in K$.

2.2.2.2 Rectangular Beliefs

Given the results from the previous section, we can define beliefs over the general state space $\Omega \times (A \cup \tilde{A})$. Defining beliefs over this general state space allows the Receiver to form a joint belief about the risky state $\omega \in \Omega$ and the message $m \in M$, i.e., the Receiver forms beliefs of the events "the state is ω, and I observe message m." Then, the probability of this event depends on the risky state $\omega \in \Omega$ and the ambiguous communication device that generates the message.

Definition 2.1. *For a strategy Π of the Sender, we define the* ***set of ex-ante beliefs*** *of the Receiver as*

$$\Phi_\Pi^0 = \Bigg\{ \rho^k \in \Delta(\Omega \times (A \cup \tilde{A})) \, : \, \exists \pi_k \in \Pi \text{ with}$$
$$\rho^k(\omega, m) = \begin{cases} p_0(\omega)\pi_k(m|\omega) & \text{if } m \in supp(\Pi), \\ 0 & \text{otherwise.} \end{cases} \Bigg\}.$$

Please note that the strategy of the Sender generates the information structure of the persuasion games. Hence, it has to influence the joint belief over states and messages.

At the interim stage the Receiver observes a message $m \in \text{supp}(\Pi)$. The information structure at the ex-ante stage ($t = 0$) and interim stage ($t = 1$) can be represented by the following partitions

$$\mathcal{F}_0 = \Omega \times (A \cup b(A)),$$
$$\mathcal{F}_1 = \Big\{\{\Omega \times m\}_{m \in A \cup b(A)}\Big\}.$$

Then, given an observation $\hat{m} \in \text{supp}(\Pi)$, the Receiver updates her ex-ante belief set prior-by-prior using Bayes' formula, i.e., she updates each prior belief in Φ^0_Π with Bayes' formula

$$\rho^k|_{\hat{m}} = \rho^k((\omega, m)|\hat{m}) = \frac{p_0(\omega)\pi_k(m|\omega)}{\int_\Omega p_0(\omega')\pi_k(m|\omega')\,\mathrm{d}\omega'},$$

if $m = \hat{m}$ and zero otherwise. Then, the set of updated beliefs given $\hat{m} \in \text{supp}(\Pi)$ is

$$\text{Bay}(\Phi^0_\Pi|\hat{m}) = \{\rho^k|_{\hat{m}} \,:\, \rho^k \in \Phi^0_\Pi\}.$$

Remark 2.1. *Note that* $\rho^k((\omega, m)|\hat{m}) = 0$ *for* $\hat{m} \notin$ *supp*(Π) *and* $\rho^k((\omega, \hat{m})|\hat{m}) = p^{\pi_k}_{\hat{m}}(\omega)$ *for all* $\omega \in \Omega$.

To define rectangularity let us first look at the case without ambiguity, i.e., if $\Pi = \{\pi\}$ and $\Phi^0_\Pi = \{\rho\}$ is singleton. After observing message m the updated belief is given by $\rho|_m$. Furthermore, the marginal beliefs of observing $m \in A \cup \tilde{A}$ under ρ is

$$\rho(\Omega, m) = \int_\Omega \rho(\omega, m)\,\mathrm{d}\omega = \int_\Omega p_0(\omega)\pi(m|\omega)\,\mathrm{d}\omega.$$

Then, the structure of Bayes' formula implies that multiplying the updated belief after observing message m with the marginal probability of observing m leads to the prior belief restricted to the events that the message is m. This holds for all messages m and, therefore, for all information sets of the partition defined above. Hence, integrating over all $m \in \text{supp}(\Pi)$ leads to the prior belief

$$\rho(\omega, m) = \int_{\text{supp}(\Pi)} \rho(\Omega, m')\rho|_{m'}(\omega, m)\,\mathrm{d}m'.$$

Now, we generalize these considerations to an ambiguous setting, i.e., Π is not a singleton. Rectangularity requires that any combination of marginal belief and updated belief is a prior belief that the agent considers as possible. The Receiver knows which messages she could receive and, thus, which updated beliefs potentially exist. Taking this knowledge into account, rectangularity requires that any combination of marginal and updated belief is an element of the ex-ante belief set.

Definition 2.2. *The pasting of an ex-ante belief* $\bar{\rho} \in \Phi^0_\Pi$ *and a collection of updated beliefs* $(\rho|_{\hat{m}})_{\hat{m}} \in \bigtimes_{\hat{m}\in supp(\Pi)} Bay(\Phi^0_\Pi|\hat{m})$ *is defined as*[5]

$$\begin{aligned}\bar{\rho} \circ (\rho|_{\hat{m}})_{\hat{m}}(\omega, m) &:= \int_{supp(\Pi)} \bar{\rho}(\Omega, \hat{m})\rho(\omega, m|\hat{m})\, \mathrm{d}\hat{m} \\ &= \Big(\int_\Omega p_0(\omega')\bar{\pi}(m|\omega')\, \mathrm{d}\omega' \Big) \frac{p_0(\omega)\pi(m|\omega)}{\int_\Omega p_0(\omega')\pi(m|\omega')\, \mathrm{d}\omega'}.\end{aligned}$$

The set of ex-ante beliefs is called ***rectangular*** *(or stable under pasting) if it contains all pastings of an ex-ante belief* $\bar{\rho} \in \Phi^0_\Pi$ *and interim beliefs* $(\rho|_{\hat{m}})_{\hat{m}}$*, i.e.,*

$$\bar{\rho} \circ (\rho|_{\hat{m}})_{\hat{m}}(\cdot) \in \Phi^0_\Pi$$

for all $\bar{\rho} \in \Phi^0_\Pi$ *and* $(\rho|_{\hat{m}})_{\hat{m}} \in \bigtimes_{\hat{m}\in supp(\Pi)} Bay(\Phi^0_\Pi|\hat{m})$*.*

If Φ^0_Π is not rectangular, one can always construct the smallest set, which is rectangular and contains Φ^0_Π by backward induction. We call this set the rectangular hull and denote it with $\mathrm{rect}(\Phi^0_\Pi)$. Simple calculations show that $\mathrm{Bay}(\Phi^0_\Pi|\hat{m}) = \mathrm{Bay}(\mathrm{rect}(\Phi^0_\Pi)|\hat{m})$. The same holds for the set of marginal beliefs under Φ^0_Π and $\mathrm{rect}(\Phi^0_\Pi)$. For a more detailed explanation of the construction and the properties of the rectangular hull, we refer to Chapter 1.

So far, we focused on the beliefs of the Receiver. The Sender only chooses an action at the ex-ante stage. Therefore, the interim beliefs of the Sender do not influence the equilibria of the game. For technical completeness, we can always find an information structure of the Sender that does not influence the ex-ante decision of the Sender but ensures that the ex-ante belief set of the Sender is rectangular for any strategy Π. For example, the Sender could observe which communication device generated the observed message at the interim stage. However, this chapter aims to find a belief formation process that ensures dynamically consistent behavior. Since the Sender can never behave dynamically inconsistently, we do not go into details.

2.3 Dynamic Consistency and PBE

Finally, we show that rectangularity implies dynamically consistent behavior of the Receiver and, therefore, the existence of a perfect Bayesian equilibrium.

Definition 2.3. *A perfect Bayesian equilibrium (PBE) with rectangular beliefs consists of a strategy* Π^* *of the Sender, a strategy* $(\hat{a}_m)_{m\in M}$ *of the Receiver and a belief system* Ψ *for each player. Strategies and belief systems have to satisfy the following conditions:*

[5] Please note, that the pasting is always well defined due to the common support assumption. Furthermore, the second equality follows since $\rho(\omega, m|\hat{m}) = 0$ if $m \neq \hat{m}$.

- *The belief systems of both players consist of an ex-ante belief set* Ψ_i^0 *and interim belief set* Ψ_i^m *for each message* $m \in A \cup \tilde{A}$ *such that*

$$\Psi_R^0 = \mathrm{rect}(\Phi_{\Pi^*}^0) \quad \textit{and} \quad \Psi_S^0 = \Phi_{\Pi^*}^0.$$

  Furthermore, the interim belief sets are derived by Bayes rule whenever possible, i.e., $\Psi_i^m = Bay(\Psi_i^0 | m)$ *for all* $m \in \mathrm{supp}(\Pi^*)$.

- *The equilibrium strategy of the Sender* Π^* *with supp*$(\Pi^*) \subseteq A \cup \tilde{A}$ *maximizes his ex-ante worst-case expected utility*

$$\min_{\rho \in \Psi_S^0} \mathbb{E}_\rho \left[\nu(\hat{a}_m, \omega) \right].$$

- *The equilibrium strategy of the Receiver maximizes her interim worst-case expected utility for all* $m \in \mathrm{supp}(\Pi^*)$

$$\min_{\rho|_m \in \Psi_R^m} \mathbb{E}_{\rho|_m}(u(a_m, \omega)),$$

  and her ex-ante worst-case expected utility given the ex-ante belief set Ψ_R^0

$$\min_{\rho \in \Psi_R^0} \mathbb{E}_\rho(u(a_m, \omega)).$$

The following proposition shows that we can generalize any ex-ante best response of the Sender and interim best response of the Receiver to a perfect Bayesian equilibrium using rectangularity.

Proposition 2.2. *Let* (M, Π) *be the optimal ex-ante choice of the Sender and* $(a_m)_{m \in M}$ *the optimal interim choice of the Receiver as in BLL. Then, there exists* (M^*, Π^*)*, with* $M^* \subseteq A \cup \tilde{A}$ *and* $|\Pi^*| = 2$ *that generates the same value of the Sender as* (M, Π)*. Furthermore,* Π^**,* $(\hat{a}_m)_{m \in M^*}$*, and*

$$\begin{aligned} \Psi_R^0 &= \mathrm{rect}(\Phi_{\Pi^*}^0), \\ \Psi_S^0 &= \Phi_{\Pi^*}^0, \\ (\Psi_i^m)_{m \in M^*} &= (Bay(\Psi_i^0 | m))_{m \in M^*} \end{aligned}$$

are a PBE with rectangular beliefs.

Proof. The first part of the proof follows from Proposition 2.1. Furthermore, the Sender never behaves dynamically inconsistently. We only have to show that the Receivers interim best response of BLL is an interim and ex-ante best response given rectangular beliefs. Remember that $p_{\hat{m}}^{\pi_k}(\cdot) = \rho^k((\cdot, \hat{m}) | \hat{m})$ for all $\hat{m} \in \mathrm{supp}(\Pi)$ and that the set of Bayesian updates given Φ_Π^0 or $\mathrm{rect}(\Phi_\Pi^0)$ are the same. Therefore, the interim best response given the state beliefs of BLL is an interim best response given rectangular beliefs, as well. Furthermore, we can rewrite the ex-ante expected utility of the Receiver as

$$\min_{\rho \in \mathrm{rect}(\Phi^0_{\Pi^*})} \int_{\mathrm{supp}(\Pi)} \rho(\Omega, \hat{m}) \mathbb{E}_{\rho|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m},$$

where $\rho|_{\hat{m}}$ is the Bayesian update of ρ given message $\hat{m}$. We first show the relation between ex-ante and interim worst-case expected utility. Let ρ^* denote the ex-ante worst-case belief given rectangular beliefs. Then,

$$\int_{\mathrm{supp}(\Pi^*)} \rho^*(\Omega, \hat{m}) \mathbb{E}_{\rho^*|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$
$$= \int_{\mathrm{supp}(\Pi^*)} \rho^*(\Omega, \hat{m}) \min_{\rho|_{\hat{m}} \in \mathrm{Bay}(\mathrm{rect}(\Phi^0_{\Pi^*})|\hat{m})} \mathbb{E}_{\rho|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}. \tag{2.1}$$

To prove Equation (2.1), we first show that the left-hand side is greater or equal than the right-hand side using the inequality

$$\mathbb{E}_{\rho^*|_{\hat{m}}}(u(a_{\hat{m}}, \omega)) \geq \min_{\rho|_{\hat{m}} \in \mathrm{Bay}(\mathrm{rect}(\Phi^0_{\Pi^*})|\hat{m})} \mathbb{E}_{\rho|_{\hat{m}}}(u(a_{\hat{m}}, \omega))$$

for all $\hat{m} \in \mathrm{supp}(\Pi^*)$.

To prove the other direction, let $\rho'|_{\hat{m}}$ be the worst-case belief given that she observed $\hat{m}$. Then, due to rectangularity, there exists $\bar{\rho} \in \mathrm{rect}(\Phi^0_{\Pi^*})$ such that $\rho^* \circ (\rho'|_{\hat{m}})_{\hat{m}} = \bar{\rho}$. Furthermore, rectangularity implies, that $\bar{\rho}(\cdot|\hat{m}) = \rho'(\cdot|\hat{m})$ and $\bar{\rho}(\Omega, \hat{m}) = \rho^*(\Omega, \hat{m})$ for all $\hat{m}$. Hence,

$$\int_{\mathrm{supp}(\Pi^*)} \rho^*(\Omega, \hat{m}) \mathbb{E}_{\rho^*|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m} \leq \int_{\mathrm{supp}(\Pi^*)} \bar{\rho}(\Omega, \hat{m}) \mathbb{E}_{\bar{\rho}|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$
$$= \int_{\mathrm{supp}(\Pi^*)} \rho^*(\Omega, \hat{m}) \mathbb{E}_{\rho'|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$
$$= \int_{\mathrm{supp}(\Pi^*)} \rho^*(\Omega, \hat{m}) \min_{\rho|_{\hat{m}} \in \mathrm{Bay}(\mathrm{rect}(\Phi^0_{\Pi^*})|\hat{m})} \mathbb{E}_{\rho|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}.$$

Combining both directions proves Equation (2.1). Finally, we show that an interim best response of the Receiver is an ex-ante best response, as well. We denote the (sender-preferred) interim best response of the Receiver given message $\hat{m}$ by $\hat{a}_{\hat{m}}$, i.e.,

$$\min_{\rho|_{\hat{m}} \in \mathrm{Bay}(\Phi^0_{\Pi^*}|\hat{m})} \mathbb{E}_{\rho|_{\hat{m}}}(u(\hat{a}_{\hat{m}}, \omega)) \geq \min_{\rho|_{\hat{m}} \in \mathrm{Bay}(\Phi^0_{\Pi^*}|\hat{m})} \mathbb{E}_{\rho|_{\hat{m}}}(u(a_{\hat{m}}, \omega))$$

for any arbitrary $a_{\hat{m}} \in A$ and all $\hat{m} \in \mathrm{supp}(\Pi^*)$. We have to show that $(\hat{a}_{\hat{m}})_{\hat{m} \in \mathrm{supp}(\Pi^*)}$ is ex-ante optimal. Since $\rho(\Omega, \hat{m}) \geq 0$ for all $\hat{m} \in \mathrm{supp}(\Pi^*)$ and $\rho(\Omega, \hat{m}) = 0$ for all $\hat{m} \notin \mathrm{supp}(\Pi^*)$, Equation (2.1) implies

$$\min_{\rho \in \mathrm{rect}(\Phi^0_{\Pi^*})} \int_{\mathrm{supp}(\Pi^*)} \rho(\Omega, \hat{m}) \mathbb{E}_{\rho|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$
$$= \min_{\rho \in \mathrm{rect}(\Phi^0_{\Pi^*})} \int_{\mathrm{supp}(\Pi^*)} \rho(\Omega, \hat{m}) \min_{\rho'|_{\hat{m}} \in \mathrm{Bay}(\Phi^0_{\Pi^*}|\hat{m})} \mathbb{E}_{\rho'|_{\hat{m}}}(u(a_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$

$$\leq \min_{\rho \in \text{rect}(\Phi^0_{\Pi^*})} \int_{\text{supp}(\Pi^*)} \rho(\Omega, \hat{m}) \min_{\rho'|_{\hat{m}} \in \text{Bay}(\Phi^0_{\Pi^*}|\hat{m})} \mathbb{E}_{\rho'|_{\hat{m}}}(u(\hat{a}_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$
$$= \min_{\rho \in \text{rect}(\Phi^0_{\Pi^*})} \int_{\text{supp}(\Pi^*)} \rho(\Omega, \hat{m}) \mathbb{E}_{\rho|_{\hat{m}}}(u(\hat{a}_{\hat{m}}, \omega))\, \mathrm{d}\hat{m}$$

for any arbitrary $(a_{\hat{m}})_{\hat{m} \in \text{supp}(\Pi)}$. The inequality follows from the interim optimality of $(\hat{a}_{\hat{m}})_{\hat{m} \in \text{supp}(\Pi^*)}$ and the last equality from Equation (2.1).

Hence, the Receivers' ex-ante best response equals the interim best response and the interim equilibrium of Beauchêne et al. (2019) satisfies ex-ante optimality. □

Remark 2.2. *In the proof of Proposition 2.1, we show how (M^*, Π^*) can be constructed. The construction is similar as for Bayesian persuasion in Kamenica and Gentzkow (2011). Intuitively, any two messages m and m' that are not synonyms of each other but induce the same optimal strategy, i.e., $\hat{a}_m = \hat{a}_{m'}$, are replaced by the same message $\bar{m}$. This implies that $M^* \subseteq A \cup \tilde{A}$. The construction shows that even if the message sets M and M^* are different, the Receiver's actions do not change.*

To illustrate the previous results, we come back to our example from Section 2.2.1

Example 2.2 (Example 2.1 cont.). *Remember that the optimal ambiguous communication device was given by $\Pi = \{\pi, \pi'\}$ with*

$$\pi(i|\omega_i) = 1 = 1 - \pi(e|\omega_i), \qquad \pi(i|\omega_e) = 0 = 1 - \pi(e|\omega_e),$$
$$\pi'(i|\omega_i) = 0 = 1 - \pi'(e|\omega_i), \qquad \pi'(i|\omega_e) = 1 = 1 - \pi'(e|\omega_e).$$

Then, the set of ex-ante beliefs of the Receiver is $\Phi^0_\Pi = \{\rho, \rho'\}$ with

$$\rho(\omega, m) = \begin{cases} p_0 & \text{if } m = i, \omega = \omega_i, \\ 1 - p_0 & \text{if } m = e, \omega = \omega_e, \\ 0 & \text{otherwise}, \end{cases} \qquad \rho'(\omega, m) = \begin{cases} p_0 & \text{if } m = e, \omega = \omega_i, \\ 1 - p_0 & \text{if } m = i, \omega = \omega_e, \\ 0 & \text{otherwise}. \end{cases}$$

To construct the rectangular hull, we need to calculate all interim beliefs

$$\rho(\omega, m|i) = \begin{cases} ,1 & \text{if } m = i, \omega = \omega_i, \\ 0 & \text{otherwise}, \end{cases} \qquad \rho(\omega, m|e) = \begin{cases} 1 & \text{if } m = e, \omega = \omega_e, \\ 0 & \text{otherwise}, \end{cases}$$
$$\rho'(\omega, m|i) = \begin{cases} 1 & \text{if } m = i, \omega = \omega_e, \\ 0 & \text{otherwise}, \end{cases} \qquad \rho'(\omega, m|e) = \begin{cases} 1 & \text{if } m = e, \omega = \omega_i, \\ 0 & \text{otherwise}, \end{cases}$$

and and marginal beliefs

$$marg(\rho(\cdot, i)) = p_0, \qquad marg(\rho(\cdot, e)) = 1 - p_0,$$
$$marg(\rho'(\cdot, e)) = p_0, \qquad marg(\rho'(\cdot, i)) = 1 - p_0.$$

Then, we obtain the rectangular hull $\mathrm{rect}(\Phi^0_\Pi) = \{\rho, \rho', \hat{\rho}, \bar{\rho}\}$ *by combining any marginal and interim belief, where* ρ *and* ρ' *are as before and*

$$\bar{\rho}(\omega, m) = \begin{cases} 1 - p_0 & \text{if } m = i, \omega = \omega_i, \\ p_0 & \text{if } m = e, \omega = \omega_e, \\ 0 & \text{otherwise,} \end{cases} \qquad \hat{\rho}(\omega, m) = \begin{cases} 1 - p_0 & \text{if } m = e, \omega = \omega_i, \\ p_0 & \text{if } m = i, \omega = \omega_e, \\ 0 & \text{otherwise.} \end{cases}$$

Given the rectangular hull, the worst-case belief of the Receiver, if she plans to choose the generic drug after message m*, is* $\mathbb{P}(\omega = \omega_i, m) = 1 - p_0 > \frac{3}{4}$*. Therefore, always prescribing the brand name drug is ex-ante optimal, and the Receiver behaves dynamically consistently.*

2.4 Value of Information

Our example shows that the Receiver is better off by making her decision based on p_0. Therefore, she would prefer getting no additional information than getting ambiguous information. This result is consistent with the recent literature on the (negative) value of information under ambiguity, e.g., Li (2020) or Hill (2020). However, BLL show in their Subsections 6.3 and 6.4 that the Receiver may benefit from listening to an ambiguous device.

We denote the ex-ante expected utility of action a of the Receiver without any additional information by $U^0(a)$, i.e.,

$$U^0(a) = \int_\Omega u(a, \omega) p_0(\omega) \, \mathrm{d}\omega.$$

Definition 2.4. *A communication device* Π *has a positive value of information for the Receiver if*

$$\max_{(a_m)_{m \in \operatorname{supp} \Pi} \in A^{|\operatorname{supp} \Pi|}} \min_{\rho \in \mathrm{rect}(\Phi^0_\Pi)} \mathbb{E}_\rho(u(a_m, \omega)) \geq \max_{a \in A} U^0(a).$$

Ambiguous information induces two effects. On the one hand, an ambiguous communication device generates ambiguous beliefs and decreases the worst-case expected utility of the Receiver. On the other hand, the communication device still reveals information about the state. This information allows the Receiver to choose an action that is better suited for the state and increases her expected utility. Then, the value of information is positive if the second effect exceeds the negative effect of ambiguity and ambiguity aversion.

BLL say that a communication device satisfies a *participation constraint* if

$$\max_{(a_m)_{m\in \operatorname{supp}\Pi}\in A^{|\operatorname{supp}\Pi|}} \min_{\pi\in\Pi} \int_\Omega \int_M \pi(m|\omega)u(a_m,\omega)\,\mathrm{d}m\, p_0(\omega)\,\mathrm{d}\omega \geq \max_{a\in A} U^0(a).$$

They call this condition a participation constraint since it ensures that the Receiver is willing to pay attention to the communication device. If the participation constraint is not satisfied, the Receiver would be better off by ignoring the communication device, ex-ante. Since $\Phi^0_\Pi \subseteq \operatorname{rect}(\Phi^0_\Pi)$, it follows that

$$\begin{aligned}&\max_{(a_m)_{m\in \operatorname{supp}\Pi}\in A^{|\operatorname{supp}\Pi|}} \min_{\pi\in\Pi} \int_\Omega \int_M \pi(m|\omega)u(a_m,\omega)\,\mathrm{d}m\, p_0(\omega)\,\mathrm{d}\omega \\ = &\max_{(a_m)_{m\in \operatorname{supp}\Pi}\in A^{|\operatorname{supp}\Pi|}} \min_{\rho\in\Phi^0_\Pi} \mathbb{E}_\rho(u(a_m,\omega)) \geq \max_{(a_m)_{m\in \operatorname{supp}\Pi}\in A^{|\operatorname{supp}\Pi|}} \min_{\rho\in\operatorname{rect}(\Phi^0_\Pi)} \mathbb{E}_\rho(u(a_m,\omega)).\end{aligned}$$

Thus, any communication device with a positive value of information satisfies the participation constraint of BLL.

BLL characterize a condition that guarantees that the Receiver benefits from listening to a communication device (see BLL Proposition 8). We now translate this condition to our setting. We denote the default actions by a_0, i.e., the action that maximizes $U^0(a)$.

Definition 2.5. *Let $\hat{a}_m$ denote the interim optimal action of the Receiver with rectangular beliefs Bay$(\operatorname{rect}(\Phi^0_\Pi)|m)$. A message m is value-increasing (to the Receiver) if $\mathbb{E}_{\rho|_m}(u(\hat{a},\omega)) \geq U^0(a_0)$ for all $\rho|_m \in$ Bay$(\operatorname{rect}(\Phi^0_\Pi)|m)$.*

BLL show that a communication device Π satisfies the participation constraint if Π only uses value-increasing messages. The next proposition proves a stronger and very intuitive result: A communication device that increases the worst-case expected utility of the Receiver for any message has a positive value of information.

Proposition 2.3. *If Π only uses value-increasing messages, Π has a positive value of information for the Receiver.*

Proof. Since $\mathbb{E}_{\rho|_m}(u(\hat{a},\omega)) \geq U^0(a_0)$ for all $\rho|_m \in \operatorname{Bay}(\operatorname{rect}(\Phi^0_\Pi)|m)$, it follows that

$$\min_{\rho|_m\in\operatorname{Bay}(\operatorname{rect}(\Phi^0_\Pi)|m)} \mathbb{E}_{\rho|_m}(u(\hat{a},\omega)) \geq U^0(a_0). \tag{2.2}$$

Then, rectangularity and Equation (2.2) imply

$$\begin{aligned}&\max_{(a_m)_{m\in \operatorname{supp}\Pi}\in A^{|\operatorname{supp}\Pi|}} \min_{\rho\in\operatorname{rect}(\Phi^0_\Pi)} \mathbb{E}_\rho(u(a_m,\omega)) \\ &= \min_{\rho\in\operatorname{rect}(\Phi^0_{\Pi^*})} \int_{\operatorname{supp}(\Pi^*)} \rho(\Omega,m) \min_{\rho'|_m\in\operatorname{Bay}(\Phi^0_{\Pi^*}|m)} \mathbb{E}_{\rho'|_m}(u(\hat{a}_m,\omega))\,\mathrm{d}m \\ &\geq \min_{\rho\in\operatorname{rect}(\Phi^0_{\Pi^*})} \int_{\operatorname{supp}(\Pi^*)} \rho(\Omega,m)U^0(a_0)\,\mathrm{d}m = U^0(a_0).\end{aligned}$$

□

2.5 Conclusion and Discussion

We show that the gain of ambiguous persuasion arises due to ambiguity and ambiguity aversion and not due to dynamically inconsistent behavior. First, we show that we can restrict without loss of generality to straightforward messages and synonyms, i.e., $M \subset A \cup \tilde{A}$. Given this result, we can introduce beliefs over the more general state space $\Omega \times A \cup \tilde{A}$. This state space allows for the dependence of the risky state and ambiguous signals. Therefore, the Receiver can take the potentially ambiguous information structure at the ex-ante stage into account. Then, rectangular beliefs ensure dynamically consistent behavior in ambiguous persuasion and the existence of a perfect Bayesian equilibrium. This shows that ambiguity induces new equilibria in persuasion settings, even if the players behave dynamically consistently. To conclude, we discuss some related issues and literature.

Endogenous Ambiguity in Cheap Talk Kellner and Le Quement (2018) show that in a cheap talk setting, an ambiguous strategy of the Sender can lead to an interim equilibrium that improves the ex-ante expected payoff of Sender and Receiver. In their setting, players face a risky state $\omega \in \Omega$. The Sender can commit his signal on an ambiguous payoff-irrelevant state $\theta \in \Theta$, which leads to an ambiguous posterior belief for the Receiver. As in our ambiguous persuasion setting, the equilibrium strategy of the Receiver is not ex-ante optimal. However, similarly to the procedure described above, defining beliefs and rectangularity over the general state space $\Omega \times \Theta$, leads to a perfect Bayesian equilibrium with the same strategies as in the interim equilibrium of Kellner and Le Quement (2018).

Preferences for Partial Information of Li (2020) Li (2020) characterizes aversion to partial information under ambiguity aversion. He shows that an ambiguity-averse decision maker (DM) with maxmin preferences is always (weakly) averse to partial information. Furthermore, the DM is neutral to a specific information partition π if and only if his ex-ante belief set is rectangular with respect to π. More formally, let Π denote the set of all partitions of a state space S and let $\mathcal{F}$ be the set of acts which are maps from states to consequences $f : S \to X$. Li defines ex-ante preferences on an extended choice domain, which is the product space of information partitions $\pi \in \Pi$ and acts $f \in \mathcal{F}$. Given an information partition π, the DM anticipates the possible future information and constructs his ex-ante preference $\succeq$ recursively from the interim preferences at each event $E \in \pi$. Then, Li says that a preference relation $\succeq$ exhibits aversion to partial information, if $(\pi^0, f) \succeq (\pi, f)$ for all acts $f \in \mathcal{F}$ and partition $\pi \in \Pi$, where π^0 denotes the information partition where no information is learned, i.e., $\pi^0 = \{S\}$.

It is important to note, that the definition of aversion to partial information requires $(f, \pi^0) \succeq (f, \pi)$ for all acts $f \in \mathcal{F}$. In the ambiguous persuasion setting, the Receiver cannot condition his action on a state or message without any additional information. But anticipating that he will receive ambiguous information allows him at the ex-ante stage to condition his action on the messages that he could observe. In Li's setting, this would imply that given π^0 the DM can only choose from constant acts. Given an ambiguous communication device, the DM can choose any act that is measurable with respect to the information partition induced by the communication device. These are exactly the two effects we describe after Definition 2.4. On the one hand, an ambiguous information device induces ambiguity, which decreases the utility of an ambiguity avers Receiver. On the other hand, anticipating this information at the ex-ante stage allows the Receiver to choose an action for each message that could occur with positive probability. Li focuses only on the first effect since the set of acts $\mathcal{F}$ is the same under π^0 and π. Therefore, his result about partial information aversion of maxmin preferences does not contradict our result about a positive value of information (Proposition 2.3).

Subjective Trees of Hill (2020) Hill (2020) formulates a dynamic consistency axiom in a model with so-called subjective trees. Roughly speaking, subjective trees are information structures that are not necessarily represented by a partition of the state space. He argues that using his version of the dynamic consistency axiom resolves the conflict between dynamic consistency and ambiguity. In Appendix A, he defines a different setup with an extended state space. Under the extended state space, his formulation of the dynamic consistency axiom is equivalent to the standard formulation. In this work, we use a general state space instead of a non-partitional information structure, as in the Appendix of Hill (2020). Furthermore, our definition of a positive value of information is similar to the definition of Hill (2020).

Other Approaches Dealing with Dynamic Inconsistency Cheng (2020) uses a model similar to BLL but focuses on the Receiver's ex-ante optimization problem. He shows that if the Receiver can commit to his ex-ante optimal choice, the Sender cannot gain from ambiguous persuasion. The same results can be archived without commitment if the Receiver uses the updating rule of Hanany and Klibanoff (2007). These updating rules restrict the interim belief set to beliefs that maintain the ex-ante optimality.

Our approach follows the idea of Riedel et al. (2018). They discuss why a dynamically consistent agent expands his ex-ante belief set to a rectangular hull. Intuitively, an agent who knows that he receives further information before deciding should take his knowledge about the information structure into account. Therefore, different information structures may induce different ex-ante belief sets. In our setting, this occurs if one compares the ex-ante belief (set) without any additional information and the ex-ante belief set in the presence of an ambiguous communication device.

The consistent planning approach of Siniscalchi (2011) is another way to deal with dynamically inconsistent behavior. Following the idea of Strotz (1955), a player considers that his future selves will have different worst-case beliefs. However, even if the interpretation is different, it would lead to similar optimal actions but different beliefs of the Receiver as in our setting.

Chapter 3

Dynamic Consistency in Ambiguous Dutch Auctions

3.1 Introduction

In the canonical model with subjective expected utility maximizers and independent private values, it is well-known that the descending price (or Dutch) auction and the first-price sealed-bid auction generate the same equilibrium outcomes. However, this result breaks down if buyers are non-expected utility maximizers. In practice, buyers usually have only little information about the valuation of their opponents. Therefore, it is an important and interesting question how buyers bid if they are faced with ambiguous beliefs about the valuation of their opponents.

Bose and Daripa (2009) address this question and analyze a discrete decreasing price auction with two ambiguity-averse buyers with valuation $v \in [0,1]$. In their setting, buyers can not fix one subjective belief about the valuation of the other buyer. The beliefs of both buyers are given by a set of density functions. More precisely, there is one underlying density function f with full support in $[0,1]$ and the set of beliefs is an ϵ-contamination of f:

$$\Phi^{-1} := \{(1-\epsilon)f(\cdot) + \epsilon l(\cdot) \, : \, l \in \mathcal{P}\},$$

where $\mathcal{P}$ denotes the set of all density functions on $[0,1]$. Intuitively, $1-\epsilon$ can be interpreted as the confidence in the subjective belief f, or ϵ as the degree of ambiguity.[1] Further, buyers are ambiguity-averse and maximize their worst-case expected utility á la Gilboa and Schmeidler (1989). Bose and Daripa (2009) consider the following modified Dutch mechanism (MDM): The seller starts with a price close to one. At the

[1]For more details and an axiomatization of ϵ-contamination see Kopylov (2016).

beginning of a period k, the seller chooses secretly and randomly one buyer. This buyer is approached first and can either accept or reject the offer at price p_k. If he rejects the offer, the other buyer can accept or reject the same offer. If one of the buyers accepts, the game ends immediately. If both buyers reject the price p_k, the game proceeds to the next period $k+1$ with a lower price p_{k+1}. The procedure is repeated until either a buyer accepts or the last price p_n is reached. Buyers never learn who received which offer first.

For a given MDM, the buyers face an incomplete information game with two players and ambiguous information about the valuation of the other buyer. Ambiguity may lead to dynamically inconsistent behavior. Bose and Daripa (2009) use the consistent planning approach of Siniscalchi (2011). They show that for any degree of ambiguity ϵ, the seller can always design an MDM such that he can extract almost all surplus. Intuitively, consider a price p_k and buyer i with value $v^i > p_k$ and suppose buyer i gets the price offer p_k. His ex-post utility of accepting this price is $v^i - p_k$. Since buyers are ambiguity-averse, the expected utility of waiting one period is $v^i - p_{k+1}$ times the worst-case belief of receiving the offer p_{k+1} if he rejects the current price p_k. In the worst-case, it is relatively likely that the opponent buyer j ends the game before buyer i receives the offer p_{k+1}. If the price difference is small enough, the expected utility of waiting becomes smaller than the utility of accepting the current price $v^i - p_k$. Therefore, the buyers accept a price that is very close to their valuation. Then, by making the price difference small enough, the seller can extract almost all surplus.

However, almost all results of Bose and Daripa (2009) are based on an incorrect worst-case belief. In this chapter, we first correct the worst-case belief of Bose and Daripa (2009). Then, we show that the seller can still extract almost all surplus even if buyers behave dynamically consistently. In contrast to Bose and Daripa (2009), we assume that buyers have rectangular beliefs instead of using the consistent planning approach of Siniscalchi (2011). To our knowledge, we are the first who analyze sequential auctions with ambiguity-averse buyers and rectangular beliefs.

Even if the definition of rectangularity is similar to Chapter 1, we can not apply the results of Chapter 1. Since the buyers never know which buyer receives the current offer first, the incomplete information game induced by an MDM cannot be represented by a multistage game.

It is often conjectured that rectangularity and consistent planning lead to similar equilibrium outcomes. However, we show that this is not the case for the incomplete information game induced by an MDM. Independently of the approach, the seller can extract almost all surplus by making the price difference arbitrarily small. However, we show that the equilibrium strategies of buyers with rectangular worst-case beliefs differ from the equilibrium strategies of buyers, which use the consistent planning approach.

This chapter is organized as follows: First, we discuss the related literature. In Section 3.2, we define the modified Dutch mechanism, strategies, beliefs, and formulate the information structure of the game. Further, in Section 3.2.5, we explain the mistake of the worst-case belief of Bose and Daripa (2009) and derives the corrected worst-case belief. In Section 3.3, we analyze the modified Dutch mechanism with dynamically consistent buyers. We first define rectangularity and derive the rectangular worst-case belief. Then, we show the surplus extraction result with rectangular beliefs in Section 3.3.3. Section 3.4 illustrates the results with a numerical example. The different equilibrium predictions of rectangularity and consistent planning are discussed in Section 3.5. Finally, Section 3.6 concludes and discusses future research.

Related Literature There is vast literature on auction design with expected utility maximizers. Among others Myerson (1981) and Riley and Samuelson (1981) discuss auctions with risk-neutral and Matthews (1983) and Maskin and Riley (1984) with risk-averse buyers.

Karni (1988) shows that first-price sealed-bid auctions and decreasing price auctions are equivalent if and only if buyers are dynamically consistent. Our result does not contradict the work of Karni (1988). Even if buyers with rectangular beliefs behave dynamically consistently, the rectangular belief sets depend on the information structure. Therefore, in a Dutch auction, rectangularity will lead to different belief sets than a first-price auction. Dynamically consistent behavior can occur, even if the different auctions are not equivalent.

Lucking-Reiley (1999) provides an interesting field experiment. He compares the revenue of two dynamic auctions (the English and Dutch auction) with the revenue of two static auctions (first- and second-price auction). The canonical model with expected utility maximizers predicts that the English auction (Dutch auction) and the second-price auction (first-price auction) are strategically equivalent. In his experiment, the English auction and the second-price auction generate almost the same revenues for the seller. In contrast, the Dutch auction generates 30 percent higher revenues than the first-price auction. This result is in line with our theory.

The literature on ambiguous auction increased in the last years. Among others, Lo (1998) analyzes first and second-price auction with ambiguity-averse buyers. Bose et al. (2006) study static auctions with ambiguity-averse buyers and seller. Di Tillio et al. (2016) consider a screening model with one agent and one principle. The valuation is privately known to the agent. Ambiguity arises due to an ambiguous mechanism. The principal can design a set of mechanisms and commit to one without revealing it to the buyer. Di Tillio et al. (2016) show that a seller can increase his profit by using the ambiguous mechanism.

Ghosh and Liu (2020) and Auster and Kellner (2020) investigate sequential auction settings using the consistent planning approach of Siniscalchi (2011). The setting of Ghosh and Liu (2020) differs from our setting. In their model, multiple units of a good are sold to multiple buyers. In each period, each buyer submits a sealed bid simultaneously. The buyer with the highest bid gets one unit of the good and leaves the auction. This procedure is repeated until all units are sold. If two buyers submit the same bid, ties are broken with a coin toss.[2] Similar to our setting Auster and Kellner (2020) analyze a Dutch auction but in continuous time. Further, they allow more general belief sets than the ϵ-contamination. In their setting, the seller cannot extract almost all surplus. However, the Dutch auction still generates a higher surplus than a first-price sealed-bid auction. Further, due to the continuous-time structure, the timing becomes less complex, and consistent planning and rectangular beliefs lead to equivalent equilibrium outcomes in their setting (see Section 3.6).

3.2 Ambiguous Dutch Auction

Our basic setting and the definition of an ambiguous Dutch auction mostly follow the setting of Bose and Daripa (2009).

3.2.1 Basic Setting

There is one seller who wants to sell one indivisible object. The seller's valuation of this object is normalized to zero. Two buyers with valuation $v_i \in [0,1]$ for $i = 1,2$ compete for the object. The own valuation is private information of each buyer. The seller is risk and ambiguity-neutral. He believes that each buyers valuation is drawn from a distribution F with density $f(v) > 0$ for all $v \in [0,1]$. In contrast, the buyers are less confident about the opponent's valuation. Each buyer is risk-neutral, but ambiguity avers about the valuation of the other buyer and maximizes his worst-case expected payoff à la Gilboa and Schmeidler (1989). The set of priors of each buyer is given by an ϵ-contamination of the density f, i.e.,

$$\Phi_i^{-1} := \{g(\cdot) = (1-\epsilon)f(\cdot) + \epsilon l(\cdot) \,:\, l \in \mathcal{P}\},$$

where $\mathcal{P}$ denotes the set of all density functions with support $\text{supp}(l) \subset [0,1]$. Throughout the whole chapter, we will denote with capital letter, F, G, L, and M the distribution functions corresponding to densities f, g, l, and m. Please note, that the

[2]Even if it seems to be a minor difference if ties are broken before or after buyers submit their bits, it influences the equilibrium outcome. We discuss this issue in more detail in Section 3.6.

ϵ-contamination structure of $g \in \Phi_i^{-1}$ implies an ϵ-contamination structure of the corresponding distribution function $G(\cdot) = (1-\epsilon)F(\cdot) + \epsilon L(\cdot)$.

3.2.2 The Modified Dutch Mechanism (MDM)

The timing of the modified Dutch mechanism (MDM) is as follows: At the beginning the seller publicly announces a price sequence $\{p_1, p_2, \ldots, p_n\}$, where p_k denotes the asked price in period k. In period k:

1) The seller tosses a fair coin to decide which buyer to approach first.

2) The buyer chosen in Step 1) gets the offer p_k. If he accepts, he receives the object at the price p_k, and the game is over.

3) If he rejects, the second buyer gets the offer p_k. If he accepts, he receives the object at the price p_k, and the game is over.

4) If the second buyer rejects and $k < n$, the game proceeds to period $k+1$.

This procedure is repeated in each period until either one of the buyers accepts a price or period n is reached. If the buyers do not accept any price $p_1, p_2, \ldots, p_n$, the object remains unsold. The procedure and the price sequence are common knowledge, but the buyers never know the result of the coin toss at Step 1), i.e., they do not know who is approached first.

We will consider the same price sequence as Bose and Daripa (2009), which depends on the degree of ambiguity ϵ and a parameter $\delta \in (0,1)$. We will see later, that the seller can use δ to influence the difference of two consecutive prices and therefore the surplus of the buyers. For $\delta > 0$, let $\{p_0, p_1, p_2, \ldots, p_n\}$ be the price sequence, where

$$p_0 = 1 \quad \text{and}$$
$$p_k = \frac{(1-\delta)^k}{\left(1-\delta+\frac{\epsilon\delta}{2}\right)^{k-1}} \quad \text{for any } k > 0.$$

It is important to note that p_k and the price difference

$$\Delta_0 := p_0 - p_1 = \delta,$$
$$\Delta_k := p_k - p_{k+1} = \left(\frac{1-\delta}{1-\delta+\frac{\epsilon\delta}{2}}\right)^k \frac{\epsilon\delta}{2}$$

are decreasing in k. Bose and Daripa (2009) show that $\lim_{n\to\infty} \sum_{k=0}^{n-1} \Delta_k = 1$. Furthermore, for any given $\eta \in (0,1)$ there exists an integer T such that $\sum_{k=0}^{T} \Delta_k \geq 1-\eta$. Then, let n be the smallest integer for which this inequality is satisfied and p_n the last offered price. This choice of n ensures, that all types above η participate at the auction. The intuition of the surplus extraction result is as follows: The seller can use η and δ to design the price sequence in such a way that almost all types participate and each type obtains a surplus smaller than δ. Hence, by making η and δ arbitrarily small the seller can extract almost all surplus.

3.2.3 Information and Strategies of Buyers

Given an MDM, the strategy $\sigma_i : [0,1] \to \{A, R\}^n$ of a buyer i consists of a strategy for each type. A strategy $\sigma^i(v^i)$ of type v^i is a plan to accept or reject the seller's offer at every price offer of the price sequence given the history of the game so far, i.e., $\sigma^i(v^i) \in \{A, R\}^n$, where A denotes accepting and R rejecting.

Ex-post, the seller's payoff is p_k if the object is sold in period k and zero otherwise. The payoff of buyer i of type v^i is $u_i(\sigma_i(v^i), \sigma_j(v^j)) = v^i - p_k$ if he buys the object in period k and zero if he does not obtain the object. We assume that ex-ante the seller can commit to the mechanism described above, including the price sequence $\{p_1, p_2, \ldots, p_n\}$. Given the mechanism, the setting reduces to an incomplete information game with two players (buyer i and j).

Similar to Bose and Daripa (2009), we define interior cut-off strategies as follows.[3]

Definition 3.1. *A strategy of buyer i, $i \in \{1, 2\}$, is called an interior cut-off strategy if there exists a vector $\mathbf{v}^i = (v_1^i, \ldots, v_n^i), 0 \leq v_n^i \leq v_{n-1}^i \leq \cdots \leq v_1^i \leq 1$, such that for $k \geq 1$, the highest price accepted by the interval of types $[v_k^i, v_{k-1}^i)$ is p_k, where $v_0^i = 1$.*

We will prove later that without loss of generality, we can restrict to interior cut-off strategies. Furthermore, like Bose and Daripa (2009), we assume two simplifying assumptions to solve indifference. If a buyer is indifferent between accepting and rejecting, he accepts the price. Further, if a buyer is indifferent between buying in two different periods, he buys in the earlier period. Thus, buyers choose the seller-preferred action in case of indifference.

As described above, the valuation of each buyer is private information, and buyers are ambiguity-averse. Receiving a price offer p_k reveals information about the valuation of the other buyer. Let us assume that both buyers play interior cut-off strategies. If buyer i gets the price offer p_k, he does not know if buyer j already got the offer p_k and rejected it or if he is asked first. Getting the offer p_k only reveals that buyer j rejected all prices before p_k. Therefore, the type of buyer j has to be smaller than v_{k-1}^j, which is the lowest type who accepts p_{k-1}. The information partition of the state space $\Omega = [0, 1]$ which is induced by the MDM and an interior cut-off strategy of the

[3] Please note, that our definition differs from Bose and Daripa (2009) since we do not require a strict inequality $v_k^i < v_{k-1}^i$ for all $k = 1, \ldots, n$.

opponent $\mathbf{v}^j$ is as follows

$$
\begin{aligned}
I^i_{-1} = I^i_0 = \Omega&,\\
I^i_1 &= \left\{[0, v^j_0), [v^j_0, 1]\right\} = \{[0,1), \{1\}\},\\
I^i_k &= \left\{[0, v^j_{k-1}), [v^j_{k-1}, v^j_{k-2}), \ldots, [v^j_1, v^j_0), \{1\}\right\},
\end{aligned}
$$

where $k = -1$ denotes the ex-ante stage after the MDM is announced but before the first price is offered.

3.2.4 Updating Beliefs and Worst-Case Belief

Assume player i gets the offer p_k. To decide whether he should reject or accept the price, he compares the expected payoff of waiting one period and the payoff of accepting p_k. The expected payoff of waiting one period depends on the worst-case belief that he gets the offer p_{k+1} given that he rejects p_k.

Before we derive and correct the worst-case belief of Bose and Daripa (2009), we have to specify how ambiguity avers players update their beliefs. We assume that both buyers use prior-by-prior Bayesian updating, i.e., they update each belief in Φ_i^{-1} using Bayes' rule. The prior-by-prior Bayesian update of Φ_i^{-1} in period k given the event $E \in I^i_k$ is

$$\Phi_i^{k,E} = \{\mathrm{Bay}(g|E) \,:\, g \in \Phi_i^{-1}\},$$

where $\mathrm{Bay}(g|E)$ denotes the Bayesian update of g given the event E

$$\mathrm{Bay}(g|E)(\cdot) = \frac{g(\cdot)}{g(E)}.$$

Note, that

$$g_i^{k-1}(E) = 0$$

for all $E \in \left\{[v^j_{l-2}, v^j_{l-3})\right\}_{l=3,\ldots,k} \cup \{\{1\}\}$ and $g_i^{k-1} \in \Phi_i^{k-1,E'}$ with $E' \in I^i_{k-1}$. However, these are the events that imply that the game already ended before period k. Therefore, the information sets that influence the decision in period k are only the first two sets of I^i_k, i.e., $[0, v^j_{k-1})$ and $[v^j_{k-1}, v^j_{k-2})$. Due to the full support assumption on f and since $\epsilon < 1$, these two information sets have a strictly positive probability for any $g_i^{k-1} \in \Phi_i^{k-1,E'}$ and $E' \in I^i_{k-1}$ and Bayes' rule is well defined.

Formally, there exists a set of interim beliefs $\Phi_i^{k,E}$ in period k for each event $E \in I^i_k$. However, for all events except $[0, v^j_{k-1})$, the game already ended before period k. To simplify notation, we sometimes denote with Φ_i^k the set of interim beliefs given the event $[0, v^j_{k-1})$.

Lemma 3.1. *The ϵ-contamination structure is maintained under Bayesian updating. Let $E \subseteq \Omega$. Then,*

$$\begin{aligned}\Psi^1 &:= \left\{ Bay(g|E) \,:\, g \in \Phi_i^{-1} \right\} \\ &= \left\{ (1-\epsilon_k) Bay(f|E) + \epsilon_k Bay(l|E) \,:\, l \in \mathcal{P} \right\} =: \Psi^2 \end{aligned} \tag{3.1}$$

with $\epsilon_k = \frac{\epsilon}{(1-\epsilon)f(E)+\epsilon} > \epsilon$ for all $k \geq 1$ and $\epsilon_0 = \epsilon$.

The proof of Lemma 3.1 can be found in Section 3.7.2.1 in the Appendix.

Corollary 3.1. *In each period $k \geq 1$, the set of updated beliefs is given by*

$$\Phi_i^k = \left\{ (1-\epsilon_k) f^k(\cdot) + \epsilon_k l^k(\cdot) \,:\, l \in \mathcal{P} \right\}$$

with

$$f^k(\cdot) := \frac{f(\cdot)}{F(v_{k-1}^j)}, \qquad l^k(\cdot) := \frac{l(\cdot)}{L(v_{k-1}^j)}$$

and $\epsilon_k = \frac{\epsilon}{(1-\epsilon)F(v_{k-1}^j)+\epsilon}$. Furthermore, $\epsilon_k \geq \epsilon_{k+1}$ for all $k = 1, \ldots n$.

Proof. The result follows immediately from Lemma 3.1 with $E = [0, v_{k-1}^j)$. Furthermore, $[0, v_k^j) \subset [0, v_{k-1}^j)$ and the full support assumption on f imply $F(v_k^j) < F(v_{k-1}^j)$. Therefore, $\epsilon_k \geq \epsilon_{k+1}$ for all $k = 1, \ldots n$. □

3.2.5 Worst-Case Belief of Bose and Daripa (2009)

The entire analysis of Bose and Daripa (2009) is based on the following lemma. It formalizes the observation that in each period given that buyer i gets the offer p_k he compares the expected payoff of accepting p_k with the expected payoff of waiting and accepting p_{k+1}.

Lemma 3.2 (Lemma 1 Bose and Daripa (2009)). *Suppose the item has not been sold in periods $1, \ldots, k-1$ and in period $k < n$ the seller offers the item to buyer i at price p_k. Suppose j follows an interior cut-off strategy that gives rise to a vector of cut-offs $\mathbf{v}^j = (v_1^j, \ldots, v_n^j)$. For any type v of i, the difference in payoff from buying immediately at price p_k versus waiting one period to buy at price p_{k+1} is*

$$G_k^i(v) = v - p_k - (v - p_{k+1})\bar{H}_k^i,$$

where $\bar{H}_k^i$ is the worst-case belief of buyer i that he will get the offer p_{k+1} if he rejects p_k.

Bose and Daripa (2009) claim that the worst-case belief $\bar{H}_k^i$ of getting offer p_{k+1} if he rejects p_k is given by $(1-\epsilon)H_k^i$, where

$$H_k^i = \frac{F(v_k^j) + F(v_{k+1}^j)}{F(v_k^j) + F(v_{k-1}^j)}.$$

However, this worst-case belief is not correct. Let A denote the event i *obtains the item at* p_{k+1} and B the event i *refuses the current offer of* p_k. Then, Bose and Daripa (2009) derive H_k^i by claiming that[4]

$$\min_{g^k\in\Phi_i^k} \mathbb{P}^{g^k}(A|B) = \min_{l^k\in \text{Bay}(\mathcal{P}|v<v_{k-1}^j)} (1-\epsilon)\mathbb{P}^{f^k}(A|B) + \epsilon\mathbb{P}^{l^k}(A|B). \tag{3.2}$$

But since $A, B \not\subset \Omega$, Lemma 3.1 and Equation (3.2) do not hold.

Lemma 3.3. *The worst-case belief* $\bar{H}_k^i$ *of Lemma 3.2 is given by*

$$\bar{H}_k^i = \frac{(1-\epsilon)(F(v_k^j) + F(v_{k+1}^j))}{(1-\epsilon)(F(v_k^j) + F(v_{k-1}^j)) + \epsilon}.$$

Further, there exists no density f *with full support on* $[0,1]$ *such that* $\bar{H}_k^i = (1-\epsilon)H_k^i$ *for all* $k = 1, \ldots, n-1$.

Proof. We first derive the worst-case belief $\bar{H}_k^i = \min_{g\in\Phi_i^k} \mathbb{P}^{g^k}(A|B)$. Then, we show that $\bar{H}_k^i$ differs from $(1-\epsilon)H_k^i$. For an arbitrary fixed $g^k \in \Phi_i^k$ one can calculate $\mathbb{P}^{g^k}(A|B)$ analogously to H_k^i in Appendix A.1. of Bose and Daripa (2009).[5] Using $g^k(\cdot) = \frac{g(\cdot)}{g(v<v_{k+1}^j)}$, we can rewrite $\min_{g\in\Phi_i^k} \mathbb{P}^{g^k}(A|B)$ to

$$\min_{g\in\Phi_i^{-1}} \mathbb{P}^{g}(A|B) = \min_{g\in\Phi_i^{-1}} \frac{G(v_k^j) + G(v_{k+1}^j)}{G(v_k^j) + G(v_{k-1}^j)},$$

where G denotes the distribution function of g. The ϵ-contamination structure of G gives

$$\mathbb{P}^{g}(A|B) = \frac{(1-\epsilon)(F(v_k^j) + F(v_{k+1}^j)) + \epsilon(L(v_k^j) + L(v_{k+1}^j))}{(1-\epsilon)(F(v_k^j) + F(v_{k-1}^j)) + \epsilon(L(v_k^j) + L(v_{k-1}^j))}. \tag{3.3}$$

Then, a worst-case belief $g^* = \arg\min_{g\in\Phi_i^{-1}} \mathbb{P}^g(A|B)$ has to satisfy $L^*(v_{k+1}^j) = L^*(v_k^j) = 0 < 1 = L^*(v_{k-1}^j)$, which is well defined since $v_{k-1}^j \geq v_k^j \geq v_{k+1}^j$. $L^*(v_{k+1}^j) = 0$ and $L^*(v_{k-1}^j) = 1$ follows immediately from Equation (3.3). Furthermore, $v_{k-1}^j \geq v_{k+1}^j$ implies that $\mathbb{P}^g(A|B)$ is monotone increasing in $L^*(v_k^j)$. Therefore, $L^*(v_k^j) = 0$. Then,

$$\bar{H}_k^i = \mathbb{P}^{g^*}(A|B) = \frac{(1-\epsilon)(F(v_k^j) + F(v_{k+1}^j))}{(1-\epsilon)(F(v_k^j) + F(v_{k-1}^j)) + \epsilon}.$$

$\bar{H}_k^i$ is smaller than $(1-\epsilon)H_k^i$ if $F(v_k^j) + F(v_{k-1}^j) < 1$ and greater than $(1-\epsilon)H_k^i$ if $F(v_k^j) + F(v_{k-1}^j) > 1$. Hence, $\bar{H}_k^i = (1-\epsilon)H_k^i$ if and only if $F(v_k^j) + F(v_{k-1}^j) = 1$. But, this condition is cannot be satisfied for all $k = 1, \ldots, n$. □

[4]Please note, that the worst-case expected payoff of waiting equals $\min_{G\in\mathcal{P}_B} \mathbb{P}^g(A|B)(v - p_{k+1}) + (1 - \mathbb{P}^g(A|B)) \cdot 0$. Therefore, minimizing the expected payoff is equivalent to minimize $\mathbb{P}^g(A|B)$.

[5]For completeness we derive $\mathbb{P}^{g^k}(A|B)$ in Section 3.7.1 ind the Appendix.

Almost all proofs of Bose and Daripa (2009) build upon the fact that the worst-case belief is given by $(1-\epsilon)H_k^i$ and $H_k^i < 1$ for all $k = 1, \dots, n-1$. However, there might exist $k \in 1, \dots, n-1$ such that the corrected worst-case belief $\bar{H}_k^i > 1-\epsilon$. Therefore, the correction of the proofs of Bose and Daripa (2009) is not straightforward. A more detailed discussion on the correction of the results of Bose and Daripa (2009) can be found in Section 3.7.3 in the Appendix.

3.3 Dynamically Consistent Buyers

Epstein and Schneider (2003) define rectangularity (or stability under pasting) as a condition on ambiguous beliefs that ensures dynamically consistent behavior. We will see that rectangular beliefs will lead to different equilibrium strategies as consistent planning. However, we show that the seller can still extract almost all surplus. In Section 3.5, we discuss the different implications of rectangularity and the consistent planning approach of Siniscalchi (2011) in more detail.

Now, we define rectangular beliefs for the incomplete information game induced by an MDM. Rectangularity ensures that a buyer takes the possible future worst-case beliefs into account and therefore behaves dynamically consistently. Then, we derive the rectangular worst-case belief of a buyer and discuss surplus extraction with rectangular beliefs.

3.3.1 Rectangularity

The basic idea of rectangularity is that agents take their possible future worst-case beliefs into account. Let us first consider the case without ambiguity, i.e., Φ_i^{-1} consists of a single prior belief g. For an arbitrary event $E \in I_k^i$, $g(E)$ is the marginal probability of the event E. Further, remember that $\text{Bay}(g|E)(\cdot)$ denotes the Bayesian update of g given the event E. Then, Bayes' rule implies that multiplying the marginal probability $g(E)$ with the updated belief $\text{Bay}(g|E)(\cdot)$ and taking the sum over all events in I_k^i, generates the prior belief g

$$g(\cdot) = \sum_{E \in I_k^i : g(E) > 0} g(E)\text{Bay}(g|E)(\cdot).$$

Under ambiguity, the set of beliefs Φ_i^{-1} is not a singleton. Rectangularity generalizes the above considerations to ambiguous settings. It ensures that the combination of any marginal and updated beliefs are an element of the ex-ante belief set.

Definition 3.2. *The pasting of an ex-ante belief* $\bar{g} \in \Phi_i^{-1}$ *and a collection of updated beliefs* $(g_E)_{E \in I_k^i} \in \times_{E \in I_k^i} \Phi_i^{k,E}$ *is defined as*

$$\bar{g} \circ (g_E)_{E \in I_k^i}(\cdot) := \sum_{E \in I_k^i} \bar{g}(E) g_E(\cdot).$$

The set of ex-ante beliefs is called ***rectangular*** *(or stable under pasting) if it contains any pasting of an ex-ante belief and interim beliefs, i.e.,*

$$\bar{g} \circ (g_E)_{E \in I_k^i} \in \Phi_i^{-1}$$

for all $\bar{g} \in \Phi_i^{-1}$ *and* $(g_E)_{E \in I_k^i} \in \times_{E \in I_k^i} \Phi_i^{k,E}$.

An arbitrary ex-ante belief set does not have to satisfy rectangularity. However, Epstein and Schneider (2003) and Chapter 1 show that given an arbitrary ex-ante belief set Φ_i^{-1} there always exists a rectangular hull of Φ_i^{-1}, denoted by $\text{rect}(\Phi_i^{-1})$. The rectangular hull is the smallest set of density functions that contains Φ_i^{-1} and satisfies rectangularity. To construct the rectangular hull, one starts with the set of beliefs at the terminal period n and period $n-1$ and constructs the rectangular hull by combining any marginal probabilities at $n-1$ with any updated belief at n. Then, given the rectangular hull in period $n-1$, one proceeds by backward induction. Trivially, if Φ_i^{-1} is rectangular, then $\Phi_i^{-1} = \text{rect}(\Phi_i^{-1})$.

There are two important properties of rectangularity, that are also used for the construction of the rectangular hull. First, the set of Bayesian updates under the rectangular hull $\text{rect}(\Phi_i^{-1})$ equals the set of Bayesian updates under Φ_i^{-1}, i.e., for all $k = 0, \dots, n$ and $F \in I_k^i$

$$\{\text{Bay}(g|F)(\cdot) \, : \, g \in \Phi_i^{-1}\} = \{\text{Bay}(g|F)(\cdot) \, : \, g \in \text{rect}(\Phi_i^{-1})\}.$$

Second, an analogous statement holds for the set of marginal beliefs $g(F)$ for all $F \in I_k^i$ and $k = -1, \dots n$ of the rectangular hull $\text{rect}(\Phi_i^{-1})$ and Φ_i^{-1}. Epstein and Schneider (2003) and Chapter 1 discuss the construction and properties of the rectangular hull in more detail.

3.3.2 Rectangular Worst-Case Beliefs

The probability $\mathbb{P}^g(A|B)$ of getting the offer p_{k+1} if buyer i rejects the current price p_k, depends on two events. First, with positive probability buyer j gets the offer p_k after buyer i rejected p_k. If buyer j accepts the offer p_k, the game is over and i does not get the offer p_{k+1}. Second, if buyer j rejects p_k, period $k+1$ is reached. In period $k+1$, with positive probability buyer j gets the offer p_{k+1} first. If he accepts p_{k+1} the game ends and buyer i does not receive the offer p_{k+1}. Hence, getting the offer p_{k+1}

if buyer i rejects the current price p_k depends on the probability that j does neither accept p_k nor p_{k+1} (given that period $k+1$ is reached).

Therefore, the worst-case probability $\mathbb{P}^g(A|B)$ depends on the worst-case belief in period k and the worst-case belief in period $k+1$. But, with ambiguous beliefs the worst-case belief in period $k+1$ is in general not the Bayesian update of the worst case belief in period k. Intuitively, in period k if buyer j plays an interior cut-off strategy, the worst-case of buyer i is that buyer j accepts the current price p_k. Therefore, the worst-case belief in period k is $g^{*,k}(v^j) = (1-\epsilon_k)f^k(v^j) + \epsilon_k l^k(v^j)$ with $l^k([v^j_k, v^j_{k-1}]) = 1$. Similar, in period $k+1$ the worst-case belief is $g^{*,k+1}(v^j) = (1-\epsilon_{k+1})f^{k+1}(v^j) + \epsilon_{k+1}\bar{l}^{k+1}(v^j)$ with $\bar{l}^{k+1}([v^j_{k+1}, v^j_k]) = 1$. Then, $g^{*,k+1}$ does not equal the Bayesian update of $g^{*,k}$ in period $k+1$.

The consistent planning approach of Bose and Daripa (2009) and the worst-case belief of Lemma 3.2 assume that the worst-case probability $\mathbb{P}^g(A|B)$ only depends on the worst-case beliefs in period k. Therefore, it neither takes into account that $\mathbb{P}^g(A|B)$ depends on the worst-case belief in period k and $k+1$ nor that the worst-case belief in period $k+1$ does not equal the Bayesian update of the worst-case belief in period k.

Rectangularity takes this change of the worst-case belief into account. The rectangular worst-case probability of getting the offer p_{k+1} if he rejects p_k depends on the worst-case belief of period k and the worst-case belief of period $k+1$. Therefore, it allows for a change in the worst-case belief between period k and $k+1$. The following proposition formally characterizes the rectangular worst-case belief if the buyers follow an interior cut-off strategy.

Proposition 3.1. *Suppose the ex-ante belief set of buyer j is given by rect(Φ_i^{-1}). Further, suppose the item has not been sold in periods $1,\dots,k-1$ and in period $k<n$ the seller offers the item to buyer i at price p_k. Suppose j follows an interior cut-off strategy that gives rise to a vector of cut-offs $\mathbf{v}^j = (v^j_1,\dots,v^j_n)$. For any type v of i the difference in payoff from buying immediately at price p_k versus waiting one period to buy at price p_{k+1} is*

$$G^i_k(v) = v - p_k - (v - p_{k+1})\hat{H}^i_k, \tag{3.4}$$

where

$$\hat{H}^i_k = (1-\epsilon_k)\frac{F^k(v^j_k) + (1-\epsilon_{k+1})F^k(v^j_k)F^{k+1}(v^j_{k+1})}{1+(1-\epsilon_k)F^k(v^j_k)}. \tag{3.5}$$

Proposition 3.1 is important for the equilibrium analysis. To characterize equilibrium strategies, we only have to compare the payoffs from buying immediately with the payoff of waiting one period and accept the next price. Type v accepts the price p_k if

$G_k^i(v) \geq 0$. Since $G_k^i(v)$ is monotone increasing in v, a type v who does not accept p_{k+1} would never accept the higher price p_k. Thus, we do not need to consider any strategy that involves waiting for more than one period.

Proof. The payoff of accepting p_k is $v - p_k$. If buyer i rejects the offer p_k, he gets a payoff of $v - p_{k+1}$ if he gets and accepts the offer p_{k+1}. If the game ends before he gets the offer p_{k+1}, his payoff is zero. Further, $\hat{H}_k^i$ is the probability that i gets the offer p_{k+1} if he rejects p_k. Therefore, $G_k^i(v)$ is given by Equation (3.4). It is left to show that $\hat{H}_k^i$ is specified as above.

In Section 3.7.1 in the Appendix, we show that for a fixed belief $g \in \Phi_i^{-1}$ the probability of getting offer p_{k+1} if he rejects p_k is

$$\mathbb{P}^g(A|B) = \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)}.$$

The rectangular worst-case belief is then given by

$$\hat{H}_k^i = \min_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)} \tag{3.6}$$

where G^k is the distribution function corresponding to g^k and G^{k+1} the distribution function corresponding to the Bayesian update $\mathrm{Bay}(g^k | v < v_k^j)$ at $k+1$. Since $\Phi_i^{-1} \subseteq \mathrm{rect}(\Phi_i^{-1})$, we first have to show that the distribution functions G^k and G^{k+1} can be represented by an ϵ-contamination of the distribution function F.

Note that for any arbitrary $g^k \in \mathrm{rect}(\Phi_i^k)$, $G^k(v_k^j) = g^k(E)$ with $E = [0, v_k^j] \in I_{k+1}^i$ and $G^{k+1}(v_{k+1}^j) = \mathrm{Bay}(g|E)(E')$ with $E' = [0, v_{k+1}^j] \in I_{k+2}^i$. The sets of marginal and updated beliefs given the rectangular hull of Φ_i^k equal the sets of marginal and updated beliefs given Φ_i^k. Therefore, there exists $m^k \in \Phi_i^k$ and $l^k \in \mathrm{Bay}(\mathcal{P}|E)$ such that

$$G^k(v_k^j) = g^k(E) = m^k(E) = M^k(v_k^j) = (1 - \epsilon_k)F^k(v_k^j) + \epsilon_k L^k(v_k^j), \tag{3.7}$$

where the last step follows since any density function in Φ_i^k can be represented as ϵ-contamination of f. Similar, for G^{k+1}, there exists $\bar{m}^k \in \Phi_i^k$ and $\bar{l}^k \in \mathrm{Bay}(\mathcal{P}|E')$ such that

$$\begin{aligned} G^{k+1}(v_{k+1}^j) &= \mathrm{Bay}(g^k|E)(E') = \mathrm{Bay}(\bar{m}^k|E)(E') = \bar{M}^{k+1}(v_{k+1}^j) \\ &= (1 - \epsilon_{k+1})F^{k+1}(v_{k+1}^j) + \epsilon_{k+1}\bar{L}^{k+1}(v_{k+1}^j). \end{aligned} \tag{3.8}$$

Now, dividing denominator and nominator by $G^k(v_k^j)$, we can rewrite Equation (3.6):

$$\hat{H}_k^i = \min_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)}$$

$$
\begin{aligned}
&= \min_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{1 + G^{k+1}(v_{k+1}^j)}{\frac{1}{G^k(v_k^j)} + 1} \\
&= \min_{l \in \mathrm{Bay}(\mathcal{P}|E), \bar{l} \in \mathrm{Bay}(\mathcal{P}|E')} \frac{1 + (1-\epsilon_{k+1})F^{k+1}(v_{k+1}^j) + \epsilon_{k+1}\bar{L}^{k+1}(v_{k+1}^j)}{\frac{1}{(1-\epsilon_k)F^k(v_k^j) + \epsilon_k L^k(v_k^j)} + 1} \\
&= \frac{1 + (1-\epsilon_{k+1})F^{k+1}(v_{k+1}^j)}{\frac{1}{(1-\epsilon_k)F^k(v_k^j)} + 1} \\
&= \frac{(1-\epsilon_k)F^k(v_k^j) + (1-\epsilon_k)(1-\epsilon_{k+1})F^k(v_k^j)F^{k+1}(v_{k+1}^j)}{1 + (1-\epsilon_k)F^k(v_k^j)} \\
&= (1-\epsilon_k)\frac{F^k(v_k^j) + (1-\epsilon_{k+1})F^k(v_k^j)F^{k+1}(v_{k+1}^j)}{1 + (1-\epsilon_k)F^k(v_k^j)},
\end{aligned}
$$

where the third step follows from the ϵ-contamination structure of Equation (3.7) and Equation (3.8) and the fourth step since $\bar{L}^{k+1}(v_{k+1}^j) = L^k(v_k^j) = 0$ minimizes the equation. □

The following lemma shows that rectangular beliefs lead to dynamically consistent behavior.

Lemma 3.4. *With rectangular beliefs the worst-case belief in period $k+1$ is the Bayesian update of the worst-case belief in period k. Therefore,*

$$
\min_{g^k \in rect(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)} = \min_{g^k \in \Phi_i^k, \bar{g}^{k+1} \in \Phi_i^{k+1}} \frac{G^k(v_k^j) + \bar{G}^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)}.
$$

Proof. First, we show that the left-hand side is greater or equal than the right-hand side. Let

$$
g^* \in \operatorname*{arg\,min}_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)}.
$$

denote a rectangular worst-case belief in period k. Further, let $G^{*,k}(v_k^j) = g^*(E)$ with $E = [0, v_k^j] \in I_{k+1}^i$ and $G^{*,k+1}(v_{k+1}^j) = \mathrm{Bay}(g^*|E)(E')$ with $E' = [0, v_{k+1}^j] \in I_{k+2}^i$. Then, similar to the proof of Proposition 3.1, there exists $m^k, \bar{m}^k \in \Phi_i^k$ such that

$$
\begin{aligned}
G^{*,k}(v_k^j) &= g^*(E) = m^k(E) = M^k(v_k^j), \\
G^{*,k+1}(v_{k+1}^j) &= \mathrm{Bay}(g^*|E)(E') = \mathrm{Bay}(\bar{m}^k|E)(E') = \bar{M}^{k+1}(v_{k+1}^j)
\end{aligned}
$$

and

$$
\begin{aligned}
\min_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)} &= \frac{M^k(v_k^j) + \bar{M}^{k+1}(v_{k+1}^j)M^k(v_k^j)}{1 + M^k(v_k^j)} \\
&\geq \min_{g^k \in \Phi_i^k, \bar{g}^{k+1} \in \Phi_i^{k+1}} \frac{G^k(v_k^j) + \bar{G}^{k+1}(v_{k+1}^j)G^k(v_k^j)}{1 + G^k(v_k^j)}.
\end{aligned}
$$

To prove that the right-hand side is greater or equal than the left-hand side, let

$$(\bar{g}^{*,k}, \bar{g}^{*,k+1}) \in \underset{g^k \in \Phi_i^k, \bar{g}^{k+1} \in \Phi_i^{k+1}}{\arg\min} \frac{G^k(v_k^j) + \bar{G}^{k+1}(v_{k+1}^j) G^k(v_k^j)}{1 + G^k(v_k^j)}.$$

The definition of rectangularity implies that $\bar{g}^{*,k} \circ (\bar{g}_E^{*,k+1})_{E \in I_{k+1}^i} \in \mathrm{rect}(\Phi_i^k)$. Then

$$\min_{g^k \in \Phi_i^k, \bar{g}^{k+1} \in \Phi_i^{k+1}} \frac{G^k(v_k^j) + \bar{G}^{k+1}(v_{k+1}^j) G^k(v_k^j)}{1 + G^k(v_k^j)} \geq \min_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j) G^k(v_k^j)}{1 + G^k(v_k^j)}.$$

Combining both directions gives

$$\begin{aligned} \hat{H}_k^i &= \min_{g^k \in \Phi_i^k, \bar{g}^{k+1} \in \Phi_i^{k+1}} \frac{G^k(v_k^j) + \bar{G}^{k+1}(v_{k+1}^j) G^k(v_k^j)}{1 + G^k(v_k^j)} \\ &= \min_{g^k \in \mathrm{rect}(\Phi_i^k)} \frac{G^k(v_k^j) + G^{k+1}(v_{k+1}^j) G^k(v_k^j)}{1 + G^k(v_k^j)}. \end{aligned}$$

□

Hence, due to rectangularity, the interim worst-case belief is the Bayesian update of the ex-ante worst-case belief. Therefore, rectangularity implies dynamically consistent behavior. The next subsection shows that the seller can extract almost all surplus, even if the buyers behave dynamically consistently.

3.3.3 Surplus Extraction under Dynamic Consistency

The rectangular worst-case belief differs from the worst-case belief of Bose and Daripa (2009). This subsection shows that the main result of Bose and Daripa (2009) still holds if buyers behave dynamically consistently. We first characterize and define equilibrium strategies under rectangular beliefs if δ is sufficiently small. The equilibrium existence follows from the results of Bose and Daripa (2009). In contrast to Bose and Daripa (2009), under rectangular beliefs, there might exist prices such that no type of buyer i or j plan to buy at these prices. However, given the characterization of the equilibrium strategies, we can still show that the seller can extract almost all surplus even if the buyers behave dynamically consistently.

Let us start with the definition and characterization of equilibrium strategies. Remember, that a strategy σ_i of buyer i specifies for each type a plan to reject or accept the seller's offer at every information set, $\sigma_i : [0,1] \to \{A, R\}^n$, where A denotes accepting and R rejecting. Further, $u_i(\sigma_i(v^i), \sigma_j(v^j))$ is the ex-post payoff of buyer i of the strategy profile (σ_i, σ_j) if the types are given by v^i and v^j.

Definition 3.3. *A (perfect Bayesian) equilibrium with rectangular beliefs consists of a set of beliefs for each information set and a strategy profile* (σ_1^*, σ_2^*) *such that*

- *the ex-ante belief sets equal the rectanuglar hull* $rect(\Phi_i^{-1})$ *for* $i = 1, 2$,
- *prior-by-prior Bayesian updating is applied whenever possible, i.e., the set of interim beliefs conditional on the price offer* p_k *are* $rect(\Phi_i^k)$ *for* $k = 0, \ldots, n$ *and* $i = 1, 2$*, and*
- *given the rectangular belief sets at each information set the equilibrium strategies are a best response for each type, i.e., for all* $i \in \{1, 2\}$ *and* $v^i \in [0, 1]$

$$\min_{g \in rect(\Phi_i^k)} \mathbb{E}^g(u_i(\sigma_i^*(v^i), \sigma_j^*)) \geq \min_{g \in rect(\Phi_i^k)} \mathbb{E}^g(u_i(\sigma_i(v^i), \sigma_j^*)) \quad \forall \sigma_i(v^i) \in \{A, R\}^n$$

for all $k = -1, \ldots n$.

Further, we call an equilibrium symmetric if $\sigma_i = \sigma_j$.[6]

So far, our definition of an interior cut-off strategy does not specify any out-off equilibrium behavior. However, as in Bose and Daripa (2009), we can extend an interior cut-off strategy to a perfect cut-off strategy as follows.

Remark 3.1. *Let* $p(v)$ *denote the highest price that a buyer of type* v *accepts. By monotonicity, type* v *would also accept any price smaller than* $p(v)$*. A perfect cut-off strategy is an interior cut-off strategy with the additional requirement that a type who accepts* p_k *does accept every price* p_l *with* $l \geq k$.

The following proposition shows that for δ sufficiently small, we can restrict the equilibrium analysis without loss of generality to cut-off strategies.

Proposition 3.2. *There exists* $\bar{\delta} \in (0, 1)$ *such that for all* $\delta < \bar{\delta}$ *the equilibrium strategies of both players are perfect cut-off strategies and* $v_k^i < v_{k-2}^i$ *for all* $k = 2, \ldots, n$*. Furthermore, if no types of* j *buy at a price* p_k*, then no types of* i *buy at price* p_k *and vice versa.*

The formal proof can be found in Section 3.7.2.2 in the Appendix. To simplify notation we explain the intuition of the proof in terms of interior cut-off strategies. Statements as "buy p_k" mean that the buyer accepts p_k and any lower price but does reject any price higher than p_k. Then, one can extend the interior cut-off strategy to a perfect cut-off strategy as described by Remark 3.1. The idea of the proof is as follows. We first show that there exists a non-degenerate interval of types who buy at price p_n and a non-degenerate interval of types who buy at p_1. Suppose in equilibrium there exists

[6] Note that $\sigma_i = \sigma_j$ implies that both buyers have the same information partitions $I_k^i = I_k^j$ and therefore the same rectangular belief sets $\text{rect}(\Phi_i^{-1}) = \text{rect}(\Phi_j^{-1})$ for all $k = -1, \ldots n$.

prices p_{k-l} through p_n such that no types of j would buy at prices $\{p_{k-l+1}, \dots p_{n-1}\}$, but there are types of j who buy at p_{k-l} and, of course, at p_n. Then, buyer i will never accept prices in $\{p_{k-l+1}, \dots p_{n-2}\}$, because he could profit from deviating to price p_{n-1}. But, there could exist types of i who buy at p_{n-1}. Since no type of buyer j buys at prices $\{p_{k-l+1}, \dots p_{n-1}\}$, there has to exist a type, v^j_{k-l}, who is indifferent between buying at p_{k-l} and waiting till p_n.

The proof of Proposition 3.2 shows that if there are types of i who buy at p_{n-1}, then for δ sufficiently small there are types of j just below v^j_{k-l} who would profit from deviating and buying at p_{n-1}. If no type of i buys at p_{n-1} we can still show that for δ sufficiently small there are types of j just below v^j_{k-l} who would profit from deviating and buying at p_{n-1} if $l > 2$. Therefore, if no type of j buys at p_{n-1}, then there is a non-degenerate interval of types who buy at p_n and a non-degenerate interval of types who buy at p_{n-2}. Then, Proposition 3.2 follows from iteration.

The assumption $l > 2$ is essential. If no type of i buys at p_{n-1} and $l = 2$, the proof of Proposition 3.2 cannot rule out "price gaps" in the following sense: Let p_k be such a "price gap". Then, there are no types of buyer i or j such that p_k is the highest price accepted by a type of buyer i and j. However, "price gaps" can only occur, if $l = 2$. Therefore, a "price gap" contains at most one price and we get $v^i_k < v^i_{k-2}$. For example, a cut-off strategy such that for both players, there exist types who accept as a highest price $p_1, p_3, p_5, \dots$, and p_n could be an equilibrium. But a strategy with a "price gap" containing more than one price can never be an equilibrium.

Finally, the existence of "price gaps" does not contradict the definition of perfect cut-off strategies. As described above, the "price gaps" only require that there are no types such that a price p_k is the highest price that these types accept. But it is still possible that there exists types who accept prices $p_{k-l} > p_k$ and therefore also p_k.

Given the previous results, we can now characterize equilibrium strategies.

Proposition 3.3. *For $\delta < \bar{\delta}$, in any equilibrium the strategy of any buyer i is a perfect cut-off strategy $\mathbf{v}^i = (v^i_1, \dots, v^i_n)$, where $v^i_n = p_n$. Further, for $1 \leq k \leq (n-1)$, $v^i_k \in (p_k, v^i_{k-1}]$, where $v_0 = 1$ and v^i_k is given by*

i) if there exists types of i who buy at p_k

$$v^i_k = p_k + \Delta_k \frac{\hat{H}^i_k}{1 - \hat{H}^i_k}$$

with $\hat{H}^i_k$ as in Equation (3.5) of Proposition 3.1.

ii) $v^i_k = v^i_{k-1}$ if no types of i buy at p_k.

Proof. The result follows immediately from Proposition 3.2, Proposition 3.1 and Remark 3.1. Proposition 3.2 and Remark 3.1 imply that for sufficiently small δ, any equilibrium strategy is a perfect cut-off strategy as described by Proposition 3.2 and Remark 3.1. If no type of buyer i buys at p_k, then the interval $[v_k^i, v_{k-1}^i)$ must be degenerated, i.e., $v_k^i = v_{k-1}^i$. If there exists a type who buys at price p_k and a type who buys at p_{k-1}, then v_k^i is the lowest type who accepts p_k. This type has to be indifferent between waiting until p_{k-1} and buying at p_k, i.e., $G_k^i(v_k^i) = 0$. Furthermore, $G_k^i(p_k) < 0$ and since there exist types who buy at p_{k-1}, $G_k^i(v_{k-1}) > 0$. $G_k^i(v)$ is strictly increasing and continuous in v. Therefore, there exists $v_k^i \in (p_k, v_{k-1}^i]$ such that $G_k^i(v_k^i) = 0$ and Equation (3.4) implies

$$\begin{aligned} & v_k^i - p_k = (v_k^i - p_{k+1})\hat{H}_k^i \\ \Leftrightarrow \quad & v_k^i = p_k + (p_k - p_{k+1})\frac{\hat{H}_k^i}{1-\hat{H}_k^i} = p_k + \Delta_k \frac{\hat{H}_k^i}{1-\hat{H}_k^i}. \end{aligned}$$

If no type of buyer i buys at p_{k-1}, then a type who buys at p_{k-2} has to exist. Thus, the same consideration can be repeated with $v_{k-2}^i = v_{k-1}^i$. □

The existence of a symmetric equilibrium for sufficiently small δ follows from Bose and Daripa (2009). They show that a symmetric equilibrium exists even if the best response mapping is discontinuous at $v_k^i = v_{k-1}^i$. Hence, there exists a symmetric equilibrium for sufficiently small δ even if "price gaps" with at most one price occur. We focus on symmetric equilibria for the rest of the chapter and therefore omit the superscripts i or j.

Now, we can state the main result. For any preference parameter ϵ, the seller can design an MDM such that he can extract almost all surplus of almost all buyers.

Proposition 3.4. *For any preference parameter $\epsilon > 0$, there exists $\eta > 0$ and $\delta^*(\epsilon) > 0$ such that for any $\delta < \delta^*(\epsilon)$ there is an MDM such that in any equilibrium of the game induced by this MDM, the item is sold if at least one buyer has a valuation greater than η and no type obtains an ex-post surplus greater than δ.*

We have to ensure that the ex-post surplus of both buyers is still smaller than δ even if "price gaps" occur. The first part of the proof is analogous to Bose and Daripa (2009). The second part considers rectangular beliefs and takes "price gaps" into account.

Proof. From the previous results it follows that for any $\epsilon > 0$, there is a $\delta^*(\epsilon) > 0$ such that whenever $\delta < \delta^*(\epsilon)$, an equilibrium exists and all equilibria can be characterized by

$$v_k = p_k + \Delta_k \frac{\hat{H}_k}{1-\hat{H}_k}.$$

Further, as noted in Section 3.2.2, for any $\eta \in (0,1)$ there exists an integer T such that by choosing $n = T$, the price sequence of the MDM covers at least a fraction $(1-\eta)$ of types. The item is not sold to at most types in $[0,\eta]$. Thus, it only remains to show that no type that buys gets an ex-post surplus greater than δ. This part has to be adjusted. For any price p_k with $k > 2$ such that there exists types who accept the price p_k, we have one of the following cases.

Case 1) There are types who buy at p_{k-1}:[7]

The highest type who accepts p_k is v_{k-1}. Therefore, the maximum rent of a buyer who accepts p_k, is

$$v_{k-1} - p_k = p_{k-1} - p_k + \Delta_{k-1}\frac{\hat{H}_{k-1}}{1-\hat{H}k-1} = \Delta_{k-1} + \Delta_{k-1}\frac{\hat{H}_{k-1}}{1-\hat{H}_{k-1}}$$
$$< \Delta_{k-1}\left(1 + \frac{1-\epsilon}{\epsilon}\right) = \frac{1}{\epsilon}\frac{\delta\epsilon}{2}\left(\frac{1-\delta}{1-\delta+\frac{\epsilon\delta}{2}}\right)^{k-1} < \frac{\delta}{2} < \delta,$$

where the first inequality follows since $\hat{H}_{k-1} < 1-\epsilon$ and the second inequality since $\frac{1-\delta}{1-\delta+\frac{\epsilon\delta}{2}} < 1$.

Case 2) There are no types who buy at p_{k-1}:

Then $v_{k-1} = v_{k-2}$ and similar to Case 1), we get

$$v_{k-2} - p_k = p_{k-2} - p_{k+1} + \Delta_{k-2}\frac{\hat{H}_{k-2}}{1-\hat{H}_{k-2}} = \Delta_{k-2} + \Delta_{k-1} + \Delta_{k-2}\frac{\hat{H}_{k-2}}{1-\hat{H}_{k-2}}$$
$$< \Delta_{k-2}\left(2 + \frac{1-\epsilon}{\epsilon}\right) = \frac{1+\epsilon}{\epsilon}\frac{\delta\epsilon}{2}\left(\frac{1-\delta}{1-\delta+\frac{\epsilon\delta}{2}}\right)^{k-1} < (1+\epsilon)\frac{\delta}{2} < \delta,$$

where the first inequality follows since, $\Delta_{k-2} > \Delta_{k-1}$ and $\hat{H}_{k-2} < 1-\epsilon$. The second and third inequality follow since, $\frac{1-\delta}{1-\delta+\frac{\epsilon\delta}{2}} < 1$ and $\epsilon < 1$.

Finally, we have to consider $k = 2$ and $k = 1$. Suppose, there are types who accept p_2. Then, by Lemma 3.6 in Section 3.7.2.2 in the Appendix, it follows that a positive mass of types accepts p_1. Similar to Case 1) we get

$$v_1 - p_2 < \delta$$

[7] Case 1) is analogously to Bose and Daripa (2009). However, Bose and Daripa (2009) uses the fact that $\bar{H}_{k-1} < 1-\epsilon$ which does not hold for the corrected worst-case belief of Section 3.2.5. Therefore, it is not straightforward, that the main result of Bose and Daripa (2009) is still satisfied with the corrected worst-case belief. However, in Section 3.7.3 in the Appendix we prove that the result still holds by using the specific expression of $\bar{H}_{k-1}$.

and the maximum rent of buyers who accept p_2 is smaller than δ. For $k = 1$, by definition of v_0 and p_1, the maximum rent of a type who accepts p_1 is

$$v_0 - p_1 = 1 - (1 - \delta) = \delta.$$

Hence, there exists no type who can extract a surplus strictly greater than δ. □

3.4 Numerical Example

The following numerical example shows how the different beliefs, the incorrect, the correct and the rectangular worst-case belief influence the surplus extraction. As in Bose and Daripa (2009), F is the uniform distribution on $[0, 1]$, $n = 7$, $\delta = 0,05$ and $\epsilon = 0.2$. Each table shows the price sequence p_k, the equilibrium cut-off values v_k, the buyers' maximum rent, and the worst-case belief. The equilibrium cut-off values v_k are derived by solving

$$v_k = p_k + \Delta_k \frac{\tilde{H}_k}{1 - \tilde{H}_k} \tag{3.9}$$

for $k = 1, \ldots, 6$ with $v_0 = 1$ and $v_7 = p_7$. $\tilde{H}_k$ is either given by the worst-case belief of Bose and Daripa (2009) $(1 - \epsilon)H_k$, the corrected worst-case belief of Section 3.2.5 $\bar{H}_k$, or the rectangular worst-case belief $\hat{H}_k$.

For all cases, $\tilde{H}_k$ and therefore the cut-off value v_k depend on the cut-off values v_{k-1} and v_{k+1}. Then, from Equation (3.9) we get a system of recursive equations with $v_0 = 1$ and $v_7 = p_7$. Solving this system we get the values $v_1, \ldots, v_6$ listed in Table 3.1 to Table 3.3.

Furthermore, the tables show the maximum rent of the buyers. The maximum rent in period k is the highest rent of a type who accepts price p_k. More formally, the maximum rent is given by $v_{k-1} - p_k$, i.e., the difference of the highest type who accepts p_k and the price p_k.

Price	Value	Max rent	H^{BD}
0.9500	0.9682	0.0500	0.9819
0.9450	0.9644	0.0232	0.9954
0.9401	0.9593	0.0243	0.9947
0.9352	0.9542	0.0241	0.9947
0.9303	0.9492	0.0240	0.9943
0.9254	0.9434	0.0238	0.9849

Table 3.1: Worst-Case Belief of Bose and Daripa (2009) with $(1 - \epsilon)H_k = (1 - \epsilon)H^{BD}$.

Table 3.1 shows the results for the (incorrect) worst-case belief of Bose and Daripa (2009)

$$\tilde{H}_k = (1-\epsilon)H_k = (1-\epsilon)\frac{F(v_k)+F(v_{k+1})}{F(v_k)+F(v_{k-1})} = (1-\epsilon)H^{BD}.$$

H^{BD} is smaller than one and therefore $(1-\epsilon)H_k = (1-\epsilon)H^{BD}$ smaller than $(1-\epsilon)$. Bose and Daripa (2009) use this fact repeatedly to prove almost all their results.

Price	Value	Max rent	H^{corr}
0.9500	0.9866	0.0500	1.1004
0.9450	0.9823	0.0416	1.1037
0.9401	0.9768	0.0423	1.1024
0.9352	0.9715	0.0417	1.1014
0.9303	0.9655	0.0412	1.0984
0.9254	0.9562	0.0402	1.0802

Table 3.2: Correct Worst-Case Belief with $\bar{H}_k = (1-\epsilon)H^{corr}$.

Table 3.2 illustrates the results with the corrected worst-case belief from Section 3.2.5, i.e.,

$$\tilde{H}_k = \bar{H}_k = \frac{(1-\epsilon)(F(v_k)+F(v_{k+1}))}{(1-\epsilon)(F(v_k)+F(v_{k-1}))+\epsilon} = (1-\epsilon)H^{corr}.$$

One can see that H^{corr} is greater than one and thus $\bar{H}_k = (1-\epsilon)H^{corr} > 1-\epsilon$. Therefore, the proofs of Bose and Daripa (2009) cannot be applied to the corrected worst-case belief. However, we show in Section 3.7.3 in the Appendix that the results of Bose and Daripa (2009) still hold with the corrected worst-case belief. Table 3.2 illustrates this result. The maximum rent of a buyer in Table 3.2 is still smaller or equal to δ. Table 3.3 shows the results with the rectangular worst-case belief

Price	Value	Max rent	H^{rect}
0.9500	0.9678	0.0500	0.9771
0.9450	0.9637	0.0228	0.9883
0.9401	0.9585	0.0236	0.9867
0.9352	0.9534	0.0234	0.9856
0.9303	0.9483	0.0232	0.9842
0.9254	0.9426	0.0229	0.9752

Table 3.3: Rectangular Worst-Case Belief with $\hat{H}_k = (1-\epsilon)H^{rect}$.

$$\tilde{H}_k = \hat{H}_k = (1-\epsilon_k)\frac{F^k(v_k)+(1-\epsilon_{k+1})F^k(v_k)F^{k+1}(v_{k+1})}{1+(1-\epsilon_k)F^k(v_k)} = (1-\epsilon)H^{rect}.$$

Comparing Table 3.2 and Table 3.3 illustrate the difference of rectangularity and consistent planning. The worst-case belief and the maximum rent are smaller for dynamically consistent buyers than for consistent planning. We compare both approaches in the next section in more detail.

3.5 Rectangularity versus Consistent Planning

Bose and Daripa (2009) analyze the dynamic Dutch auction using the consistent planning approach of Siniscalchi (2011). It is still an open question if there exist settings or conditions such that rectangularity and consistent planning lead to the same equilibrium or decision outcomes. Since rectangular beliefs and consistent planning are both based on a backward induction procedure, it is sometimes conjectured for specific settings that both approaches lead to the same equilibrium outcome, e.g., in Auster and Kellner (2020) and Chapter 1. However, the numerical example shows that this is not the case in this chapter. First, let us describe the consistent planning approach in the ambiguous Dutch auction setting before explaining the difference in equilibrium predictions.

3.5.1 Consistent Planning in the Ambiguous Dutch Auction

Consistent planning follows a backward induction procedure with an additional tie breaking rule. Let us start in period n and assume buyer i gets the offer p_n. In the last period, the expected payoff of rejecting is zero. Therefore, buyer i accepts the offer if and only if his valuation v satisfies $v \geq p_n$. In period $n-1$, if buyer i gets the offer p_{n-1} he has to compare the following plans: i) accept p_{n-1}, ii) reject p_{n-1} and accept p_n, iii) reject p_{n-1} and reject p_n. Then, consistent planning ensure, that if rejecting p_n is optimal in period n, the third plan is optimal in period $n-1$. Further, if accepting p_n is optimal in period n, only plan i) and ii) are consistent with the decision in period n. Hence, if accepting p_n is optimal in period n, buyer i compares plan i) and ii) in period $n-1$ and choose the one with higher expected payoff. This procedure is repeated until period zero. In an arbitrary period k, buyer i will reject the offer p_k if rejecting p_{k+1} is optimal in period $k+1$. If accepting p_{k+1} is optimal in period $k+1$, in period k buyer i compares the expected payoff of accepting p_k with the expected payoff of rejecting p_k and accepting p_{k+1}. Formally, this leads to Lemma 3.2, i.e., buyer i accepts p_k if and only

$$G_k^i(v) = v - p_k - (v - p_{k+1})\bar{H}_k^i > 0.$$

However, the procedure described above does not consider any strategic reasoning. Even further, to apply the consistent planning approach of Siniscalchi (2011) we have to model the ambiguous Dutch auction as a decision-theoretical setting. For a fixed strategy of buyer j, consistent planning analyzes the decision problem of buyer i in periods $n-1$ and n as depicted in Figure 3.1. Solid circles represent decision nodes of buyer i, where A denotes accepting and R rejecting the current price. Empty circles represent nature moves and E_k^1 denotes the event that buyer i does not get the offer p_k and E_k^2 the event that i gets the offer p_k with $k = n-1, n$. Then, $\bar{H}_k^i$ is given by the worst-case belief of event E_k^2.

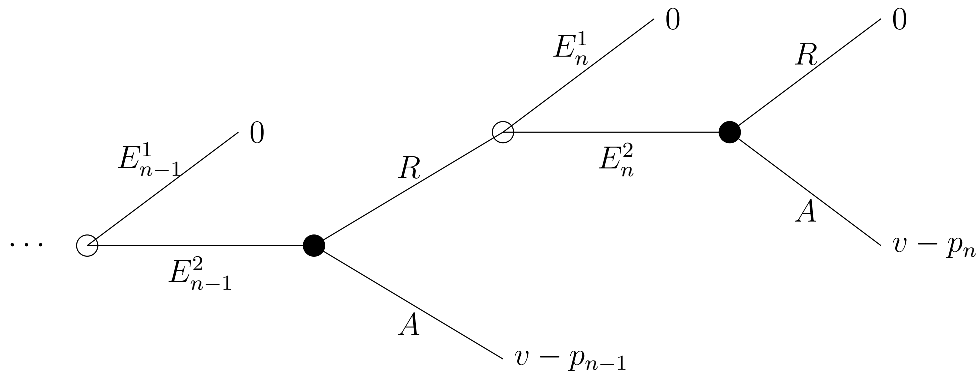

Figure 3.1: Decision Problem of Buyer i with Consistent Planning in Period n and $n-1$.

3.5.2 Consistent Planning and Rectangularity

The decision problem in Figure 3.1 does not consider that buyer i knows the auction structure. He knows that the worst-case belief of event E_n^2 depends on the probability that the object remains unsold at p_{n-1} and the probability that he gets the offer p_n given that period n is reached. Further, buyer i knows that period n is only reached if his worst-case in period $n-1$ does not occur. Taking this into account, Figure 3.2 represents the true decision setting of buyer i, where q_k^i denotes the position of buyer i, i.e., if he is asked first or second and $j \in A, B$ the action of buyer j. The decision problem as depicted in Figure 3.2 allows buyer i to take changes of the worst-case belief within the event E_n^2 into account. Further, in Figure 3.2 we can split period $n-1$ and n. In Figure 3.1 the switch from period $n-1$ to n happens within the event E_n^2. The consistent planning approach of Siniscalchi (2011) is defined for a decision-theoretical setting. It cannot take the specific timing of the ambiguous Dutch auction into account. Therefore, rectangularity and consistent planning lead to different equilibrium strategies. Formally, this follows from the different worst-case beliefs. The rectangular

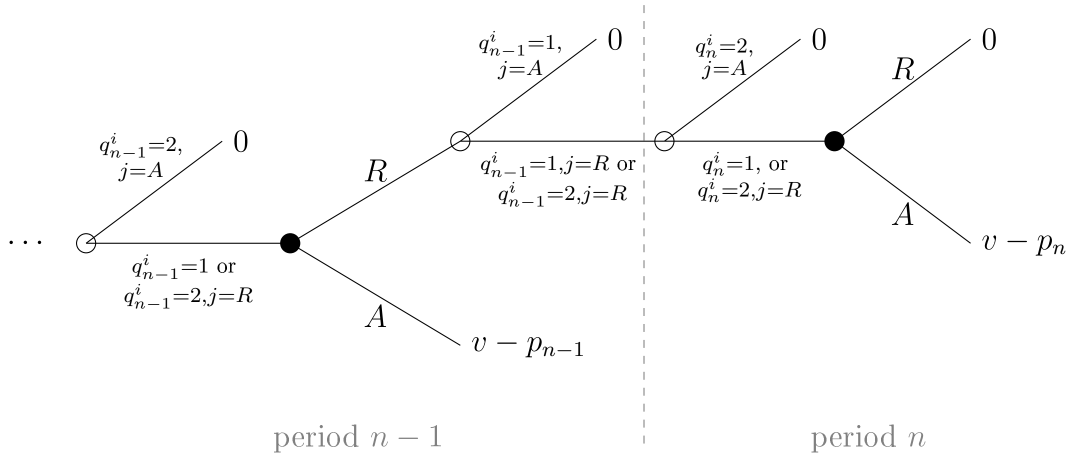

Figure 3.2: Decision Problem of Buyer i with Rectangular Beliefs in Period n and $n-1$.

worst-case belief is given by[8]

$$\hat{H}^i_k = \min_{g^k \in \Phi^k_i, g^{k+1} \in \Phi^{k+1}_i} \frac{G^k(v^j_k) + G^{k+1}(v^j_{k+1})G^k(v^j_k)}{1 + G^k(v^j_k)},$$

whereas the consistent planning worst-case belief is

$$\bar{H}^i_k = \min_{g^k \in \Phi^k_i} \frac{G^k(v^j_k) + G^{k+1}(v^j_{k+1})G^k(v^j_k)}{1 + G^k(v^j_k)},$$

where for $\bar{H}^i_k$, G^{k+1} is the distribution function corresponding to $g^{k+1} = \text{Bay}(g^k | v < v^i_k)$ the Bayesian update of g^k.

For each period k, the worst-case belief g^* of buyer i in a period k, is that buyer j accepts the current offer, i.e., $g^*(v) = (1-\epsilon)f^k(v) + \epsilon l^k(v)$ with $l^k(v \in [v^i_{k-1}.v^i_k)) = 1$. Therefore, the worst-case belief in period $k+1$ cannot be the Bayesian update of the worst-case belief in period k and for a fixed strategy of buyer j

$$\hat{H}^i_k < \bar{H}^i_k.$$

Rectangularity takes this change in the worst-case belief into account and induces a different optimal strategy than consistent planning.

3.6 Conclusion and Discussion

We consider a Dutch auction mechanism with ambiguity-averse buyers. First, we correct the worst-case belief of Bose and Daripa (2009). Then, we analyze the Dutch auction for dynamically consistent buyers and showed that the seller can still extract

[8]This representation of the rectangular worst-case belief follows from the proof of Proposition 3.1.

almost all surplus. Furthermore, we discuss the different implications of rectangularity and consistent planning. Rectangularity allows the buyers to take the structure of the auction into account. Therefore, under the rectangular worst-case belief, a buyer believes that it is less likely to receive the offer p_{k+1} if he rejects p_k. This decreases the expected payoff of waiting, and dynamically consistent buyers accept prices closer to their types.

This observation opens many future research questions. As a first step, it is essential to understand how the consistent planning approach of Siniscalchi (2011) could be extended to games. Is Figure 3.1 the right way to apply consistent planning, or should the definition be extended to take situations as described in Figure 3.2 into account? Without ambiguity, this problem does not arise since the belief in period $k+1$ is always the Bayesian update of the belief in period k.

Further, when do consistent planning and rectangularity predict the same equilibrium outcome? The timing in the Dutch auction setting is very complex. Since the buyers do not know who gets the price offer p_k first, the game cannot be represented by a multistage game. Auster and Kellner (2020) analyze an ambiguous Dutch auction in a continuous-time setting. Due to the continuous-time setting, they do not have to specify which buyer is asked first and do not need the complex structure of Bose and Daripa (2009). In their setting, rectangularity and consistent planning induce the same equilibrium strategy.

We conjecture that in multistage games, rectangularity and consistent planning will predict equivalent equilibrium outcomes. Multistage games require a specific time structure. In each period, players first observe private information (e.g., the move of one opponent from the previous period). Then, players move simultaneously. Therefore, the probability of reaching a specific information set in the next period depends only on the worst-case belief of one period. However, more detailed analysis and discussion are left for future research.

Alternatively, one could use a different time structure for the Dutch auction. In each period, both buyers simultaneously choose an action. If both buyers reject the current price, the game precedes to the next period. If buyer i accepts and buyer j rejects, buyer i gets the object and vice versa. If both buyers accept, the winner is chosen randomly. This version of the Dutch auction mechanism could be modeled as a multistage game. However, as discussed by Bose and Daripa (2009), it leads to an uncertain outcome of accepting. Therefore, the full surplus extraction result does not hold in this setting.

3.7 Appendix

3.7.1 Derivation Worst-Case Belief

The probability that buyer i obtains the item at price p_{k+1} given that he refuses the current offer p_k, $\mathbb{P}^g(A|B)$ can be derived analogously to Bose and Daripa (2009) by replacing F with G. For completeness, we repeat the derivation here. Let us fix one belief $g \in \Phi_i^{-1}$ and G be the corresponding distribution function. Further, G^k denotes the distribution function corresponding to the updated belief in period k, $g^k \in \Phi_i^k$.

We first derive the probability π_k^i that period $k+1$ is reached given that i refuses the current offer p_k. Then, we derive the probability ψ_k^i, that i gets the offer p_{k+1} given that stage $k+1$ is reached. The probability $\mathbb{P}^g(A|B)$ is then given by $\pi_k^i \psi_k^i$.

π_k^i can be derived as follows. Let $q^i \in \{1, 2\}$ denote the position of player i in period k, i.e., if he is asked first or second. Further, let A^i denote the event, that i gets the offer p_k. Then,

$$\pi_k^i = \mathbb{P}(q^i = 1|A^i)G^k(v_k^j) + \mathbb{P}(q^i = 2|A^i)$$

with

$$\begin{aligned}\mathbb{P}(q^i = 1|A^i) &= \frac{\mathbb{P}(q^i = 1)\mathbb{P}(A^i|q^i = 1)}{\mathbb{P}(q^i = 1)\mathbb{P}(A^i|q^i = 1) + \mathbb{P}(q^i = 2)\mathbb{P}(A^i|q^i = 2)} \\ &= \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}G^k(v_k^j)} = \frac{1}{1 + G^k(v_k^j)}\end{aligned}$$

and

$$\mathbb{P}(q^i = 2|A^i) = 1 - \mathbb{P}(q^i = 1|A^i) = \frac{G^k(v_k^j)}{1 + G^k(v_k^j)}.$$

Therefore,

$$\pi_k^i = \frac{2G^k(v_k^j)}{1 + G^k(v_k^j)}.$$

In period $k+1$, i gets the offer p_{k+1} for sure if he is asked first. If he is asked second, he only gets the offer if buyer j refuses p_{k+1}:

$$\psi_k^i = \frac{1}{2} + \frac{1}{2}G^{k+1}(v_{k+1}^j).$$

Then,

$$\begin{aligned}\mathbb{P}^g(A|B) = \pi_k^i \psi_k^i &= \frac{2G^k(v_k^j)}{1 + G^k(v_k^j)} \left(\frac{1}{2} + \frac{1}{2}G^{k+1}(v_{k+1}^j) \right) \\ &= \frac{G^k(v_k^j) + G^k(v_k^j)G^{k+1}(v_{k+1}^j)}{1 + G^k(v_k^j)}.\end{aligned}$$

Now, for a fixed g, if $G^k(v_k^j) = \frac{G(v_k^j)}{G(v_{k-1}^j)}$ for all $k \geq 1$ we get

$$\mathbb{P}^g(A|B) = \frac{G(v_k^j) + G(v_{k+1}^j)}{G(v_{k-1}^j) + G(v_k^j)}.$$

3.7.2 Proofs of Section 3.3

3.7.2.1 Proof of Lemma 3.1

The proof generalizes the proof of Nishimura and Ozaki (2002, page 6 - 8) and consists of four steps. Step 1) proves an observation which is used recurrently in Step 2) and 3). Step 2) and 3) show that $\Psi^2 \subseteq \Psi^1$ and $\Psi^1 \subseteq \Psi^2$. Finally, Step 4) proves $\epsilon_k > \epsilon$.

Step 1) First, we show that $g \in \Psi^1$ if and only if there exists a density function $l \in \mathcal{P}$ such that

$$\begin{aligned}
g(\cdot) &= \frac{(1-\epsilon)f(\cdot) + \epsilon l(\cdot)}{(1-\epsilon)f(E) + \epsilon l(E)} \\
&= \frac{(1-\epsilon)f(E)}{(1-\epsilon)f(E) + \epsilon l(E)} \frac{f(\cdot)}{f(E)} + \frac{\epsilon l(E)}{(1-\epsilon)f(E) + \epsilon l(E)} \frac{l(\cdot)}{l(E)} \\
&= \frac{(1-\epsilon)f(E)}{(1-\epsilon)f(E) + \epsilon l(E)} \text{Bay}(f|E)(\cdot) + \frac{\epsilon l(E)}{(1-\epsilon)f(E) + \epsilon l(E)} \text{Bay}(l|E)(\cdot).
\end{aligned}$$

We can assume, without loss of generality, that $l(E) > 0$. Otherwise, $g(x) = \frac{f(x)}{f(E)}$ for all $x \in E$ which immediately implies that $g \in \Psi^2$.

Step 2) Let g be an arbitrary density function in Ψ^2. Then, the definitions of Ψ^2 and ϵ_k imply that there exists an $l \in \mathcal{P}$ such that

$$\begin{aligned}
g(\cdot) &= ((1-\epsilon_k)\text{Bay}(f|E)(\cdot) + \epsilon_k \text{Bay}(l|E)(\cdot)) \\
&= \frac{(1-\epsilon)f(E)}{(1-\epsilon)f(E) + \epsilon} \text{Bay}(f|E)(\cdot) + \frac{\epsilon}{(1-\epsilon)f(E) + \epsilon} \text{Bay}(l|E)(\cdot). \qquad (3.10)
\end{aligned}$$

Since $\mathcal{P}$ consists of all density functions on $[0,1]$ there exists a density function $\bar{l} \in \mathcal{P}$ such that

$$\bar{l}(x) = \text{Bay}(l|E)(x) \quad \forall x \in E$$

and $\bar{l}(x) = 0$ otherwise. Then, $\bar{l}(E) = 1$ and

$$\text{Bay}(\bar{l}|E)(\cdot) = \text{Bay}(l|E)(\cdot).$$

Using this properties of $\bar{l}$ we can rewrite Equation (3.10):

$$g(\cdot) = \frac{(1-\epsilon)f(E)}{(1-\epsilon)f(E) + \epsilon} \text{Bay}(f|E)(\cdot) + \frac{\epsilon \bar{l}(E)}{(1-\epsilon)f(E) + \epsilon \bar{l}(E)} \text{Bay}(\bar{l}|E)(\cdot).$$

Then by Step 1) it follows that $g \in \Psi^1$. Hence, $\Psi^2 \subseteq \Psi^1$.

Step 3) For the other direction, i.e., $\Psi^1 \subseteq \Psi^2$, let g be an arbitrary element of Ψ^1. Then, Step 1) implies

$$\begin{aligned}
g(\cdot) =& \frac{(1-\epsilon)f(E)}{(1-\epsilon)f(E)+\epsilon l(E)}\text{Bay}(f|E)(\cdot) + \frac{\epsilon l(E)}{(1-\epsilon)f(E)+\epsilon l(E)}\text{Bay}(l|E)(\cdot) \\
=& \frac{(1-\epsilon)f(E)\text{Bay}(f|E)(\cdot)}{(1-\epsilon)f(E)+\epsilon l(E)} \frac{(1-\epsilon)f(E)+\epsilon+\epsilon l(E)-\epsilon l(E)}{(1-\epsilon)f(E)+\epsilon} \\
&+ \frac{\epsilon l(E)\text{Bay}(l|E)(\cdot)}{(1-\epsilon)f(E)+\epsilon l(E)} \frac{(1-\epsilon)f(E)+\epsilon}{(1-\epsilon)f(E)+\epsilon} \\
=& (1-\epsilon)f(E)\text{Bay}(f|E)(\cdot)\frac{(1-\epsilon)f(E)+\epsilon l(E)+\epsilon(1-l(E))}{((1-\epsilon)f(E)+\epsilon l(E))((1-\epsilon)f(E)+\epsilon)} \\
&+ \underbrace{\frac{\epsilon}{(1-\epsilon)f(E)+\epsilon}}_{=\epsilon_k} \frac{((1-\epsilon)f(E)+\epsilon)l(E)\text{Bay}(l|E)(\cdot)}{(1-\epsilon)f(E)+\epsilon l(E)} \\
=& \epsilon_k \underbrace{\frac{((1-\epsilon)f(E)+\epsilon)l(E)}{(1-\epsilon)f(E)+\epsilon l(E)}}_{:=\tilde{\epsilon}} \text{Bay}(l|E)(\cdot) + (1-\epsilon)f(E)\text{Bay}(f|E)(\cdot) \\
&\cdot \left(\frac{1}{(1-\epsilon)f(E)+\epsilon} + \frac{\epsilon(1-l(E))}{((1-\epsilon)f(E)+\epsilon l(E))((1-\epsilon)f(E)+\epsilon)}\right) \\
=& \epsilon_k\tilde{\epsilon}\text{Bay}(l|E)(\cdot) + \underbrace{\frac{(1-\epsilon)f(E)}{(1-\epsilon)f(E)+\epsilon}}_{=1-\epsilon_k}\text{Bay}(f|E)(\cdot) \\
&+ \underbrace{\frac{\epsilon}{(1-\epsilon)f(E)+\epsilon}}_{=\epsilon_k} \frac{(1-l(E))(1-\epsilon)f(E)\text{Bay}(f|E)(\cdot)}{((1-\epsilon)f(E)+\epsilon l(E))} \\
=& \epsilon_k\tilde{\epsilon}\text{Bay}(l|E)(\cdot) + (1-\epsilon_k)\text{Bay}(f|E)(\cdot) \\
&+ \epsilon_k \underbrace{\frac{(1-l(E))(1-\epsilon)f(E)\text{Bay}(f|E)(\cdot)}{((1-\epsilon)f(E)+\epsilon l(E))}}_{=1-\tilde{\epsilon}} \text{Bay}(f|E)(\cdot) \\
=& (1-\epsilon_k)\text{Bay}(f|E)(\cdot) + \epsilon_k\Big(\tilde{\epsilon}\text{Bay}(l|E)(\cdot) + (1-\tilde{\epsilon})\text{Bay}(f|E)(\cdot)\Big).
\end{aligned}$$

Since $\tilde{\epsilon} \in (0,1]$, $\bar{l}(x) := \tilde{\epsilon}\text{Bay}(l|E)(x) + (1-\tilde{\epsilon})\text{Bay}(f|E)(x)$ for all $x \in E$ and zero otherwise is a density function, i.e., $\bar{l} \in \mathcal{P}$. Furthermore, $\text{Bay}(\bar{l}|E) = \bar{l}$, since $\bar{l}(E) = 1$. Hence, $g(\cdot) \in \Psi^2$ and $\Psi^1 \subseteq \Psi^2$.

Step 4) The full support assumption of f implies that $f(E) < 1$ for all $E \neq \Omega$. Therefore, it follows immediately that

$$\epsilon_k = \frac{\epsilon}{(1-\epsilon)f(E)+\epsilon} > \epsilon.$$

Furthermore, if $E = \Omega$, then $f(\Omega) = 1$ and $\epsilon_k = \epsilon$.

3.7.2.2 Proof of Proposition 3.2

The proof follows a similar idea as the proof of Bose and Daripa (2009). However, since most parts explicitly use the form of the worst-case belief, we have to adjust parts of the proof. It consists of several lemmata. First, Lemma 3.5 shows that both buyers have a positive measure of types who plan to buy at p_n. Lemma 3.6 proves that in any equilibrium, a positive measure of types plans to buy at p_1. Then, the idea is as follows. Suppose there is a "price gap" in the sense that there are no types of buyer j who plan to buy at prices in $\{p_{n-l+1}, \ldots, p_{n-1}\}$. Then, there are no types of i who plan to buy at prices $\{p_{n-l+1}, \ldots, p_{n-2}\}$. Lemma 3.8 shows, that if there exists types of i who buy at p_{n-1}, then there exists types of j who strictly prefer to buy at p_{n-1} as well. Further, if no type of i buys at p_{n-1}, then there still exist types of j who strictly prefer to buy at p_{n-1} if $l > 2$. Hence, there might exist a "price gap", but the "price gap" contains at most one price. Then, Proposition 3.2 follows by iteration.

Lemma 3.5. *In any equilibrium, $v_n^i = v_n^j = p_n$. Further, a positive measure of types of both buyers plan to buy at price p_n but not at any earlier price.*

The proof is analogously to the proof of Lemma 3 of Bose and Daripa (2009). The idea of the proof is, that all types $v \in (p_n, p_{n-1})$ do not buy at price $p \geq p_{n-1}$, since $v - p < 0$. But, since the payoff of not buying is zero, buyers of type $v \in (p_n, p_{n-1})$ prefer to buy at p_n.

Lemma 3.6. *In any equilibrium, a positive measure of types of each buyer plans to buy at price p_1.*

Lemma 3.6 can be proven analogously to the proof of Lemma 4 in Bose and Daripa (2009) using the fact that $\hat{H}_k^i < 1 - \epsilon$ and that the ϵ-contamination structure still holds for marginal probabilities of the rectangular hull. For completeness, we repeat it here.

Proof. Suppose j does not plan to buy at prices $p_1, \ldots, p_k$ for $1 \leq k < n$ and p_{k+1} is the first price at which j buys. (This is denoted as $v_1^j = \cdots = v_k^j = 1$ and $v_{k+1}^j < 1$.) Clearly, the best response of i is not to buy at prices $p_1, \ldots, p_{k-1}$. If i refuses p_k the probability that the game reaches p_{k+1} is one. Therefore,

$$\hat{H}_k^j = \min_{g^{k+1} \in \mathrm{rect}(\Phi_j^k)} \frac{1}{2} + \frac{1}{2} g^{k+1}(v \leq v_{k+1}^i) = \frac{1}{2} + \frac{1}{2}(1 - \epsilon_{k+1}) F^{k+1}(v_{k+1}^i), \tag{3.11}$$

where the last equality follows since, $g^{k+1}(v \leq v_{k+1}^i) = G^k(v_{k+1}^i)$ and the set of marginal distributions given the rectangular hull, $\mathrm{rect}(\Phi_j^{k,\cdot})$, equals the set of marginal distributions given Φ_j^k. Therefore, the ϵ-contamination structure is maintained for marginal distributions of the rectangular hull.

Further, j does not plan to buy at prices $p_1, \ldots, p_k$. Therefore, $v_k^j = 1$ and $F^{k+1}(v_{k+1}^i) = F(v_{k+1}^i)$ and $\epsilon_{k+1} = \epsilon$. Then, the payoff from refusing p_k is $(\frac{1}{2} + \frac{1}{2}(1-\epsilon)F(v_{k+1}^j))(v - p_{k+1})$. Define the following function:

$$\begin{aligned}\hat{G}_k^i(v) :=& v - p_k - \left(\frac{1}{2} + \frac{1}{2}(1-\epsilon)F(v_{k+1}^j)\right)(v - p_{k+1})\\ =& \frac{1}{2}(v-p_k)(1-(1-\epsilon)F(v_{k+1}^j)) - \frac{1}{2}(1+(1-\epsilon)F(v_{k+1}^j))\Delta_k.\end{aligned}$$

Note that

$$\begin{aligned}2\hat{G}_k^i(1) &= (1-p_k)(1-(1-\epsilon)F(v_{k+1}^j)) - (1+(1-\epsilon)F(v_{k+1}^j))\Delta_k\\ &> \delta\epsilon - (2-\epsilon)\Delta_k \geq \delta\epsilon - (2-\epsilon)\Delta_1 = \frac{\delta\epsilon^2}{2(1-\delta)+\delta\epsilon} > 0,\end{aligned}$$

where the second step follows from the fact that $(1-p_k) \geq (1-p_1) = \delta$, and the fact that $F(v_{k+1}^j) < 1$, and the third step uses $\Delta_1 \geq \Delta_k$. Since $\hat{G}_k^i(v)$ is continuous, increasing in v, and negative at $v = p_k$, there exists v_k^i such that $\hat{G}_k^i(v) > 0$ for $v > v_k^i$ and $\hat{G}_k^i(v_k^i) = 0$. Since we know that i does not plan to buy at any earlier price than p_k, it must be that types $[v_k^i, 1]$ of buyer i plan to buy at p_k. The second part shows that there exist types of j close to one who would deviate and buy at p_k. First, remember that

$$G_k^j(v) = v - p_k - (v - p_{k+1})\hat{H}_k^j$$

with

$$\hat{H}_k^j = (1-\epsilon_k)\frac{F^k(v_k^j) + (1-\epsilon_{k+1})F^k(v_k^j)F^{k+1}(v_{k+1}^j)}{1+(1-\epsilon_k)F^k(v_k^j)} \leq (1-\epsilon_k) \leq 1-\epsilon.$$

Then,

$$\begin{aligned}G_k^j(1) &= (1-p_k) - (1-p_{k+1})\hat{H}_k^j = (1-p_k)(1-\hat{H}_k^j) - \Delta_k\hat{H}_k^j\\ &\geq \delta(1-\hat{H}_k^j) - \Delta_1\hat{H}_k^j > \delta\epsilon - \Delta_1 = \delta\epsilon - \frac{\delta\epsilon}{2}\frac{1-\delta}{(1-\delta)+\frac{\delta\epsilon}{2}} > 0,\end{aligned}$$

where the first inequality follows since $1-p_k \geq 1-p_1 = \delta$ and $\Delta_k \leq \Delta_1$. Further, $\hat{H}_k^j \leq 1-\epsilon < 1$ implies the third inequality. The last inequality follows since $\frac{1-\delta}{(1-\delta)+\frac{\delta\epsilon}{2}} < 1$. Since $G_k^j(v)$ is increasing and continuous, there are types of j of positive measure near one who would deviate and buy at p_k. Contradiction. □

Lemma 3.7. *Suppose there are types of buyer i who plan to buy at p_{n-l-t} and p_{n-l} but not at prices in between. Then $v_{n-l-t}^i - v_n^i < \delta(l+t)$*

Similar to Lemma 3.6, the proof is analogously to Lemma 5 of Bose and Daripa (2009) using the fact that $\hat{H}_k^i < 1 - \epsilon$ and Equation (3.11). For more details see Lemma 3.10 in Section 3.7.3.

Remember that we assume an equilibrium strategy where no types of j plan to buy at prices in $\{p_{n-l+1}, \ldots, p_{n-1}\}$ with $l \geq 2$. Then, no types of i plan to buy at prices $\{p_{n-l+1}, \ldots, p_{n-2}\}$.

Lemma 3.8. *There exists $\bar{\delta} > 0$ such that for all $\delta < \bar{\delta}$ then there are types (of positive measure) of j who buy at p_{n-1} if*

i) there are types of i who buy at p_{n-1} or

ii) no types of i buy at p_{n-1} and $l > 2$.

The proof of Lemma 3.8 explicitly uses the expression of the worst-case belief and therefore has to be adjusted.

Proof. In the proposed equilibrium, types $v > v_{n-l}^j$ of j buy at prices $p \geq p_{n-l}$, with type v_{n-l}^j and some types just above v_{n-l}^j buying at price p_{n-l}. But since j does not buy at prices $\{p_{n-l+1}, \ldots, p_{n-1}\}$, types just below v_{n-l}^j must buy at p_n and not before. Therefore, in the proposed equilibrium, it must be that v_{n-l}^j is indifferent between buying at p_{n-l} or p_n. So we have, for buyer j,

$$v_{n-l}^j - p_{n-l} = (v_{n-l}^j - p_n)\hat{H}_{n-l}^j, \tag{3.12}$$

where $\hat{H}_{n-l}^j$ is the worst-case belief, that j gets the offer p_n if he refuses p_{n-l}. One can derive $\hat{H}_{n-l}^j$ as follows. Similar to Section 3.7.1, denote with π_{n-l}^j, the probability that the object stays unsold in period $n-l$, with π_{n-1}^j, the probability that the object stays unsold in period $n-1$ and with ψ_n^j the probability that j gets the offer p_n. Then, $\hat{H}_{n-l}^j = \min \pi_{n-l}^j \pi_{n-1}^j \psi_n^j$ with

$$\pi_{n-l}^j = \frac{2G^{n-l}(v_{n-l}^i)}{1 + G^{n-l}(v_{n-l}^i)},$$
$$\pi_{n-1}^j = G^{n-1}(v_{n-1}^i),$$
$$\psi_n^j = \frac{1}{2} + \frac{1}{2}G^n(v_n^i).$$

Similar to the proof of Proposition 3.1, the properties of rectangularity impliy

$$\hat{H}_{n-l}^j = \frac{(1-\epsilon_{n-l})F^{n-l}(v_{n-l}^i)(1-\epsilon_{n-1})F^{n-1}(v_{n-1}^i)(1+(1-\epsilon_n)F^n(v_n^i))}{1+(1-\epsilon_{n-l})F^{n-l}(v_{n-l}^i)}.$$

Rewriting Equation (3.12) gives,

$$v_{n-l}^j - p_{n-l} = \frac{(p_{n-l} - p_n)\hat{H}_{n-l}^j}{1 - \hat{H}_{n-l}^j}. \tag{3.13}$$

To establish that contrary to what has been supposed, i.e., that there are types of j who will in fact buy at price p_{n-1}, we show that there are types v such that

$$G^j_{n-1}(v) = v - p_{n-1} - (v - p_n)\hat{H}^j_{n-1} > 0,$$

where $\hat{H}^j_{n-1}$ is the usual rectangular worst-case belief, that j gets the offer p_n if he refuses p_{n-1}. It is useful to break up the analysis into several cases.

Case 1) l and t are fixed positive integers:

Intuitively, this is the case where both i and j follow strategies where they do not buy for some finite number of prices. Note that in this case, $\delta(l+t) \to 0$ as $\delta \to 0$.

i) Some types of i buy at p_{n-1}:

In the proposed equilibrium no types of j buy at p_{n-1}. Consider the value of $G^j_{n-1}(\cdot)$ at v^j_{n-l}. We have

$$\begin{aligned} G^j_{n-1}(v^j_{n-l}) &= v^j_{n-l} - p_{n-1} - (v^j_{n-l} - p_n)\hat{H}^j_{n-1} \\ &= (v^j_{n-l} - p_{n-l}) + (p_{n-l} - p_n) - \Delta_{n-1} - \left((v^j_{n-l} - p_{n-l}) + (p_{n-l} - p_n)\right)\hat{H}^j_{n-1} \\ &= (p_{n-l} - p_n)\frac{1 - \hat{H}^j_{n-1}}{1 - \hat{H}^j_{n-l}} - \Delta_{n-1} \\ &> \left(2\frac{1 - \hat{H}^j_{n-1}}{1 - \hat{H}^j_{n-l}} - 1\right)\Delta_{n-1}, \end{aligned}$$

where the second step follows from Equation (3.13) and the third step follows from the fact that $p_{n-l} - p_n \geq p_{n-2} - p_n = \Delta_{n-2} + \Delta_{n-1} > 2\Delta_{n-1}$. From the previous lemma, we know that as $\delta \to 0$: $v^i_{n-l-t} - v^i_n \to 0$, $v^i_{n-l} - v^i_n < v^i_{n-l-t} - v^i_n \to 0$ and $v^i_{n-1} - v^i_n < v^i_{n-l-t} - v^i_n \to 0$. Then, using

$$\epsilon_{n-l} = \frac{\epsilon}{(1-\epsilon)F(v^i_{n-l-t})) + \epsilon}, \qquad F^{n-l} = \frac{F(v^i_{n-l})}{F(v^i_{n-l-t})},$$

$$\epsilon_{n-1} = \frac{\epsilon}{(1-\epsilon)F(v^i_{n-l})) + \epsilon}, \qquad F^{n-1} = \frac{F(v^i_{n-1})}{F(v^i_{n-l})},$$

$$\epsilon_{n} = \frac{\epsilon}{(1-\epsilon)F(v^i_{n-1})) + \epsilon}, \qquad F^{n} = \frac{F(v^i_{n})}{F(v^i_{n-1})},$$

and Lemma 3.7, one can show that

$$\lim_{\delta \to 0} \frac{1 - \hat{H}^j_{n-1}}{1 - \hat{H}^j_{n-l}} > \frac{1}{2}.$$

Hence for sufficiently small δ, we have $G^j_{n-1}(v^j_{n-l}) > 0$.

ii) No types of i buy at p_{n-1} and $l > 2$:

In this case, if buyer j refuses p_{n-1}, he knows that the game proceeds to the next stage. Then, analogously to Equation (3.11) we get

$$\hat{G}^j_{n-1}(v) = v - p_{n-1} - \left(\frac{1}{2} + \frac{1}{2}(1-\epsilon_n)F^n(v^i_n)\right)(v - p_n).$$

It follows that[9]

$$\begin{aligned}
2\hat{G}^j_{n-1}(v^j_{n-l}) &= (v^j_{n-l} - p_{n-1})\left(1-(1-\epsilon_n)F^n(v^i_n)\right) - \Delta_{n-1}\left(1+(1-\epsilon_n)F^n(v^i_n)\right)\\
&= (v^j_{n-l} - p_{n-l} + p_{n-l} - p_n + p_n - p_{n-1})\left(1-(1-\epsilon_n)F^n(v^i_n)\right)\\
&\quad - \Delta_{n-1}\left(1+(1-\epsilon_n)F^n(v^i_n)\right)\\
&= (v^j_{n-l} - p_{n-l})\left(1-(1-\epsilon_n)F^n(v^i_n)\right)\\
&\quad + (p_{n-l} - p_n)\left(1-(1-\epsilon_n)F^n(v^i_n)\right) - 2\Delta_{n-1}\\
&= (p_{n-l} - p_n)\left(1-(1-\epsilon_n)F^n(v^i_n)\right)\left(\frac{\hat{H}^j_{n-l}}{1-\hat{H}^j_{n-l}} + 1\right) - 2\Delta_{n-1}\\
&= (p_{n-l} - p_n)\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\hat{H}^j_{n-l}} - 2\Delta_{n-1}\\
&> \left(l\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\hat{H}^j_{n-l}} - 2\right)\Delta_{n-1},
\end{aligned}$$

where the fourth step follows from Equation (3.13) and the final inequality follows, as before, from the fact that Δ_k is decreasing in k and $p_{n-l} - p_n = p_{n-l} - p_{n-l+1} + p_{n-l+1} - \cdots - p_n > l\Delta_{n-1}$. Similar to Case 1.1) one can show that as $\delta \to 0$

$$\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\hat{H}^j_{n-l}} \to 1$$

and we have for δ sufficiently small $\hat{G}^j_{n-1}(v^j_{n-l}) > 0$ since $l > 2$.[10]

Analogously to Bose and Daripa (2009), we can show that in Case 2) (t is arbitrary and l varies with n) and Case 3) (l is a fixed integer and t varies with n) $G^j_{n-1}(v^j_{n-l}) > 0$ and $\hat{G}^j_{n-1}(v^j_{n-l}) > 0$. For completeness, we repeat it here:

Case 2) t is arbitrary and l varies with n:

In this case the gap $p_{n-l} - p_{n-1}$ does not vanish as $\delta \to 0$. Then, since for any given $\eta > 0$, $\hat{H}^j_{n-l}$ is bounded away from zero, it follows from that $v^j_{n-l} - p_{n-l}$ does not vanish.

[9]Please note that, due to a calculation error in Bose and Daripa (2009), our expression slightly deviates from their expression.

[10]The condition $l > 2$ is needed, since without further conditions on F and ϵ it is not clear if $\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\hat{H}^j_{n-l}} \nearrow 1$ or $\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\hat{H}^j_{n-l}} \searrow 1$. Therefore, for $l = 2$ we get $\hat{G}^j_{n-1}(v^j_{n-l}) \geq 0$ and it could be possible, that j is indifferent between buying at p_{n-2} or buying at p_{n-1} or waiting till p_n.

Since $p_{n-1} < p_{n-l}$, also $v^j_{n-l} - p_{n-1}$ does not vanish. However, $p_n - p_{n-1} \to 0$ as $\delta \to 0$ and $\hat{H}^j_{n-1} < 1$ (for Case 1.i)) and $(1 - \epsilon_n F^n(v^i_n) < 1$ (for Case 1.ii)). Therefore, for δ small enough, $G^j_{n-1}(v^j_{n-l}) > 0$ and $\hat{G}^j_{n-1}(v^j_{n-l}) > 0$.

Case 3) l is a fixed integer and t varies with n:

This is the case when as $\delta \to 0$, $\delta(l+t)$ does not go to zero because t (and n) become arbitrarily large as δ becomes small. However, this is analogous to the case we have analyzed before with i, and j roles switched. We know that in equilibrium, both buyers have types who plan to buy at the price p_{n-l}. If i plans to buy at prices p_{n-l-t} and p_{n-l}, but not to buy in between, the best response of j should involve not buying at prices $\{p_{n-l-t+1}, \ldots, p_{n-k-2}\}$. If $p_{n-l-t} - p_{n-l-1}$ does not go to zero, we can use the arguments from Case 2) above to argue that contrary to what is being supposed, for small δ, buyer i will have some types of positive measure who buy at p_{n-l-1} rather than waiting till p_{n-l}

Now, since $G^j_{n-1}(\cdot)$ (and $\hat{G}^j_{n-1}(\cdot) > 0$) is strictly increasing, continuous, and negative at p_{n-1}, there is $v^j_{n-1} \in (p_{n-1}, v^j_{n-l})$ such that $G^j_{n-1}(v) > 0$ (and $\hat{G}^j_{n-1}(v) > 0$) for $v \in (v^j_{n-1}, v^j_{n-l})$. Since types below v^j_{n-l} do not buy at any price greater than or equal to p_{n-l}, these types (of positive measure) strictly prefer to stop at p_{n-1} rather than wait till p_n. This contradicts the supposed equilibrium. □

3.7.3 Correction Results of Bose and Daripa (2009)

Using $\bar{H}^i_k$ instead of $(1-\epsilon)H^i_k$ leads to several problems in the proofs of Bose and Daripa (2009). Many proofs use the fact that $(1-\epsilon)H^i_k < 1-\epsilon$, but this does not hold for $\bar{H}^i_k$. More precisely, there exists f with full support on $[0,1]$ and $\mathbf{v}^j$, such that $\bar{H}^i_k > (1-\epsilon)$. The numerical example in Section 3.4 illustrates this problem.

We now summarize the results of Bose and Daripa (2009) and show how the proofs have to be adapted when using the corrected worst-case belief.

Proposition 1 of Bose and Daripa (2009) Bose and Daripa (2009) first show that there exists $\bar{\delta} > 0$ such that for all $\delta < \bar{\delta}$, the equilibrium strategy of both players are interior cut-off strategy. The formal proof consists of several lemmata and uses the fact that $H^i_k < 1$ frequently. However, the result still holds under the corrected worst-case belief. Even if the idea of the proof stays the same, many technical details have to be adjusted. First, as in Bose and Daripa (2009), one can show that there exists a positive measure of types of both buyers who plan to buy at price p_n but not at any earlier price.

Lemma 3.9 (Lemma 4 of Bose and Daripa (2009))**.** *In any equilibrium, a positive measure of types of each buyer plan to buy at price* p_1.

Proof. The first part of the proof still holds given the corrected worst-case belief. Suppose j does not plan to buy at prices $p_1, \ldots, p_k$ for $1 \leq k < n$ and p_{k+1} is the first price at which j buys. (This is denoted as $v_1^j = \cdots = v_k^j = 1$ and $v_{k+1}^j < 1$.) Clearly, the best response of i is not to buy at prices $p_1, \ldots, p_{k-1}$. If i refuses p_k the probability that the game reaches p_{k+1} is one, i.e., $\pi_k^i = 1$. Therefore,

$$\begin{aligned}\bar{H}_k^i &= \min_{g \in \Phi_i^k} \frac{1}{2} + \frac{1}{2} G^{k+1}(v_{k+1}^j) \\ &= \frac{1}{2} + \frac{1}{2}(1 - \epsilon_{k+1}) F^{k+1}(v_{k+1}^j) = \frac{1}{2} + \frac{1}{2}(1-\epsilon) F(v_{k+1}^j), \end{aligned} \tag{3.14}$$

where the last step follows from $v_k^j = 1$. Therefore, the payoff from refusing p_k is $(\frac{1}{2} + \frac{1}{2}(1-\epsilon)F(v_{k+1}^j))(v - p_{k+1})$. Define the following function:

$$\begin{aligned}\hat{G}_k^i(v) :=& v - p_k - \left(\frac{1}{2} + \frac{1}{2}(1-\epsilon)F(v_{k+1}^j)\right)(v - p_{k+1}) \\ =& \frac{1}{2}(v - p_k)(1 - (1-\epsilon)F(v_{k+1}^j)) - \frac{1}{2}(1 + (1-\epsilon)F(v_{k+1}^j))\Delta_k. \end{aligned} \tag{3.15}$$

Note that

$$\begin{aligned} 2\hat{G}_k^i(1) &= (1 - p_k)(1 - (1-\epsilon)F(v_{k+1}^j)) - (1 + (1-\epsilon)F(v_{k+1}^j))\Delta_k \\ &> \delta\epsilon - (2-\epsilon)\Delta_k \geq \delta\epsilon - (2-\epsilon)\Delta_1 = \frac{\delta\epsilon^2}{2(1-\delta) + \delta\epsilon} > 0, \end{aligned}$$

where the second step follows from the fact that $(1 - p_k) \geq (1 - p_1) = \delta$, and the fact that $F(v_{k+1}^j) < 1$, and the third step uses $\Delta_1 \geq \Delta_k$. Since $\hat{G}_k^i(v)$ is continuous, increasing in v, and negative at $v = p_k$, there exists v_k^i such that $\hat{G}_k^i(v) > 0$ for $v > v_k^i$ and $\hat{G}_k^i(v_k^i) = 0$. Since we know that i does not plan to buy at any earlier price than p_k, it must be that types $[v_k^i, 1]$ of buyer i plan to buy at p_k. For the second part, Bose and Daripa (2009) use the fact that $(1-\epsilon)H_k^i < 1 - \epsilon$. But, this does not hold given $\bar{H}_k^i$. Therefore, we have to adjust the second part. Using the definition of $\bar{H}_k^i$, we can still show, that $G_k^j(1) > 0$. First, remember that

$$G_k^j(v) = v - p_k - (v - p_{k+1})\bar{H}_k^j$$

with the worst case belief

$$\bar{H}_k^j = \frac{(1-\epsilon)(F(v_k^i) + F(v_{k+1}^i))}{(1-\epsilon)(F(v_k^i) + F(v_{k-1}^i)) + \epsilon}.$$

Then,

$$\begin{aligned}
G_k^j(1) &= (1-p_k) - (1-p_{k+1})\bar{H}_k^j \\
&= (1-p_k)(1-\bar{H}_k^j) - \Delta_k \bar{H}_k^j \\
&\geq \delta(1-\bar{H}_k^j) - \Delta_1 \bar{H}_k^j \\
&= \delta \frac{(1-\epsilon)(F(v_{k-1}^i) - F(v_{k+1}^i)) + \epsilon}{(1-\epsilon)(F(v_k^i) + F(v_{k-1}^i)) + \epsilon} - \frac{\delta\epsilon}{2} \frac{1-\delta}{(1-\delta) + \frac{\delta\epsilon}{2}} \frac{(1-\epsilon)(F(v_k^i) + F(v_{k+1}^i))}{(1-\epsilon)(F(v_k^i) + F(v_{k-1}^i)) + \epsilon} \\
&\geq \delta \frac{(1-\epsilon)(F(v_{k-1}^i) - F(v_{k+1}^i)) + \epsilon - \epsilon(1-\epsilon)}{(1-\epsilon)(F(v_k^i) + F(v_{k-1}^i)) + \epsilon} > 0,
\end{aligned}$$

where the first inequality follows since $1 - p_k \geq 1 - p_1 = \delta$ and $\Delta_k \leq \Delta_1$. The second inequality follows since $\frac{1-\delta}{(1-\delta)+\frac{\delta\epsilon}{2}} < 1$ and $(F(v_k^i) + F(v_{k+1}^i)) \leq 2$ and the last inequality since $(F(v_{k-1}^i) - F(v_{k+1}^i)) \geq 0$. Since $G_k^j(v)$ is increasing and continuous, there are types of j of positive measure near 1 who would deviate and buy at p_k. Contradiction. □

Lemma 3.10 (Lemma 5 of Bose and Daripa (2009)). *Suppose there are types of buyer i who plan to buy at p_{n-l-t} and p_{n-l} but not at prices in between. Then $v_{n-l-t}^i - v_n^i < \delta(l+t)$*

Proof. Since there are no types who buy at prices between p_{n-l-t} and p_{n-l}, buyer j will never accept a price in $\{p_{n-l-t+1}, \ldots, p_{n-l-2}\}$.
Case 1) Some types of buyer j buy at prices p_{n-l-t} and/or p_{n-l-1}, $t \geq 1$:

In this case, if i refuses p_{n-l-t} it is possible, that the game ends before, p_{n-l} is offered. We know, that v_{n-l-t}^i is given by $G_{n-l-t}^i(v) = 0$, i.e.,

$$\begin{aligned}
& v_{n-l-t}^i - p_{n-l-t} = (v_{n-l-t}^i - p_{n-l})\bar{H}_{n-l-t}^i \\
& \qquad = (v_{n-l-t}^i - p_{n-l-t} + \Delta_{n-l-t} + \cdots + \Delta_{n-l-1})\bar{H}_{n-l-t}^i \\
\Leftrightarrow \quad & v_{n-l-t}^i - p_{n-l-t} = (\Delta_{n-l-t} + \cdots + \Delta_{n-l-1}) \frac{\bar{H}_{n-l-t}^i}{1 - \bar{H}_{n-l-t}^i}. \qquad (3.16)
\end{aligned}$$

Here $\bar{H}_{n-l-t}^i$ is the worst-case probability that i gets the offer p_{n-l} if he rejects the current price p_{n-l-t}. Since there are types of buyer i who plan to buy at p_{n-l-1}, $\bar{H}_{n-l-t}^i$ can be derived from the following three probabilities. Let π_{n-l-t}^i denote the probability that the object remains unsold at p_{n-l-t}, $\hat{\pi}_{n-l-1}^i$ denotes the probability that the object remains unsold at p_{n-l-1} and ψ_{n-l}^i the probability that i gets the offer p_{n-l}. Similar to Section 3.7.1 one can show that

$$\begin{aligned}
\pi_{n-l-t}^i &= \frac{2G(v_{n-l-t}^j)}{G(v_{n-l-1}^j) + G(v_{n-l-t}^j)}, \\
\hat{\pi}_{n-l-1}^i &= \frac{G(v_{n-l-1}^j)}{G(v_{n-l-t}^j)},
\end{aligned}$$

$$\psi^i_{n-l} = \frac{1}{2} + \frac{1}{2}\frac{G(v^j_{n-l})}{G(v^j_{n-l-1})}. \tag{3.17}$$

The worst-case probability is then given by

$$\begin{aligned}\bar{H}^i_{n-l-t} &= \min_{g\in\Phi_i^{-1}} \pi^i_{n-l-t}\hat{\pi}^i_{n-l-1}\psi^i_{n-l} = \min_{g\in\Phi_i^{-1}} \frac{G(v^j_{n-l-1}) + G(v^j_{n-l})}{G(v^j_{n-l-t-1}) + G(v^j_{n-l-t})}\\ &= \frac{(1-\epsilon)(F(v^j_{n-l-1}) + F(v^j_{n-l}))}{(1-\epsilon)(F(v^j_{n-l-t-1}) + F(v^j_{n-l-t})) + 2\epsilon}. \end{aligned}\tag{3.18}$$

Further,

$$\begin{aligned}\frac{\bar{H}^i_{n-l-t}}{1-\bar{H}^i_{n-l-t}} &= \frac{(1-\epsilon)(F(v^j_{n-l-1}) + F(v^j_{n-l}))}{(1-\epsilon)(F(v^j_{n-l-t-1}) + F(v^j_{n-l-t}) - F(v^j_{n-l-1}) - F(v^j_{n-l})) + 2\epsilon}\\ &\leq \frac{2(1-\epsilon)}{2\epsilon} = \frac{1-\epsilon}{\epsilon}, \end{aligned}\tag{3.19}$$

where the inequality holds since, $F(v^j_{n-l-1})+F(v^j_{n-l}) \leq 2$ and $F(v^j_{n-l-t-1})+F(v^j_{n-l-t})-F(v^j_{n-l-1}) - F(v^j_{n-l}) \geq 0$. The rest of the proof follows as in Bose and Daripa (2009). We repeat it for completeness. Let $\alpha := \frac{1-\delta}{1-\delta+\frac{\delta\epsilon}{2}}$. Note, that $\alpha > 1$ and from the definition of Δ_k, we have $\Delta_k = \frac{1}{2}\delta\epsilon\alpha^k < \frac{1}{2}\delta\epsilon$. Therefore,

$$\begin{aligned} v^i_{n-l-t} - p_n &= v^i_{n-l-t} - p_{n-l-t} + \Delta_{n-l-t} + \cdots + \Delta_{n-1}\\ &= (\Delta_{n-l-t} + \cdots + \Delta_{n-l-1})\frac{\bar{H}^i_{n-l-t}}{1-\bar{H}^i_{n-l-t}} + \Delta_{n-l-t} + \cdots + \Delta_{n-1}\\ &\leq (\Delta_{n-l-t} + \cdots + \Delta_{n-l-1})\frac{1-\epsilon}{\epsilon} + \Delta_{n-l-t} + \cdots + \Delta_{n-1}\\ &= (\Delta_{n-l-t} + \cdots + \Delta_{n-l-2})\frac{1-\epsilon}{\epsilon} + \frac{\Delta_{n-l-1}}{\epsilon} + \Delta_{n-l} + \cdots + \Delta_{n-1}\\ &< \frac{1}{2}\left(\delta(1-\epsilon)(t-1) + \delta + \delta\epsilon l\right) < \delta(t + \epsilon(1-t) + \epsilon l) < \delta(l+t), \end{aligned}$$

where the second step follows from Equation (3.16), the third step from Equation (3.19) and the fifth step uses $\Delta_k < \frac{1}{2}\delta\epsilon$ for all $k = n-l-t, \ldots, n-1$. Finally, since $v^i_n = p_n$, $v^i_{n-l-t} - v^i_n < \delta(l+t)$.[11]

The proof of the Case 2), follows as in Bose in Daripa. For completeness, we repeat it here.

Case 2) No types of buyer j buy at prices p_{n-l-t} and p_{n-l-1}, $t \geq 1$:

We know that types of i buy at p_{n-l-t} and at p_{n-l} but not at the prices in between. If $t > 1$, any type of i who buys atp_{n-l-t} can deviate profitably and buy atp_{n-l-1} instead. Contradiction. Therefore in this case the only possibility is $t = 1$.

[11] For $t = 1$, or if j only buys at p_{n-l-t} OR p_{n-l-1}, $v^j_{n-l-t} = v^j_{n-l-1}$ and the worst-case belief changes slightly. However, similar to above it follows that $\frac{\bar{H}^i_{n-l-t}}{1-\bar{H}^i_{n-l-t}} \leq \frac{2(1-\epsilon)}{\epsilon}$ and $v^i_{n-l-t} - p_{n-l-t} \leq \delta(1-\epsilon)(t-1) + \delta + \delta\epsilon l < \delta(l+t)$.

So it remains to prove the inequality when $t = 1$ and no type of j buys at p_{n-l-1}. In this case, analogously to Equation (3.15) v^i_{n-l-1} is given by

$$\hat{G}^i_{n-l-1}(v) = v - p_{n-l-1} - \left(\frac{1}{2} + \frac{1}{2}R^i_{n-l-1}\right)(v - p_{n-l}) = 0,$$

where R^i_{n-l-1} is the conditional probability that j rejects p_{n-l}.[12] Using the fact, that $v - p_{n-l} = v - p_{n-l-1} + \Delta_{n-l-1}$, and solving

$$v^i_{n-l-1} - p_{n-l} = \Delta_{n-l-1}\frac{1 + R^i_{n-l-1}}{1 - R^i_{n-l-1}} < \Delta_{n-l-1}\frac{2-\epsilon}{\epsilon},$$

where the inequality uses the fact that $1 - \epsilon_n < 1 - \epsilon$. Similar to Case 1), it follows

$$\begin{aligned} v^i_{n-l-1} - p_n &= v^i_{n-l-1} - p_{n-l-1} + \Delta_{n-l-1} + \cdots + \Delta_{n-1} \\ &< \Delta_{n-l-1}\left(\frac{2-\epsilon}{\epsilon} + 1\right) + \Delta_{n-l} + \cdots + \Delta_{n-1} \\ &= \Delta_{n-l-1}\frac{2}{\epsilon} + \Delta_{n-l} + \cdots + \Delta_{n-1} < \delta + \frac{\epsilon\delta}{2}l < \delta(l+1). \end{aligned}$$

□

Suppose in equilibrium no types of j plan to buy at prices in $\{p_{n-l+1}, \ldots, p_{n-1}\}$. Then, there are no types of i who plan to buy at prices $\{p_{n-l+1}, \ldots, p_{n-2}\}$. The next lemma shows that given this equilibrium strategy, there are types of buyer j who can profit from deviating and buy at p_{n-1}. Hence, the strategy described above cannot be an equilibrium. Then, Proposition 1 follows from iterating this result.

Lemma 3.11 (Lemma 6 of Bose and Daripa (2009)). *There is $\bar{\delta} > 0$ such that for all $\delta < \bar{\delta}$ there are types (of positive measure) of j who buy at p_{n-1}.*

Proof. Again, the structure is similar to Bose and Daripa (2009). Case 2) and 3) follow analogeously to Bose and Daripa (2009) and the proof of Lemma 3.8. Only Case 1) (l and t are fixed positive integers) requires some adjustments In the proposed equilibrium, types $v > v^j_{n-l}$ of j buy at prices $p \geq p_{n-l}$, with type v^j_{n-l} and some types just above buying at price p_{n-l}. But since j does not buy at prices $\{p_{n-l+1}, \ldots, p_{n-1}\}$, types just below v^j_{n-l} must buy at p_n and not before. Therefore, in the proposed equilibrium, it must be that v^j_{n-l} is indifferent between buying at p_{n-l} or p_n. So we have, for buyer j,

$$v^j_{n-l} - p_{n-l} = (v^j_{n-l} - p_n)\bar{H}^j_{n-l}, \tag{3.20}$$

[12] Suppose the lowest price higher that p_{n-l} at which some types of j buy is $p_{n-l-1-s}$. Then $R^i_{n-l-1} = (1 - \epsilon_{n-l-1})\frac{F(v^j_{n-l})}{F(v^j_{n-l-1-s})}$.

where $\bar{H}^j_{n-l}$ is the worst-case belief, that j gets the offer p_n if he refuses p_{n-l}. One can derive $\bar{H}^j_{n-l}$ similar to Equation (3.18)

$$\bar{H}^j_{n-l} = \frac{(1-\epsilon)(F(v^i_{n-1} + F(v^i_n))}{(1-\epsilon)(F(v^i_{n-l} + F(v^i_{n-l-t})) + 2\epsilon}$$

if there are types of i who buy at p_{n-1}. If no types of i plan to buy at p_{n-1}, then $v^i_{n-1} = v^i_{n-l}$ and 2ϵ in the denominator has to be replaced by ϵ.

Rewriting Equation (3.20) gives,

$$v^j_{n-l} - p_{n-l} = \frac{(p_{n-l} - p_n)\bar{H}^j_{n-l}}{1 - \bar{H}^j_{n-l}}. \tag{3.21}$$

To establish that there are types of j who will in fact want to buy at price p_{n-1}, we show that there are types v such that

$$G^j_{n-1}(v) = v - p_{n-1} - (v - p_n)\bar{H}^j_{n-1} > 0,$$

where $\bar{H}^j_{n-1}$ is the usual worst-case belief, that j gets the offer p_n if he refuses p_{n-1}.

Case 1.1) Some types of i buy at p_{n-1}:

In the proposed equilibrium, no types of j buy at p_{n-1}. Therefore, it must be that $G^j_{n-1}(v)$ is not strictly positive for any $v \in [p_{n-1}, v^j_{n-l}]$. Consider the value of $G^j_{n-1}(\cdot)$ at v^j_{n-l}. We have

$$\begin{aligned}
G^j_{n-1}(v^j_{n-l}) &= v^j_{n-l} - p_{n-1} - (v^j_{n-l} - p_n)\bar{H}^j_{n-1} \\
&= (v^j_{n-l} - p_{n-l}) + (p_{n-l} - p_n) - \Delta_{n-1} - \left((v^j_{n-l} - p_{n-l}) + (p_{n-l} - p_n)\right)\bar{H}^j_{n-1} \\
&= (p_{n-l} - p_n)\frac{1 - \bar{H}^j_{n-1}}{1 - \bar{H}^j_{n-l}} - \Delta_{n-1} \\
&> \left(2\frac{1 - \bar{H}^j_{n-1}}{1 - \bar{H}^j_{n-l}} - 1\right)\Delta_{n-1},
\end{aligned}$$

where the third step follows from the fact that $p_{n-l} - p_n \geq p_{n-2} - p_n = \Delta_{n-2} + \Delta_{n-1} > 2\Delta_{n-1}$. Then,

$$\begin{aligned}
&\frac{1 - \bar{H}^j_{n-1}}{1 - \bar{H}^j_{n-l}} \\
&= \frac{\left((1-\epsilon)(F(v^i_{n-l}) - F(v^j_n)) + \epsilon\right)\left((1-\epsilon)(F(v^i_{n-l}) + F(v^j_{n-l+t})) + 2\epsilon\right)}{\left((1-\epsilon)(F(v^i_{n-l}) + F(v^j_{n-l+t}) - F(v^i_{n-1}) - F(v^i_n)) + 2\epsilon\right)} \\
&\quad \cdot \frac{1}{\left((1-\epsilon)(F(v^i_{n-l})F(v^j_{n-1})) + \epsilon\right)}.
\end{aligned}$$

From the previous lemma, we know that as $\delta \to 0$: $v^i_{n-l-t} - v^i_n \to 0$, $v^i_{n-l} - v^i_n < v^i_{n-l-t} - v^i_n \to 0$ and $v^i_{n-1} - v^i_n < v^i_{n-l-t} - v^i_n \to 0$. Therefore, as $\delta \to 0$

$$\frac{1-\bar{H}^j_{n-1}}{1-\bar{H}^j_{n-l}} \to \frac{\epsilon(1-\epsilon)F(v^i_n)+\epsilon^2}{2\epsilon(1-\epsilon)F(v^i_n)+\epsilon^2} > \frac{\epsilon(1-\epsilon)F(v^i_n)+\epsilon^2}{2\left(\epsilon(1-\epsilon)F(v^i_n)+\epsilon^2\right)} = \frac{1}{2}.$$

Hence for sufficiently small δ, we have $G^j_{n-1}(v^j_{n-l}) > 0$.

Case 1.2) No types of i buy at p_{n-1}:

In this case, if buyer j refuses p_{n-1}, he knows that the game proceeds to the next stage and with probability $\frac{1}{2}$ he gets the offer next period. Then,

$$\hat{G}^j_{n-1}(v) = v - p_{n-1} - \left(\frac{1}{2} + \frac{1}{2}(1-\epsilon_n)F^n(v^i_n)\right)(v - p_n).$$

It follows that[13]

$$\begin{aligned}
2\hat{G}^j_{n-1}(v^j_{n-l}) &= (v^j_{n-l} - p_{n-1})\left(1-(1-\epsilon_n)F^n(v^i_n)\right) - \Delta_{n-1}\left(1+(1-\epsilon_n)F^n(v^i_n)\right) \\
&= (v^j_{n-l} - p_{n-l} + p_{n-l} - p_n + p_n - p_{n-1})\left(1-(1-\epsilon_n)F^n(v^i_n)\right) \\
&\quad - \Delta_{n-1}\left(1+(1-\epsilon_n)F^n(v^i_n)\right) \\
&= (v^j_{n-l} - p_{n-l})\left(1-(1-\epsilon_n)F^n(v^i_n)\right) \\
&\qquad + (p_{n-l} - p_n)\left(1-(1-\epsilon_n)F^n(v^i_n)\right) - 2\Delta_{n-1} \\
&= (p_{n-l} - p_n)\left(1-(1-\epsilon_n)F^n(v^i_n)\right)\left(\frac{\bar{H}^j_{n-l}}{1-\bar{H}^j_{n-l}} + 1\right) - 2\Delta_{n-1} \\
&= (p_{n-l} - p_n)\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\bar{H}^j_{n-l}} - 2\Delta_{n-1} \\
&> \left(2\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\bar{H}^j_{n-l}} - 2\right)\Delta_{n-1}.
\end{aligned}$$

Now, similar to Case 1.1) we know that as $\delta \to 0$: $v^i_{n-l-t} - v^i_n \to 0$, $v^i_{n-l} - v^i_n < v^i_{n-l-t} - v^i_n \to 0$ and $v^i_{n-1} - v^i_n < v^i_{n-l-t} - v^i_n \to 0$. Therefore, as $\delta \to 0$

$$\frac{1-(1-\epsilon_n)F^n(v^i_n)}{1-\bar{H}^j_{n-l}} \to \frac{2(1-\epsilon)F(v^i_n)+\epsilon}{(1-\epsilon)F(v^i_n)+\epsilon} > 1.$$

Hence, for sufficiently small δ, $\hat{G}^j_{n-1}(v^j_{n-l}) > 0$.

Analogously to Bose and Daripa (2009), we can show that in Case 2) (t is arbitrary and l varies with n) and Case 3) (l is a fixed integer and t varies with n) $G^j_{n-1}(v^j_{n-l}) > 0$

[13] Please note that, due to a calculation error in Bose and Daripa (2009), our expression slightly deviates from their expression.

and $\hat{G}^j_{n-1}(v^j_{n-l}) > 0$. In the proof of Lemma 3.8 we explain Case 2) and 3) in more detail.

Now, since $G^j_{n-1}(\cdot)$ (and $\hat{G}^j_{n-1}(\cdot)$) is strictly increasing, continuous, and negative at p_{n-1}, there is $v^j_{n-1} \in (p_{n-1}, v^j_{n-l})$ such that $G^j_{n-1}(v) > 0$ (and $\hat{G}^j_{n-1}(v) > 0$) for $v \in (v^j_{n-1}, v^j_{n-l})$. Since types below v^j_{n-l} do not buy at any price greater than or equal to p_{n-l}, these types (of positive measure) strictly prefer to stop at p_{n-1} rather than wait till p_n. This contradicts the supposed equilibrium. □

Lemma 2 of Bose and Daripa (2009) Lemma 2 of Bose and Daripa (2009) generalize the cut-off strategies to perfect cut-off strategies. If $p_k(v)$ is the highest price that type v would accept. Then, monotonicity implies that v accepts all prices lower than $p_k(v)$. The result and the proof do not depend on $\bar{H}^i_k$ and therefore still hold with the corrected worst-case belief.

Proposition 2 of Bose and Daripa (2009) Next, Bose and Daripa (2009) characterize the cut-off types in an equilibrium. Using our corrected worst-case belief the notation of the statement changes slightly. However, the intuition and the proof follow analogously to the proof of Bose and Daripa (2009).

Proposition 3.5 (Proposition 2 of Bose and Daripa (2009)). *For $\delta < \bar{\delta}$, in any equilibrium the strategy of any buyer i is a perfect cut-off strategy $\mathbf{v}^i = (v^i_i, \ldots, v^i_n)$, where $v^i_n = p_n$. Further, for $1 \leq k \leq n-1$, $v^i_k \in (p_k, v^i_{k-1})$, where $v^i = 0$ and v^i_k is given by*

$$v^i_k = p_k + \Delta_k \frac{\bar{H}^i_k}{1 - \bar{H}^i_k},$$

where $\bar{H}^i_k$ is the corrected worst-case belief. Further, for any given $\mathbf{v}^j$, v^i_k is unique.

Proposition 3 of Bose and Daripa (2009) Proposition 3 of Bose and Daripa (2009) shows the existence of a symmetric equilibrium for all $\delta < \underline{\delta}$ with $\underline{\delta} > 0$. The proof only uses the fact that $\bar{H}^i_k < 1$. Since this is satisfied for the corrected belied, existence follows from the proof of Bose and Daripa (2009) which is an application of Brouwer's fixed point theorem.

Proposition 4 of Bose and Daripa (2009) The main result of Bose and Daripa (2009) shows that the seller can extract almost all surplus

Proposition 3.6 (Proposition 4 of Bose and Daripa (2009)). *For any preference parameter $\epsilon > 0$, there exists $\delta^*(\epsilon) > 0$ such that for ans $\delta < \delta^*(\epsilon)$ and $\eta > 0$, there is an MDM such that in any equilibrium of the game induced by the MDM, the item is* *sold if at least one buyer has valuation greater than η and no type obtains an ex-post surplus greater than δ.*

Proof. The first part can be proven as in Bose and Daripa (2009). From the previous results it follows that for any $\epsilon > 0$, there is a $\delta^*(\epsilon) > 0$ such that whenever $\delta < \delta^*(\epsilon)$, an equilibrium exists and all equilibria can be characterized by

$$v_k = p_k + \Delta_k \frac{\bar{H}_k}{1-\bar{H}_k}.$$

Further, as noted in Section 3.2.2 for any $\eta \in (0,1)$, there exists an integer T such that by choosing $n = T$, the price sequence of the MDM covers at least a fraction $(1-\eta)$ of types. The item is not sold to at most types in $[0, \eta]$. Thus, it only remains to show that no type that buys gets an ex-post surplus greater than δ. This part has to be adjusted.

Since types in $[v_k, v_{k-1}]$ buys at price p_k, the ex-post surplus of any type buying at p_k is at most $v_{k-1} - p_k$. Furthermore, the characterization of the equilibrium cut-off types implies

$$\begin{aligned}
v_{k-1} - p_k &= \Delta_{k-1}\left(1 + \frac{\bar{H}_k}{1-\bar{H}_k}\right) = \Delta_{k-1}\frac{1}{1-\bar{H}_k} \\
&= \frac{\epsilon\delta}{2}\underbrace{\left(\frac{1-\delta}{(1-\delta)\frac{\delta\epsilon}{2}}\right)^{k-1}}_{<1}\frac{(1-\epsilon)\overbrace{(F(v_k)+F(v_{k-1}))}^{\leq 2}+\epsilon}{(1-\epsilon)(F(v_{k-1})-F(v_{k+1}))+\epsilon} \\
&< \frac{\epsilon\delta}{2}\frac{2(1-\epsilon)+\epsilon}{(1-\epsilon)\underbrace{(F(v_{k-1})-F(v_{k+1}))}_{>0}+\epsilon} \\
&< \frac{\epsilon\delta}{2}\frac{2}{\epsilon} = \delta.
\end{aligned}$$

Hence, no type can extract an ex-post surplus greater than δ. □

Bibliography

Aryal, G. and R. Stauber (2014). A note on Kuhn's theorem with ambiguity averse players. *Economics Letters 125*(1), 110–114.

Auster, S. and C. Kellner (2020). Robust bidding and revenue in descending price auctions. *Journal of Economic Theory*, forthcoming.

Battigalli, P., E. Catonini, G. Lanzani, and M. Marinacci (2019). Ambiguity attitudes and self-confirming equilibrium in sequential games. *Games and Economic Behavior 115*, 1–29.

Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2015). Self-confirming equilibrium and model uncertainty. *American Economic Review 105*(2), 646–77.

Beauchêne, D., J. Li, and M. Li (2019). Ambiguous persuasion. *Journal of Economic Theory 179*, 312–365.

Bewley, T. F. (2002). Knightian decision theory, part I. *Decisions in Economics and Finance 25*(2), 79–110.

Bose, S. and A. Daripa (2009). A dynamic mechanism and surplus extraction under ambiguity. *Journal of Economic Theory 144*(5), 2084–2114.

Bose, S., E. Ozdenoren, and A. Pape (2006). Optimal auctions with ambiguity. *Theoretical Economics 1*(4), 411–438.

Cheng, X. (2020). Ambiguous persuasion: An ex-ante perspective. mimeo.

Cheng, X. (2021). Relative maximum likelihood updating of ambiguous beliefs. mimeo.

Di Tillio, A., N. Kos, and M. Messner (2016). The design of ambiguous mechanisms. *The Review of Economic Studies 84*(1), 237–276.

Eichberger, J. and D. Kelsey (1999). Education signalling and uncertainty. In *Beliefs, Interactions and Preferences in Decision Making*, pp. 135–157. Springer.

Ellis, A. (2018). On dynamic consistency in ambiguous games. *Games and Economic Behavior 111*, 241–249.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Epstein, L. G. and M. Schneider (2003). Recursive multiple-priors. *Journal of Economic Theory 113*(1), 1–31.

Etner, J., M. Jeleva, and J.-M. Tallon (2012). Decision theory under ambiguity. *Journal of Economic Surveys 26*(2), 234–270.

Fudenberg, D. and J. Tirole (1991a). *Game theory*, Volume 1. MIT Press.

Fudenberg, D. and J. Tirole (1991b). Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory 53*(2), 236–260.

Ghosh, G. P. and H. Liu (2020). Sequential auctions with ambiguity. mimeo.

Gilboa, I. (2009). *Theory of decision under uncertainty*, Volume 1. Cambridge university press Cambridge.

Gilboa, I. and M. Marinacci (2016). Ambiguity and the bayesian paradigm. In *Readings in Formal Epistemology*, pp. 385–439. Springer.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Gilboa, I. and D. Schmeidler (1993). Updating ambiguous beliefs. *Journal of Economic Theory 59*(1), 33–49.

Greenberg, J. (2000). The right to remain silent. *Theory and Decision 48*(2), 193–204.

Hanany, E. and P. Klibanoff (2007). Updating preferences with multiple priors. *Theoretical Economics 2*(3), 261–298.

Hanany, E. and P. Klibanoff (2009). Updating ambiguity averse preferences. *The BE Journal of Theoretical Economics 9*(1), 1–53.

Hanany, E., P. Klibanoff, and S. Mukerji (2020). Incomplete information games with ambiguity averse players. *American Economic Journal: Microeconomics 12*(2), 135–187.

Harsanyi, J. C. (1967). Games with incomplete information played by "bayesian" players, part I. The basic model. *Management Science 14*(3), 159–182.

Harsanyi, J. C. (1968a). Games with incomplete information played by "bayesian" players, part II. Bayesian equilibrium points. *Management Science 14*(5), 320–334.

Harsanyi, J. C. (1968b). Games with incomplete information played by "bayesian" players, part III. The basic probability distribution of the game. *Management Science 14*(7), 486–502.

Hill, B. (2020). Dynamic consistency and ambiguity: A reappraisal. *Games and Economic Behavior 120*, 289–310.

Ju, N. and J. Miao (2012). Ambiguity, learning, and asset returns. *Econometrica 80*(2), 559–591.

Kajii, A. and T. Ui (2005). Incomplete information games with multiple priors. *The Japanese Economic Review 56*(3), 332–351.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Karni, E. (1988). On the equivalence between descending bid auctions and first price sealed bid auctions. *Theory and Decision 25*(3), 211–217.

Kellner, C. and M. T. Le Quement (2018). Endogenous ambiguity in cheap talk. *Journal of Economic Theory 173*, 1–17.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). Recursive smooth ambiguity preferences. *Journal of Economic Theory 144*(3), 930–976.

Knight, F. H. (1921). *Risk, uncertainty and profit.* Boston, MA: Houghton Mifflin Co.

Kops, C. and I. Pasichnichenko (2020). A test of information aversion. mimeo.

Kopylov, I. (2016). Subjective probability, confidence, and bayesian updating. *Economic Theory 62*(4), 635–658.

Kreps, D. M. and R. Wilson (1982). Sequential equilibria. *Econometrica 50*(4), 863–894.

Li, J. (2020). Preferences for partial information and ambiguity. *Theoretical Economics 15*(3), 1059–1094.

Liu, H. and W. Xiong (2016). Dynamic consistency in incomplete information games under ambiguity. *International Journal of Approximate Reasoning 76*, 63–79.

Lo, K. C. (1998). Sealed bid auctions with uncertainty averse bidders. *Economic Theory 12*(1), 1–20.

Lo, K. C. (1999). Extensive form games with uncertainty averse players. *Games and Economic Behavior 28*(2), 256–270.

Lucking-Reiley, D. (1999). Using field experiments to test equivalence between auction formats: Magic on the internet. *American Economic Review 89*(5), 1063–1080.

Maskin, E. and J. Riley (1984). Optimal auctions with risk averse buyers. *Econometrica 52*(6), 1473–1518.

Matthews, S. A. (1983). Selling to risk averse buyers with unobservable tastes. *Journal of Economic Theory 30*(2), 370–400.

Muraviev, I., F. Riedel, and L. Sass (2017). Kuhn's theorem for extensive form ellsberg games. *Journal of Mathematical Economics 68*, 26–41.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research 6*(1), 58–73.

Nishimura, K. G. and H. Ozaki (2002). Economics of self-feeding fear. mimeo.

Ortoleva, P. and D. Shishkin (2020). Ambiguous information and dilation: An experiment. mimeo.

Pires, C. P. (2002). A rule for updating ambiguous beliefs. *Theory and Decision 53*(2), 137–152.

Riedel, F. (2009). Optimal stopping with multiple priors. *Econometrica 77*(3), 857–908.

Riedel, F. and L. Sass (2014). Ellsberg games. *Theory and Decision 76*(4), 469–509.

Riedel, F., J.-M. Tallon, and V. Vergopoulos (2018). Dynamically consistent preferences under imprecise probabilistic information. *Journal of Mathematical Economics 79*, 117–124.

Riley, J. G. and W. F. Samuelson (1981). Optimal auctions. *The American Economic Review 71*(3), 381–392.

Sarin, R. and P. Wakker (1998). Dynamic choice and nonexpected utility. *Journal of Risk and Uncertainty 17*(2), 87–120.

Savage, L. J. (1954). *The Foundations of Statistics.* John Wiley & Sons.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica 57*(3), 571–587.

Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics 6*(3), 379–421.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies 23*(3), 165–180.

Short Curriculum Vitae

Marieke Pahlke

Academic Education

2015 – today	**Bielefeld University, Germany** PhD student, Member of the Bielefeld Graduate School of Economics and Management (BiGSEM)
2018	**Northwestern University's Kellogg School of Management, US** Visiting Research Student invited by Peter Klibanoff (6 months)
2013 – 2015	**Bielefeld University, Germany** Master of Science Wirtschaftsmathematik (Mathematical Economics)
2010 – 2013	**Bielefeld University, Germany** Bachelor of Science Wirtschaftsmathematik (Mathematical Economics)

Working Paper

- Dynamic Consistency in Incomplete Information Games with Multiple Priors
- Dynamic Consistency in Ambiguous Persuasion

Conferences and Seminar Presentations

2021	World Congress of the Game Theory Society (virtual)
2020	TOM seminar PSE (virtual), World Congress of the ES (virtual)
2019	TOM seminar PSE (Paris), ESEM (Manchester), 19th Annual SAET Conference (Ischia), 12th RGS Doctoral Conference in Economics (Bochum)
2018	20th Workshop of Central European Program in Economic Theory (Udine), Workshop Mathematics of Subjective Probability (Milan), 13th BIGSEM Doctoral Workshop on Economic Theory (poster, Bielefeld)
2016	11th BIGSEM Doctoral Workshop on Economic Theory (organizer, Bielefeld)