



universität
wien

DISSERTATION

Titel der Dissertation

„Foundations of Dynamic Games“

Verfasser

Mag. Michael Greinecker

Angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, 2012

Studienkennzahl lt. Studienblatt:
Dissertationsgebiet lt. Studienblatt:
1. Betreuer:
2. Betreuer

094 140
Volkswirtschaftslehre (Economics)
ao. Univ.-Prof. Mag. Dr. Konrad Podczeck
Dipl.-Math. Dr. Martin Meier

Contents

1	Acknowledgements	5
2	Introduction	7
2.1	Payoff-Relevance	8
2.2	Commitment in Extensive Form Games	9
2.3	Strategic Stochastic Processes	10
3	Payoff-Relevance	12
3.1	Introduction	12
3.2	Environment	14
3.3	Markov Perfect Equilibrium	17
3.4	Limitations of MPE and SMPE	19
3.5	Discussion	22
3.6	Appendix	23
4	Commitment in Extensive Form Games	27
4.1	Introduction	27
4.2	Literature	28
4.3	Environment	29
4.4	Modeling Commitment	32
4.5	Commitment Outcomes	34
4.6	Examples	36
4.7	Discussion	40

5 Strategic Stochastic Processes	42
with Klaus Ritzberger	
5.1 Introduction	42
5.2 Literature	44
5.3 Information	46
5.4 Measurable Extensive Forms	54
5.5 Strategies	56
5.6 Kuhn's Theorem	62
5.7 Discussion	66
5.8 Appendix I: General Preliminaries	68
5.9 Appendix II: Kernels and Conditional Probability	71
References	76
Abstract	82
Zusammenfassung	83
Curriculum Vitae	84

Chapter 1

Acknowledgements

My advisors, Konrad Podczeck and Martin Meier made this thesis possible. Without their encouragement, support, criticism and feedback, I would have gotten lost on the way. They gave me the freedom to develop my own ideas and helped me to avoid many blind alleys.

Klaus Ritzberger has supported my career on countless occasions, provided valuable feedback for my work and coauthored the paper “Strategic Stochastic Processes”, a great experience for me. He also lead a reading group at the Institute of Advanced Studies (Vienna), which attended I for a long time and which strongly shaped my views on game theory and economics in general. Two of the papers in this thesis, “Payoff-Relevance” and “Strategic Stochastic Processes” grew out of discussions that started there.

Many more people have given me insightful feedback on some of my papers and supplied me with additional material and ideas. I want to thank Guilherme Carmona, Venkataraman Bhaskar, Lawrence Blume, Egbert Dierker, Ehud Lehrer, Christina Pawlowitsch, Kirk Sturtz, Philip Reny, Ariel Rubinstein, and Nik Weaver.

People at the Economics Department at the University of Vienna, the Institute of Advanced Studies (Vienna), and the Vienna Graduate School of Economics have always been very supportive. For the paper “Payoff-Relevance”,

I received financial support from the Graf Hardegg Foundation, which is appreciated.

I want to thank Zvika Neeman for making it possible to spend the winter semester of 2011 at the economics department of Tel Aviv University, an extremely inspiring environment for doing game theory.

Writing a thesis proved to be more than just a purely intellectual challenge. Many a friend helped me to to make it through the process and I'm feeling blessed for not being able to mention them individually. Thank you!

Chapter 2

Introduction

Non-cooperative game theory has become the language for most of microeconomic modeling and a significant part of macroeconomic modeling. Describing a strategic environment explicitly amounts to writing down an extensive form game.¹ Applied game theory often applies the conceptual tools from game theory with calculus and recursive optimization methods. Applied games tend to be infinite. Prices and quantities can take on a continuum of values. Abstract existence theorems are of much less relevance in applied work, equilibria are shown to exist by actually finding them. On the other hand, applied game theory makes often strong conceptual simplifications. The focus is often on symmetric equilibria and on equilibria in pure strategies. Dynamic games are often analyzed by restrictive solution concepts such as Markov perfect equilibrium or open-loop equilibrium. The papers in this thesis share the aim of understanding applied game theoretic modeling through the lens of classical game theory. They also share a methodological approach that tries to isolate the relevant information for understanding concepts. These two themes will run through the present discussion of the three papers that make up this thesis.

¹In epistemic game theory, beliefs are often treated as part of the strategic environment too. See, for example, (Aumann & Dreze, 2008) for a strong statement of this view.

2.1 Payoff-Relevance

Many strategic interactions occur over time, often without any natural ending. So dynamic games are popular for modeling many phenomena, from models of fishing to monetary policy. Dynamic games allow for very complex strategies, there is much a player can condition on. The predictive power of standard solution concepts such as subgame perfect equilibrium is weak. Folk theorem style reasoning can often be used to show that anything goes. In order to obtain sharper predictions, significant restrictions on strategies are often imposed. The most popular such approach relies on stationary strategies, strategies are only allowed to depend on exogenously given states. A profile of stationary strategies with the property that every strategy is optimal, in its restricted class, at every state is known as a Markov perfect equilibrium, a closed-loop equilibrium, or a stationary equilibrium. Using standard dynamic programming methods, one can show that (under rather mild assumptions) such a strategy profile is actually a subgame perfect equilibrium without any restrictions (see, for example, 14.2.3. in (Amir, 2006)). The official reason usually given for using Markov perfect equilibria is that one wants to have a model in which behavior depends only on the “payoff-relevant states of the world”. But what is payoff-relevant should be determined by the extensive form game itself, and not be a modeling choice. For this reason, Maskin and Tirole proposed a way of deriving the payoff-relevant states in terms of the extensive form in (Maskin & Tirole, 2001). The basic idea is the following: If one starts with an exogenously given set of states such that the game has a stationary structure with respect to this set of states, one can identify each state with the set of histories leading to the state. This way, one gets a partition of the nonterminal histories with the property that any two histories in the same cell allow for the same continuations and players induced preferences over strategies are the same at both histories, provided all other players use stationary strategies. So Maskin and Tirole took the coarsest such partition to be the set of payoff-relevant states. To show that such a partition actually exists, they made two restrictive assumptions. First, they assumed that there are finitely many players

who can at each history choose only among finitely many actions. Second, players can explicitly condition on calendar time. This rules out the majority of applications of Markov perfect equilibria, which are in games with infinite action spaces and use time-independent strategies. Luckily, as I show, one can prove their result without any assumptions by using some tools from lattice theory. This makes their notions applicable to a large literature of applied dynamic games in economics. But there is a deeper conceptual problem in their formulation of payoff-relevance. Equivalent histories are required to allow for equivalent continuations. But continuations are not defined in terms of deep strategic structure. They are sequences of names of actions and names are not payoff-relevant. Their reformulation of Markov perfect equilibrium does not have any bite when one allows for a relabeling of actions. I do provide a partial fix: I show that for a large class of games that are “not too symmetric”, there exists a canonical labeling of actions in terms of deep strategic structure that ensures that Markov perfect equilibria behave the way they should and induce isomorphic behavior in isomorphic subgames, thus satisfying what is commonly known as “subgame consistency”.

I want to emphasize that I analyze the notion of payoff-relevance in order to understand a widely employed solution concept better. I do not advocate its use and see a lot of merit in considering history-dependent behavior.

2.2 Commitment in Extensive Form Games

Another popular solution concept for dynamic games is open-loop equilibrium. In an open-loop equilibrium, the behavior of a player depends only on time and not on what a player learns about the behavior of other players. Open-loop equilibria are often taken as the “commitment solution of a game” (see, for example, (Van Long, 2010)). They are usually not subgame perfect, but Nash. Now, most commitment problems in economics strike me as problems of conditional commitments, as problems of threats and promises. In the paper, I provide a notion of modeling commitment

that strikes me as a more faithful rendition of the idea of commitment. A different reason for studying commitments was motivated by my first exposure to game theory. There I learned about Nash equilibrium and later that Nash equilibrium relies on threats and promises that are not credible, which motivated more demanding solution concepts such as subgame perfect equilibrium and sequential equilibrium. But this was not a formal statement. If it were true, one would assume that one obtains Nash equilibria when incorporating all possible commitments in a game and then apply an equilibrium notion that does not rely on some implicit commitments. In my framework, this turned out not to be true. Nash equilibrium behavior (in pure strategies) can be supported by commitments, but not only Nash behavior.

The actual formal approach I take is in the spirit of (Schelling, 1960). Players commit by burning their bridges. They can, at each round, publicly dispose of some available choices at no cost. I start with an extensive form game with perfect recall in which the notion of “every round” is meaningful and construct a new extensive form game that takes the original game and allows for all possible commitments. To analyze behavior with commitments, I examine the sequential equilibria in the extended game. Every such equilibrium induces a distribution over the outcomes in the original game and I take such distributions to be commitment outcomes. The central result says that every play induced by a Nash equilibrium in pure strategies is a commitment solution. The result fails for mixed strategy Nash equilibria or if commitments are not completely costless, as two examples show.

2.3 Strategic Stochastic Processes

The last paper of the thesis is joint work with Klaus Ritzberger. Together with Carlos Alós-Ferrer, he developed an extremely general framework for studying extensive forms in (Alós-Ferrer & Ritzberger, 2005) and (Alós-Ferrer & Ritzberger, 2008). They clarified many foundational questions such as when exactly strategy profiles induce outcomes uniquely. Something that was still missing in their framework was a theory of strategic randomization; their framework was essentially deterministic. We are working on closing

this gap.

This paper contains essential building blocks for a general theory of randomization in extensive form games. Ultimately, we want to have a very general framework for studying strategic randomization in extensive forms. The paper focuses on two areas: The modeling of information by certain quotients of measurable spaces and the general measure theoretic machinery that has to run in the background. The class of extensive forms is restricted to ones in which there is perfect recall and players are aware of calendar time, a case general enough to develop the main themes. These themes are also the part where most of my individual contribution lies. We plan on developing the framework in full generality. It is shown that a form of Kuhn's theorem holds in the present framework: For every mixed strategy, there is an outcome-equivalent behavior strategy. The proof builds on previous work in (Aumann, 1964) and (Mertens *et al.*, 1994), but delivers, within the class of games considered, a stronger result.

There are many ways to represent finite extensive forms that are essentially equivalent. Moreover, showing this equivalence is usually straightforward. The extensive forms we consider in this paper are in the spirit of (von Neumann & Morgenstern, 1947), making a choice amounts to reducing the set of feasible outcomes. Part of the paper is showing that various equivalent formulations are also equivalent in the measure theoretic sense, certain isomorphisms turn out to be measurable isomorphisms with the structure we use.

This paper has clearly a slightly mathematical bent and there is no point in trying to hide this fact. The organization therefore follows the logical development more than the economic ideas. To make the paper reasonably self-contained, there are two appendices to clarify notation and terminology and gather part of the mathematical background that is less widely known.

Chapter 3

Payoff-Relevance

3.1 Introduction

We prefer simple models to complicated models. In game theory, we deal with models in which the players use models themselves. A player has to have a model of the strategic environment she faces and that model has to contain everything that is relevant to her decision making. Yet what is relevant to her, will often depend on the models used by other players. A seemingly irrelevant variable may become important to her because other players make their behavior conditional on that variable and their behavior matters to the player.

The model of the strategic environment used when making a choice will usually depend on past play. Partially because past play influences what players can do and want, partially because other players take past play into account themselves. We want to focus on the former aspect, so we have to look at minimal models of the past and we have to look at the models of players jointly. The models players will use in this paper are simply sets of states, with each state summarizing the part of the past that is still relevant to today's decision making, states are what is payoff-relevant. Players employ their models when making decisions, they choose strategies that depend on states alone. We will look at how small such models can be in order to drop as much payoff-irrelevant knowledge of the past as possible.

Our starting point is the reformulation of a (stationary) Markov perfect equilibrium due to Maskin and Tirole in (Maskin & Tirole, 2001).¹ When payoff-relevant states were still just a modelling choice made on a case by case basis, they undertook the important task of defining the payoff relevant states in terms of the extensive form. In the setting of dynamic games with observable actions, they defined a Markov perfect equilibrium as a subgame-perfect equilibrium in which every player's strategy is "measurable with respect to the coarsest partition of histories for which, if all other players use measurable strategies, each player's decision-problem is also measurable." By this, they provided an explicit foundation for what is basically the most popular solution concept for dynamic and stochastic games in applied work.²

In order to show that such a partition actually exists, Maskin and Tirole made two restrictive assumptions: They assumed that all players can condition on calendar time and they assumed that players can choose only among finitely many actions when called to make a decision. The first assumption goes against the spirit of payoff-relevance, the second assumption rules out many important applications of dynamic games. Both restrictions turn out to be superfluous (Theorem 1).

The definition of payoff-relevant states by Maskin and Tirole is quite sensitive to what they take to be decision problems. Two actions are treated differently, whenever they have different names. For this reason, in a Markov perfect equilibrium, players may make non-isomorphic choices in isomorphic subgames. But names of actions are not payoff relevant, they merely decorate the game tree. In general, their notion of Markov Perfect Equilibrium may tell us very little about what the payoff-relevant states are. Actually, any subgame-perfect equilibrium becomes a stationary Markov perfect equilibrium after a simple relabeling of actions (Theorem 2). This is not just troubling on a conceptual basis, it also renders the concept inapplicable when one uses a definition of the extensive form that does not identify

¹An accessible textbook treatment of their approach can be found in (Mailath & Samuelson, 2006), section 5.6.

²See (Amir, 2006) for an overview.

different actions by names. In the classical graph-theoretic definition of the extensive form of Kuhn in (Kuhn, 1953), Markov perfect equilibrium cannot even be defined.

For a large class of dynamic games that are “sufficiently asymmetric”, there exists a canonical labeling of actions that ensures that all Markov perfect equilibria induce isomorphic choices in isomorphic subgames (Theorem 3). Moreover, this labeling uses only information included in every reasonable definition of the extensive form (payoffs and the order structure of moves). For games outside this class of games, a relabeling of this kind may not exist.

Major proofs are collected in an appendix.

3.2 Environment

We will focus exclusively on games with observable actions³ in discrete time. In the usual definition of an extensive form game, dynamics are represented by a graph and information and choices are modeled by partitions. For games with observable actions, such a definition is unnecessarily cumbersome. Information sets coincide with histories and making an arbitrary temporal ordering of moves when players move simultaneously adds only irrelevant information. For our analysis, we will take histories, which are nothing but sequences of action profiles, as the basic unit of analysis. Our formal model has three ingredients: A set of players, a set of histories and for each player, preferences over plays.

We start with a set of *players* I . We merely require I to be nonempty, we do not rule out an infinity of players. Our second ingredient is a set \mathcal{H} of *histories*. The members of \mathcal{H} are finite or infinite sequences of I -tuples.⁴ Each term of the sequence is an *action profile* and each coordinate in an action

³Such games are also known as games of almost perfect information, games with perfect monitoring, or as simultaneous move games.

⁴Formally, an I -tuple is simply a function with domain I .

profile is an *action*. We have to impose some consistency requirements on \mathcal{H} :

- (i) The empty sequence \emptyset belongs to \mathcal{H} .
- (ii) If $(\mathbf{a}^t)_{t=1}^{T_1} \in \mathcal{H}$ (with $T_1 \in \mathbb{N} \cup \{\infty\}$) and $T_2 < T_1$ then $(\mathbf{a}^t)_{t=1}^{T_2} \in \mathcal{H}$.
- (iii) If $(\mathbf{a}^t)_{t=1}^T \in \mathcal{H}$ for all $T \in \mathbb{N}$ then $(\mathbf{a}^t)_{t=1}^\infty \in \mathcal{H}$.
- (iv) Suppose h is not of maximal length. For each player $i \in I$, let

$$A_i(h) = \{a_i : \exists a_{-i} \text{ such that } (h, (a_i, a_{-i})) \in \mathcal{H}\}.$$

Then for all $\mathbf{a} \in \prod_{i \in I} A_i(h)$, $(h, \mathbf{a}) \in \mathcal{H}$.

The conditions (i) and (ii) ensure that \mathcal{H} has a well defined tree structure. Condition (iii) for infinite histories ensures that maximal histories exist even when the time horizon is infinite. Finally, condition (iv) guarantees that the actions a player can play after a certain history are independent of what the others play after that history.

We call histories of maximal length *plays* and all other histories *nonterminal histories*. We denote the set of nonterminal histories by \mathcal{H}^* , these are the histories where players can actually act. Clearly, they all have finite length. The last ingredient in our definition of the extensive form is a reflexive, complete and transitive preference ordering \preceq_i over plays for each player $i \in I$. So $(I, \mathcal{H}, (\preceq_i)_{i \in I})$ is a *game*. We may refer to the set alone \mathcal{H} also as a *game form*.

This framework is quite flexible. We can model situations in which only some players can choose after some history by allowing the other players only to "choose" one single action. This way, we can accommodate games of perfect information, games with overlapping generations of players as in (Bhaskar, 1998), and games with long- and short-run players as in (Fudenberg & Levine, 1989). Technically, we could even work with histories indexed by large ordinals and accommodate transfinite games such as long cheap talk as in (Aumann & Hart, 2003).

A (pure) *strategy* for player $i \in I$ is a function s_i that maps each nonterminal history h to an action in $A_i(h)$. We restrict ourselves to pure strategies in order to avoid measurability problems when dealing with continua of actions. When we specify a strategy for each player, we get a *strategy profile* $(s_i)_{i \in I}$. By recursion, every strategy profile determines a unique play. A *continuation* c of a nonterminal history h is a sequence of action profiles that makes the sequence (h, c) a play.⁵ So a continuation is "what can happen after h ". At each nonterminal history, every player $i \in I$ has conditional preferences \preceq_i^h over continuations such that $c \preceq_i^h c'$ if and only if $(h, c) \preceq_i (h, c')$.

Given a nonterminal history h of length T , we define an *h -subgame* $(I, \mathcal{H}^h, (\preceq_i^h)_{i \in I})$ in a natural way. The set of histories consists of all continuations and initial segments of continuations. The continuations are the plays and preferences for player i are simply \preceq_i^h . Every strategy s_i for player i induces a unique *continuation strategy* $s_i|_h$ in the h -subgame by setting $s_i|_h(h') = s_i((h, h'))$. Since every strategy profile induces a unique play, every player has preferences over her strategies, for fixed strategies of the others. By slight abuse of notation, we also write $s_i \preceq_i s'_i$. A *subgame-perfect equilibrium* of the game $(I, \mathcal{H}, (\preceq_i)_{i \in I})$ is a strategy profile such that for every player $i \in I$, every nonterminal history h , and every possible strategy s of player i , the inequality $s|_h \preceq_i^h s_i|_h$ holds.

Finally, we will need some mathematical definitions related to partitions: A partition Π of a set X is a family of nonempty, pairwise disjoint subsets of X such that $\bigcup_{P \in \Pi} P = X$. We call the elements of a partition *cells*. If Π_1 and Π_2 are partitions of X , we say that Π_1 is *coarser* than Π_2 if every cell in Π_1 is a union of cells in Π_2 . This is equivalent to every cell in Π_2 being the subset of a cell in Π_1 . Let Π be a partition of X and f be a function defined on X . We say that f is Π -*measurable*, if for every cell $P \in \Pi$ and every two elements $x, y \in P$, we have $f(x) = f(y)$. When x and y are in the same cell of the partition Π , we will say that x and y are Π -*equivalent*.

⁵The notation (h, c) denotes the concatenation of the sequences h and c . The context should prevent confusion when the same notation is used for ordered pairs. What we call continuations, are *futures* in (Maskin & Tirole, 2001).

3.3 Markov Perfect Equilibrium

A *partition profile* $(\Pi_i)_{i \in I}$ lists for each player $i \in I$, a partition Π_i of \mathcal{H}^* . A partition profile should provide a model of the relevant past for every player that encompasses the current strategic environment. For this, we need the following consistency conditions, taken from (Maskin & Tirole, 2001). A partition profile $(\Pi_i)_{i \in I}$ is *consistent* if it satisfies the following two conditions:

- (i) If the histories h_1 and h_2 are Π_i -equivalent for any player $i \in I$, then h_1 and h_2 have the same continuations.
- (ii) Suppose all players $j \neq i$ employ Π_j -measurable strategies. If h_1 and h_2 are Π_i -equivalent, then $s_i|h_1 \preceq_i^{h_1} s'_i|h_1$ if and only if $s_i|h_2 \preceq_i^{h_2} s'_i|h_2$ for all strategies s_i and s'_i of player i such that $s_i|h_1 = s_i|h_2$ and $s'_i|h_1 = s'_i|h_2$.

Condition (i) formalizes the idea that all players can behave in the same way in the future and condition (ii) that one can restrict oneself to strategies measurable with respect to ones own partition, provided everybody else choose strategies measurable with respect to their partitions. Condition (ii) does not imply that the preferences over continuations are the same. Consider two histories ending in the following one-shot games:

	L	R		L	R
T	(3,3)	(1,3)	T	(1,1)	(3,1)
B	(3,1)	(1,1)	B	(1,3)	(3,3)

Clearly, both players have different preferences over continuations. But due to the special structure of this example, a player's payoff depends in both games only on the action chosen by the other player. So both players are indifferent between their actions and hence continuation strategies. So (ii) holds trivially.

A partition profile $(\Pi_i)_{i \in I}$ is *dated* if any two histories h_1 and h_2 that are Π_i -equivalent for some player i are of the same length. In a dated partition

profile, players employing measurable strategies can condition on calendar time.

What we are really after are minimal models. We say that the partition profile $(\Pi_i)_{i \in I}$ is *coarser* than the partition profile $(\Pi'_i)_{i \in I}$ if Π_i is coarser than Π'_i for every player $i \in I$. A coarser partition profile corresponds to players using smaller models.

Theorem 1 *There exists a coarsest consistent partition profile and a coarsest consistent dated partition profile.*

The maximally coarse consistent partition profile contains for every player all payoff-relevant information, they form the state space for the player. A special case of this result has been obtained by Maskin and Tirole. They show that a coarsest consistent dated partition profile exists when no player can choose among infinitely many actions after some history and when the set of players is finite. For finitely many players, they have also shown that maximally coarse consistent partition-profiles exist in general by employing Zorn's lemma. They did not rule out the possibility of different, incomparable, maximally coarse consistent partitions in that case.

In applied work, one usually assumes that there is a single state space shared by all players. Our formulation allows for different players to have different partitions in the coarsest consistent partition profile. One could easily formulate everything in terms of a partition common to all players. It is also possible to rule out a diversity of partitions by essentially requiring that everything some player does has an impact on every other player at every time. This is the approach employed by Maskin and Tirole. The assumption is satisfied in games in which everyone's strategic possibilities depend on some aggregate stock, such as the stock of fish in the classical fishwar model in (Levhari & Mirman, 1980).

Finally, Maskin and Tirole define a *Markov Perfect Equilibrium (MPE)* as a subgame perfect strategy profile in which every player plays a strategy mea-

asurable with respect to her partition in the coarsest consistent dated partition profile. They define *stationary Markov Perfect Equilibrium (SMPE)* as a subgame perfect strategy profile in which every player plays a strategy measurable with respect to her partition in the coarsest consistent partition profile. For expositional clarity, we focus on the latter. The discussion can easily be adopted to the dated case.

3.4 Limitations of MPE and SMPE

When we defined consistent partition profiles, we required that two histories are equivalent for any player only if they allow for the same continuations. But sequences of action profiles are a poor representation of a decision problem. In a (S)MPE, players may face isomorphic subgames differing only by the names of actions and behave differently in the two subgames. So (S)MPE does not conform to reasonable notions of subgame consistency⁶. To make this precise, we need to formalize a notion of being essentially the same.

An *isomorphism* between the games $(I, \mathcal{H}, (\preceq_i)_{i \in I})$ and $(I, \mathcal{H}', (\preceq'_i)_{i \in I})$ is a bijection $\phi : \mathcal{H} \rightarrow \mathcal{H}'$ such that:

- (i) For two histories h_1 and h_2 , h_1 is an initial segment of h_2 if and only if $\phi(h_1)$ is an initial segment of $\phi(h_2)$.
- (ii) For any nonterminal history h there exists a family $(\phi_i^h)_{i \in I}$ of bijections $\phi_i : A_i(h) \rightarrow A'_i(\phi(h))$ such that

$$\phi(h, (a_i)_{i \in I}) = (\phi(h), (\phi_i^h(a_i))_{i \in I}).$$

- (iii) For any two plays h_1 and h_2 and any player i , $h_1 \preceq_i h_2$ if and only if $\phi(h_1) \preceq'_i \phi(h_2)$.

⁶The concept of subgame-consistency was explicitly introduced in (Selten & Güth, 1982). The idea itself is already in (Selten, 1965).

Condition (i) guarantees that ϕ is an order-isomorphism under the “is an initial segment of”-ordering. Condition (iii) makes it an order isomorphism on the terminal histories for the preferences of each player. Condition (ii) is required to preserve the internal product structure of histories. Isomorphic games are the same in everything that is payoff relevant. But by a mere relabeling of actions, we can change every subgame-perfect equilibrium into a stationary Markov perfect equilibrium.

Theorem 2 *Given any game, there exists an isomorphic game in which every subgame-perfect equilibrium is a stationary Markov perfect equilibrium.*

If ϕ is an isomorphism between $(I, \mathcal{H}, (\preceq_i)_{i \in I})$ and $(I, \mathcal{H}', (\preceq'_i)_{i \in I})$ and $(s_i)_{i \in I}$ is a strategy profile for $(I, \mathcal{H}, (\preceq_i)_{i \in I})$, there is also an *induced strategy profile* $(s_i^\phi)_{i \in I}$ for $(I, \mathcal{H}', (\preceq'_i)_{i \in I})$. Let h be a nonterminal history in \mathcal{H} of length T and h' be the unique history of length $T + 1$ that occurs when all players follow the strategy profile $(s_i)_{i \in I}$ at h . Now $s_i^\phi(\phi(h))$ is simply defined to be the i^{th} coordinate of the last term of $\phi(h')$.

With this definition out of the way, we can define subgame consistency. A strategy profile $(s_i)_{i \in I}$ in a game $(I, \mathcal{H}, (\preceq_i)_{i \in I})$ is *subgame-consistent* if, whenever there is an isomorphism ϕ between two subgames, then we have $\phi(s_i(h)) = s_i^\phi(\phi(h))$ for all $i \in I$ and every nonterminal history h in the first game. Informally, in a subgame-consistent strategy profile, players facing the same strategic situation behave in the same way. By our last result, there clearly exists stationary Markov perfect equilibria that are not subgame-consistent.

Suppose there are two different isomorphisms between subgames. Then some strategy profile in one subgame induces different strategy profiles under the different isomorphisms. Such a strategy profile cannot be subgame-consistent. To avoid this problem, we look at games in which two subgames can be isomorphic in only one way. Formally, we say that a game is *rigid* if the only isomorphism of the game with itself is the identity mapping.

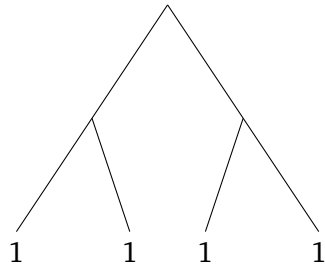
Remark 1 *A game is rigid if and only if no two subgames are isomorphic to each other in more than one way.*

Proof: Since the whole game is a subgame, if two subgames are isomorphic to each other in only one way, the game must be rigid. Conversely, we observe first that every isomorphism from a subgame to itself can be extended to an isomorphism on the whole game that maps each history outside the subgame to itself. So in a rigid game, every subgame can be isomorphic to itself only under the identity. Now, if f and f' are different isomorphisms from one subgame to another, then $f^{-1} \circ f$ and $f^{-1} \circ f'$ are different isomorphisms of the first subgame to itself, which cannot be in a rigid game. ■

For rigid games, we can relabel actions in a way that makes stationary Markov perfect equilibria subgame-consistent.

Theorem 3 *Given any rigid game, there exists an isomorphic game such that every stationary Markov perfect equilibrium is subgame-consistent.*

For games that are not rigid, a mere relabeling will not suffice. Here is a trivial example with only one player:



The two proper subgames are clearly isomorphic, but no pure strategy can make all isomorphic choices at the same time. Any isomorphism reproduces the tree structure, including the payoff-irrelevant redundancies. There is really no decision problem in this example. The example is nongeneric, but this is natural in this context. The symmetries avoided by rigid games stem from nongeneric payoff-ties.

Remark 2 *If in game with only finitely many histories, there exists a player who is not indifferent between any two plays, then the game is rigid.*

Proof: By assumption, there is a player $i \in I$ whose preferences \preceq_i over plays are a linear ordering. Isomorphisms are completely determined by what they do to plays. The identity on \mathcal{H} is the only isomorphism to itself that doesn't change any plays. Now by (iii) in the definition of an isomorphism, the restriction of every isomorphism to the set of plays is an order-isomorphism for \preceq_i . The only order isomorphism of a finite, linearly ordered set to itself is the identity. ■

3.5 Discussion

The notion of subgame consistency used here is quite strong. In particular, we assume that *all* isomorphisms induce behavior in corresponding subgames. This corresponds to the idea that an observer can hold any view on how different subgames are related to each other. A different approach would require that if two subgames are related by *some* isomorphism, there exists *some* isomorphism that relates the behavior in one subgame to the behavior in another subgame. For elementary games, these notions coincide. For other games, requiring this form of subgame-consistency would mean that players have a shared understanding of what subgames are considered to be the same. This requires them to have a more complex model of their environment that is not just based on the structure of the game.

Satisfying our notion of subgame-consistency would be easier if players could employ mixed strategies. Our notion does rule out certain pure equilibria even in one-shot games, where symmetric equilibria always exist in mixed strategies as shown in (Nash, 1951). If one prefers to work with the mixed extension, one replaces actions by behavior strategies, interpreted as transition probabilities. To justify this, one appeals to a suitable extension of Kuhn's theorem, such as the one in Chapter 5. This will require additional assumptions. The extension is relatively straightforward when only finitely many players have nontrivial choices each period and after each nonterminal history, a player can play at most countably many actions. In this case, \mathcal{H}^* will be at most countable and every consistent partition profile will only

include partitions with at most countably many cells. No special measurable or topological structure has to be imposed on the state space. Also, if we start with a dynamic game defined in terms of a given countable state space, we are guaranteed that our construction gives us less states (Proposition 5.6.2. in (Mailath & Samuelson, 2006)), so for many applications, there is no problem in admitting mixed strategies.

Finally, it is worth pointing out that one could separate all results from specific solution concepts such as SPE. Effectively, we have shown that one can represent future strategic possibilities by summary statistics, states, and that for rigid games, one can do this in a way that preserves symmetries in the strategic possibilities. If we look at purely forward-looking forms of behavior, states will contain all relevant information.

3.6 Appendix

The following proof builds on the original proof of Maskin and Tirole and work of Ore on the lattice structure of partitions. The construction of the coarsest consistent partition profile is based on a characterization of the finest common coarsening of a family of partitions due to Ore. Partitions of a set form a complete lattice with the “coarser than”-ordering and Ore laid the foundations for the theory in (Ore, 1942). The reader interested in partition lattices will find a comprehensive overview in (Grätzer, 2003), section IV.4.

Proof of Theorem 1

We show that a coarsest consistent partition profile exists. Essentially the same proof works for dated partition profiles. For each player $i \in I$, we define a relation \equiv_i on \mathcal{H}^* such that $h \equiv_i h'$ if there is a finite sequence of nonterminal histories $h, h_1, h_2, \dots, h_n, h'$ such that consecutive histories in the sequence are equivalent for the partition of i in some consistent partition profile Π_i . The relation \equiv_i is clearly an equivalence relation. Let Π_i^* be the partition of \mathcal{H}^* into \equiv_i -equivalence classes.

We will now show that $(\Pi_i^*)_{i \in I}$ is the coarsest consistent partition profile. It is clearly a partition profile and by construction coarser than any consistent partition profile. It remains to verify that it is also consistent. We begin with consistency condition (i). If h and h' are Π_i^* -equivalent, they are also \equiv_i -equivalent and that means they are connected by a finite sequence of nonterminal histories such that consecutive histories are equivalent under some partition in a consistent partition profile. But then consecutive histories in the sequence must allow for the same continuations by (i) and by transitivity, h and h' allow for the same continuations. This proves (i).

For consistency condition (ii), observe that the coarser a partition is, the less strategies are measurable. That means that whenever all players $j \neq i$ play Π_j -measurable strategies, they play strategies measurable with respect to their partition in every consistent partition profile. Now suppose all players $j \neq i$ employ Π_j^* -measurable strategies. Let h and h' be Π_i^* -equivalent and $s_i|h \preceq_i^h s'_i|h$ with $s_i|h_1 = s_i|h_2$ and $s'_i|h_1 = s'_i|h_2$. Then there exists a finite sequence h_1, \dots, h_{n+1} of nonterminal histories and a finite sequence of consistent partition profiles $(\Pi_i^1)_{i \in I}, \dots, (\Pi_i^n)_{i \in I}$ such that $h_1 = h$, $h_{n+1} = h'$ and h_k, h_{k+1} are in the same cell of Π_i^k for $k = 1, \dots, n$. Now if all players $j \neq i$ employ Π_j^* -measurable strategies, they also employ Π_j^k -measurable strategies. Since all continuations are the same, we can modify s_i and s'_i such that $s_i|h^k = s_i|h$ and $s'_i|h^k = s'_i|h$ for $k = 1, \dots, n$. Hence, $s_i|h^k \preceq_i^{h^k} s'_i|h^k$ for $k = 1, \dots, n$. By transitivity of \equiv_i , consistency condition (ii) holds. ■

Proof of Theorem 2

Proof: Let $(\mathcal{H}, (\succeq_i)_{i \in I})$ be a given game. We rename every action played at a certain history so that actions played after different histories get different names. Here is one way to do this: If $a = (a_i)_{i \in I}$ is an action profile and z some mathematical object, write $a|z$ for $((a_i, z))_{i \in I}$. Set $f(\emptyset) = \emptyset$. Now suppose $f(h)$ is already defined when h has length $T - 1$. For a finite history

$h = (a^t)_{t=1}^T$ of length T , set

$$f(h) = \left(f((a^t)_{t=1}^{T-1}), a^T | h \right).$$

If $h = (a^t)_{t=1}^\infty$ is an infinite history, let $f(h)$ be the unique sequence such that the initial segment of length T coincides with $f((a^t)_{t=1}^T)$ for all T . Set $\mathcal{H}' = f(\mathcal{H})$, which is a game form. Define $(\preceq'_i)_{i \in I}$ on elements of $f(\mathcal{H})$ of maximal length by $f(h) \preceq'_i f(h')$ if and only if $h \preceq_i h'$ for every player i . It is obvious that f is an isomorphism.

Since actions played after different histories are different, there are no two histories in \mathcal{H}' that have a common continuation. So the only stationary consistent partition consists of all singletons and every strategy is measurable with respect to this partition. So every subgame perfect equilibrium uses strategies measurable with respect to this partition and is therefore a SMPE. ■

Proof of Theorem 3

Proof: Let $G = (I, \mathcal{H}, (\preceq_i)_{i \in I})$ be a rigid game. By transforming the game as in the proof of Theorem 2, we can assume that no action can be played after different histories. For each player $i \in I$ we define a function g_i on $A_i = \bigcup_{h \in \mathcal{H}^*} A_i(h)$. By assumption, for each $a_i \in A_i$ there is a unique history h_{a_i} such that a_i can be played after h_{a_i} . Let (a_i, a_{-i}) be a profile with $(h_{a_i}, (a_i, a_{-i})) \in \mathcal{H}$. Now we let $g_i(a_i)$ be the set of all $a'_i \in A_i$ such that there is some h -subgame isomorphic to the h_{a_i} -subgame with isomorphism $f : \mathcal{H}^{h_{a_i}} \rightarrow \mathcal{H}^h$ such that a'_i is the i^{th} coordinate of the last component of $f(h_{a_i}, (a_i, a_{-i}))$. Effectively, $g_i(a_i)$ is the set of actions that serve the same role as a_i in some subgame.

We now construct a function g on \mathcal{H} recursively. We set $g(\emptyset) = \emptyset$. Suppose g is already defined for all histories of length $n-1$ and $h' = (h, (a_i)_{i \in I})$ is a history of length n . Then we set $g(h') = (g(h), (g_i(a_i))_{i \in I})$. If $h = (a^t)_{t=1}^\infty$ is an infinite history, let $g(h)$ be the unique sequence such that the initial segment of length T coincides with $g((a^t)_{t=1}^T)$ for all T . Clearly, g is an injection.

Now $g(\mathcal{H})$ is a game form. Define $(\preceq'_i)_{i \in I}$ on elements of $g(\mathcal{H})$ of maximal length by $g(h) \preceq'_i g(h')$ if and only if $h \preceq_i h'$ for every player i . Then $G' = (I, g(\mathcal{H}), (\preceq'_i)_{i \in I})$ is a game isomorphic to G under the isomorphism $g : \mathcal{H} \rightarrow g(\mathcal{H})$. Moreover, any isomorphic subgames of G' actually coincide by construction, so every SMPE of G' is necessarily subgame-consistent. ■

Chapter 4

Commitment in Extensive Form Games

4.1 Introduction

It is a fundamental tenet of decision theory that having more options can never hurt a decision maker. It also is a fundamental tenet of game theory that other players knowing a player has more options can hurt a player. Players often benefit from publicly getting rid of their options, an insight that predates game theory proper by more than two thousand years in the work of Sun Tzu:

At the critical moment, the leader of an army acts like one who has climbed up a height and then kicks away the ladder behind him. He carries his men deep into hostile territory before he shows his hand. He burns his boats and breaks his cooking-pots; like a shepherd driving a flock of sheep, he drives his men this way and that, and nothing knows whither he is going.

(Giles, 2010)

The idea that players can make strategic commitments to their benefit was introduced to the world of game theory in (Schelling, 1960). Schelling

convincingly argued that players often have no interest to carry out their threats and promises when called to act on them, so that the ability to make binding commitments can be crucial in making threats and promises work. The insight that not every Nash equilibrium is self-enforcing from a dynamic point of view was the starting point of developing extensive form refinements such as sub-game perfection in (Selten, 1965).

In this paper, we take a look at an idealized world in which players can costlessly make any commitments. We start with an extensive form game and create a larger extensive form game that explicitly allows for all possible commitments. Commitments are modeled as a player's ability to burn her bridges. At each stage, she can arbitrarily restrict her future possible choices. In the enlarged game, we study the sequential equilibria. We cannot compare equilibria in the original and the enlarged game directly. But by stripping plays in the extended game from all commitments made along them, we can compare the distributions over plays.

After a brief survey of the related literature, a formal model of extensive form games is given. The formal definition does matter, because we construct the commitment version of an extensive game explicitly in terms of the underlying game in the following section. After that, commitment outcomes are defined and it is shown that every distribution over plays induced by a pure strategy Nash equilibrium is a commitment outcome. This is followed by a number of examples, some remarks and then a concluding discussion.

4.2 Literature

The literature on commitment in which commitment is modeled in a general framework focuses mostly on static, one-shot games. (Tennenholtz, 2004), (Kalai *et al.*, 2010), and (Peters & Szentes, 2012) allow for commitment devices that condition on each other and are effective without any delay. The main results of all these papers are folk theorems to the effect that every individually rational payoff can be implemented with commitments.

(Hamilton & Slutsky, 1993) , (van Damme & Hurkens, 1996) take a

one-shot game and analyze what happens when the ordering of move gets changed. (Bade *et al.*, 2009), and (Renou, 2009) embed a one-shot game in dynamic games in which players can dispose of some actions. In (Caruana & Einav, 2008), players can announce over several periods what actions they will take in a one-shot game played at the end. The last message is binding and changing messages becomes increasingly costly, so talk is not cheap. Among these papers, the one by (Renou, 2009) is most closely related to the present work. A one-shot game gets extended to a two-stage game in which players can first choose subsets of their strategy spaces and then play the game with their chosen reduced strategy spaces. Then, the sub-game perfect equilibria of the extended game are analyzed. Our framework can be seen as a generalization of this approach that allows for explicit dynamics and the results and examples in (Renou, 2009) are of direct relevance to our framework.

4.3 Environment

There is a finite set of *players* I and, by slight abuse of notation, we will also denote the number of players by I . There exists a set of finite sequences of I -tuples \mathcal{H} , the *histories*. A history is meant to capture what has happened in a game up to a certain period. The set \mathcal{H} includes the empty history \emptyset , which we view as the origin of a tree. A history that is of maximal length is a *play*. A history that is not a play is a *nonterminal history*, and this is where players make their choices.

Every player $i \in I$ has a finite-valued *choice correspondence* C_i defined on histories. Choices shouldn't be thought of as actions, each choice incorporates an action and how informed the player taking it. See H3 below and the discussion of H3. The correspondence may be empty-valued at some histories. We write C for the Cartesian product of all choice correspondences. That is, $C(h) = \prod_{i=1}^I C_i(h)$. We make the following assumptions:

- H1 \mathcal{H} is the smallest set that includes the empty sequence \emptyset such that if $h \in \mathcal{H}$ and $c \in C(h)$, then $(h, c) \in \mathcal{H}$.

- H2 If h is a history and $C(h) = \emptyset$, then $C_i(h) = \emptyset$ for all $i \in I$.
- H3 If $C_i(h) \cap C_i(h') \neq \emptyset$ for some histories h and h' and some player $i \in I$, then $C_i(h') = C_i(h)$.
- H4 There exists a number n , such that no history is longer than n .
- H5 If the histories h and h' have different length, then $C_i(h) \neq C_i(h')$ for every player $i \in I$.

Assumption H1 allows us to construct the tree-structure of the game from the choice correspondences. Conversely, assumption H2 ensures that we can recover all choice correspondences from \mathcal{H} . In particular, $C_i(h)$ is the set of choices c_i such that $(h, (c_i, c_{-i})) \in \mathcal{H}$ for some c_{-i} . Assumption H2 allows us to use the histories as basic objects, and treat the choice correspondences as a derived object.

Assumption H3 ensures what is usually ensured by information sets. If a choice is available after different histories, the player must not be able to differentiate between the histories on the basis of the choices she faces at them. One can recover information sets in the usual sense from choice correspondences. The information sets of player $i \in I$ are maximal sets of non-terminal histories on which C_i is constant and nonempty-valued.

Assumption H4 ensures that the game has a finite horizon. Finally, assumption H5 ensures that every player is aware of the calendar time, which is, at least in the setting of finite games, the most restrictive assumption we make. At least in games of perfect recall, every player has a subjective notion of time. Along any play, information sets are linearly ordered. But this notion of time will in general differ from player to player (see (Ritzberger, 1999)). Since we want to model a players ability to make commitments at every round, rounds should be meaningful in our framework, and H5 guarantees just that.

It should be noted that extensive forms as defined here allow for simultaneous choices to be represented as truly simultaneous. The small cost this brings is that we have to introduce dummy choices, that is, a player may not have two choices available at an information set and thus no choice in

the true sense of the word. We call choice situations in which a player has at least two choices available *proper choices* and interpret players without proper choices as inactive.

To tie everything up, we endow every player $i \in I$ with a real-valued payoff-function u_i defined on plays. We call the whole list $(I, \mathcal{H}, (u_i))$ a *game*.

A pure strategy for player $i \in I$ specifies for every choice situation $C_i(h)$ she faces a choice $c \in C_i(h)$. Formally, a *pure strategy* s_i is a set that contains exactly one element from each set in $\{C_i(h) : h \in \mathcal{H}\} \setminus \{\emptyset\}$. A *behavior strategy* consists of a probability distribution over each set in $\{C_i(h) : h \in \mathcal{H}\} \setminus \{\emptyset\}$. If we want to incorporate moves of *nature*, we declare one player to be nature and fix a behavior strategy for her. We say that player $i \in I$ satisfies *perfect recall* if the following two conditions are satisfied:

- R1 If two non-terminal histories h and h' have initial segments of the same length h_n and h'_n such that $C_i(h_n) \neq C_i(h'_n)$, then $C_i(h) \neq C_i(h')$.
- R2 If two non-terminal histories h and h' differ in some period in the choice of i , then $C_i(h) \neq C_i(h')$.

R1 expresses that player i never forgets what she knew and R2 that she never forgets what she did. The game $(I, \mathcal{H}, (u_i))$ satisfies *perfect recall* if every $i \in I$ satisfies perfect recall. In analyzing games with perfect recall, we can restrict players to using behavior strategies instead of mixtures of pure strategies without loss of generality by Kuhn's theorem. That the notion of perfect recall used here coincides with the usual one, follows from the arguments in (Ritzberger, 1999).

A particularly simple class of games are games with observable actions. The game $(I, \mathcal{H}, (u_i))$ has *observable actions* if for all histories $h, h' \in \mathcal{H}$ and all players $i \in I$, one has $C_i(h) = C_i(h')$ implies that $h = h'$ or $C_i(h) = \emptyset$. In words, no player faces the same choices after different histories. After a player makes a choice, it becomes known to everybody. It is immediate that a game with observable actions satisfies perfect recall.

4.4 Modeling Commitment

Our fundamental approach to commitment is that a player can dispose of choices available to her. Let C_i be the choice-correspondence of $i \in I$. We call a correspondence C'_i a *contraction* of C_i , if the following two conditions hold.

C1 For every history h , $C'_i(h) \subseteq C_i(h)$.

C2 If $C_i(h) \neq \emptyset$ then $C'_i(h) \neq \emptyset$.

Condition C2 means that no player can “leave the game.” This condition ensures that the set of plays remains essentially the same when we allow for commitments and allows us to compare outcomes with and without commitments.

For each game G , we construct a *commitment version* $C(G)$. In order to construct the set of histories \mathcal{H}^C , it suffices by H2 to construct all finite histories. We can partition the set of finite histories according to their length and let \mathcal{H}_n^C denote the set of histories with length n . Moreover we can build up the choice correspondences C_i^C from their restrictions $C_{i,n}^C = C_i^C|_{\mathcal{H}_n^C}$. We construct both the finite histories and the restrictions of the choice correspondences recursively. Note that by H1, \mathcal{H}_n^C is determined by \mathcal{H}_{n-1}^C and C_{n-1}^C . Moreover, for each history h in \mathcal{H}_{n-1}^C let $\pi(h)$ be the projection to the corresponding history in the original game. That is, $\pi(h)$ strips h of all commitments. We could define π formally by adding a further recursion.

We let \mathcal{H}_0^C contain the empty sequence alone and for every player $i \in I$, we let

$$C_{i,0}^C(\emptyset) = \left\{ (c, C_i^{(0)}) : C_i^{(0)} \text{ is a contraction of } C_i \text{ and } c \in C_i(\emptyset) \right\}.$$

That is, initially, a player can reduce her strategic possibilities and make an arbitrary choice. Suppose now that both \mathcal{H}_{n-1}^C and $C_{i,n-1}^C$ have been constructed for all i . Now

$$\mathcal{H}_n^C = \left\{ (h, c) : h \in \mathcal{H}_{n-1}^C, c \in C_{n-1}^C(h) \right\},$$

with C_{n-1}^C being the Cartesian product of all the $C_{i,n-1}^C$. Let $C_i^{(n)}$ be the contraction chosen in the last choice played by $i \in I$ in $h \in \mathcal{H}_n^C$. For every player $i \in I$ and $h \in \mathcal{H}_n^C$, let

$$C_{i,n}^C(h) = \left\{ \left\{ (c, C_i^{(n)}) \right\} \times P(h) : C_i^{(n)} \text{ is a contraction of } C_i^{(n-1)} \right. \\ \left. \text{and } c \in C_i^{(n-1)}(\pi(h)) \right\},$$

with $P(h)$ being the public history of commitments made in history h . The function P can be defined recursively too. We have to take the Cartesian product, so that players can make their choices conditional on what commitments have been made already.

For the definition of $C(G)$, we also have to define payoffs. If h and h' are plays in \mathcal{H}^C and $i \in I$ is any player, we let $u_i^C(h) = u_i(\pi(h'))$. So players do not care about commitment moves in themselves, they only care about them instrumentally. This actually defines a game and everything is well defined:

Proposition 1 *The commitment version of a game with perfect recall is a game with perfect recall too. Moreover, π maps plays in \mathcal{H}^C onto the set of plays in \mathcal{H} .*

Proof: This is essentially a straightforward verification. We show that $C(G)$ satisfies perfect recall if G does, the rest is even more straightforward. To see that R1 holds, consider two different non-terminal histories h and h' in $C(G)$ that have initial segments h_n and h'_n of the same length n such that $C_i^C(h_n) \neq C_i(h'_n)^C$. By H3, $C_i^C(h_n) \cap C_i(h'_n)^C = \emptyset$. So h and h' differ either in the last coordinate in $\pi(h_{n+1})$ and $\pi(h'_{n+1})$ or $P(h_{n+1})$ and $P(h'_{n+1})$ are different. In the first case, $C_i^C(h) \neq C_i^C(h')$ follows from the fact that G satisfies R2, in the second case, it holds by construction. Similarly, one verifies R2 from the fact that different choices of $i \in I$ in two nonterminal

histories h and h' can only occur if $\pi(h) \neq \pi(h')$ or $P(h) \neq P(h')$. ■

In this framework, the first commitment and the first choice are made concurrently by each player. But this is without loss of generality. If one wants to allow players to make commitments before the actual game starts, one simply introduces a dummy round in which no player has a proper choice.

It should be noted that, formally, we do not allow for making commitments to behavior strategies. It is not entirely obvious how one could commit to a randomized strategy. The only verifiable way seems to be to condition on an external randomization device. But this amounts to committing to a pure strategy that depends on moves of nature. This is allowed in our framework, but needs explicit modeling. However, the finiteness assumption guarantees that, at least in any game with a non-trivial extensive form, there always exists behavior strategies that cannot be interpreted this way.

A way of modeling even more commitments would be to allow players to reduce their payoff at some plays at each stage. Getting rid of an action could then be modeled as reducing the payoff for all plays compatible with a certain action being played so that it becomes worse than the worst possible play in the original game.

4.5 Commitment Outcomes

A strategy profile for the commitment version does not necessarily induce a strategy profile for the original game. The problem is that players can condition their "ordinary choices" on the contractions chosen before, an option not available in the original game. We therefore define our equilibrium notion in terms of outcomes, the plays: Given the game G , a *commitment outcome* of G is the distribution over plays induced by a sequential equilibrium¹ of $C(G)$ under π . The first result shows that commitments do indeed make

¹A sequential equilibrium is usually defined as a behavior strategy profile and a system of conditional beliefs that are derived as the induced limit beliefs from a sequence of completely mixed behavior strategy profiles converging to the given profile. Clearly, all relevant information is contained in this sequence and we call this sequence itself a sequential equilibrium.

all pure strategy Nash equilibria feasible. It generalizes the corresponding result by (Renou, 2009) for static games.

Theorem 1 *Let G be a game and h^* be induced by a pure strategy Nash equilibrium. Then h^* is a commitment outcome of G .*

Proof: Let $(s_i)_{i \in I}$ be a pure strategy Nash equilibrium. Let $(b_i^n)_{i \in I} \rightarrow (b_i)_{i \in I}$ be an arbitrary sequential equilibrium of $C(G)$. Such an equilibrium always exists. For each $i \in I$, let c_i be the choice after history \emptyset that consists of the choice s_i makes in $C_i(\emptyset)$ and the contraction $C_i^{(0)}$ that reduces the choice correspondence to the at most single-valued correspondence that agrees with s_i . Let b'_i be the behavioral strategy that carries this choice out with probability one and agrees with b_i everywhere else. Construct a new sequence of behavioral strategies (b_i^{n*}) such that $b_i^{n*} = 1/nb_i^n + (n-1)/nb'_i$. Clearly, this sequence converges to b'_i . Moreover, conditional beliefs off the equilibrium path induced by (s_i) agree with those induced by b_i^n along the sequence. We verify that this is a sequential equilibrium. For this, we use the one deviation property of sequential rationality (see Exercise 227.1 in (Osborne & Rubinstein, 1994) or (Hendon *et al.*, 1996)).

Off the equilibrium path, no player has a reason to ever deviate, since b'_i and b_i agree there. On the equilibrium path, players can never deviate after the first round since they have committed themselves to playing according to (s_i) . In the first round, no player has a reason to deviate, since their strategies come from a Nash equilibrium and are therefore ex-ante optimal. ■

It should be noted that we made implicitly much use of commitments being costless. If every other player $j \neq i$ commits to their strategy s_j , player $i \in I$ has no reason to follow and would not do so if there were any cost associated with making a commitment.

It is generally much easier to work with sub-game perfect equilibria instead of sequential equilibria. The two equilibrium concepts agree on games with

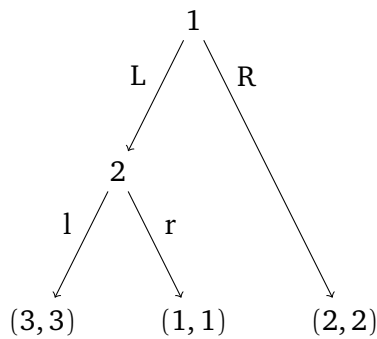
observable actions, since nonterminal histories correspond to both information sets and sub-games in such a game. So the following simple observation will be quite useful.

Proposition 2 *Let $G = (I, \mathcal{H}, (u_i))$ be a game with observable actions. Then $C(G)$ has observable actions too.*

Proof: Let h and h' be two different nonterminal histories in \mathcal{H}^C . Then either $\pi(h) \neq \pi(h')$ or $P(h) \neq P(h')$. In both cases, $C_i^C(h) \neq C_i^C(h')$ for all $i \in I$. ■

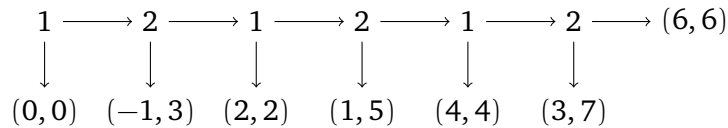
4.6 Examples

A straightforward consequence of Theorem 1 is that the ability to commit may not be to the benefit of any player. The following common interest game has a unique sub-game perfect equilibrium giving both players a payoff of 3, yet has a Nash equilibrium giving both players a mere 2.



The example also shows that Theorem 1 does not hold without the implicit assumption of costless commitments. If commitments are costly, at least one player will not commit initially to a pure strategy. If all others commit to a pure strategy, her decision problem is unaffected. Now a player committing to a strategy will only do so if she is better off with the commitment when commitments are costly. In the example, both players are worse off under commitment, so no player will make an initial commitment.

Not every commitment solution is induced by a Nash equilibrium. An interesting example is given by the centipede game of (Rosenthal, 1981). We use a slightly shortened version with the following extensive form (for simplicity, actions are not labeled; they are ‘down’ and ‘pass’):



There is a unique play induced by every Nash equilibrium and that is the play in which player 1 plays down immediately.² The play in which no player plays down is, however, a commitment solution, supported by the following reasoning. Suppose player 2 makes no initial commitment. Then player 1 can ensure a payoff of 0 by playing down. But even better, she can pass and commit to pass again the next time she is to move, which will compel player two to pass too to get a payoff of at least 5 instead of 2. In this situation, player 2 can either get a payoff of 5 or simultaneously pass and commit. Since player 1 will then pass to get a payoff of 6 instead of 3, player two gains from the commitment.

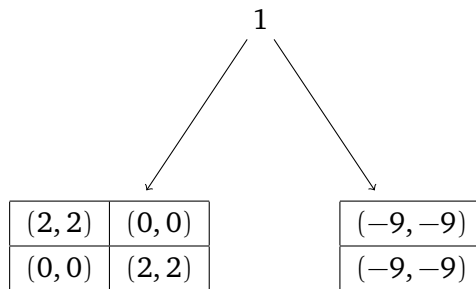
Theorem 1 showed that every pure strategy Nash equilibrium induces a commitment outcome. The proof relied crucially on strategies being pure. It is not possible to fully commit to mixed strategies, since mixtures require some choice being left open. Indeed the result fails for mixed strategy Nash equilibria, as the following example shows (this was already observed in (Renou, 2009), who provides a different example). The game consists of a dummy round without proper choices after which two players play Battle of the Sexes:

(2, 1)	(0, 0)
(0, 0)	(1, 2)

²Given any longer play, the last player that passed on playing down with positive probability has a profitable deviation.

Besides the two pure strategy equilibria, the game has a single nondegenerate mixed strategy equilibrium in which the row player plays the first row with a probability of $2/3$ and the column player plays the second column with a probability of $2/3$. In this equilibrium, both players obtain a payoff of $2/3$. This cannot be a commitment outcome. In order for the mixed strategy equilibrium to be feasible, players must not rule out any choice in the first round. But there is a profitable deviation in this case. The row player can get rid of the second row. In the resulting sub-game, the column player will then choose the first column for sure and this gives the row player a higher payoff of 2.

Making a commitment in our framework achieves actually two things. A player reduces her choices and announces a credible message that she did so. These messages allow players to implement a form of cheap talk. Consider the following game:



Player 1 will never move right initially and the right sub-game seems to be strategically irrelevant. This is actually not so. Player 1 can randomly rule out strategies in the right sub-game and use this to coordinate behavior in the left sub-game. So every correlated equilibrium of the pure coordination game on the left becomes feasible. The public randomization here is fully controlled by player 1. But it is actually possible to do something similar in a framework in which players would have an incentive to manipulate public randomization if they could. Consider the following three player game, which might be seen as a three player version of Battle of sexes: There are three friends 1,2, and 3 who can meet at three possible places 1,2, and 3.

A player who ends up alone at one place gets a payoff of 0. If at least two players meet at one of the places, they get a payoff of at least 2. The player who ends up with a friend at her place (1 at 1, 2 at 2, 3 at 3) gets a payoff of 3. The game has an efficient, symmetric correlated equilibrium in which all players meet at each of these places with a probability of 1/3. Every player would like to be able to manipulate the correlation device that has to be used. But this outcome is implementable as a commitment outcome in a game in which, similarly to the way before, an "irrelevant" sub-game exists in which player 1 and 2 have three choices, which we might denote simply by 1, 2, and 3. Initially, player 1 and 2 can implicitly send messages 1, 2, and 3. There is a sequential equilibrium in the commitment version in which player 1 and 2 choose each of the three messages with probability 1/3. If m_1 and m_2 are the random messages of player 1 and 2, then all players meet at the place $m_1 + m_2 \pmod 3$. Neither player 1 nor player 2 can effect the value of $m_1 + m_2 \pmod 3$ by changing the distribution of her messages. Also, no player has an incentive to make further commitments.

A similar approach allows us to extend a game in seemingly irrelevant ways to allow for any form of public correlation with a distribution with rational values. The construction is similar to various approaches to construct jointly controlled lotteries from the cheap talk literature, as in (Forges, 1990). The mathematical tool to use is the following simple fact:

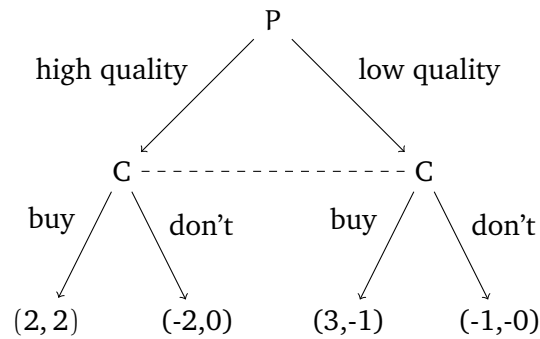
Proposition 3 *Let f and g be independent random variables with values in the set $S = \{1, 2, \dots, n\}$ and assume that f is uniformly distributed on S . Then $f + g \pmod n$ is uniformly distributed.*

Proof: Let $p_m^f = 1/n$ and p_m^g be the probability of f and g taking the value m , respectively. For each m and $k \in S$ there is a unique $l_m(k) \in S$ such that $k + l_m(k) \pmod n$ is m . So

$$\sum_{k=1}^n p_k^g p_{l_m(k)}^f = \sum_{k=1}^n p_k^g 1/n = 1/n \sum_{k=1}^n p_k^g = 1/n$$

is the total probability of choosing m under $\mu + \nu \pmod n$. ■

So far, all our examples were games with observable actions. The following extremely simple example shows how commitments can help in avoiding moral hazard problems. A producer (P) can produce a high quality or a low quality product. Quality is costless. There is a customer (C) who can then buy the product or not, but does not know the quality. The game tree should be rather clear:



Clearly, there is a unique Nash equilibrium in which low quality is produced and then not sold. Moreover, producing low quality is strictly dominant for the producer. But with an initial commitment round, the producer can commit to producing high quality and then actually sell it.

4.7 Discussion

We have modeled commitment by the ability to dispose of choices. An advantage of this approach is that the ultimate behavior in the commitment extension can be compared to behavior in the underlying game. Commitments are not contracts that can incorporate an automatic reaction to other contracts as in (Tennenholtz, 2004), (Kalai *et al.*, 2010), and (Peters & Szentes, 2012). To react to a commitment, a player has to react with her strategy to the observation of the commitment of other players. As a result, commitment solutions can be much smaller than equilibrium notions in which contracts can directly condition on other contracts. The latter approach

means implicitly that there is no temporal gap between reactions, players are patient enough not to care about the gap between observing another contract and reacting to it. As such, it is not surprising that it allows for a folk theorem and using commitment outcomes does not.

Chapter 5

Strategic Stochastic Processes

with Klaus Ritzberger

5.1 Introduction

An extensive form is meant to tell us who can do what, when, and knowing what. It specifies the rules of the game and should provide us with a complete map of the strategic environment. The theory of extensive forms, a kind of strategic cartography, needs a theory of choices, a theory of timing, and a theory of information and knowledge. In a purely deterministic framework, we have an essentially complete theory of extensive forms, developed by Carlos Alós-Ferrer and Klaus Ritzberger in (Alós-Ferrer & Ritzberger, 2005) and (Alós-Ferrer & Ritzberger, 2008). See also (Alós-Ferrer *et al.*, 2011).

But game theory has always been closely related to probability. When Pascal and Fermat lay the foundations for a mathematical treatment of probability at the end of the 17th century, they did so to understand the odds in games of dice (see for example (Hacking, 2006)). The success of game theory is not imaginable without the concept of a mixed strategy and a decision theory powerful enough to deal with risk. This paper is a step towards building an adequate theory of extensive forms that incorporates strategic randomization.

The building blocks of a mathematical treatment of probability in finite

or countably infinite games are straightforward and of a combinatorial nature. The elementary theory of discrete probability suffices. In this paper, we go beyond discrete probability and establish a framework for the treatment of probability in extensive forms based on measure theoretic probability. We do this with a minimum of structure, we avoid topological and quasi-topological regularity assumptions where possible. This gives us deeper insights on what it is that we really need to make extensive form analysis work outside the deterministic framework.

Game theory provides us with peculiar challenges in probabilistic modeling. Everything is based on counterfactuals. We cannot look only at the actual distribution of what happens along plays, we have to look at what could happen if some player changes her behavior. What is a null set will depend on the strategies chosen by other players, so a player must not ignore null sets. This rules out some of the most powerful tools measure theoretic probability theory has to offer.

In our framework, probability is countably additive. Savage's axiomatization of subjective expected utility in (Savage, 1972) provides us only with foundations for finitely-additive probabilities. But finitely additive probabilities lead to absurd conclusions in strategic settings. Consider the following simple zero-sum game, first introduced in (Wald, 1945):

Example 1 *Ann and Bob simultaneously choose a natural number. If both choose the same number, the game ends in a draw. Otherwise, the player who has chosen the larger number wins. It is easy to see that the game has no Nash equilibrium in countably additive mixed strategies. But there is an equilibrium in finitely additive strategies. Both players choose a mixed strategy that puts no mass on any single number and therefore puts no mass on any finite set. So for each number Bob can choose, the probability of Ann choosing a larger number is 1 and she wins for sure! But the game is symmetric, and the same reasoning can be used to show that Bob would win for sure. The problem is that the order in which we evaluate probabilities does matter in the finitely additive case. Fubini's theorem fails, the order of iterated integrals matters.*

Our approach is classical. We treat randomization as objective random-

ization. Players throw dice and coins and spin roulette wheels. It might be possible to interpret our results epistemologically, but we will not pursue such an approach. Also, we will treat everything in terms of an objective outside observer. For simplicity, we only study games satisfying perfect recall and in which players are aware of calendar time, a case that turns out to be challenging enough.

Here is a brief outline of what we do: We start with a brief literature review. We then discuss how we model information and knowledge. This is followed by a discussion of how we model strategies and we show how strategies actually induce outcomes. We then discuss when Kuhn's theorem holds in our framework. That is, when can a player replace mixed strategies by behavior strategies. It fails without further assumptions, there are mixed strategies that cannot be replaced by behavior strategies. We then give sufficient conditions for Kuhn's theorem to hold. Two appendices provide mathematical background information and clarify notation and terminology. It might be a good idea to browse the appendices first. For the issues we discuss, we need not specify any preferences or payoffs.

A small remark on terminology is in order: Every result that is not original is categorized as a lemma in this paper. This should not be taken as a sign of it being a minor result.

5.2 Literature

The foundations of a usable theory of randomization in extensive form games were given by Harold Kuhn in (Kuhn, 1950) and (Kuhn, 1953). Kuhn showed that in finite extensive form games, behavior strategies are as powerful as mixed strategies if and only if the game satisfies perfect recall. More precisely, for every mixed strategy of a player, there is a behavior strategy such that both strategies induce the same outcome for every profile of strategies for the other players. This important fact is now known as Kuhn's theorem. Often the term is applied to the sufficiency side only, a path we

follow here. Kuhn's theorem allows us to analyze randomization in extensive forms with strategy spaces of much smaller dimension. Also, behavior strategies are usually much easier to interpret.

In the extensive forms Kuhn considered, no play can go through the same information set twice and in such games, mixed strategies are always at least as powerful as behavior strategies. It was then shown by John Isbell in (Isbell, 1957) that, in a sufficiently general framework, mixed strategies are as least as powerful as behavior strategies if and only if no play passes twice through the same information set.

A version of Kuhn's theorem for games in which players have a continuum of actions available was given by Robert Aumann in (Aumann, 1964). Aumann showed, in his framework, that if a player has perfect recall, that player can restrict herself to using behavior strategies. Aumann's paper made it obvious that Kuhn's theorem was essentially about existence of conditional probabilities. The paper contains essentially the blueprint for all proofs of Kuhn's theorem. Aumann also pointed out problems with a naive approach to modeling mixed strategies and gave an alternative approach that we will discuss and take up in section 5.5. In spite of the great advances Aumann made, the paper had some serious weaknesses. In particular, only the problem of one player was actually modeled and randomization of other players couldn't be explicitly modeled. Outcomes were deterministic functions of the choices of the player and the profile of other players. A partial converse to Aumann's result was given by Gideon Schwarz in (Schwarz, 1974). Schwarz showed that under an additional measurability assumption, perfect recall is necessary for behavior strategies being as least as powerful as mixed strategies.

The approach of Aumann was extended to a workable framework in a series of exercises in (Mertens *et al.*, 1994). In particular, the behavior of other players was explicitly modeled and the result covers most needs of applied game theory. Our treatment follows a different approach to modeling information and makes use of stronger results on probabilistic conditioning.

It is worth pointing out that the problem of randomized decisions is not confined to game theory. Similar problems have been discussed quite early

in statistical decision theory, starting with (Wald & Wolfowitz, 1951), and in dynamic programming and gambling theory. See in particular (Maitra & Sudderth, 2000) and the references therein.

This paper differs from previous papers addressing similar questions by employing a new framework for modeling information and knowledge, allowing for a countable number of players, and by weakening the common assumption that action spaces are standard Borel.

5.3 Information

The canonical approach to modeling information and knowledge in finite games is based on partitions. But partitions do not provide us with sufficient foundations for general probabilistic settings. The domain of a probability measure is a σ -algebra and partitions do not come with natural σ -algebras. The naive approach of taking the σ -algebra generated by a partition as its informational content leads to absurd conclusions as the following example from (Billingsley, 1995) (Example 33.11) and (Dubra & Echenique, 2004) shows:

Example 2 *Consider the unit interval $[0, 1]$. Nature draws a number using the uniform distribution μ . A player observes the drawn number perfectly, her information partition is the partition into singletons. The σ -algebra \mathcal{S} generated by this partition consists of the subsets of $[0, 1]$ that are either countable or have a countable complement. Our player can choose a number in $[0, 1]$ herself, conditional on her σ -algebra. Her strategy is a function $f : [0, 1] \rightarrow [0, 1]$ measurable with respect to \mathcal{S} .*

No matter what strategy the player chooses, it will equal a constant function almost surely. This is easily shown. For each positive n , there is a closed subinterval of $[0, 1]$ with length at most $1/n$ and probability 1 under the distribution $\mu|_{\mathcal{S} \circ f^{-1}}$. Pick such an interval for each n . Their intersection contains a single point r , so

$$\mu|_{\mathcal{S} \circ f^{-1}}\{r\} = \mu \circ f^{-1}\{r\} = 1.$$

Instead of trying to derive σ -algebras from partitions, we will start with σ -algebras and derive partitions. We still want to be able to work with partitions, since they are easy to interpret. If \mathcal{S} is a σ -algebra on a set S , there is a natural partition into sets of the form

$$\bigcap \{A \in \mathcal{S} : x \in A\} \text{ for some } x \in S.$$

We denote this partition by $\mathcal{A}(\mathcal{S})$ and call its elements *atoms*. This definition is slightly unconventional. Atoms are commonly defined as measurable sets that have no proper nonempty measurable subsets. This agrees with our definition when the atoms in our sense are measurable. The partition into atoms is exactly the partition into elements that cannot be separated by measurable sets. No sensible form of reasoning allows a player to draw distinctions between elements that are finer than what the partition into atoms provides.

Atoms may fail to be measurable. Consider an uncountable product $\{0, 1\}^{\kappa}$ endowed with the product σ -algebra. The atoms are the singletons, but they are not measurable since a measurable set in the product σ -algebra is determined by only countably many coordinates. However, all atoms are measurable when \mathcal{S} is countably generated.¹

Given a probability measure μ on (S, \mathcal{S}) , we want to derive a probability measure on $\mathcal{A}(\mathcal{S})$. For this we need to endow $\mathcal{A}(\mathcal{S})$ with a σ -algebra first. Let $\pi : S \rightarrow \mathcal{A}(\mathcal{S})$ be the *projection* that maps each $x \in S$ to the unique $P \in \mathcal{A}(\mathcal{S})$ such that $x \in P$. We endow $\mathcal{A}(\mathcal{S})$ with the finest σ -algebra that makes π measurable and denote this σ -algebra by $\Sigma_{\mathcal{S}}$. We can think of $\Sigma_{\mathcal{S}}$ as generated by taking a measurable set and contracting its atoms to points. We will use the following lemma a lot:

Lemma 1 *Let (S, \mathcal{S}) be a measurable space. Then $\pi : S \rightarrow \mathcal{A}(\mathcal{S})$ is bimeasurable and*

$$\mathcal{S} = \{\pi^{-1}(B) : B \in \Sigma_{\mathcal{S}}\}.$$

¹Let \mathcal{C} be a countable, nonempty family generating \mathcal{S} . We can assume without loss of generality that \mathcal{C} is closed under complements. Then every atom is of the form $\bigcap \{A \in \mathcal{C} : x \in A\}$ for some $x \in S$.

Proof: We already know that π is measurable. By construction,

$$\Sigma_{\mathcal{S}} = \{B \subseteq \mathcal{A}(\mathcal{S}) : \pi^{-1}(B) \in \mathcal{S}\} = \left\{ B \subseteq \mathcal{A}(\mathcal{S}) : \bigcup B \in \mathcal{S} \right\}.$$

Let $A \in \mathcal{S}$. We have

$$A = \bigcup_{s \in A} \pi(s) = \bigcup \pi(A) = \pi^{-1}(\pi(A)),$$

since A is the union of atoms contained in it. This shows that $\pi(A) \in \Sigma_{\mathcal{S}}$ and that we can reconstruct \mathcal{S} from $\Sigma_{\mathcal{S}}$. ■

As a corollary, we obtain the following universal property:

Corollary 1 *Let (T, \mathcal{T}) be a measurable space. A function $f : \mathcal{A}(\mathcal{S}) \rightarrow T$ is measurable if and only if $f \circ \pi$ is measurable.*

Proof: Necessity is obvious. So let $f \circ \pi$ be measurable and $B \in \mathcal{T}$. Then $\pi(\pi^{-1} \circ f^{-1}(B)) = f^{-1}(B) \in \Sigma_{\mathcal{S}}$. ■

The following lemma shows that $\Sigma_{\mathcal{S}}$ has a very simple structure, which is usually not shared by other quotient constructions:

Lemma 2 *Let (S, \mathcal{S}) be a measurable space and let C be a choice set for $\mathcal{A}(\mathcal{S})$. That is $C \subseteq S$ and $C \cap P$ contains exactly one element for each $P \in \mathcal{A}(\mathcal{S})$. Endow C with the trace σ -algebra $C \cap \mathcal{S}$. Then the restriction $\pi|_C : C \rightarrow \mathcal{A}(\mathcal{S})$ is an isomorphism.*

Proof: By construction, $\pi|_C$ is bijective. We have $(\pi|_C)^{-1}(B) = \pi^{-1}(B) \cap C$ for all $B \in \Sigma_{\mathcal{S}}$, so $\pi|_C$ is measurable. Also, $\pi|_C(A \cap C) = \pi(A)$ for all $A \in \mathcal{S}$, so $\pi|_C$ is bimeasurable. ■

Since the trace σ -algebra of a countably generated σ -algebra is countably generated, it follows from this lemma that $\Sigma_{\mathcal{S}}$ is countably generated when \mathcal{S} is. This proof requires of course the existence of a choice set C , which is ensured by the axiom of choice. We also give a more constructive, but longer proof of this fact:

Corollary 2 *If \mathcal{S} is countably generated, then $\Sigma_{\mathcal{S}}$ is countably generated too.*

Proof: Since \mathcal{S} is countably generated, there is a function $f : S \rightarrow [0, 1]$ such that $\mathcal{S} = \sigma(f)$. The trace σ -algebra $f(S) \cap \mathcal{B}$ on $f(S)$ is clearly countably generated. Since \mathcal{B} separates points, the level sets of f are exactly the atoms of \mathcal{S} . We can therefore construct a bijection $g : f(S) \rightarrow \mathcal{A}(\mathcal{S})$ such that $g(t) = f^{-1}(\{t\})$ for all $t \in f(S)$. We are done when we can show that g is an isomorphism. We have $g \circ f = \pi$, so the following diagram commutes:

$$\begin{array}{ccc}
 S & \xrightarrow{\pi} & \mathcal{A}(\mathcal{S}) \\
 & \searrow f & \nearrow g \\
 & & f(S)
 \end{array}$$

To see that the forward image of a measurable set under g is measurable, let $B \in f(S) \cap \mathcal{B}$. We have $f^{-1}(B) \in \mathcal{S}$ and since π is bimeasurable, $\pi(f^{-1}(B)) = g(B) \in \Sigma_{\mathcal{S}}$.

We now show that the forward image of a measurable set under f is measurable. Let $A \in \mathcal{S}$. Since $\mathcal{S} = \sigma(f)$ and taking the trace on the range of f doesn't change this, there is $B \in f(S) \cap \mathcal{B}$ such that $A = f^{-1}(B)$. Since $f : S \rightarrow f(S)$ is surjective, $B = f(f^{-1}(B)) = f(A) \in f(S) \cap \mathcal{B}$. To see now that g is measurable, let $C \in \Sigma_{\mathcal{S}}$. Then $f(\pi^{-1}(C)) = g^{-1}(C) \in f(S) \cap \mathcal{B}$. ■

A consequence of the proof is that $(\mathcal{A}(\mathcal{S}), \Sigma_{\mathcal{S}})$ can be identified with a subset of $[0, 1]$ with the trace σ -algebra. The next lemma points to an important structural property of our construction:

Lemma 3 *Let (S, \mathcal{S}) and (T, \mathcal{T}) be countably generated measurable spaces and $f : S \rightarrow T$ be a measurable function. Let $\pi_S : S \rightarrow \mathcal{A}(\mathcal{S})$ and $\pi_T : T \rightarrow \mathcal{A}(\mathcal{T})$ be the canonical projections. Then there is a unique measurable function $\pi(f) : \mathcal{A}(\mathcal{S}) \rightarrow \mathcal{A}(\mathcal{T})$ such that the diagram*

$$\begin{array}{ccc}
S & \xrightarrow{f} & T \\
\pi_S \downarrow & & \downarrow \pi_T \\
\mathcal{A}(S) & \xrightarrow{\pi(f)} & \mathcal{A}(T)
\end{array}$$

commutes. Moreover, $\pi(f)$ is bimeasurable if f is.

Proof: Clearly, there is a unique function that maps each $\pi_S(s)$ to $\pi_T(f(s))$. It is measurable, since for all $B \in \Sigma_T$,

$$\pi(f)^{-1}(B) = \pi_S\left(f^{-1} \circ \pi_T^{-1}(B)\right)$$

and π_S is bimeasurable. Now let $A \in \Sigma_S$ and assume that f is bimeasurable. Then

$$\pi(f)(A) = \pi_T\left(f(\pi_S^{-1}(A))\right),$$

which is in Σ_T since π_T and f are bimeasurable. ■

From the categorical point of view, it follows from Lemma 3 that there is an endofunctor π in the category of measurable spaces and measurable functions. By Corollary 2, it is also an endofunctor in the subcategory of countably generated measurable spaces and measurable functions.

In an extensive form, players can learn over time. The increase in information over time can be modeled as a sequence of finer and finer σ -algebras, a filtration. Let $(\mathcal{S}_0, \mathcal{S}_1, \dots)$ be a sequence of σ -algebras on the same underlying set S such that $\mathcal{S}_n \subseteq \mathcal{S}_{n+1}$ for all n . It is easily seen that every cell in $\mathcal{A}(\mathcal{S}_n)$ is the union of cells in $\mathcal{A}(\mathcal{S}_{n+1})$. We can therefore define projections $\pi_{mn} : \mathcal{A}(\mathcal{S}_n) \rightarrow \mathcal{A}(\mathcal{S}_m)$ for $m \leq n$ such that π_{mn} maps a cell $P \in \mathcal{A}(\mathcal{S}_n)$ to the unique cell in $\mathcal{A}(\mathcal{S}_m)$ that contains P as a subset. Let $\pi_n : S \rightarrow \mathcal{A}(\mathcal{S}_n)$ be the usual projection for each n .

$$\mathcal{A}(\mathcal{S}_0) \xleftarrow{\pi_{01}} \mathcal{A}(\mathcal{S}_1) \xleftarrow{\pi_{12}} \mathcal{A}(\mathcal{S}_2) \xleftarrow{\pi_{23}} \dots$$

Lemma 4 The family $(\mathcal{A}(\mathcal{S}_n), \Sigma_{\mathcal{S}_n})$ and the projections $(\pi_{mn})_{m \leq n}$ form a projective system. That is:

(i) The mappings $(\pi_{mn})_{m \leq n}$ are all measurable.

(ii) For all n , π_{nn} is the identity on $\mathcal{A}(\mathcal{S}_n)$.

(iii) For all $l \leq m \leq n$, we have $\pi_{ln} = \pi_{lm} \circ \pi_{mn}$.

Moreover, if $m \leq n$, then $\pi_m = \pi_{mn} \circ \pi_n$.

Proof: Only (i) is not straightforward. Since $\pi_m = \pi_{mn} \circ \pi_n$, we have $\pi_m^{-1}(A) = \pi_n^{-1} \circ \pi_{mn}^{-1}(A) \in \mathcal{S}_m \subseteq \mathcal{S}_n$ for all $A \in \Sigma_{\mathcal{S}_m}$. Since π_n is surjective and bimeasurable, $\pi(\pi_n^{-1} \circ \pi_{mn}^{-1}(A)) = \pi_{mn}^{-1}(A) \in \Sigma_{\mathcal{S}_n}$. ■

Define a set \mathbb{P} by

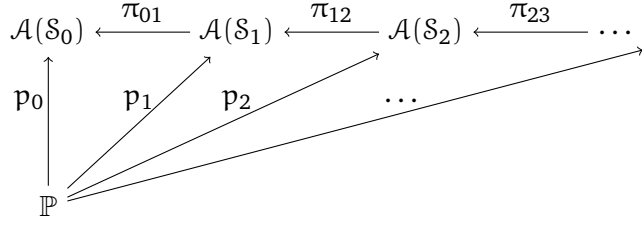
$$\mathbb{P} = \left\{ (x_0, x_1, x_2, \dots) \in \prod_{n=0}^{\infty} \mathcal{A}(\mathcal{S}_n) : m \leq n \text{ implies } x_m = \pi_{mn}(x_n) \right\}.$$

\mathbb{P} is known as the *projective limit*² of the projective system above. Its elements satisfy the restriction that a cell in a coordinate must be a subset of the cell before. We can also write the projective limit as

$$\mathbb{P} = \left\{ (\pi_0(s), \pi_1(s), \dots) : s \in S \right\},$$

the two formulations are easily seen to be equivalent. We can therefore think of the elements of \mathbb{P} as paths of learning. The family of initial segments of elements of \mathbb{P} forms a tree. Let $p_m : \prod_n \mathcal{A}(\mathcal{S}_n) \rightarrow \mathcal{A}(\mathcal{S}_m)$ be the projection onto the m^{th} factor for all m . The embedding of \mathbb{P} in a product allows us to employ mathematical tools developed for product spaces. For this, we have to verify that \mathbb{P} is actually in $\bigotimes_n \Sigma_{\mathcal{S}_n}$. This will be the case when \mathcal{S}_n is countably generated for every n .

²Another term often used is *inverse limit*. Projective systems become then *inverse systems*.



Lemma 5 \mathbb{P} is a measurable subset of $\prod_n \mathcal{A}(\mathcal{S}_n)$ if \mathcal{S}_n is countably generated for all n .

Proof: Let G_{mn} be the graph of $\pi : \mathcal{A}(\mathcal{S}_n) \rightarrow \mathcal{A}(\mathcal{S}_m)$. Since $\Sigma_{\mathcal{S}_m}$ is countably generated and separates points of $\mathcal{A}(\mathcal{S}_m)$, the graph G_{mn} is a measurable subset of $\mathcal{A}(\mathcal{S}_n) \otimes \mathcal{A}(\mathcal{S}_m)$. So $G'_{mn} = \{(y, x) : (x, y) \in G_{mn}\}$ is a measurable subset of $\mathcal{A}(\mathcal{S}_m) \otimes \mathcal{A}(\mathcal{S}_n)$. Now let

$$X_n = \prod_{k=0}^{n-1} \mathcal{A}(\mathcal{S}_k) \times G'_{n,n+1} \times \prod_{k=n+2}^{\infty} \mathcal{A}(\mathcal{S}_k)$$

for every n . Then $\mathbb{P} = \bigcap_{n=0}^{\infty} X_n$, so \mathbb{P} is measurable. ■

We let $\mathcal{P} = \mathbb{P} \cap \bigotimes_n \Sigma_{\mathcal{S}_n}$ be the trace σ -algebra on \mathbb{P} . By the last lemma, every probability measure on \mathbb{P} can be extended to $\prod_n \mathcal{A}(\mathcal{S}_n)$ if \mathcal{S}_n is countably generated for all n . We will from now on assume that \mathcal{S}_n is countably generated for all n . All the information that can be obtained from the sequence (\mathcal{S}_n) is encoded in \mathbb{P} . Another way to analyze the combined information of this sequence is by pooling the σ -algebras. Let $\mathcal{S} = \bigvee_n \mathcal{S}_n$ be the pooled σ -algebra. It is worth pointing out that $\bigvee_n \mathcal{S}_n \neq \bigcup_n \mathcal{S}_n$ unless (\mathcal{S}_n) is an eventually constant sequence. This has been shown in (Broughton & Huff, 1977). Every cell in $\mathcal{A}(\mathcal{S}_n)$ is the union of cells in $\mathcal{A}(\mathcal{S})$ for all n . There is a natural relation between $(\mathcal{A}(\mathcal{S}), \Sigma_{\mathcal{S}})$ and $(\mathbb{P}, \mathcal{P})$.

Lemma 6 The function $\xi : \mathbb{P} \rightarrow \mathcal{A}(\mathcal{S})$ given by $\xi(P_1, P_2, \dots) = \bigcap_n P_n$ is an isomorphism.

Proof: It is easily seen that ξ is bijective. Let $\pi_m^{\vee} : \mathcal{A}(\mathcal{S}) \rightarrow \mathcal{A}(\mathcal{S}_m)$ be the natural projection for each $m \in \mathbb{N}$. That π_m^{\vee} is measurable for each m

follows as in the proof of Lemma 4. It is straightforward to check that the following diagram commutes for all $m \in \mathbb{N}$:

$$\begin{array}{ccc}
 \mathbb{P} \cap \prod_n \mathcal{A}(\mathcal{S}_n) & \begin{array}{c} \xleftarrow{\xi} \\ \xrightarrow{\xi^{-1}} \end{array} & \mathcal{A}(\mathcal{S}) \\
 \searrow \text{p}_m|_{\mathbb{P}} & & \swarrow \pi_m^\vee \\
 & \mathcal{A}(\mathcal{S}_m) &
 \end{array}$$

We show that ξ is measurable, the proof that ξ^{-1} is measurable is analogous. Since $\Sigma_{\mathcal{S}}$ is generated by the sets of the form $\pi_m^{\vee^{-1}}(A)$ with $A \in \Sigma_{\mathcal{S}_m}$ for some m , it suffices to show that $\xi^{-1}(\pi_m^{\vee^{-1}}(A)) \in \mathcal{P}$. Since $\text{p}_m|_{\mathbb{P}}$ is measurable, $\text{p}_m|_{\mathbb{P}^{-1}}(A) \in \mathcal{P}$. Since the diagram commutes, $\text{p}_m|_{\mathbb{P}^{-1}}(A) = \xi^{-1}(\pi_m^{\vee^{-1}}(A))$. ■

Call a family of partitions (Π_i) *stable*, if $\bigcap_i P_i \neq \emptyset$ for all $(P_i) \in \prod_i \Pi_i$. Stable partitions occur naturally in the modeling of simultaneous choices:

Example 3 Consider a normal form game with player set I and strategy spaces (X_i) . For each player $i \in I$, we can identify her strategies with a certain partition of $\prod_{i \in I} X_i$. For each $s \in X_i$, we let

$$P_s = \prod_{j \neq i} X_j \times \{s\}$$

and let Π_i be the partition into sets of the form P_s . The family (Π_i) is then stable.

Note that we made no use of the restriction $\mathcal{S}_n \subseteq \mathcal{S}_{n+1}$ in the proof of Lemma 6 besides the fact that the intersection of elements of \mathbb{P} is nonempty. Therefore, we have essentially also proven the following:

Lemma 7 Let (\mathcal{S}_n) be a sequence of countably generated σ -algebras on the set S . Then the function $\psi : \prod_n \mathcal{A}(\mathcal{S}_n) \rightarrow \bigvee_n \mathcal{A}(\mathcal{S}_n)$ given by $\psi(P_1, P_2, \dots) = \bigcap_n P_n$ is an isomorphism.

It is hard to interpret increases in information if one can not point at the uncertainty resolved. Let us therefore define a *proper filtration* on a set S to

be a sequence (\mathcal{S}_n) of countably generated σ -algebras such that $\mathcal{S}_n \subseteq \mathcal{S}_{n+1}$ and whenever $B \in \mathcal{S}_{n+1}$ is a union of atoms in \mathcal{S}_n , we have $B \in \mathcal{S}_n$. In a proper filtration, a player can only learn by making distinctions between elements previously lying in the same atom.

Ideally, we would like to recover all σ -algebras in a proper filtration from their atoms and $\bigvee_n \mathcal{S}_n$. This is in general not possible. Nik Weaver has contributed the following counterexample on mathoverflow.com.

Example 4 *First, let \mathcal{B} be the Borel σ -algebra on \mathbb{R} and let \mathcal{B}' be the σ -algebra generated by \mathcal{B} together with one non-Borel set E . Note that E is a union of atoms of \mathcal{B} . Now for each n let \mathcal{S}_n be the σ -algebra of subsets of $\mathbb{R} \times \mathbb{N}$ generated by sets of the form $A \times [n, \infty)$ for $A \in \mathcal{B}$ and sets of the form $B \times \{k\}$ for $B \in \mathcal{B}'$ and $k < n$. Since \mathcal{B} is countably generated, so is each \mathcal{S}_n .*

The atoms of \mathcal{S}_n are the singletons $\{(x, k)\}$ for $x \in \mathbb{R}$ and $k < n$ and the sets $\{x\} \times [n, \infty)$ for $x \in \mathbb{R}$. There are no new sets in \mathcal{S}_{n+1} that are unions of atoms of \mathcal{S}_n . However, $E \times \mathbb{N}$ appears in the σ -algebra generated by $\bigcup_n \mathcal{S}_n$, and this is a union of atoms of \mathcal{S}_0 .

5.4 Measurable Extensive Forms

We are now able to put the various ingredients together to arrive at a general model of measurable extensive forms with perfect recall. The formulation is in the spirit of the extensive forms in (von Neumann & Morgenstern, 1947), which it generalizes. Making a choice amounts to refining the set of outcomes that remain feasible.

A measurable extensive form with perfect recall consists of a nonempty countable set of *players* I and for each player $i \in I$ two proper filtrations (\mathcal{J}_n^i) and (\mathcal{C}_n^i) of a nonempty set W . They represent information and choices, respectively. For all $i \in I$ and $n \in \mathbb{N}$ let I_n^i be the set of \mathcal{J}_n^i -atoms and C_n^i be the set of \mathcal{C}_n^i -atoms. We endow them with the σ -algebras $\Sigma_{\mathcal{J}_n^i} = \mathcal{J}_n^i$ and $\Sigma_{\mathcal{C}_n^i} = \mathcal{C}_n^i$, respectively. We require the following conditions to be satisfied:

M1 $\mathcal{S}_0^i = \{W, \emptyset\}$ for all $i \in I$.

M2 \mathcal{S}_{n+1}^i is finer than \mathcal{C}_n^i for all $i \in I$ and $n \in \mathbb{N}$.

M3 \mathcal{S}_n^i is coarser than \mathcal{C}_n^i for all $i \in I$ and $n \in \mathbb{N}$.

M4 \mathcal{S}_{n+1}^i is coarser than $\bigvee_{i \in I, m \leq n} \mathcal{C}_n^i$.

M5 For all $i \in I$ and natural numbers $m < n$, if $x \in I_m^i \cap I_n^i$, then $x \in C_n^i$.

M6 The atoms of $\bigvee_{n,i} \mathcal{C}_n^i$ are singletons.

We call W the set of *plays* and endow it with the σ -algebra $\mathcal{W} = \bigvee_{i,n} \mathcal{C}_n^i$. The σ -algebra \mathcal{S}_n^i represents the information available to player $i \in I$ at period n and \mathcal{C}_n^i the choices available to player $i \in I$ at period n . M1 says that initially, no player knows anything. M2 says that a player can recall what she did in the period before. M3 says that choices refine information available. Together, M2 and M3 imply that a player can recall everything she ever did or knew and thus represent the assumption of perfect recall. M4 says that players cannot learn anything that cannot be reconstructed from the choices of players before. M5 says that an information set that player i can meet twice along a play must be an information set where she cannot actually do something. M6 says that two plays are only different if some player made somewhere different choices along the play. Since $\mathcal{W} = \bigvee_{n,i} \mathcal{C}_n^i$ is countably generated, M6 implies that W has at most the cardinality of the continuum.³ M6 is essentially without loss of generality from the game theoretic point of view. If M6 would fail, we would call the atoms of $\bigvee_{n,i} \mathcal{C}_n^i$ plays instead of the elements of W . In that case, we could replace W by a smaller set without changing the strategic environment so that M6 holds.

In a measurable extensive form, every player has to choose every period. There is no last period, but this is essentially without loss of generality. If the game is supposed to actually end, we can reduce the choices of every player from then on to a single choice. Nothing interesting can happen afterwards. Implicitly, players can condition on calendar time.

³There must be a function $f : W \rightarrow [0, 1]$ such that $\mathcal{W} = \sigma(f)$. Since the atoms of \mathcal{W} are singletons, and the Borel σ -algebra on $[0, 1]$ separates points, f must be injective.

5.5 Strategies

We are now able to discuss what strategies should be. For much of the discussion, we will drop time and player indices. This allows us to focus on conceptual problems that have to be dealt with. We will provide complete definitions afterwards. So we will deal with two countably generated σ -algebras \mathcal{I} and \mathcal{C} on a set W and let I and C be the corresponding sets of atoms endowed with the appropriate σ -algebras \mathcal{J} and \mathcal{C} . By assumption \mathcal{C} is finer than \mathcal{J} .

Primarily, a pure strategy should map pieces of information to available choices. So it is first and foremost a function $s : I \rightarrow C$. A player who knows the piece of information $x \in I$ also knows that she can only choose a choice available at x ; choices refine information. This imposes the condition that $s(x) \subseteq x$ for all $x \in I$. Now if we have a distribution over I , a strategy should also induce a distribution over C and this means that s has to be measurable. So a pure strategy has to be a measurable function $s : I \rightarrow C$ such that $s(x) \subseteq x$ for all $x \in I$.

In finite games, mixed strategies are usually defined as probability distributions over pure strategies. Let S be the family of all pure strategies. There is a function $e : S \times I \rightarrow C$, called the *evaluation*, defined by $e(s, x) = s(x)$. To construct probability distributions on S , we need a σ -algebra \mathcal{S} on S . Now, if we have a probability distribution μ on S and a probability distribution ν on I , we want to be able to talk about the resulting probability distribution on C . If $B \subseteq C$ is measurable, the probability of a choice lying in B would be $e^{-1}(B) \subseteq S \times I$. This set should be measurable, so we want $e^{-1}(B) \in \mathcal{S} \otimes \mathcal{J}$ for every measurable B , which amounts to e being jointly measurable. But Aumann has shown in (Aumann, 1961)⁴ that even in the most well-behaved case, this is generally impossible:

⁴The arguments in (Aumann, 1961) are not very accessible. Easier proofs of the same results can be found in (Rao, 1971), which relies strongly on the machinery of classical descriptive set theory.

Example 5 Let $W = [0, 1] \times [0, 1]$. Let I be the set of lines of the form

$$\{(x, y) : y \in [0, 1]\}$$

for some $x \in [0, 1]$. Let C be the set of singleton subsets of W . We can identify the elements of I with points in the unit interval and of course elements in C with points in the unit square and endow these sets with the corresponding σ -algebras. This allows us to identify strategies with measurable functions $f : [0, 1] \rightarrow [0, 1]$, and S with the set of such measurable functions. Now the set of all measurable functions from the unit interval to itself is not of bounded Borel class.⁵ It follows from Theorem D in (Aumann, 1961) that there is no σ -algebra \mathcal{S} on S such that e is $\mathcal{S} \otimes \mathcal{I}$ -measurable.

As a consequence, Aumann uses a different notion of mixed strategy. He takes mixed strategies to be pure strategies that may condition on an external randomization device. In our context, a mixed strategy will be a jointly measurable function $s : \Omega \times I \rightarrow C$ with (Ω, Σ, μ) being an exogenous probability space such that $f(\omega, x) \subseteq x$ for all $x \in I$ and all $\omega \in \Omega$. The probability space Ω serves as a randomization device and may be interpreted as a roulette wheel.

If $s : I \rightarrow C$ is a pure strategy, we can treat it as a mixed strategy s' such that $s'(\omega, x) = s(x)$ for all x and the randomization device being arbitrary. For every measurable set $B \subseteq C$, $s'^{-1}(B) = \Omega \times s^{-1}(B)$, so this is indeed a mixed strategy. We can therefore treat pure strategies as degenerate mixed strategies.

In finite games, behavior strategies randomize independently at each information set. Let us examine what the equivalent in a more general setting would be. If (Ω, Σ, μ) is a probability space that incorporates all uncertainty

⁵One can construct all measurable functions from a separable metric space to another separable metric space using a transfinite recursion in ω_1 steps. A family of measurable functions such that for some countable ordinal α , all function in it are generated in at most α steps is of *bounded Borel class*. Aumann actually uses the notion of Banach class as defined in (Aumann, 1961). It is pointed out in (Rao, 1971) that a set of functions is of bounded Borel class if and only if it is of bounded Banach class.

in a game, (I, \mathcal{J}) the information sets and (C, \mathcal{C}) the choices taken there, then we have a natural function $f : \Omega \times I \rightarrow C$. If choices are chosen independently at each information set, the family $f(x, \cdot)_{x \in I}$ should be independent. Proofs of the following result⁶ can be found in ((Sun, 1998), Proposition 1.1), and in ((Podczeck, 2010), Remark 3):

Lemma 8 *Let (Ω, Σ, μ) and (I, \mathcal{J}, ν) be probability spaces. Let (C, \mathcal{C}) be a countably generated measurable space such that the atoms of \mathcal{C} are singletons, and $f : \Omega \times I \rightarrow C$ be jointly measurable such that $(f(\cdot, x))_{x \in I}$ is a family of independent random variables. Then for ν -almost all $x \in I$, the function $f(\cdot, x)$ is μ -almost surely equal to a constant.*

If nature picks a number from the uniform distribution on $[0, 1]$ and a player observes this choice perfectly, the player cannot independently randomize at each information set. But the problem is purely artificial. A play goes through only countably many information sets. So most information sets will not lie on the same play. If two information sets can never be passed by the same play, the joint distribution of behavior at these information sets is irrelevant for the distribution over plays. A very simple example will clarify this:

Example 6 *A single player has to move twice, in both cases she can choose left or right. She wants to play each of her four pure strategies with probability $1/4$. She can do this in the following way: She initially flips a fair coin and, depending on the outcome, goes left or right. Afterwards, she flips the coin again to decide again whether she goes left or right. In this case, she chooses independently at each node, since the coin flips she makes the second time are independent.*

Another possibility would be to independently flip two coins initially, but look only at the first coin flipped, which she uses to decide whether to go left or right. Then, she takes a look at the second coin and the outcome determines

⁶These papers contain actually a stronger result with a weaker independence condition. The present version is sufficient for our purposes. Also, these papers assume the product probability spaces to be completed, but the proofs do not require this assumption.

her second choice. In this case, the choices at two of the nodes will not be independent, since choices there are determined by the second coin. But the dependence is clearly of no relevance.

Relying on this insight, we will only strive for independence along plays and even that, we achieve only implicitly. We will treat a behavior strategy as a function that maps information sets to probability measures on available choices. For technical reasons, we will use a technically more convenient construction that is equivalent according to Lemma 9 in Appendix II. A behavioral strategy will be a certain kernel (see Appendix II). The kernel $s : I \times \mathcal{C} \rightarrow [0, 1]$ is a behavior strategy if $s(x, C_x) = 1$ for all $x \in I$ with $C_x = \{c \in C : c \subseteq x\}$. Note that $C_x \in \mathcal{C}$. Since x is an atom of \mathcal{I} and \mathcal{I} is countably generated, $x \in \mathcal{I} \subseteq \mathcal{C}$. Now $\bigcup C_x = x$, hence $C_x \in \mathcal{C}$. It is explained in Appendix II that we can identify measurable functions with certain kernels, and this allows us to see behavior strategies also as generalizations of pure strategies. We now list the exact definitions:

A *pure strategy* for player $i \in I$ is a sequence (s_n) of measurable functions

$$s_n : I_n^i \rightarrow C_n^i$$

such that for all $n \in \mathbb{N}$ and $x \in I_n^i$, we have $s_n(x) \subseteq x$.

A *mixed strategy* for player $i \in I$ given the probability space $(\Omega_i, \Sigma_i, \mu_i)$ is a sequence (s_n) of jointly measurable functions

$$s_n : \Omega_i \times I_n^i \rightarrow C_n^i$$

such that for all $n \in \mathbb{N}$, $\omega \in \Omega_i$ and $x \in I_n^i$, we have $s_n(\omega, x) \subseteq x$.

A *behavior strategy* for player $i \in I$ is a sequence (b_n) of kernels

$$b_n : I_n^i \times C_n^i \rightarrow [0, 1]$$

such that for all $n \in \mathbb{N}$ and $x \in I_n^i$, we have $b_n(x, C_x^n) = 1$ with

$$C_x^n = \{c \in C_n^i : c \subseteq x\} \in \mathcal{C}_n^i.$$

For the sake of convenience, we introduce a fourth kind of strategy that generalizes all previous notions of strategy:

A *generalized strategy* for player $i \in I$ given the probability space $(\Omega_i, \Sigma_i, \mu_i)$ is a sequence (s_n) of kernels

$$s_n : \Omega_i \times \prod_{m=0}^n I_m^i \times \mathcal{C}_n^i \rightarrow [0, 1]$$

such that for all $n \in \mathbb{N}$ and $(x_0, \dots, x_n) \in \prod_{m=0}^n I_m^i$, we have

$$\mu_i \left\{ \omega : s_n(\omega, x_0, \dots, x_n, C_x^n) = 1 \right\} = 1.$$

We consider generalized strategies merely as a technically useful generalization of the previous notions of strategies. Since we have perfect recall, the space $\prod_{m=0}^n I_m^i$ contains a lot of redundancy, the information is already contained in I_n^i . It is straightforward how one can interpret all kinds of strategies as special forms of generalized strategies. In particular, we only have to clarify how generalized strategies induce a distribution over plays in order to show how all kinds of strategies induce distributions over plays.

Consider a measurable extensive form with perfect recall with player set I , and specify for every player $i \in I$ a probability space $(\Omega_i, \Sigma_i, \mu_i)$ and a generalized strategy (s_n^i) given the corresponding probability space. Also, let $\Omega = \prod_i \Omega_i$ and $\Sigma = \otimes_i \Sigma_i$ and the measure $\mu = \otimes_i \mu_i$. For each $n \in \mathbb{N}$, let

$$C_n = \prod_{i \in I} C_n^i$$

and endow it with the σ -algebra $\mathcal{C}_n = \bigotimes_i \mathcal{C}_n^i$. Similarly, for each $n \in \mathbb{N}$, let

$$I_n = \prod_{i \in I} I_n^i$$

and endow it with the σ -algebra $\mathcal{J}_n = \bigotimes_i \mathcal{J}_n^i$. By Lemma 10 from Appendix II, we can define kernels

$$s_n : \Omega \times \prod_{m=0}^n I_m \times \mathcal{C}_n \rightarrow [0, 1]$$

by taking the product of all the s_n^i . Our next step is to dispose of the I_m . By M2-M4 in the definition of measurable extensive forms with perfect recall and Lemma 7, we have a natural, measurable projection

$$\iota : C_{n-1} \rightarrow \prod_{m=0}^n I_m$$

for each $n > 1$ that has a right inverse ι_r . That is, $\iota \circ \iota_r$ is the identity on $\prod_{m=0}^n I_m$. We do not claim that ι_r is measurable. We use ι to define a kernel that does not depend on the spaces (I_n^i) . Define a kernel

$$\tilde{s}_n : \Omega \times C_{n-1} \times \mathcal{C}_n \rightarrow [0, 1]$$

by

$$\tilde{s}_n(\omega, c, B) = s_n(\omega, \iota(c), B)$$

for all $\omega \in \Omega$, $c \in C_{n-1}$, $B \in \mathcal{C}_n$, and $n > 0$. For $n = 0$ we let

$$\tilde{s}_0 : \Omega \times \mathcal{C}_0 \rightarrow [0, 1]$$

be defined by

$$\tilde{s}_0(\omega, B) = s_0\left(\omega, \prod_i W, B\right)$$

for all $\omega \in \Omega$ and $B \in \mathcal{C}_0$. By applying Lemma 11 from Appendix II to μ and the sequence (\tilde{s}_n) , we obtain a unique probability measure ν on

$(\prod_n C_n, \otimes_n \mathcal{C}_n)$, that is supported on the underlying projective limit. The latter is ensured by the condition we imposed on generalized strategies. By applying Lemma 6 and Lemma 7, we obtain a unique measure on the atoms of $\bigvee_{i,n} \mathcal{C}_n^i$. By M5, these atoms are singletons, so we have actually defined a probability measure on (W, \mathcal{W}) , which we denote by τ . We call this probability measure the *outcome* of the strategy profile.

5.6 Kuhn's Theorem

In this section, we analyze the relationship between mixed and behavior strategies. Consider a game with player set I , and specify for every player $i \in I$ a probability space $(\Omega_i, \Sigma_i, \mu_i)$.

We say that player $i \in I$ has the *Kuhn property* if for every mixed strategy (s_n^i) of i , there exists a behavior strategy (b_n^i) such that for every profile (s_n^{-i}) of generalized strategies for all players but i , the outcomes of (s_n^i, s_n^{-i}) and (b_n^i, s_n^{-i}) are the same. The fact that in a finite game of perfect recall, every player has the Kuhn property, is Kuhn's theorem. The following example shows that this may fail without further assumptions in our framework:

Example 7 *There exists a set S , two countably generated σ -algebras \mathcal{S} and $\mathcal{S}' \subseteq \mathcal{S}$ that separate points, and a probability measure ν on \mathcal{S} such that there exists no kernel $k : S \times S \rightarrow [0, 1]$ with the first coordinate \mathcal{S}' measurable, such that*

$$\mu(A \cap A') = \int_{A'} k(\cdot, A) d\nu$$

for all $A \in \mathcal{S}$ and $A' \in \mathcal{S}'$ (see Example 10.4.19 in (Bogachev, 2007)). Our game will have a single player, so we drop player indices. Her randomization device is $\Omega = S \times S$ endowed with the σ -algebra $\mathcal{S}' \otimes \mathcal{S}$ and the probability measure μ specified by

$$\mu(A \times B) = \nu(A \cap B).$$

We let $W = S \times S$. Obviously, $\mathcal{I}_0 = \{W, \emptyset\}$. We let

$$\mathcal{C}_0 = \mathcal{I}_1 = \{A \times S : A \in \mathcal{S}'\}.$$

Finally, $\mathcal{C}_1 = \mathcal{C}_n = \mathcal{I}_n = \mathcal{S} \otimes \mathcal{S}$ for $n > 1$. The game effectively ends after two rounds and is of perfect information. So we have $G = \emptyset$. We let

$$s_0((s_1, s_2), W) = \{s_1\} \times S$$

for all $(s_1, s_2) \in \Omega$. Also, we let

$$s_n((s_1, s_2), x) = \{(s_1, s_2)\}$$

for all $n > 0$, $(s_1, s_2) \in \Omega$ and $x \in I_n$. The outcome τ of the strategy is supported on the diagonal of $S \times S$ and $\tau(A \times A) = \nu(A)$ for all $A \times A \in \mathcal{W}$. We show that there is no behavioral strategy that gives the same outcome. Suppose for the sake of contradiction that (b_n) would be such a behavioral strategy. Clearly, we must have

$$b_0(W, A \times S) = \nu(A)$$

for all $A \times S \in \mathcal{C}_0$. Also, b_1 must satisfy

$$\tau((A' \times S) \times (S \times A)) = \nu(A \cap A') = \int_{A' \times S} b_1(\cdot, A) d\tau_0$$

for all $A' \times S \in \mathcal{I}_0$ and $A \in \mathcal{C}_1$ with τ_0 being the marginal of τ on $\mathcal{C}_0 = \mathcal{I}_1$. But this means that we can define a kernel $k : S \times S \rightarrow [0, 1]$ by

$$k(s, B) = b_1(\{s\} \times S, S \times B)$$

that satisfies the conditions no kernel can satisfy by assumption.

To simplify the discussion for what it takes for a player to have the Kuhn property, we show that it suffices to analyze a certain class of two-player games.

Proposition 1 Consider a game with player set I , and specify for every player $i \in I$ a probability space $(\Omega_i, \Sigma_i, \mu_i)$. Let $i \in I$ be any player. There exists a two-player game with players i^* and j^* , such that i has the Kuhn property in

the original game if i^* has the Kuhn property in the new game. Moreover, j^* is perfectly informed about the past.

Proof: We construct a player j^* who represents the pool of all players but i . The underlying set of plays W will be the same. Also, we will simply let $i^* = i$. Her randomization device will still be $(\Omega_i, \Sigma_i, \mu_i)$ and she will still have the filtrations (\mathcal{I}_n^i) and (\mathcal{C}_n^i) . The filtrations of player j^* will be given by $\mathcal{I}_0^{j^*} = \{W, \emptyset\}$ and for all $n > 0$, $\mathcal{I}_n^{j^*} = \bigvee_{k \neq i, m < n} \mathcal{C}_m^k$. For all $n \in \mathbb{N}$, we let $\mathcal{C}_n^{j^*} = \bigvee_{k \neq i} \mathcal{C}_n^k$. The randomization device of j^* will be $\prod_{k \neq i} \Omega_k$ with σ -algebra $\bigotimes_{k \neq i} \Sigma_k$ and probability measure $\bigotimes_{k \neq i} \mu_k$. It is easily verified that this provides us with two-player game with players i^* and j^* . As in our construction of outcomes, we can combine generalized strategies of all players other than i in the original game to one big generalized strategy of j^* in the new game. ■

Since player j^* in the proof does not have to use the Ω_k independently and can condition on more information, player i having the Kuhn property in the new game will generally be more demanding than having it in the original game. For the rest of the section, we will study two player games played by Ann and Bob under the assumption that Bob is informed of every choice made by someone and study when Ann has the Kuhn property. We will drop the index for Ann's mixed or behavior strategies.

We let (s_n) be a mixed strategy of Ann. When looking for an equivalent behavior strategy, it is clear what b_0 has to be, since b_0 is essentially a probability measure on C_0 that has to equal the distribution of s_0 on C_0 . The hard step is finding b_1 . Finding the rest of the sequence (b_n) is then analogous. A sufficient condition is that the randomization device (Ω, Σ, μ) of Ann is perfect. This is the only place where we need a regularity assumption.

Theorem 1 *Ann has the Kuhn property if her randomization device (Ω, Σ, μ) is perfect.*

Proof: We start under the assumption that her randomization device is perfect and countably generated. It suffices to construct b_1 , constructing the

remaining (b_2, \dots) works analogously by iterating the proof. Let $\pi : I_1 \rightarrow C_0$ be the canonical projection and $\kappa_\pi : I_1 \times \mathcal{C}_0 \rightarrow [0, 1]$ its rendition as a kernel. Let $\kappa_{s_1} : \Omega \times \mathcal{C}_1 \rightarrow [0, 1]$ be the rendition of s_1 as a kernel. We define a measure ν on $C_0 \times \Omega$ by

$$\nu(B \times A) = \mu(s_0^{-1}(B) \cap A)$$

for all $B \in \mathcal{C}_0$ and $A \in \Sigma$. By slight abuse of notation, s_0 is treated as a function $s_0 : \Omega \rightarrow C_0$. By Lemma 12, there is a kernel $\kappa : C_0 \times \Sigma \rightarrow [0, 1]$ such that $\nu = \mu s_0^{-1} \otimes \kappa$. There is now a kernel $\kappa_I : I_1 \times \Sigma \rightarrow [0, 1]$ given by $\kappa_I = \kappa_\pi \otimes \kappa$. The probability measure $\kappa_I(x, \cdot)$ gives the conditional probability of measurable subsets of Ω , given that Ann has observed the piece of information x . That the piece of information x may also contain information on Bob's choices in the first period is irrelevant. We now let $b_1 : I_1 \times \mathcal{C}_1 \rightarrow [0, 1]$ be given by

$$b_1(x, B) = \kappa_I\left(x, \{\omega : s_1(\omega, x) \in B\}\right).$$

To see that b_1 is actually a kernel, note that you can first take the pointwise product of κ_I and the identity kernel on I_1 to obtain a kernel from I_1 to $\Omega \times I_1$ and then apply the kernel κ_{s_1} . The kernel thus obtained is exactly b_1 .

We now show that (b_0, b_1) gives rise to the same behavior in the periods 0 and 1 as (s_0, s_1) , no matter what Bob is doing. So let \mathcal{C}_B be a countably generated σ -algebra on W with set of atoms C_B and the σ -algebra \mathcal{C}_B on C_B . The set C_B represents the period 0 choices of Bob. We require $\mathcal{I}_1 \subseteq \mathcal{C}_b \vee \mathcal{C}_0$. Bob's behavior in period 0 is fully described by a probability measure β on (C_B, \mathcal{C}_B) . Now we can construct the joint distribution on $\Omega \times C_0 \times C_B \times C_1$ by combining μ , (s_0, s_1) , and β . But by construction, we get the same distribution from (b_0, b_1) , β and κ . So the marginals on $C_0 \times C_B \times C_1$ are the same. For b_n with $n > 1$, the same argument works, but one has to use a kernel for Bob and take his randomization device into account.

Finally, we dispose of the assumption that (Ω, Σ, μ) is countably generated. We used Lemma 12 to find for each $c \in C_0$ the conditional probabilities of all events in Σ given c . But we do not need to know all these probabilities. We only need to know the probabilities of events of the form $s_0^{-1}(A)$ for $A \in \mathcal{C}_0$ or $\{\omega : s_1(\omega, x) \in B\}$ for $B \in \mathcal{C}_1$ and $x \in I_1$. But all these events lie in a countably generated sub- σ -algebra. This is obvious for the events of the form $s_0^{-1}(A)$, we just have to take preimages of a countable set of generators for \mathcal{C}_0 . We now show that there is a countably generated σ -algebra $\Sigma' \subseteq \Sigma$ such that $\{\omega : s_1(\omega, x) \in B\} \in \Sigma'$ for all $x \in I_1$. Note that these are actually x -sections of the sets $s_1^{-1}(B)$ with $B \in \mathcal{C}_1$, so it suffices to find a countably generated $\Sigma' \subseteq \Sigma$ such that s_1 is $\Sigma' \otimes \mathcal{J}_1$ -measurable. Let $\{F_0, F_1, F_2, \dots\}$ be a countable set of generators of \mathcal{C}_1 . For each $n \in \mathbb{N}$, there is a countable family $\mathcal{G}_n \subseteq \Sigma$ such that

$$s_1^{-1}(F_n) \in \sigma\left(\{A \times B : A \in \mathcal{G}_n, B \in \mathcal{J}_1\}\right),$$

so we can just let $\Sigma' = \sigma(\bigcup_n \mathcal{G}_n)$. Since (Ω, Σ, μ) is perfect, the probability space (Ω, Σ', μ) is perfect too. ■

5.7 Discussion

It should be clear from both Example 7 and Theorem 1 that the validity of Kuhn's theorem in the general case depends on the possibility to represent conditional probabilities as kernels, so called regular conditional probabilities. It is not clear whether the assumption that Ann's randomization device is perfect can be seriously weakened. It is an open question whether there exists any probability space that is not perfect, but for which regular conditional probabilities exist for every sub- σ -algebra (see page 779 in (Ramachandran, 2002)).

Even though the frameworks are not directly comparable, as far as the conditioning aspect goes, Theorem 1 is stronger than the corresponding versions of Kuhn's theorem obtained in (Aumann, 1964) and (Mertens *et al.*,

1994). We require neither choices nor randomization devices to be standard Borel. It is not entirely clear whether the assumption that choices are standard Borel is essential to the proof in (Mertens *et al.*, 1994). The formal statement there (Theorem II.1.4) certainly requires them, but the authors seem to prove a slightly stronger theorem in their proof and use the assumption only there. They want to construct equivalent behavior strategies for a class of strategies even larger than our class of generalized strategies. They allow a different randomization device to be used in every period. The standard Borel assumption on choices allows them to replace these randomization devices by a single randomization device.

A strong point of our formulation of extensive forms is that our formulation of perfect recall does not require extraneous objects such as the measurable recall functions used by Aumann and Mertens, Sorin, Zamir. The projections serve as recall functions and their measurability is automatic in our framework.

The biggest problem in the current formulation is of course that extensive forms without perfect recall or without a common notion of time for all player cannot even be formulated. For now, we provide a sketch of how one can extend the framework. Essentially, we can “glue” information sets together. We glue the information sets $x \in I_n^i$ and $y \in I_m^i$ together by requiring player i to behave the same way at both information sets. For this, we need a notion of having the same choices available at x and y . Formally, we take a function $g : x \rightarrow y$ that is $x \cap \mathcal{C}_n^i / y \cap \mathcal{C}_m^i$ -bimeasurable and has the property that for every \mathcal{C}_n^i -atom $c \subseteq x$, there is a \mathcal{C}_m^i -atom $c' \subseteq y$ such that $g(c) \subseteq c'$. By Lemma 3, the function $\pi(g)$ gives us an isomorphism between the choices available at x and the choices available at y . This way of gluing together information sets allows us to model much more general extensive forms. The open question is how to characterize perfect recall in terms of such functions g . This question is however independent of the measurability questions that have guided us thus far.

5.8 Appendix I: General Preliminaries

The set of natural numbers $\{0, 1, 2, \dots\}$ is denoted by \mathbb{N} . A set is *countable* if it can be mapped injectively into \mathbb{N} . The power set of a set S is denoted by 2^S . We will often use index-free notation for set-theoretic operations. If S is a set and $\mathcal{F} \subseteq 2^S$, then

$$\bigcup \mathcal{F} = \{x \in S : x \in F \text{ for at least one } F \in \mathcal{F}\}$$

and

$$\bigcap \mathcal{F} = \{x \in S : x \in F \text{ for all } F \in \mathcal{F}\}.$$

Note that $\bigcup \emptyset = \emptyset$. The family \mathcal{F} is *closed under countable unions* if $\bigcup \mathcal{C} \in \mathcal{F}$ whenever \mathcal{C} is a countable subset of \mathcal{F} . The family \mathcal{F} is *closed under complements* if $S - F = \{x \in S : x \notin F\} \in \mathcal{F}$ whenever $F \in \mathcal{F}$.

A σ -algebra \mathcal{S} on a set S is a family of subsets of S closed under complements and countable unions. There always is a largest such σ -algebra, the power-set 2^S . We will usually endow finite sets with this σ -algebra. The elements of a σ -algebra \mathcal{S} are *measurable sets* or, if it needs clarification, *\mathcal{S} -measurable sets*. In a probabilistic context, we also call them *events*. If \mathcal{S}_1 and \mathcal{S}_2 are both σ -algebras on S then we say that \mathcal{S}_1 is *finer than* \mathcal{S}_2 or \mathcal{S}_2 *coarser than* \mathcal{S}_1 if $\mathcal{S}_2 \subseteq \mathcal{S}_1$. If \mathcal{F} is any family of subsets of S , there is a coarsest σ -algebra on S that contains all elements of \mathcal{F} . We call it the σ -algebra *generated* by \mathcal{F} and denote it by $\sigma(\mathcal{F})$. A useful and easily proven fact is that whenever $A \in \sigma(\mathcal{F})$, there is countable set $\mathcal{C} \subseteq \mathcal{F}$ such that $A \in \sigma(\mathcal{C})$. The σ -algebra \mathcal{S} is *countably generated* if there is a countable family \mathcal{C} such that $\mathcal{S} = \sigma(\mathcal{C})$. If $(\mathcal{S}_i)_{i \in I}$ is a family of σ -algebras on S , we let $\bigvee_{i \in I} \mathcal{S}_i = \sigma(\bigcup_{i \in I} \mathcal{S}_i)$.

A *measurable space* is a pair (S, \mathcal{S}) with \mathcal{S} being a σ -algebra on S . If it is clear from the context which σ -algebra is employed, we might dare to speak of the measurable space S . We will always endow the unit interval $[0, 1]$ with the σ -algebra generated by its closed subsets, the *Borel σ -algebra* \mathcal{B} . It is countably generated, since \mathcal{B} is also generated by the closed subintervals with rational endpoints. If (S, \mathcal{S}) is a measurable space and $A \subseteq S$, the *trace*

σ -algebra $\mathcal{A} \cap \mathcal{S}$ on A is given by $\mathcal{A} \cap \mathcal{S} = \{A \cap B : B \in \mathcal{S}\}$. It is countably generated if \mathcal{S} is.

A family \mathcal{C} of subsets of set S is a π -system if $A \cap B \in \mathcal{C}$ for all $A, B \in \mathcal{C}$. A family \mathcal{D} of subsets of set S is a \mathcal{D} -system if $S \in \mathcal{D}$, $A - B \in \mathcal{D}$ whenever $A, B \in \mathcal{D}$ and $B \subseteq A$, and $\bigcup_n B_n \in \mathcal{D}$ whenever (D_n) is a sequence in \mathcal{D} such that $D_n \subseteq D_{n+1}$ for all $n \in \mathbb{N}$. The *monotone class theorem* says that if \mathcal{C} is a π -system and \mathcal{D} a \mathcal{D} -system on a set S such that $\mathcal{C} \subseteq \mathcal{D}$, then $\sigma(\mathcal{C}) \subseteq \mathcal{D}$. For a proof, see Theorem 1.1 in (Kallenberg, 2002).

A function $f : S \rightarrow T$ with (S, \mathcal{S}) and (T, \mathcal{T}) being measurable spaces is \mathcal{S}/\mathcal{T} -measurable or just *measurable* if $f^{-1}(B) \in \mathcal{S}$ for all $B \in \mathcal{T}$. A sufficient condition for f being measurable is that $f^{-1}(B) \in \mathcal{S}$ for all $B \in \mathcal{F}$, with \mathcal{F} being a family such that $\sigma(\mathcal{F}) = \mathcal{T}$. If S is a set, (T, \mathcal{T}) a measurable space, and $f : S \rightarrow T$ a function, there is a coarsest σ -algebra on S that makes f measurable. We denote it by $\sigma(f)$. It is given by

$$\{f^{-1}(B) : B \in \mathcal{T}\}.$$

Every countably generated σ -algebra is of the form $\sigma(f)$ for some f taking values in $[0, 1]$ (Theorem 6.5.5. in (Bogachev, 2007)). If T is a set, (S, \mathcal{S}) a measurable space and $f : S \rightarrow T$ a function, there is a finest σ -algebra on T that makes f measurable. It is given by

$$\{B \subseteq T : f^{-1}(B) \in \mathcal{S}\}.$$

If $(S_i, \mathcal{S}_i)_{i \in I}$ is a family of measurable spaces, we let $\bigotimes_{i \in I} \mathcal{S}_i$ be the *product σ -algebra* on $\prod_{i \in I} S_i$, the coarsest σ -algebra that makes all coordinate projections measurable. Alternatively, it is generated by the *measurable rectangles*. A measurable rectangle is the cartesian product of a family $(A_i) \in \prod_i \mathcal{S}_i$ such that $A_i = S_i$ for all but finitely many indices i .

Let (S, \mathcal{S}) and (T, \mathcal{T}) be measurable spaces and $A \in \mathcal{S} \otimes \mathcal{T}$, $s \in S$, and $t \in T$. We call the set

$$\{(s, y) \in A : \text{for some } y \in T\}$$

the *s*-section of A and define the *t*-section of A similarly. The *s*-section of A is in \mathcal{T} and the *t*-section of A is in \mathcal{S} . To see this, note that the set of subsets of $S \times T$ such that the *s*-section is in \mathcal{S} and the *t*-section is in \mathcal{S} forms a σ -algebra that contains all measurable rectangles.

We always endow products of measurable spaces with their product σ -algebra. We sometimes call a measurable function defined on a product *jointly measurable* when we want to emphasize it is measurable with respect to the product σ -algebra. If (S, \mathcal{T}) and (T, \mathcal{S}) are measurable spaces and $f : S \rightarrow T$ is measurable, then the *graph* $\Gamma(f)$ is a measurable subset in $\mathcal{S} \otimes \mathcal{T}$ if and only if there is a countably generated σ -algebra $\mathcal{T}' \subseteq \mathcal{T}$ such that $\{f(s)\} \in \mathcal{T}'$ for all $s \in S$ and such that f is \mathcal{S}/\mathcal{T}' -measurable (Proposition 2.1 in (Musiał, 1980)).

A (measurable) *isomorphism* between two measurable spaces is a measurable bijection with a measurable inverse. Two measurable spaces connected by an isomorphism are identical up to a relabeling of points. We call them *isomorphic*. A measurable function such that the forward image of every measurable set is measurable, is *bimeasurable*. Note that an isomorphism is exactly a bijective, bimeasurable function. A measurable space is a *standard Borel space* if it is isomorphic to $[0, 1]$ or a countable subset thereof with the trace σ -algebra. Standard Borel spaces occur frequently in nature according to the following result of Kazimierz Kuratowski: Any complete and separable metric space endowed with the σ -algebra generated by the open sets is standard Borel. For a proof, see Theorem 13.1.1 in (Dudley, 2002).

We will assume that the reader is familiar with probability measures and the rudiments of Lebesgue integration.

A *partition* Π on a set S is a family of disjoint nonempty sets such that $\bigcup \Pi = S$. The elements of a partition are *cells*. The partition comes with a natural *projections* $\pi : S \rightarrow \Pi$ that maps each $x \in S$ to the unique element $P \in \Pi$ such that $x \in P$. If Π_1 and Π_2 are both partitions on S , we say that Π_1 is *finer* than Π_2 or that Π_2 is *coarser* than Π_1 if every cell in Π_2 is the union of cells in Π_1 .

5.9 Appendix II: Kernels and Conditional Probability

Kernels are ubiquitous in probability theory. They are variously known as Markov kernels, transition probabilities, conditional probabilities, random mappings, random measures, Young measures, and probably a lot of other names.

Let (S, \mathcal{S}) and (T, \mathcal{T}) be measurable spaces. A *kernel* from S to T is a function $\kappa : S \times \mathcal{T} \rightarrow [0, 1]$ such that the following conditions hold:

- (i) $\kappa(\cdot, B)$ is measurable for all $B \in \mathcal{T}$.
- (ii) $\kappa(s, \cdot)$ is a probability measure for all $s \in S$.

We interpret $\kappa(s, B)$ as the probability of going from s to an element of B . We can also interpret kernels als functions with probability measures as values. To make this precise, let $\Delta(T)$ be the set of probability measures on (T, \mathcal{T}) and endow $\Delta(T)$ with the σ -algebra generated by events of the form

$$\{\mu \in \Delta(T) : \mu(B) \leq r\}$$

for some $B \in \mathcal{T}$ and $r \in [0, 1]$.

Lemma 9 *Let (S, \mathcal{S}) and (T, \mathcal{T}) be measurable spaces. Let $p : S \rightarrow \Delta(T)$ be a measurable function. Then $\kappa : S \times \mathcal{T} \rightarrow [0, 1]$ given by $\kappa(s, B) = p(s)(B)$ is a kernel. Moreover, every kernel can be obtained this way.*

Proof: See Lemma 1.40 in (Kallenberg, 2002). ■

It should be noted that kernels generalize ordinary measurable functions. If $f : S \rightarrow T$ is measurable, we can identify f with the kernel $\kappa_f : S \times \mathcal{T} \rightarrow [0, 1]$ given by

$$\kappa_f(s, B) = \begin{cases} 1 & \text{if } f(s) \in B, \\ 0 & \text{if } f(s) \notin B. \end{cases}$$

To see that this is actually a kernel, note that $\kappa_f(\cdot, B)^{-1}(\{1\}) = f^{-1}(B)$. This identification is injective if and only if all atoms in \mathcal{T} that contain points in

$f(S)$ are singletons. Similarly, we can view probability measures as kernels that are constant in the first argument.

The arbitrary product of kernels is again a kernel:

Lemma 10 *Let $((S_i, \mathcal{S}_i), (T_i, \mathcal{T}_i))_{i \in I}$ be an arbitrary family of pairs of measurable spaces and let $\kappa_i : S_i \times \mathcal{T}_i \rightarrow [0, 1]$ be a kernel for every $i \in I$. Then the function*

$$\kappa : \prod_i S_i \times \bigotimes_i \mathcal{T}_i \rightarrow [0, 1]$$

given by

$$\kappa(s, B) = \bigotimes_i \kappa_i(s, \cdot)(B)$$

is a kernel.

Proof: We only have to verify that $\kappa(\cdot, B)$ is $\bigotimes_i \mathcal{S}_i$ -measurable for every $B \in \bigotimes_i \mathcal{T}_i$, which we do by employing the monotone-class theorem. Let $\mathcal{D} \subseteq \bigotimes_i \mathcal{T}_i$ be the family of all measurable sets B such that $\kappa(\cdot, B)$ is $\bigotimes_i \mathcal{S}_i$ -measurable.

We first show that every measurable rectangle is in \mathcal{D} . Let $F \subseteq I$ be a set with n elements and

$$B = B_1 \times \dots \times B_n \times \prod_{i \in I \setminus F} T_i$$

be a measurable rectangle. Then

$$\kappa(s, B) = \kappa_1(s_1, B_1) \times \dots \times \kappa_n(s_n, B_n)$$

is a product of measurable functions and therefore measurable.

Now let $A, B \in \mathcal{D}$ with $B \subseteq A$. Then

$$\kappa(\cdot, A \setminus B) = \kappa(\cdot, A) - \kappa(\cdot, B),$$

a measurable function as the difference of two measurable functions.

Finally, Let (B_n) be an increasing sequence in \mathcal{D} with limit $B = \bigcup_n B_n$ and let $r \in [0, 1]$. Then

$$\kappa(\cdot, B)^{-1}([0, r]) = \bigcap_n \kappa(\cdot, B_n)^{-1}([0, r]),$$

by the continuity of measures. Since measurable rectangles form a π -system, we conclude by the monotone-class theorem that $\mathcal{D} = \bigotimes_i \mathcal{J}_i$. ■

It is possible to chain kernels. Let (S, \mathcal{S}) , (T, \mathcal{T}) and (U, \mathcal{U}) be measurable spaces and $\kappa_1 : S \times \mathcal{T} \rightarrow [0, 1]$ and $\kappa_2 : S \times T \times \mathcal{U} \rightarrow [0, 1]$ be kernels. Then one can define a kernel $\kappa_1 \otimes \kappa_2 : S \times \mathcal{T} \otimes \mathcal{U} \rightarrow [0, 1]$ by letting

$$(\kappa_1 \otimes \kappa_2)(s, B) = \iint \mathbf{1}_B(t, u) \, d\kappa_2(s, t, \cdot) \, d\kappa_1(s, \cdot)$$

for all $B \in \mathcal{T} \otimes \mathcal{U}$. For a proof that $\kappa_1 \otimes \kappa_2$ is actually a kernel, see Lemma 1.41(iii) in (Kallenberg, 2002). The kernel $\kappa_1 \otimes \kappa_2$ corresponds to the intuitive notion that given s , an element t of T is randomly chosen according to $\kappa_1(s, \cdot)$ and then an element $u \in U$ is randomly chosen according to $\kappa_2(s, t, \cdot)$.

It is worth pointing out that there is a category with measurable spaces as objects and kernels as morphisms. Composition is done by chaining kernels, and the identities are the kernels corresponding to the identity function. This category was termed the category of probabilistic mappings in (Lawvere, 1962) and termed the category of statistical decisions in (Čencov, 1982).

The following, incredibly useful, result is known as the Ionescu-Tulcea theorem. For a proof, see Theorem 6.17 in (Kallenberg, 2002).

Lemma 11 *Let (S_n, \mathcal{S}_n) be a sequence of measurable spaces and let (κ_n) be a sequence of kernels such that*

$$\kappa_n : S_0 \times \dots \times S_n \times \mathcal{S}_{n+1} \rightarrow [0, 1].$$

Then there exists a unique kernel

$$\kappa : S_0 \times \bigotimes_{n=1}^{\infty} S_n \rightarrow [0, 1]$$

such that

$$\kappa(s, B_1 \times \dots \times B_n \times S_{n+1} \times \dots) = \kappa_0 \otimes \dots \otimes \kappa_{n-1}(s, B_1 \times \dots \times B_n)$$

for all $s \in S_0$ and $B_1 \times \dots \times B_n \in S_1 \times \dots \times S_n$.

For each $s \in S_0$, the Ionescu-Tulcea theorem provides us with a stochastic process on $\prod_{n=1}^{\infty} S_n$. We use the result to show that randomized strategies induce distributions over outcomes. We can also use the Ionescu-Tulcea theorem to provide a different proof of Lemma 10 for countably many kernels.

Given two measurable spaces (S, \mathcal{S}) and (T, \mathcal{T}) , it is possible to construct a unique probability measure $\mu \otimes \kappa$ on $S \times T$ from a probability measure μ on (S, \mathcal{S}) and a kernel $\kappa : S \times \mathcal{T} \rightarrow [0, 1]$ such that

$$\mu \otimes \kappa(A \times B) = \int_A \kappa(s, B) d\mu$$

for all $A \times B \in \mathcal{S} \times \mathcal{T}$. Clearly, the marginal of $\mu \otimes \kappa$ is μ . That is, $\mu \otimes \kappa(A \times T) = \mu(A)$ for all $A \in \mathcal{S}$. In that case, we can view κ as giving us for every s , the conditional probability of every measurable subset of T . Not every probability measure on a product can be decomposed into the marginal and a kernel.

A probability space (Ω, Σ, μ) is *perfect* if for each random variable $f : \Omega \rightarrow \mathbb{R}$, there exists a Borel set $B \subseteq f(\Omega)$ such that $\mu f^{-1}B = 1$. Clearly, if (Ω, Σ, μ) is perfect and Σ' is a sub- σ -algebra of Σ , the space (Ω, Σ', μ) is perfect too. The class of perfect measure spaces has been introduced by Kolmogorov and Gnedenko in (Gnedenko & Kolmogorov, 1968) in order “to achive complete harmony between the abstract theory of measure and the theory of measure in metric spaces.”

We have the following useful result due to Arnold Faden in (Faden, 1985):

Lemma 12 *Let (S, \mathcal{S}, μ) be a countably generated probability space. Then (S, \mathcal{S}, μ) is perfect if and only if for every measurable space (T, \mathcal{T}) and every probability measure ν on $(T \times S, \mathcal{T} \otimes \mathcal{S})$ with marginal μ on T , there is a kernel $\kappa : T \times S \rightarrow [0, 1]$ such that $\nu = \mu \otimes \kappa$.*

Lemma 12 will be our tool of choice for constructing conditional probabilities. The kernel κ will in general not be unique, but two such kernels can only differ on μ -null set.

References

- ALÓS-FERRER, CARLOS, & RITZBERGER, KLAUS. 2005. Trees and decisions. *Econom. Theory*, **25**(4), 763–798.
- ALÓS-FERRER, CARLOS, & RITZBERGER, KLAUS. 2008. Trees and extensive forms. *Journal of Economic Theory*, **143**(1), 216 – 250.
- ALÓS-FERRER, CARLOS, KERN, JOHANNES, & RITZBERGER, KLAUS. 2011. Comment on "Trees and extensive forms" [J. Econ. Theory 143 (1) (2008) 216-250]. *Journal of Economic Theory*, **146**(5), 2165 – 2168.
- AMIR, RABAH. 2006. Dynamic Games in Economics. *Pages 443–477 of: DANA, ROSE-ANNE, LE VAN, CUONG, MITRA, TAPAN, & NISHIMURA, KAZUO (eds), Handbook on Optimal Growth 1.* Springer Berlin Heidelberg.
- AUMANN, ROBERT J. 1961. Borel structures for function spaces. *Illinois J. Math.*, **5**, 614–630.
- AUMANN, ROBERT J. 1964. Mixed and behavior strategies in infinite extensive games. *Pages 627–650 of: Advances in Game Theory.* Princeton, N.J.: Princeton Univ. Press.
- AUMANN, ROBERT J., & DREZE, JACQUES H. 2008. Rational Expectations in Games. *American Economic Review*, **98**(1).
- AUMANN, ROBERT J., & HART, SERGIU. 2003. Long Cheap Talk. *Econometrica*, **71**(6), 1619–1660.

- BADE, SOPHIE, HAERINGER, GUILLAUME, & RENOU, LUDOVIC. 2009. Bilateral commitment. *Journal of Economic Theory*, **144**(4), 1817 – 1831.
- BHASKAR, V. 1998. Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems. *The Review of Economic Studies*, **65**(1), pp. 135–149.
- BILLINGSLEY, PATRICK. 1995. *Probability and measure*. Third edn. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons Inc. A Wiley-Interscience Publication.
- BOGACHEV, VLADIMIR I. 2007. *Measure theory. Vol. I, II*. Berlin: Springer-Verlag.
- BROUGHTON, ALLEN, & HUFF, BARTHEL W. 1977. A comment on unions of sigma-fields. *Amer. Math. Monthly*, **84**(7), 553–554.
- CARUANA, G., & EINAIV, L. 2008. A theory of endogenous commitment. *Review of Economic Studies*, **75**(1), 99–116.
- ČENCOV, N. N. 1982. *Statistical decision rules and optimal inference*. Translations of Mathematical Monographs, vol. 53. Providence, R.I.: American Mathematical Society. Translation from the Russian edited by Lev J. Leifman.
- DUBRA, JUAN, & ECHENIQUE, FEDERICO. 2004. Information is not about measurability. *Math. Social Sci.*, **47**(2), 177–185.
- DUDLEY, R. M. 2002. *Real analysis and probability*. Cambridge Studies in Advanced Mathematics, vol. 74. Cambridge: Cambridge University Press. Revised reprint of the 1989 original.
- FADEN, ARNOLD M. 1985. The existence of regular conditional probabilities: necessary and sufficient conditions. *Ann. Probab.*, **13**(1), 288–298.
- FORGES, FRANÇOISE. 1990. Universal Mechanisms. *Econometrica*, **58**(6), pp. 1341–1364.

- FUDENBERG, DREW, & LEVINE, DAVID K. 1989. Reputation and Equilibrium Selection in Games with a Patient Player. *Econometrica*, **57**(4), pp. 759–778.
- GILES, L. 2010. *Sun Tzu on the Art of War*. CreateSpace.
- GNEDENKO, B. V., & KOLMOGOROV, A. N. 1968. *Limit distributions for sums of independent random variables*. Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu. Revised edition. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills., Ont.
- GRÄTZER, GEORGE. 2003. *General lattice theory*. Basel: Birkhäuser Verlag. With appendices by B. A. Davey, R. Freese, B. Ganter, M. Greferath, P. Jipsen, H. A. Priestley, H. Rose, E. T. Schmidt, S. E. Schmidt, F. Wehrung and R. Wille, Reprint of the 1998 second edition [MR1670580].
- HACKING, IAN. 2006. *The emergence of probability*. Second edn. Cambridge: Cambridge University Press. A philosophical study of early ideas about probability, induction and statistical inference.
- HAMILTON, JONATHAN H., & SLUTSKY, STEVEN M. 1993. Endogenizing the order of moves in matrix games. *Theory and Decision*, **34**(1), 47–62.
- HENDON, EBBE, JACOBSEN, HANS JØRGEN, & SLOTH, BIRGITTE. 1996. The one-shot-deviation principle for sequential rationality. *Games Econom. Behav.*, **12**(2), 274–282.
- ISELL, J. R. 1957. Finitary games. Pages 79–96 of: *Contributions to the theory of games, vol. 3*. Annals of Mathematics Studies, no. 39. Princeton, N. J.: Princeton University Press.
- KALAI, ADAM TAUMAN, KALAI, EHUD, LEHRER, EHUD, & SAMET, DOV. 2010. A commitment folk theorem. *Games Econom. Behav.*, **69**(1), 127–137.
- KALLENBERG, OLAV. 2002. *Foundations of modern probability*. Second edn. Probability and its Applications (New York). New York: Springer-Verlag.

- KUHN, H. W. 1950. Extensive games. *Proc. Nat. Acad. Sci. U.S.A.*, **36**, 570–576.
- KUHN, H. W. 1953. Extensive games and the problem of information. *Pages 193–216 of: Contributions to the theory of games, vol. 2.* Annals of Mathematics Studies, no. 28. Princeton, N. J.: Princeton University Press.
- LAWVERE, FRANCIS WILLIAM. 1962. The category of probabilistic mappings. *preprint.*
- LEVHARI, DAVID, & MIRMAN, LEONARD J. 1980. The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution. *The Bell Journal of Economics*, **11**(1), pp. 322–334.
- MAILATH, GEORGE J., & SAMUELSON, LARRY. 2006. *Repeated Games and Reputations: Long-Run Relationships.* Oxford University Press, USA.
- MAITRA, A., & SUDDERTH, W. 2000. Randomized strategies and terminal distributions. *Pages 39–52 of: Game theory, optimal stopping, probability and statistics.* IMS Lecture Notes Monogr. Ser., vol. 35. Beachwood, OH: Inst. Math. Statist.
- MASKIN, ERIC, & TIROLE, JEAN. 2001. Markov perfect equilibrium. I. Observable actions. *J. Econom. Theory*, **100**(2), 191–219.
- MERTENS, J.F., SORIN, S., & ZAMIR, S. 1994. *Repeated games.* Univ. Catholique de Louvain, Center for Operations Research & Econometrics.
- MUSIAŁ, KAZIMIERZ. 1980. Projective limits of perfect measure spaces. *Fund. Math.*, **110**(3), 163–189.
- NASH, JOHN. 1951. Non-cooperative games. *Ann. of Math. (2)*, **54**, 286–295.
- ORE, OYSTEIN. 1942. Theory of equivalence relations. *Duke Math. J.*, **9**, 573–627.

- OSBORNE, MARTIN J., & RUBINSTEIN, ARIEL. 1994. *A course in game theory*. Cambridge, MA: MIT Press.
- PETERS, MICHAEL, & SZENTES, BALÁZS. 2012. Definable and Contractible Contracts. *Econometrica*, **80**(1), 363–411.
- PODCZECK, KONRAD. 2010. On existence of rich Fubini extensions. *Econom. Theory*, **45**(1-2), 1–22.
- RAMACHANDRAN, DORAISWAMY. 2002. Perfect measures and related topics. *Pages 765–786 of: Handbook of measure theory, Vol. I, II*. Amsterdam: North-Holland.
- RAO, B. V. 1971. Borel structures for function spaces. *Colloq. Math.*, **23**, 33–38.
- RENOU, LUDOVIC. 2009. Commitment games. *Games Econom. Behav.*, **66**(1), 488–505.
- RITZBERGER, KLAUS. 1999. Recall in extensive form games. *Internat. J. Game Theory*, **28**(1), 69–87.
- ROSENTHAL, ROBERT W. 1981. Games of perfect information, predatory pricing and the chain-store paradox. *J. Econom. Theory*, **25**(1), 92–100.
- SAVAGE, LEONARD J. 1972. *The foundations of statistics*. revised edn. New York: Dover Publications Inc.
- SCHELLING, THOMAS C. 1960. *The Strategy of Conflict*. Cambridge, Mass.: Harvard Uni.
- SCHWARZ, GIDEON. 1974. Ways of randomizing and the problem of their equivalence. *Israel J. Math.*, **17**, 1–10.
- SELTEN, REINHARD. 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit: Teil I: Bestimmung des dynamischen Preisgleichgewichts. *Zeitschrift für die gesamte Staatswissenschaft*, **121**(2), 301–324.

- SELTEN, REINHARD, & GÜTH, WERNER. 1982. Game theoretical analysis of wage bargaining in a simple business cycle model. *Journal of Mathematical Economics*, **10**(2-3), 177 – 195.
- SUN, YENENG. 1998. The almost equivalence of pairwise and mutual independence and the duality with exchangeability. *Probab. Theory Related Fields*, **112**(3), 425–456.
- TENNENHOLTZ, MOSHE. 2004. Program equilibrium. *Games Econom. Behav.*, **49**(2), 363–373.
- VAN DAMME, ERIC, & HURKENS, SJAAK. 1996. Commitment robust equilibria and endogenous timing. *Games Econom. Behav.*, **15**(2), 290–311.
- VAN LONG, NGO. 2010. *A survey of dynamic games in economics*. Vol. 1. World Scientific Pub Co Inc.
- VON NEUMANN, JOHN, & MORGENSTERN, OSKAR. 1947. *Theory of Games and Economic Behavior*. Princeton, N. J.: Princeton University Press. 2d ed.
- WALD, ABRAHAM. 1945. Generalization of a theorem by v. Neumann concerning zero sum two person games. *Ann. of Math. (2)*, **46**, 281–286.
- WALD, ABRAHAM, & WOLFOWITZ, JACOB. 1951. Two methods of randomization in statistics and the theory of games. *Ann. of Math. (2)*, **53**, 581–586.

Abstract

This dissertation consists of three papers dealing with the analysis of games in which time is essential.

The first paper, "Payoff-Relevance" deals with the foundations behind the solution concept of *Markov perfect equilibrium* in games with observable actions. This solution concept is based on a notion of payoff-relevant states. Maskin and Tirole have shown that one can derive the states from the extensive form under certain finiteness assumptions. This paper shows that these assumptions can be disposed of. Moreover, it is shown that the approach of Maskin and Tirole is not invariant with respect to payoff-irrelevant relabellings of actions. For a large class of games, a fix is provided in the form of a canonical labeling that respects certain symmetries within the game.

In the second paper, "Commitment in Extensive Form Games", I analyze what can happen if players in an extensive form game are allowed to make binding public commitments at each stage. It is shown that every outcome induced by a pure-strategy Nash equilibrium is compatible with these commitment possibilities. A series of examples illustrates the approach.

The third paper, "Strategic Stochastic Processes" is joint work with Klaus Ritzberger. We formulate and prove a version of Kuhn's theorem in a general measure theoretic framework: In a game of perfect recall, there exists for each mixed strategy of a player a behavior strategy such that the induced distribution over outcomes is the same, no matter how other players behave. In the process, we introduce a clean way of modeling partitional information in a measure theoretic framework.

Zusammenfassung

Diese Dissertation besteht aus drei Arbeiten zur Analyse von Spielen in denen Zeit eine essentielle Rolle spielt.

Die erste Arbeit, "Payoff-Relevance", analysiert wie sich der Zustandsraum der dem Gleichgewichtskonzept von Markovgleichgewichten zugrundeliegt von der Extensivform ableiten lässt. In einer früheren Arbeit haben Maskin und Tirole gezeigt dass sich payoffrelevante Zustände unter gewissen Endlichkeitsannahmen aus der Extensivform ableiten lassen. Hier wird gezeigt dass diese Annahmen überflüssig sind. Weiters wird gezeigt dass die Methode von Maskin und Tirole nicht invariant unter dem Umbenennen von Aktionen ist. Dieses Problem wird für eine weite Klasse von Spielen gelöst.

In der zweiten Arbeit, "Commitment in Extensive Form Games", analysiere ich was passiert wenn Spieler in einem Extensivformspiel die Möglichkeit haben in jeder Runde bindende Verpflichtungen einzugehen. Jedes Nashgleichgewicht in reinen Strategien kann auf diese Weise gespielt werden. Eine Reihe von Beispielen illustriert meinen Zugang.

Die dritte Arbeit wurde zusammen mit Klaus Ritzberger verfasst. Wir formulieren und beweisen eine Variante des Satzes von Kuhn in einem allgemeinen maßtheoretischen Rahmen: In einem Spiel, in dem sich jede Spielerin daran erinnern kann was sie getan und gewusst hat, gibt es für jede gemischte Strategie eine Verhaltensstrategie, die für jedes Verhalten der anderen Spieler die gleiche Verteilung über die Partien induziert.

Curriculum Vitae

Personal

Birth: April 26, 1982, Hallein, Austria

Nationality: Austrian

Education

Currently, Ph.D. in Economics, University of Vienna,
Since 2010 as part of the Vienna Graduate School in Economics.

M.A. in Economics (Magister) 2009, University of Vienna, with honors.

Undergraduate thesis: *Generic core-emptiness in spatial models of voting*

Advisor: K. Podczeck

Teaching Experience

TA for *Advanced Microeconomics* by K. Podczeck,
summer terms 2008 to 2012

TA for *Mathematical Analysis* by M. Meier,
fall terms 2009-2010 and 2010-2011

TA for *Advanced Microeconomics* by E. Dierker,
summer term 2010

TA for *Microeconomics* by M. Nermuth,
fall terms 2009-2010 and 2010-2011