GAME THEORY: Analysis of Strategic Thinking

Pierpaolo Battigalli, Emiliano Catonini, Nicodemo De Vito

October 2023

Abstract

This textbook introduces some concepts of the theory of games: rationality, dominance, rationalizability, and several notions of equilibrium (Nash, randomized, correlated, self-confirming, subgame perfect, Bayesian perfect equilibrium). For each of these concepts, the interpretative aspect is emphasized. Even though no advanced mathematical knowledge is required, the reader should nonetheless be familiar with the concepts of set, function, probability, and more generally be able to follow abstract reasoning and formal arguments.

Milano,
P. Battigalli, E. Catonini, and N. De Vito

Preface

This textbook provides an introduction to game theory, the formal analysis of strategic interaction. Game theory now pervades most non-elementary models in microeconomic theory and many models in the other branches of economics and in other social sciences. We introduce the necessary analytical tools to be able to understand these models, and illustrate them with some economic applications.

We also aim at developing an abstract analysis of strategic thinking, and a critical and open-minded attitude toward the standard game-theoretic concepts as well as new concepts.

Most of this textbook rely on relatively elementary mathematics. Yet, our approach is formal and rigorous. The reader should be familiar with mathematical notation about sets and functions, with elementary linear algebra and topology in Euclidean spaces, and with proofs by mathematical induction. Elementary calculus is sometimes used in examples.

Contents

1

Introduction

Game theory is the formal analysis of the behavior of interacting individuals. The crucial feature of an interactive situation is that the consequences of the actions of an individual depend (also) on the actions of other individuals. This is typical of many games people play for fun, such as chess or poker. Hence, interactive situations are called "games" and interactive individuals are called "players." If a player's behavior is intentional and he is aware of the interaction (which is not always the case), he should try and anticipate the behavior of other players. This is the essence of strategic thinking. In the rest of this chapter we provide a semi-formal introduction of some key concepts in game theory.

1.1 Decision Theory and Game Theory

Decision theory is a branch of applied mathematics that analyzes the decision problem of an individual (or a group of individuals acting as a single decision unit) in isolation. The external environment is a primitive of the decision problem. Decision theory provides simple decision criteria characterizing an individual's preferences over different courses of action, provided that these preferences satisfy some "rationality" properties, such as completeness (any two alternatives are comparable) and transitivity (if a is preferred to b and b is preferred to c, then a is preferred to c). These criteria are used to find optimal (or rational) decisions.

Game theory could be more appropriately called **interactive decision theory**. Indeed, game theory is a branch of applied mathematics that analyzes interactive decision problems: There are several individuals, called **players**, each facing a decision problem whereby the "external environment" (from the point of view of this particular player) is given by the other players' behavior (and possibly some random variables). In other words, the welfare (utility, payoff, final wealth) of each player is affected not only by his own behavior, but also by the behavior of other players. Therefore, in order to figure out the best course of action each player has to guess which course of action the other players are going to take.

1.2 Why Economists Should Use Game Theory

We are going to argue that game theory should be the main analytical tool used to build formal economic models. More generally, game theory should be used in all formal models in the social sciences that adhere to methodological individualism, i.e., try to explain social phenomena as the result of the actions of many agents, which in turn are freely chosen according to some consistent criterion.

The reader familiar with the pervasiveness of game theory in economics may wonder why we want to stress this point. Isn't it well known that game theory is used in countless applications to model imperfect competition, bargaining, contracting, political competition, and, in general, all social interactions where the action of each individual has a non-negligible effect on the social outcome? Yes, indeed! And yet it is often explicitly or implicitly suggested that game theory is not needed to model situations where each individual is negligible, such as perfectly competitive markets. We are going to explain why this is —in our view— incorrect: the bottom line will be that *every "complete" formal model of an economic (or social) interaction must be a game;* economic theory has analyzed perfect competition by taking shortcuts that have been very fruitful, but must be seen as such, just shortcuts.[1]

If we subscribe to methodological individualism, as mainstream economists claim to do, every social or economic observable phenomena we are interested in analyzing should be reduced to the actions of the

[1] Our view is very similar to the original motivations for the study of game theory offered by the "founding fathers" von Neumann and Morgenstern in their seminal book *The Theory of Games and Economic Behavior* [68].

individuals who form the social or economic system. For example, if we want to study prices and allocations, then we should specify which actions the individuals in the system can choose and how prices and allocations depend on such actions: if I is the set of agents, p is the vector of prices, y is the allocation, and $a = (a_i)_{i \in I}$ is the profile[2] of actions, one for each agent, then we should specify relations $p = f(a)$ and $y = g(a)$. This is done in all models of auctions. For example, in a sealed-bid, first-price, single-object auction, a is the profile of bids for the object on sale, $f(a) = \max\{a_i\}_{i \in I}$ (the object is sold for a price equal to the highest bid) and $g(a)$ is such that the object is allocated to the highest bidder,[3] who has to pay his bid. To be more general, we have to allow the variables of interest to depend also on some exogenous shocks x, as in the functional forms $p = f(a, x)$, $y = g(a, x)$. Furthermore, we should account for dynamics when choices and shocks take place over time, as in $y^t = g^t(a^1, x^1, ..., a^t, x^t)$. Of course, all the constraints on agents choices (such as those determined by technology) should also be explicitly specified. Finally, if we are to explain choices according to some rationality criterion, we should include in the model the preferences of each individual i over possible outcomes. This is what we call a "complete model" of the interactive situation.[4] We call variables, such as y, that depend on actions (and exogenous shocks) "endogenous." (Actions themselves are "endogenous" in a trivial sense.) The rationale for this terminology is that we try to analyze/explain actions and variables that depend on actions.

At this point, you may think that this is just a trite repetition of some abstract methodology of economic modelling. Well, think twice! The most standard concept in the economist's toolkit, Walrasian equilibrium, is not based on a complete model and is able to explain prices and allocations only by taking a two-step shortcut: (1) The modeler "pretends" that prices are observed and taken as parametrically given by *all* agents (including all firms) before they act, hence before they can affect such prices; this is a

[2] Given a set I of individuals, we always call "profile" a list of elements from a given set (e.g., a set of actions), one for each individual i.

[3] Ties can be broken at random.

[4] The model is still in some sense incomplete: we have not even addressed the issue of what the individuals know about the situation and about each other. But the elements sketched in the main text are sufficient for the discussion. Let us stress that what we call "complete model" does not include the modeler's hypotheses on how the agents choose, which are key to provide explanations or predictions of economic and social phenomena.

kind of logical short-circuit, but it allows to determine demand and supply functions $D(p)$, $S(p)$. Next, (2) market-clearing conditions $D(p) = S(p)$ determine equilibrium prices. Well, this can only be seen as a (clever) reduced-form approach; absent an explicit model of price formation (such as an auction model), the modeler postulates that somehow the choices-prices-choices feedback process has reached a rest point and he describes this point as a market-clearing equilibrium. In many applications of economic theory to the study of competitive markets, this is a very reasonable and useful shortcut, but it remains just a shortcut, forced by the lack of what we call a complete model of the interactive situation.

So, what do we get when instead we have a complete model? As we are going to show a bit more formally in the next section, we get what game theorists call a "game." This is why game theory should be a basic tool in economic modelling, even if one wants to analyze perfectly competitive situations. To illustrate this point, we will present a purely game-theoretic analysis of a perfectly competitive market, showing not only how such an analysis is possible, but also that it adds to our understanding of how equilibrium can be reached.

1.3 Abstract Game Models

A completely formal definition of the mathematical object called "(non-cooperative) game" will be given in due time. We start with a semi-formal introduction to the key concepts, illustrated by a very simple example, a seller-buyer mini-game. Consider two individuals, S (Seller) and B (Buyer). Let S be the owner of an object and B a potential buyer. For simplicity, consider the following bargaining protocol: S can ask one Euro (1) or two Euros (2) to sell the object, B can only accept (a) or reject (r). The monetary value of the object for individual i ($i = S, B$) is denoted V_i. This situation can be analyzed as a game which can be represented with a rooted tree with utility numbers attached to terminal nodes (leaves), player labels attached to non-terminal nodes, and action labels attached to arcs; see, for instance, Figure 1.1.

The game tree represents the formal elements of the analysis: the set of players (or roles in the game, such as seller and buyer), the actions, the rules of interaction, the consequences of complete sequences of actions and how players value such consequences.

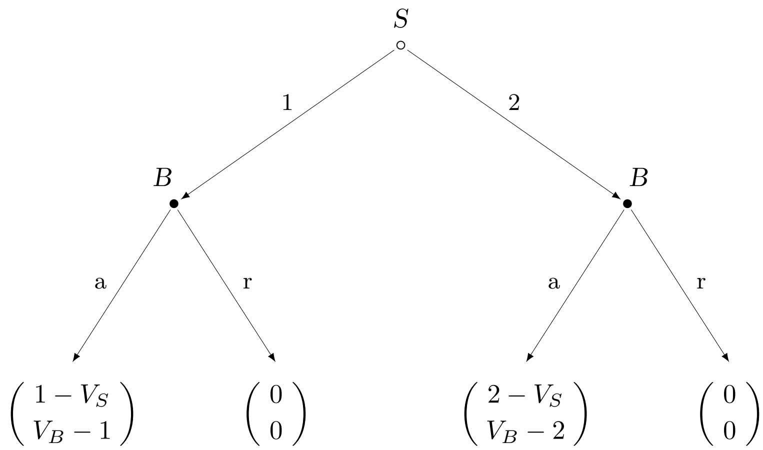


Figure 1.1: A seller-buyer mini-game.

Formally, a **game form** is given by the following elements:

- I, a set of **players**;
- For each player $i \in I$, a set A_i of **actions** which could conceivably be chosen by i at some point of the game as the play unfolds.
- Y, a set of **outcomes,** or **consequences**;
- $\mathcal{E}$, an **extensive-form** structure, that is, a mathematical representation of the rules saying whose turn it is to move, what a player knows, i.e., his or her information about past moves and random events, and what are the feasible actions at each point of the game; this determines a set Z of possible paths of play (sequences of feasible actions); a path of play $z \in Z$ may also contain some random events such as the outcomes of throwing dice;
- $g : Z \to Y$, an **outcome** (or consequence) **function** which assigns to each play z a consequence $g(z)$ in Y.

The above elements represent what the layperson would call the "rules of the game." To complete the description of the actual interactive situation (which may differ from how the players perceive it) we have to add players' preferences over consequences and—*via* expected utility calculations—lotteries over consequences:

- $(v_i)_{i \in I}$, where each $v_i : Y \to \mathbb{R}$ is a von Neumann-Morgenstern **utility function** representing player i's preferences over outcomes (consequences); preferences over lotteries are obtained *via* expected utility comparisons.

With this, we obtain the kind of mathematical structure called "game" in the technical sense of game theory. *This formal description does not say how such a game would (or should) be played.* The description of an interactive decision problem is only the first step to make a prediction (or a prescription) about players' behavior.

To illustrate, the mathematical description of the seller-buyer mini-game is as follows: $I = \{S, B\}$; $A_S = \{1, 2\}$, $A_B = \{a, r\}$; Y is a specification of the set of final allocations of the object and of monetary payments (for example, $Y = \{o_S, o_B\} \times \{0, 1, 2\}$, where (o_i, k) means object to i and k euros are given by B to S); $\mathcal{E}$ is represented by the game tree, without the utility numbers at endnodes; g is the rule stipulating that if B accepts the ask price p then B gets the object and gives p euros to S, if B rejects then S keeps the object and B keeps the money ($g(p, a) = (o_B, p)$, $g(p, r) = (o_S, 0)$); v_i is a risk-neutral (quasi-linear) utility function normalized so that the utility of no exchange is zero for both players ($v_S(o_B, p) = p - V_S$, $v_B(o_B, p) = V_B - p$, $v_i(o_S, 0) = 0$ for $i = S, B$). Game theory provides predictions on the behavior of S and B in this game based on hypotheses about players' knowledge of the rules of the game and of each other preferences, and on hypotheses about strategic thinking. For example, B is assumed to accept an ask price if and only if it is below his valuation. Whether the ask price is high or low depends on the valuation of S and what S knows about the valuation of B. If S knows the valuation of B, he can anticipate B's response to each offer and how much surplus he can extract.

A caveat on theoretical language It is common in game-theoretic work to refer to the *mathematical structure* describing an interactive situation (the "real game") as if it were the situation itself. Such abuse of language is usually innocuous, but sometimes the distinction is lost and this leads to unclear or confusing language. It is therefore useful to keep in mind the distinction between the real world objects and phenomena we are interested in and the mathematical structures used in the theoretical analysis. For example, the distinction is important in our analysis and discussion of games with incomplete or asymmetric information in Section 8.6 of Chapter 8.

1.4 Terminology and Classification of Games

Games come in different varieties and are analyzed with different methodologies. The same strategic interaction may be represented in a detailed way, using the mathematical objects described above, or in a much more parsimonious way. The amount of details in the formal representation constrains the methods used in the analysis. The terminology used to refer to different kinds of strategic situations and to different formal objects used to represent them may be confusing: Some terms that have almost the same meaning in the daily language, such as “perfect information” and “complete information,” have very different meanings in game theory; other terms, such as “non-cooperative game,” may be misleading. Furthermore, there is a tendency to confuse the substantive properties of the situations of strategic interaction that game theory aims at studying and the formal properties of the mathematical structures used in this endeavour. Here we briefly summarize the terminology and classification of game theory doing our best to dispel such confusion.

1.4.1 Cooperative vs Non-Cooperative Games

Suppose that the players (or at least some of them) could meet before the game is played and sign a binding agreement specifying their course of action (an external enforcing agency—e.g., the courts system—will force each player to follow the agreement). This could be mutually beneficial for them and if this possibility exists we should take it into account. But how? The theory of **cooperative games** does not model the process of bargaining (offers and counteroffers) which takes place before the real game starts; this theory considers instead a simple and parsimonious representation of the situation, that is, how much of the total surplus each possible coalition of players can guarantee to itself by means of some binding agreement. For example, in the seller-buyer situation discussed above, the (normalized) surplus that each player can guarantee to itself is zero, while the surplus that the $\{S, B\}$ coalition can guarantee to itself (the "gains from trade") is the difference between the lowest and the highest valuation, that is, $V_B - V_S$ in the standard case where $V_B \geq V_S$. This simplified representation is called **coalitional game**. For every given coalitional game the theory tries to figure out which division of the surplus could result, or, at least, the set of allocations that are not excluded by strategic considerations; see, for example, Osborne and Rubinstein [53, Part IV].

On the other hand, the theory of **non-cooperative games** assumes that either binding agreements are not feasible (e.g., in an oligopolistic market they could be forbidden by antitrust laws), or that the bargaining process which could lead to a binding agreement on how to play a game G is appropriately formalized as a sequence of moves in a "larger game" $\Gamma(G)$.[5] The players cannot agree on how to bargain in $\Gamma(G)$ and this is the game to be analyzed with the tools of non-cooperative game theory. It is therefore argued that non-cooperative game theory is more fundamental than cooperative game theory. For example, in the seller-buyer situation, $\Gamma(G)$ would be represented by the game-tree displayed in Figure 1.1 to illustrate the formal objects comprised by the mathematical representation of a game. The analysis of this game reveals that if S knows V_B he can use his first-mover advantage to ask for the highest price below V_B. Of course, the results of the analysis depend on details of the bargaining rules that may be unknown to an external observer and analyst. Neglecting such details, cooperative game theory typically gives weaker, but more robust predictions. Yet, in our view, the main advantage of non-cooperative game theory is not so much its sharper predictions, but rather the conceptual clarity of working with *explicit* assumptions and showing exactly how different assumptions lead to different results.

Cooperative game-theory is an elegant and useful analytical toolkit. It is especially appropriate when the analyst has little understanding of the true rules of interaction and wishes to derive some robust results from parsimonious information about the outcomes that coalitions of players can

[5] Here we are using symbol G to represent the actual interactive situation being modeled. As we move on to the formal analysis of games we will use the same symbol to denote the *mathematical structure* representing such situations. The meaning of G and similar symbols will be either explicit or clear from the context.

implement. However, non-cooperative game theory is much more pervasive in modern economics, and more generally in modern formal social sciences. We will therefore focus on non-cooperative game theory.

One thing should be clear, focusing on this part of the theory does not mean that cooperation is neglected. *Non-cooperative game theory is* ***not*** *the study of un-cooperative behavior*, but rather a method of analysis. Indeed we find the name "non-cooperative game theory" misleading, but it is now entrenched and we will conform to it.

1.4.2 Static and Sequential Games

A game is **static** if each player moves only once and all players move *simultaneously* (or at least without any information about other players' moves). Examples of static games are: Matching Pennies, Stone-Scissor-Paper, and sealed bid auctions.

A game has a **multistage** structure if some moves are *sequential*, players always know the number of previous moves, and some player may observe (at least partially) the previous behavior of his or her co-players. Examples of multistage games are: Chess, Poker, and open outcry auctions. Not all games with sequential moves have a multistage structure. We focus on this special case because it is simpler and it covers many applications.

Games with sequential moves are sometimes analyzed *as if* the players moved only once and simultaneously. The trick is to pretend that each player chooses in advance his **strategy**, i.e., a contingent plan of action specifying how to behave in every circumstance that may arise while playing the game. Each profile of strategies $s = (s_i)_{i \in I}$ determines a particular path, viz., $\zeta(s)$, hence an outcome $g(\zeta(s))$, and ultimately a profile of utilities, or "payoffs," $(v_i(g(\zeta(s))))_{i \in I}$. This mapping $s \longmapsto (v_i(g(\zeta(s))))_{i \in I}$ is called the **normal form**, or **strategic form** of the game. The strategic form of a game can be seen as a static game where players simultaneously choose strategies in advance. For example, in the seller-buyer mini-game the strategies of S (seller) coincide with his possible ask prices; on the other hand, the set of strategies of B (buyer) contains four response rules: $S_B = \{a_1 a_2, a_1 r_2, r_1 a_2, r_1 r_2\}$ where a_p (respectively, r_p) is the instruction "accept (resp., reject) price p." The strategic form is as follows:

$S \setminus B$	a_1a_2	a_1r_2	r_1a_2	r_1r_2
$p = 1$	$1 - V_S, V_B - 1$	$1 - V_S, V_B - 1$	$0, 0$	$0, 0$
$p = 2$	$2 - V_S, V_B - 2$	$0, 0$	$2 - V_S, V_B - 2$	$0, 0$

Figure 1.2: Strategic form of the seller-buyer minigame.

A strategy in a static game is just the plan to take a specific action, with no possibility to make this choice contingent on the actions of others, as such actions are being simultaneously chosen. Therefore, the mathematical representation of a static game and its normal form must be the same. For this reason, static games are often called "normal-form games," or "strategic-form" games. We are going to avoid this widespread abuse of language. In a correct language there should not be "normal-form games"; rather, looking at the normal form of a game is a *method of analysis.*

1.4.3 Assumptions about Information

Perfect Information and Observable Actions

A multistage game has **perfect information** if players move one at a time and each player—when it is his or her turn to move—is informed of all the previous moves (including the realizations of chance moves). If some moves are simultaneous, but each player can observe all past moves, we say that the game has **observable actions (or perfect monitoring, or almost perfect information)**. Examples of games with perfect information are the seller-buyer mini-game, Chess, Backgammon, and Tic-Tac-Toe. An example of a game with observable actions is the repeated Cournot oligopoly, with simultaneous choice of outputs in every period and perfect monitoring of past outputs. Note, perfect information is an assumption about the *rules* of the game.

Asymmetric Information

A game with imperfect information features **asymmetric information** if different players get different pieces of information about past moves and in particular about the realizations of chance moves. Poker and Bridge are games with asymmetric information because players observe their cards, but not the cards of others. Like perfect information, also asymmetric information is entailed by the *rules* of the game.

Complete and Incomplete Information

An event E is **common knowledge** if everybody knows E, everybody knows that everybody knows E, and so on for all iterations of "everybody knows that." To use a suggestive semi-formal expression: for every $m = 1, 2, ...$ it is the case that [(everybody knows that)^m E is the case].

There is **complete information** in an interactive decision problem represented by $G = \langle I, A, Y, \mathcal{E}, g, (v_i)_{i \in I} \rangle$ if it is common knowledge that the actual interactive situation is the one represented by G. Conversely, there is **incomplete information** if for some player i either it is *not* the case that [i knows the actual game], or for some $m = 1, 2, ...$ it is *not* the case that [i knows that (everybody knows that)^m the actual interactive situation]. Most economic situations feature incomplete information because either the outcome function (represented by g), or players' preferences (represented by the utility functions $(v_i)_{i \in I}$) are not common knowledge. Note, complete (or incomplete) information is *not* an assumption *about* the *rules* of the game, it is an assumption on players' "interactive knowledge" concerning the rules and preferences. For example, in the seller-buyer mini-game it can be safely assumed that there is common knowledge of the outcome function (who gets the object and monetary transfers), but the valuations V_S and V_B need not be commonly known. For some types of objects on sale, V_i is known only to i; for other types of objects, the seller may know the quality of the object better then the buyer, so V_B could be given by an idiosyncratic component known to B plus a quality component known to S.

Although situations with *asymmetric* information about chance moves, such as Poker, are conceptually different from situations with *incomplete* information, we will see that there is a formal similarity which allows (at least to some extent) to use the same analytical tools to study both situations.

Examples: In games like chess and poker there is—presumably—complete information; indeed, the "rules of the game" are common knowledge, and it may be taken for granted that it is common knowledge that players like to win. On the other hand, auctions typically feature incomplete information because the competitors' valuations of the objects on sale are not commonly known.

1.5 Rational Behavior

The decision problem of a single individual i can be represented as follows: i can take an action in a set of feasible alternatives, or decisions, D_i; i's welfare (payoff, utility) is determined by his decision d_i and an external state $\omega_{-i} \in \Omega_{-i}$ (which represents the realizations of a vector of variables beyond i's control, e.g., the outcome of a random event and/or the decisions of other players) according to the outcome function

$$g : D_i \times \Omega_{-i} \to Y$$

and the utility function

$$v_i : Y \to \mathbb{R}.$$

We assume that i's choice cannot affect ω_{-i} and that i does not have any information about ω_{-i} when he chooses beyond the fact that $\omega_{-i} \in \Omega_{-i}$. The choice is made once and for all. (We will show in the part on sequential games how this model can be generalized.) Assume, just for simplicity, that D_i and Ω_{-i} are both finite and let, for notational convenience, $\Omega_{-i} = \{\omega_{-i}^1, \omega_{-i}^2, ..., \omega_{-i}^n\}$. In order to make a "rational" choice, i has to assess the probabilities of the different states in Ω_{-i}. Suppose that his beliefs about ω_{-i} are represented by a probability vector $\mu^i \in \Delta(\Omega_{-i})$ where

$$\Delta(\Omega_{-i}) = \left\{ \mu^i \in \mathbb{R}_+^n : \sum_{k=1}^{n} \mu^i(\omega_{-i}^k) = 1 \right\}.$$

Then i's **rational decision given** μ^i is to take any alternative that maximizes i's expected payoff, i.e., any d_i^* such that

$$d_i^* \in \arg\max_{d_i \in D_i} \mathbb{E}_{\mu^i} \left(v_i \left(g \left(d_i, \cdot \right) \right) \right),$$

where $\mathbb{E}_{\mu^i}$ denotes the expectation with respect to μ^i,

$$\mathbb{E}_{\mu^i} \left(v_i \left(g \left(d_i, \cdot \right) \right) \right) = \sum_{k=1}^{n} \mu^i(\omega_{-i}^k) v_i \left(g(d_i, \omega_{-i}^k) \right).$$

Decision d_i^* is also called a **best response**, or **best reply**, to μ^i. The probability vector could be exogenously given (roulette lottery) or simply represent i's **subjective beliefs** about ω_{-i} (horse lottery, game against an opponent). The composite function $v_i \circ g$ is denoted u_i, that is, $u_i(d_i, \omega_{-i}) = v_i\left(g(d_i, \omega_{-i})\right)$; u_i is called the **payoff function**, because in many interesting games the rules of the decision problem (or game) attach monetary payoffs to the possible outcomes of the game and it is taken for granted that the decision maker maximizes his expected monetary payoff. In the finite case, u_i is often represented as a matrix with (k, ℓ) entry $u_i^{k\ell} = u_i(d_i^k, \omega_{-i}^\ell)$: see Figure 1.3.

	ω_{-i}^1	ω_{-i}^2	$\ldots$
d_i^1	u_i^{11}	u_i^{12}	$\ldots$
d_i^2	u_i^{21}	u_i^{22}	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$

Figure 1.3: Matrix representation of the payoff function.

This representation of a decision maker's choice, of course, fits very well the case of static games. In this case $D_i = A_i$ is simply the set of feasible actions and Ω_{-i} may be the set A_{-i} of the opponents' feasible action profiles (or a larger space if there is incomplete information).

However, it has been argued that the representation is more general:[6] Consider a game with sequential moves, e.g., a multistage game. The rules of interaction specifies all the circumstances under which player i may have to choose an action and the information i would have in those circumstances. Then player i can formulate in advance a plan of action, or strategy, which specifies how he would behave in any such circumstance given the available information. Assuming that such plan is incentive compatible, that is, that player i has no incentive to deviate from it, then player i expects to follow his plan. His uncertainty concerns the contingent behavior of others, hence the strategies they are going to follow. Every profile of strategies, when carried out, induces a particular outcome. Therefore, the planning problem faced by a rational agent is the problem of selecting an incentive-compatible strategy under uncertainty about the

[6]See von Neumann and Morgenstern [68].

strategies used by the opponents (and possibly about other unknowns). In the part of this textbook devoted to multistage games we will discuss some subtleties related to the concepts of "strategy" and "plan."

1.6 Assumptions and Solution Concepts in Game Theory

Game theory provides a mathematical language to formulate assumptions about the "rules of the game" and players' preferences, that is, all the elements listed in Section 1.3. But such assumptions are not enough to derive conclusions about how the players behave in a given game. The central behavioral assumption in game theory is that players are rational (see Section 1.5). However, without further assumptions about what players believe about the variables which affect their payoffs, but are beyond their control (in particular their opponents' behavior), the rationality assumption has little behavioral content: we can only say that a rational player's choice must be **justifiable**, i.e., it it must be a best response to *some* belief. In most games, this is too little to derive interesting results.[7]

Then, how can we obtain interesting results from our assumptions about the rules of the game and preferences? The standard approach used in game theory is analogous to the one used in the undergraduate economics analysis of competitive markets, whereby we formulate assumptions about preferences and technology and then we assume that economic agents' plans are mutually consistent best responses to equilibrium prices.

As explained in Section 1.2, there is an important difference: the textbook analysis of competitive markets does not specify a price-formation mechanism and uses equilibrium market-clearing as a theoretical shortcut to overcome this problem. On the contrary, in a game-theoretic model all the observable variables we try to explain depend on players' actions (and exogenous shocks) according to an explicitly specified function, as in auction models.

[7] There are important exceptions. For example, in many situations where each individual in a group decides how much to contribute to a public good, it is strictly dominant to give a minimal contribution (see Example 3 in Chapter 3).

Yet, there are also similarities with the analysis of competitive markets. One could say that the role of prices is played (sic) by players' beliefs since we assume that they are, in some sense, mutually consistent. The precise meaning of the statement "beliefs are mutually consistent" is captured by a **solution concept.** The simplest and most widely used solution concept in game theory, Nash equilibrium, assumes that players' beliefs about each other strategies are correct and each player best responds to her or his beliefs; as a result, each player uses a strategy that is a best response to the strategies used by other players.

Nash equilibrium is not the only solution concept used in game theory. Recent developments made clear that solution concepts implicitly capture expressible[8] *assumptions about players' rationality and beliefs*, and some assumptions that are appropriate in some contexts are too weak, or simply inappropriate, in different contexts. Therefore, it is very important to provide convincing motivations for the solution concept used in a specific application.

Let us somewhat vaguely define as **"strategic thinking"** what intelligent agents do when they are fully aware of participating in an interactive situation and form conjectures by putting themselves in the shoes of other intelligent agents. As the title of this book suggests, we mostly (though not exclusively) present game theory as an *analysis of strategic thinking.* We will often follow the traditional "textbook" game theory and present different solution concepts providing informal motivations for them, sometimes with a lot of hand-waving. However, the reader should always be aware that a more fundamental approach is being developed, where game theorists formulate explicit assumptions not only about the rules of the game and preferences, but also about players' rationality and initial beliefs, and also about how beliefs change as the play unfolds. Implications about choices and observables can be derived directly from such assumptions, without the mediation of solution concepts. Then, solution concepts become mere "shortcuts" to characterize the behavioral implications of assumptions about rationality and beliefs.

For example, a standard solution concept in game theory, subgame perfect equilibrium, requires players' strategies to form an equilibrium not

[8] Roughly, "expressible" means "something that can be expressed in a clear, precise, and not self-referential language."

only for the game itself, but also in every subgame.[9] In finite games with (complete and) perfect information, if there are no relevant ties, there is a unique subgame perfect equilibrium that can be computed with a backward induction algorithm. For two-stage games with perfect information, subgame perfect equilibrium can be derived from simple assumptions about rationality and beliefs: players are rational and the first mover believes that the second mover is rational. To illustrate, consider the seller-buyer mini-game and assume players have *complete information.* Specifically, suppose that it is common knowledge that $V_S < 1$and $V_B > 2$. The assumption of rationality of the buyer B, denoted with R_B, implies that B accepts every price below valuation V_B; since $V_B > 2$, the only strategy consistent with R_B is a_1a_2, that is, to accept both $p = 1$ and $p = 2$. Rationality of the seller S, denoted with R_S, implies that S attaches subjective probabilities to the four strategies of B and chooses the ask price that maximizes S's expected utility; if S has deterministic beliefs, S asks for the highest prices he expects to be accepted, e.g., S asks $p = 1$ if he is certain of strategy a_1r_2. But if S is certain that B is rational, given the complete information assumption, S must assign probability 1 to strategy a_1a_2. Then the rationality of S and S's in the rationality of B imply that S asks for the highest price, i.e., S uses the bargaining power derived from the knowledge that $V_B > 2$ and the first-mover advantage. In symbols, write $\mathrm{B}_i(E)$ for "i believes (with probability 1) E," where E is some event; then $R_S \cap \mathrm{B}_S(R_B)$ implies that S asks for $p = 2$; R_B implies that $p = 2$ is accepted; $R_S \cap \mathrm{B}_S(R_B) \cap R_B$ implies that the action/strategy of S is $p = 2$ and the strategy of B is a_1a_2, which is the subgame perfect equilibrium obtained *via* backward induction.

Two-stage games with complete and perfect information are so simple that the analysis above may seem trivial. On the other hand, the *rigorous* analysis of more complicated multistage games based on higher levels of mutual belief in rationality presents conceptual difficulties that will be addressed only in the second part of this textbook. Here we further illustrate strategic thinking with a semi-formal analysis of a static game representing a perfectly competitive market. We hope this will also have the side effect to make the reader understand that *game theory is useful to*

[9]A precise definition of a subgame will be given in Part II. For the time being, you should think of a subgame as a "smaller"game that starts with a non-terminal node of the original game.

analyze every "complete" model of economic interaction, including models of perfect competition.

1.6.1 Game Theoretic Analysis of a Perfectly Competitive Market

Let us analyze the decisions of a large number n of small, identical agricultural firms producing a crop (say, corn). For reasons that will be clear momentarily, it is convenient to index firms as equally spaced rational numbers in the interval $(0, 1]$, that is, $I = \{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\} \subset (0, 1]$ with generic element $i \in I$. Each firm i decides in advance how much to produce. This quantity is denoted $q(i)$. The output of corn $q(i)$ will be ready to be sold only in six months. Markets are incomplete: it is impossible to agree in advance on a price for corn to be delivered in six months. So, each firm i has to guess the price p for corn in six months. It is common knowledge that the demand for corn (not explicitly "micro-founded" in this example) is given by function $D(p) = n \max\{(\alpha - \beta p), 0\}$, that is, $D(p) = n(\alpha - \beta p)$ if $p \leq \frac{\alpha}{\beta}$ and $D(p) = 0$ otherwise. It is also common knowledge that each firm will fetch the highest price at which the market can absorb the total output for sale, $Q = \sum_{i=1}^{n} q(i)$. Thus, each firm will sell each unit of output at the uniform price $P\left(\frac{Q}{n}\right) = \frac{1}{\beta}\left(\alpha - \frac{Q}{n}\right)$ if $\frac{Q}{n} \leq \alpha$ and $p = 0$ if $\frac{Q}{n} > \alpha$ (for notational convenience, we express inverse demand as a function of the *average* output $\frac{Q}{n}$). Each firm i has total cost function $C(q) = \frac{1}{2m}q^2$, hence the marginal cost function is $MC(q) = \frac{q}{m}$. Each firm has obvious preferences, it wants to maximize the difference between revenues and costs. The technology (cost function) and preferences of each firm are common knowledge.[10]

Now, we want to represent mathematically the assumption that each firm is "negligible" with respect to the (maximum) size of the market α. Instead of assuming a large, but finite set of firms given by a fine grid I in the interval $(0, 1]$, we use a quite standard mathematical idealization and assume that there is a *continuum* of firms, normalized to the unit interval $I = (0, 1]$. Thus, the average output is $q = \int_0^1 q(i)\mathrm{d}i$ instead of $\frac{1}{n}\sum_{i=1}^{n} q(i)$. Each firm understands that its decision cannot affect the total output and

[10] The analysis can be generalized assuming heterogeneous firms, where each firm i is characterized by the marginal cost parameter $m(i)$. Then it is enough to assume that each i knows $m(i)$ and that the average $m = \frac{1}{n}\sum_{i=1}^{n} m(i)$ is common knowledge.

the market price.[11]

Given that all the above is common knowledge, we want to derive the price implications of the following assumptions about rationality and beliefs:

R (*rationality*): each firm has a conjecture about (q and) p and chooses a best response to such conjecture, for brevity, each firm is rational;[12]

$\mathrm{B}(R)$ (*mutual belief in rationality*): all firms believe (with certainty) that all firms are rational;

$\mathrm{B}^2(R) = \mathrm{B}(\mathrm{B}(R))$ (*mutual belief of order 2 in rationality*): all firms believe that all firms believe that all firms are rational;

...

$\mathrm{B}^k(R) = \mathrm{B}(\mathrm{B}^{k-1}(R))$ (*mutual belief of order k in rationality*): [all firms believe that]^k all firms are rational;

...

Note, we are using symbols for these assumptions. In a more formal mathematical analysis, these symbols would correspond to *events* in a space of states of the world. Here they are just useful abbreviations. Conjunctions of assumptions will be denoted by the symbol $\cap$, as in $R \cap \mathrm{B}(R)$. It is a good thing that you get used to these symbols, but if you find them baffling, just ignore them and focus on the words.

What are the consequences of *rationality* (R) in this market? Let $p(i)$ denote the price *expected* by firm i.[13] Then i solves

$$\max_{q \geq 0} \left\{ p(i)q - \frac{1}{2m} q^2 \right\}$$

which yields

$$q(i) = m\ p(i).$$

Since firms know the inverse demand function $P(q) = \max\{(\alpha - q)/\beta, 0\}$, each firm i's expectation of the price is below the upper bound $\overline{p}^0 = \frac{\alpha}{\beta}$: $p(i) \leq \frac{\alpha}{\beta}$. It follows that

$$q = \int_0^1 mp(i)\mathrm{d}i = m \int_0^1 p(i)\mathrm{d}i \leq \overline{q}^1 := m\frac{\alpha}{\beta},$$

[11] This example is borrowed from Guesnerie [39].

[12] The symbols R, $\mathrm{B}(R)$ etc. below should be read as abbreviations of sentences. They can also be formally defined using mathematics, but this goes beyond the scope of this textbook.

[13] This is, in general, the mathematical expectation of p according to the subjective probabilistic conjecture of firm i.

where $\overline{q}^1$ denotes the upper bound on average output when firms are rational. (We use ":=" to mean "equal by definition", when this is not obvious from the context.)

By *mutual belief in rationality*, each firm understands that $q \leq \overline{q}^1$ and $p \geq \max\{(\alpha - m\frac{\alpha}{\beta})/\beta, 0\}$. If $m \geq \beta$, this is not very helpful: assuming belief in rationality does not reduce the span of possible prices and does not yield additional implications for the endogenous variables we are interested in. It is not hard to understand that rationality and mutual belief in rationality of every order yields the same result $q \leq \frac{m\alpha}{\beta}$, where $\frac{m\alpha}{\beta} \geq \alpha$ because $m \geq \beta$.

So, let us assume from now on that $m < \beta$. Note, this is the so called "cobweb stability" condition. Given belief in rationality, $\mathrm{B}(R)$, each firm expects a price at least as high as the lower bound

$$\underline{p}^1 := \frac{1}{\beta}\left(\alpha - \overline{q}^1\right) = \frac{\alpha}{\beta}\left(1 - \frac{m}{\beta}\right) > 0.$$

Therefore $R \cap \mathrm{B}(R)$ implies that $q = m\int_0^1 p(i)di \geq m\underline{p}^1$, that is, average output is above the lower bound

$$\underline{q}^2 := m\underline{p}^1 = \alpha\frac{m}{\beta}\left(1 - \frac{m}{\beta}\right)$$

and hence the price is below the upper bound

$$\overline{p}^2 := \frac{1}{\beta}(\alpha - \underline{q}^2) = \frac{\alpha}{\beta}\left(1 - \frac{m}{\beta}\left(1 - \frac{m}{\beta}\right)\right) = \frac{\alpha}{\beta}\sum_{k=0}^{2}\left(-\frac{m}{\beta}\right)^k.$$

By $\mathrm{B}(R)$ and $\mathrm{B}^2(R)$ (*mutual belief of order 2 in rationality*), each firm predicts that $p \leq \overline{p}^2$. Rational firms with such expectations choose an output below the upper bound

$$\overline{q}^3 := m\overline{p}^2 = \alpha\frac{m}{\beta}\left(1 - \frac{m}{\beta}\left(1 - \frac{m}{\beta}\right)\right) = \alpha\frac{m}{\beta}\sum_{k=0}^{2}\left(-\frac{m}{\beta}\right)^k.$$

Thus, $R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$ implies that $q \leq \overline{q}^3$ and the price must be above the lower bound

$$\underline{p}^3 := \frac{1}{\beta}(\alpha - \overline{q}^3) = \frac{\alpha}{\beta}\sum_{k=0}^{3}\left(-\frac{m}{\beta}\right)^k.$$

Can you guess the consequences for price and average output of assuming *rationality and common belief in rationality* (mutual belief in rationality of every order)? Draw the classical Marshallian cross of demand-supply, trace the upper and lower bounds found above and go on. You will find the answer.

More formally, define the following sequences of upper bounds and lower bounds on the price:

$$\begin{aligned}\overline{p}^L &= \frac{\alpha}{\beta}\sum_{k=0}^{L}\left(-\frac{m}{\beta}\right)^k, \, L \text{ even} \\ \underline{p}^L &= \frac{\alpha}{\beta}\sum_{k=0}^{L}\left(-\frac{m}{\beta}\right)^k, \, L \text{ odd}\end{aligned}$$

It can be shown by mathematical induction that $R \cap \bigcap_{k=1}^{L-1} B^k(R)$ (rationality and mutual belief in rationality of order $k = 1, \ldots, L-1$, with $L \geq 2$) implies $p \leq \overline{p}^L$ if L is even, and $p \geq \underline{p}^L$ if L is odd. Since $\frac{m}{\beta} < 1$, the sequence $\overline{p}^{2\ell} = \frac{\alpha}{\beta}\sum_{k=0}^{2\ell}\left(-\frac{m}{\beta}\right)^k$ is decreasing in ℓ, the sequence $\underline{p}^{2\ell+1} = \frac{\alpha}{\beta}\sum_{k=0}^{2\ell+1}\left(-\frac{m}{\beta}\right)^k$ is increasing in ℓ, and they both converge to the competitive equilibrium price $p^* = \frac{\alpha}{\beta+m}$, see Figure 1.4.

A Comment on "Rational" *versus* "Rationalizable" Expectations

This price p^* is often called the **"rational expectations"** price, because it is the price that the competitive firms would expect if they performed the same equilibrium analysis as the modeler. *We refrain from using such terminology.* Game theory allows a much more rigorous and precise terminology, the one that we have been using above. In this textbook, "rationality" is a joint property of choices and beliefs and it means *only* that agents best respond to their beliefs. The phrase "rational beliefs," or "rational expectations" was coined at a time when theoretical economists did not have the formal tools to be precise and rigorous in the analysis of rationality and interactive beliefs. Now that these tools exist, such phrases as "rational expectations" should be avoided, at least in game

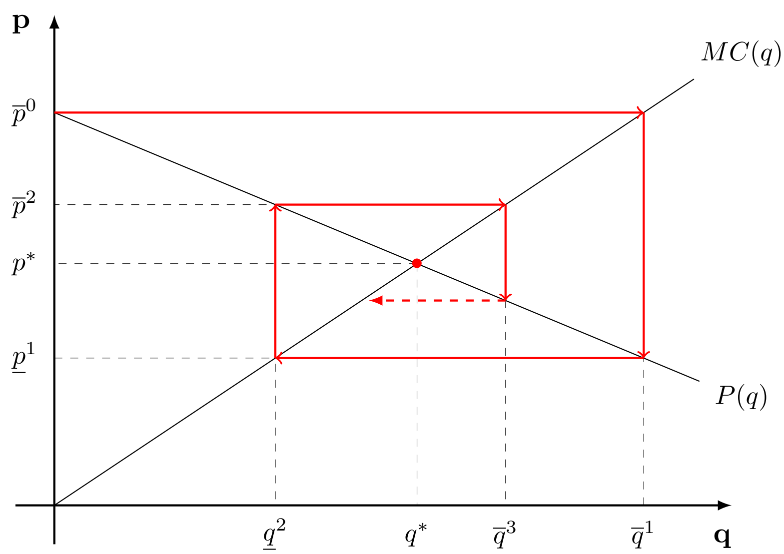


Figure 1.4: Strategic thinking in a perfectly competitive market.

theoretic analysis. On the other hand, expectations may be consistent with common belief in rationality (e.g., under common knowledge of the game), and it makes sense to call such expectations **"rationalizable**." Note well, the fact that there is only one rationalizable expectations price, p^*, is a *conclusion* of the analysis, it has not been assumed! Also, this conclusion holds only under the "cobweb stability" condition $m < \beta$. This contrasts with standard equilibrium analysis, whereby it is assumed at the outset that expectations are correct, whether or not this follows from assumptions like common belief in rationality.

What do we learn from this example? First, it shows that certain *assumptions about rationality and beliefs* (plus common knowledge of the game) may yield interesting implications or not according to the given model and parameter values. Here, rationality and common belief in rationality have sharp implications for the market price and average output if the "cobweb stability" condition holds. Otherwise they only yield a very weak upper bound on average output.

Second, recall that **strategic thinking** is what intelligent agents do when they are fully aware of participating in an interactive situation and form conjectures by putting themselves in the shoes of other intelligent agents; in this sense, the above is as good an analysis of strategic thinking as any, and yet it is the analysis of a perfectly competitive market!

Third, the example shows how solution concepts can be interpreted as shortcuts. Here we presented an iterated deletion of prices and conjectures, where one deletes prices that cannot occur when firms best respond to conjectures and delete conjectures that do not rule out deleted prices. Note, mutual beliefs of order k do not appear explicitly in this iterative deletion procedure. The procedure is a solution concept of a kind and it can be used as a shortcut to obtain the implications of rationality and common belief in rationality. Under "cobweb stability," the rational expectation equilibrium price p^* results. This shows that under appropriate conditions we can use the equilibrium concept to draw the implications of rationality and common belief in rationality. This theme will be played time and again.

Part I

Static Games

2

Static Games: Formal Representation and Interpretation

A **static game**, or one-stage game, or simultaneous-moves game, is an interactive situation where all players move simultaneously and only once.[1] The key features of a static game are formally represented by a mathematical structure (a list of sets and functions)

$$G = \langle I, Y, (A_i)_{i \in I}, g, (v_i)_{i \in I} \rangle ,$$

where:

- I is the set of individuals or **players**, typically denoted by i, j;
- A_i is the *nonempty* set of possible **actions** for player i; a_i, a_i^*, a_i', $\hat{a}_i$ are alternative action labels we frequently use;

[1]Sometimes static games are also called "normal-form games" or "strategic-form games." As we mentioned in Chapter 1, this terminology is somewhat misleading. The normal, or strategic form of a game Γ has the same structure of a static game, but the game Γ itself may have a sequential structure. The normal form of Γ shows the payoffs induced by any combination of plans of actions of the players, if such plans are implemented. Some game theorists, including the founding fathers von Neumann and Morgenstern [68], argue that from a theoretical point of view all strategically relevant aspects of a game are contained in its normal form. Be as it may, here by "static game" we specifically mean a game where players move *simultaneously*.

- $g : \bigtimes_{i\in I} A_i \to Y$ is the **outcome** (or consequence) **function** which captures the essence of the rules of the game, beyond the assumption of simultaneous moves;

- $v_i : Y \to \mathbb{R}$ is the Von Neumann-Morgenstern utility function of player i.

Structure $\langle I, Y, (A_i)_{i\in I}, g\rangle$, that is, the game without the utility functions, is called **game form**. The game form represents the essential features of the rules of the game. A game is obtained by adding to the game form a profile of utility functions $(v_i)_{i\in I}$, which represents players' preferences over lotteries of consequences, according to expected utility calculations.

From the consequence function g and the utility function v_i of player i, we obtain a function that assigns to each $a = (a_j)_{j\in I}$ the utility $v_i(g(a))$ for player i of consequence $g(a)$. This function

$$u_i = v_i \circ g : \bigtimes_{i\in I} A_i \to \mathbb{R}$$

is called **payoff function** of player i. The reason why u_i is called payoff function is that the early work on game theory assumed that consequences are distributions of monetary payments, or payoffs, and that players are risk neutral, so that it is sufficient to specify, for each player i, the monetary payoff implied by each action profile. But in modern game theory $u_i(a) = v_i(g(a))$ is interpreted as the von Neumann-Morgenstern utility of outcome $g(a)$ for player i. If there are monetary consequences, then

$$g = (g_i)_{i\in I} : \bigtimes_{i\in I} A_i \to \mathbb{R}^I,$$

where $m_i = g_i(a)$ is the net gain of player i when a is played. Assuming that player i is selfish, function v_i is strictly increasing in m_i and constant in each m_j with $j \neq i$ (note that selfishness is not an assumption of game theory, it is an economic assumption that may be adopted in game theoretic models). Thus, function v_i captures the risk attitudes of player i. For example, i is strictly risk averse if and only if v_i is strictly concave.

In some games the outcome function is *stochastic*, that is, each $a = (a_j)_{j\in I}$ corresponds to a lottery. For example, in a sealed-bid auction the object on sale is assigned at random to one of the high bidders if there is a tie. In the stochastic outcome case, u_i has to be interpreted as an expected payoff: if Y is finite,

$$u_i(a) = \sum_{y \in Y} g(a)(y) v_i(y),$$

where $g(a)(y)$ is the probability of outcome y according to lottery $g(a)$.

Games are typically represented in the reduced form

$$G = \langle I, (A_i, u_i)_{i \in I} \rangle$$

which shows only the payoff functions. We often do the same. However, it is conceptually important to keep in mind that payoff functions are derived from a consequence function and utility functions.

For ease of exposition, we often assume that all the sets A_i $(i \in I)$ are *finite*; when the sets of actions are infinite, we explicitly say so. This simplifies the analysis of probabilistic beliefs. A static game with finite action sets is called a **finite static game** (or just a finite game, if "static" is clear from the context). The following assumption is a generalization of finiteness:

Definition 1. *A (reduced form) static game $G = \langle I, (A_i, u_i)_{i \in I} \rangle$ is **compact-continuous** if, for each $i \in I$, A_i is a compact subset of a Euclidean space $\mathbb{R}^{k_i}$ and $u_i : \times_{j \in I} A_j \to \mathbb{R}$ is continuous.*[2]

We can always assume without loss of generality that, if A_i is finite, then it is a finite subset of some Euclidean space. Every such set is compact, and every function with a finite domain is continuous. Therefore, a finite game is compact-continuous. Most of our results hold for compact-continuous games. In some cases (for example, to prove the existence of an equilibrium) we may have to add other assumptions about the actions sets and the payoff functions.

We call **profile** a list of objects $(x_i)_{i \in I}$, one object x_i from some set X_i for each player $i \in I$. In particular, an **action profile** is a list

$$a = (a_i)_{i \in I} \in A = \times_{i \in I} A_i.$$

[2] A subset of a Euclidean space is compact if and only if it is closed and bounded. For most of the results that rely on compactness and continuity, it is sufficient to assume that each action set A_i is a compact subset of a metrizable topological space.

The payoff function u_i numerically represents the preferences of player i among the different action profiles $a, a', a'', \ldots \in A$. The strategic interdependence is due to the fact that the outcome function g depends on the entire action profile a and, consequently, the utility that a generic individual i can achieve depends not only on his choice, but also on those of other players.[3] To stress how the payoff of i depends on a variable under i's control as well as on a profile of variables controlled by other individuals, we denote by $-i = I\backslash\{i\}$ the set of individuals different from i, we define

$$A_{-i} = \bigtimes_{j \in I\backslash\{i\}} A_j,$$

and we write the payoff of i as a function of the two arguments $a_i \in A_i$ and $a_{-i} \in A_{-i}$, that is, $u_i : A_i \times A_{-i} \to \mathbb{R}$.

In order to be able to reach some conclusions regarding players' behavior in a game G, we impose two *minimal assumptions* (further assumptions will be introduced later on as necessary):

(H1) Each player i knows A_i, A_{-i} and his own payoff function $u_i : A_i \times A_{-i} \to \mathbb{R}$;

(H2) Each player is rational (see Chapter 3 for a more formal definition of rationality).

Assumption (H1) (i knows u_i) may seem almost tautological: is it not obvious that every individual should know his preferences over A? However, action profiles are not consequences (outcomes). We assume that each i knows his own preferences over (lotteries of) consequences,

[3]Note, in our formulation, players' utility depends only on the consequence determined by action profile a through outcome function g. Thus, we are assuming a form of *consequentialism*. If you think that consequentialism is an obvious assumption, think twice. Substantial experimental evidence suggests agents also care about other players' *intentions*: agent i may evaluate the same action profile differently depending on his belief about opponents' intentions. To provide a formal analysis of players' intentions, we need to specify what an agent thinks other agents will do, what an agent thinks other agents think other agents will do, what an agent thinks other agents think other agents think that other agents will do and so on. Although the analysis of these issues is beyond the scope of this textbook, the interested reader is referred to the growing literature on *psychological games* (see Geanakoplos *et al* [35], Battigalli and Dufwenberg [11], Battigalli, Corrao and Dufwenberg [9], and Battigalli and Dufwenberg [12]).

represented by utility function $v_i : Y \to \mathbb{R}$, but he may ignore the true outcome function $g : A \to Y$, hence he may ignore his payoff function $u_i = v_i \circ g$. The following example clarifies this point.

Example 1. *(Knowledge of the utility and payoff functions)* Players 1 and 2 work in a team for the production of a public good. Their action is their effort $a_i \in [0, 1]$. The output, Q, depends on efforts according to a Cobb-Douglas production function[4]

$$Q = K(a_1)^{\theta_1}(a_2)^{\theta_2}.$$

The cost of effort measured in terms of public good is $C_i = (a_i)^2$. Then, the outcome function is

$$g(a_1, a_2) = (K(a_1)^{\theta_1}(a_2)^{\theta_2}, (a_1)^2, (a_2)^2).$$

The utility function of player i is $v_i(Q, C_1, C_2) = Q - C_i$. The payoff function is

$$u_i(a_1, a_2) = v_i(g(a_1, a_2)) = K(a_1)^{\theta_1}(a_2)^{\theta_2} - (a_i)^2.$$

If i does not know all the parameters K, θ_1 and θ_2 (as well as the functional forms above), then he does not know the payoff function u_i, even if he knows the utility function v_i. ▲

Example 1 also clarifies the difference between "game" in the technical sense of game theory, and the "rules of the game" as understood by the layperson (see the Chapter 1). In the example above, the "rules of the game" specify the action set $[0, 1]$ and the outcome function

$$(a_1, a_2) \longmapsto (Q, C_1, C_2) = \left(K(a_1)^{\theta_1}(a_2)^{\theta_2}, (a_1)^2, (a_2)^2\right).$$

In order to complete the description of the game in the technical sense, we also have to specify the utility functions $v_i(Q, C_1, C_2) = Q - C_i$ $(i = 1, 2)$.[5] Also, the example clarifies that the outcome map $a \mapsto g(a)$ may depend on personal features of the agents playing the game: in this case, the

[4] In some applications players exert effort in a joint project and output q is interpreted as the probability that the project is successful.

[5] This specification assumes that players are risk neutral with respect to their consumption of public good.

productivity θ_i of each agent affects how the pair of efforts (a_1, a_2) (say, the hours worked) is mapped to the output Q. Therefore, the outcome function g is not always fully determined by what we would call "rules of the game" in the natural language.

Unlike the games people play for fun (and those that experimental subjects play in most game theoretic experiments), in many economic games the outcome function may not be fully known. In Example 1, for instance, it is possible that player i knows his own productivity parameter θ_i, but does not know K or θ_{-i}. Thus, assumption (H1) is substantive: i may not know g and hence he may not know $u_i = v_i \circ g$.

The **complete information** assumption that we will consider later on (e.g., in Chapter 4) is much stronger; recall from the Chapter 1 that there is complete information if the rules of the game and players' preferences over (lotteries over) consequences are common knowledge. Although, as we explained above, the rules of the game may be only partially known, there are still many interactive situations where it is reasonable to assume that they are not only known, but indeed commonly known. Yet, assuming common knowledge of players' *preferences* is often far-fetched. Thus, complete information should be thought of as an idealization that simplifies the analysis of strategic thinking. Chapter 8 will introduce the formal tools necessary to model the absence of complete information.

3

Rationality and Dominance

In this chapter we analyze static games from the perspective of *decision theory*. We take as given the conjecture[1] of a player about opponents' behavior. This conjecture is, in general, probabilistic, i.e., it can be expressed by assigning subjective probabilities to the different action profiles of the other players. A player is said to be rational if he maximizes his *expected payoff* given his conjecture. The concept of *dominance* allows to characterize which actions are consistent with the rationality assumption without knowing a player's conjecture.

3.1 Probabilities and Expected Payoff

Payoff function u_i, which is a derived utility function, implicitly represents the preferences of i over different probability measures on A, which induce corresponding **lotteries**, that is, probability measures over consequences. The preferred lotteries are those that yield a higher expected payoff. We explain this in detail below.

We first introduce some preliminary definitions about sets and functions. For any set X, we let 2^X denote the **power set** of X, i.e., the collection of all subsets of X, including the empty set $\emptyset$ and X itself. If X is finite, then $|2^X| = 2^{|X|}$, where $|\cdot|$ denotes the "cardinality of $\cdot$." For any pair of nonempty sets X and Y, we let Y^X denote the set of

[1] In general, we use the word "conjecture" to refer to a player's belief about variables that affect his payoff and are beyond his control, such as the actions of other players in a static game.

functions with domain X and codomain Y. If X and Y are both finite, then $|Y^X| = |Y|^{|X|}$.

The set of probability measures over a generic domain X is denoted by $\Delta(X)$. If X is finite[2]

$$\Delta(X) = \left\{ \mu \in \mathbb{R}_+^X : \sum_{x \in X} \mu(x) = 1 \right\}.$$

The definition above is the simplest one for finite domains. But there is an alternative, equivalent, definition, that can be more easily generalized to infinite domains. Since sometimes we will consider probability measures on infinite domains, we present here also the alternative. First consider the definition above for a finite X, and fix $\mu \in \Delta(X)$. Every subset of a finite uncertainty space X is called "event." For each event $E \subseteq X$, μ determines the probability of E as follows:

$$\mu(E) = \sum_{x \in E} \mu(x).$$

Thus, the map $E \mapsto \mu(E)$ satisfies the following properties:

- (*normalization*) $\mu(X) = 1$,
- (*additivity*) for all $E, F \in 2^X$, $E \cap F = \emptyset \Rightarrow \mu(E \cup F) = \mu(E) + \mu(F)$.

Note that additivity implies that $\mu(\emptyset) = 0$, because $\emptyset \cap E = \emptyset$, hence $\mu(E) = \mu(E \cup \emptyset) = \mu(E) + \mu(\emptyset)$, which implies $\mu(\emptyset) = 0$. These are the characterizing properties of a probability measure, defined as a function $\mu : 2^X \to [0, 1]$. Therefore, the alternative definition of $\Delta(X)$ is

$$\begin{aligned} \Delta(X) = \Big\{ \mu \in [0,1]^{2^X} &: \mu(X) = 1, \\ &\forall E, F \in 2^X, E \cap F = \emptyset \Rightarrow \mu(E \cup F) = \mu(E) + \mu(F) \Big\}. \end{aligned}$$

When X is infinite, we typically restrict our attention to a subcollection $\mathcal{X} \subseteq 2^X$ of subsets of X, called **events** (e.g., the Borel sets, when X is a subset of a Euclidean space, or—more generally—a metric space), and we strengthen the additivity property as follows:

[2]Recall that $\mathbb{R}_+^X$ is the set of nonnegative real-valued functions defined over the domain X. If X is the finite set $\{x_1, \ldots, x_n\}$, $\mathbb{R}_+^X$ is isomorphic to $\mathbb{R}_+^n$, the positive orthant of the Euclidean space $\mathbb{R}^n$.

- (*countable additivity*) if $(E_k)_{k=1}^{\infty} \in \mathcal{X}^{\mathbb{N}}$ is a sequence of pairwise disjoint events and $\bigcup_{k=1}^{\infty} E_k \in \mathcal{X}$, then

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

Whenever we consider infinite sets, we take for granted that there is a well-defined collection of events $\mathcal{X}$ (such as the Borel subsets of a metric space) and $\Delta(X)$ is the set of functions $\mu : \mathcal{X} \to [0,1]$ that satisfy normalization and countable additivity.

Note, *X can be regarded as a subset of* $\Delta(X)$: a point $x \in X$ corresponds to the degenerate probability measure δ_x, called **Dirac measure** (or deterministic, or dogmatic belief), that satisfies $\delta_x(E) = 1$ if and only if $x \in E$. Hence, with a slight abuse of notation, we can write $X \subseteq \Delta(X)$ (more abuses of notation will follow).[3]

Definition 2. *Consider a probability measure $\mu \in \Delta(X)$, where X is a* finite *set. The subset of the elements $x \in X$ to which μ assigns a positive probability is called **support** of μ and is denoted by*

$$\text{supp}\mu = \{x \in X : \mu(x) > 0\}.$$

If X is a closed subset of $\mathbb{R}^m$, we define the support of a probability measure $\mu \in \Delta(X)$ as the intersection of all closed sets C such that $\mu(C) = 1$.[4]

Suppose that there is a function $f : X \to Y$, where Y is a finite set of outcomes, or consequences,[5] and an individual has a preference relation $\succsim$ over the set of lotteries $\Delta(Y)$ represented by a Von Neumann-Morgenstern utility function $v : Y \to \mathbb{R}$, that is, for all lotteries $\lambda, \lambda' \in \Delta(Y)$,

$$\lambda \succsim \lambda' \Leftrightarrow \sum_{y \in Y} \lambda(y)v(y) \geq \sum_{y \in Y} \lambda'(y)v(y).$$

[3] If X has at least two elements, the inclusion is strict.

[4] If $X \subset \mathbb{R}^m$ is finite, this definition is equivalent to the previous one.

[5] If X is infinite, then it has to be endowed with a sigma-algebra of events $\mathcal{X} \subseteq 2^X$ (see Appendix 3.4) and f has to be $\mathcal{X}$-measurable, that is, $f^{-1}(y) \in \mathcal{X}$ for each outcome y in the finite set Y.

Then we can derive a corresponding preference relation on X as follows. First we define the **pushforward** function $\hat{f} : \Delta(X) \to \Delta(Y)$ that determines the lottery induced, through f, by a probability measure on X:

$$\hat{f}(\mu)(y) = \mu\left(f^{-1}(y)\right) = \sum_{x \in f^{-1}(y)} \mu(x),$$

where, of course, the second equality holds when X is finite. With this, we obtain a preference relation on $\Delta(X)$: for all $\mu, \mu' \in \Delta(X)$,

$$\mu \succsim \mu' \Leftrightarrow \sum_{y \in Y} \hat{f}(\mu)(y)v(y) \geq \sum_{y \in Y} \hat{f}(\mu')(y)v(y).$$

Consider the particular case where $X = A$ is the set of action profiles in a game, let $f = g : A \to Y$ be the consequence function, and fix any player i. Then, for all $\mu, \mu' \in \Delta(A)$,

$$\mu \succsim_i \mu' \Leftrightarrow \sum_{y \in Y} \hat{g}(\mu)(y)v_i(y) \geq \sum_{y \in Y} \hat{g}(\mu')(y)v_i(y).$$

Since $\hat{g}(\mu)(y) = \mu\left(g^{-1}(y)\right)$ and we have defined the payoff function as $u_i = v_i \circ g$, we have

$$\sum_{y \in Y} \hat{g}(\mu)(y)v_i(y) = \sum_{y \in Y} \sum_{a \in g^{-1}(y)} \mu(a)v_i(g(a)) = \sum_{a \in A} \mu(a)u_i(a)$$

for each $\mu \in \Delta(A)$ (to understand the second equality, note that $\{g^{-1}(y)\}_{y \in Y}$ is a partition of A, and recall that $u_i(a) = v_i(g(a))$). Therefore

$$\mu \succsim_i \mu' \Leftrightarrow \sum_{a \in A} \mu(a)u_i(a) \geq \sum_{a \in A} \mu'(a)u_i(a).$$

In other words, probability measures on A are ranked according to the corresponding expected payoffs.

Probabilities and expected payoffs in compact-continuous games
The foregoing analysis can be extended to infinite games.[6] In particular, we focus on compact-continuous games. For any compact subset X of a Euclidean space, we let $\mathcal{X} = \mathcal{B}(X)$ be the smallest sigma-algebra

[6] We elaborate on this topic in Section 3.4.1 of the Appendix.

containing all the closed subsets of X; $\Delta(X)$ is the set of all the probability measures $\mu : \mathcal{B}(X) \to [0,1]$. Let $A = \bigtimes_{i\in I} A_i$ and Y be compact subsets of Euclidean spaces. If outcome function $g : A \to Y$ is continuous and each utility function $v_i : Y \to \mathbb{R}$ $(i \in I)$ is also continuous, then

$$\left\langle I, (A_i, u_i)_{i\in I}\right\rangle = \left\langle I, (A_i, v_i \circ g)_{i\in I}\right\rangle$$

is a compact-continuous game. Since g is continuous, then it is also measurable, that is, $E \in \mathcal{B}(Y)$ implies $g^{-1}(E) \in \mathcal{B}(A)$ for each $E \subseteq Y$. With this, the pushforward function, $\hat{g} : \Delta(A) \to \Delta(Y)$, is defined as follows: for all $\mu \in \Delta(A)$ and $E \in \mathcal{B}(Y)$,

$$\hat{g}(\mu)(E) = \mu\left(g^{-1}(E)\right);$$

more compactly, the pushforward map is described as follows

$$\mu \mapsto \hat{g}(\mu) = \mu \circ g^{-1}.$$

When the probability measures $\mu, \mu' \in \Delta(A)$ have finite or countable supports, the expected utility formulas are almost as above, with $y \in Y$ replaced by $y \in g(\mathrm{supp}\mu)$ and $y \in g(\mathrm{supp}\mu')$ respectively, and $a \in A$ replaced by $a \in \mathrm{supp}\mu$ and $a \in \mathrm{supp}\mu'$ respectively. Otherwise, expected values are expressed as integrals:

$$\mathbb{E}_\mu(u_i) = \int_A u_i(a)\,\mu(\mathrm{d}a) = \int_Y v_i(y)\,\hat{g}(\mu)(\mathrm{d}y) = \mathbb{E}_{\hat{g}(\mu)}(v_i). \qquad (3.1.1)$$

Example 2. To get some intuition about the generalized integrals in eq. (3.1.1), suppose that player i has only one opponent, whose action space is an interval $A_{-i} = \left[\underline{a}_{-i}, \overline{a}_{-i}\right] \subset \mathbb{R}$. Also suppose that player i plans to choose action $\hat{a}_i$, therefore his belief $\mu \in \Delta(A_i \times A_{-i})$ assigns probability one to the segment $\{\hat{a}_i\} \times \left[\underline{a}_{-i}, \overline{a}_{-i}\right]$ and μ corresponds to a probability measure μ^i on the interval $\left[\underline{a}_{-i}, \overline{a}_{-i}\right]$, the "conjecture" of i about $-i$. Such probability measures can equivalently be represented by cumulative distribution functions (cdf). So, let F_μ be the cdf representing μ^i; in particular,

$$F_\mu(a_{-i}) = \mu^i\left(\left[\underline{a}_{-i}, a_{-i}\right]\right) = \mu\left(\{\hat{a}_i\} \times \left[\underline{a}_{-i}, a_{-i}\right]\right).$$

Then $\mathbb{E}_\mu(u_i)$ can be expressed as a Riemann-Stieltjes integral:

$$\int_A u_i(a)\,\mu(\mathrm{d}a) = \int_{\underline{a}_{-i}}^{\overline{a}_{-i}} u_i(\hat{a}_i, a_{-i})\,\mathrm{d}F_\mu(a_{-i}).$$

If μ^i has a finite support, then F_μ is piecewise constant and $\mathrm{supp}\mu^i$ is the set of points where F_μ is discontinuous (from the left). In this case,

$$\int_{\underline{a}_{-i}}^{\overline{a}_{-i}} u_i\left(\hat{a}_i, a_{-i}\right) \mathrm{d}F_\mu\left(a_{-i}\right) = \sum_{a_{-i}\in\mathrm{supp}\mu^i} u_i\left(\hat{a}_i, a_{-i}\right) \mu^i\left(a_{-i}\right).$$

If instead $\mathrm{supp}\mu^i$ is an interval (or a union of intervals) and F_μ is differentiable with integrable derivative f_μ, then the differential formula $\mathrm{d}F_\mu\left(a_{-i}\right) = f_\mu\left(a_{-i}\right)\mathrm{d}a_{-i}$ applies, and

$$\int_{\underline{a}_{-i}}^{\overline{a}_{-i}} u_i\left(\hat{a}_i, a_{-i}\right) \mathrm{d}F_\mu\left(a_{-i}\right) = \int_{\underline{a}_{-i}}^{\overline{a}_{-i}} u_i\left(\hat{a}_i, a_{-i}\right) f_\mu\left(a_{-i}\right) \mathrm{d}a_{-i}.$$

▲

3.2 Conjectures

The reason why one needs to introduce preferences over lotteries and expected payoffs is that player i cannot observe other players' actions (a_{-i}) before making his own choice. Hence, he needs to form a conjecture about such actions. If i were certain of the other players' choices, then one could represent i's conjecture simply with an action profile $a_{-i} \in A_{-i}$. However, in general, i might be uncertain about other players' actions and assign a strictly positive (subjective) probability to several profiles a_{-i}, a'_{-i}, etc.

Definition 3. *A **conjecture** of player i is a (subjective) probability measure $\mu^i \in \Delta(A_{-i})$. A* **deterministic conjecture** *is a probability measure $\mu^i \in \Delta(A_{-i})$ that assigns probability one to a particular action profile, that is, $\mu^i = \delta_{a_{-i}}$ for some $a_{-i} \in A_{-i}$.*

Note, we call "conjecture" a (probabilistic) belief about the behavior of other players, while we use the term "belief" to refer to a more general type of uncertainty.

Remark 1. *The set of deterministic conjectures of player i essentially coincides with the set of other players' action profiles, so that we can write $A_{-i} \subseteq \Delta(A_{-i})$.*

One of the most interesting aspect of game theory consists in determining players' conjectures, or, at least, in narrowing down the set of possible conjectures, combining some general assumptions about players' rationality and beliefs with specific assumptions about the given game G. However, in this chapter we will not try to "explain" players' conjectures. This is left for the following chapters.

For any given conjecture μ^i, the choice of a particular action $\bar{a}_i$ corresponds to the choice of the measure on A that assigns probability $\mu^i(a_{-i})$ to each action profile $(\bar{a}_i, a_{-i})$ $(a_{-i} \in A_{-i})$ and probability zero to the profiles (a_i, a_{-i}) with $a_i \neq \bar{a}_i$. Therefore, if a player i has conjecture μ^i and chooses (or plans to choose) action $\bar{a}_i$, the corresponding (subjective) expected payoff is

$$u_i(\bar{a}_i, \mu^i) = \mathbb{E}_{\mu^i}\left(u_i\left(\bar{a}_i, \cdot\right)\right),$$

where

$$\mathbb{E}_{\mu^i}\left(u_i\left(\bar{a}_i, \cdot\right)\right) = \sum_{a_{-i} \in \text{supp}\mu^i} \mu^i(a_{-i}) u_i(\bar{a}_i, a_{-i})$$

if μ^i has a finite support.

There are many different ways to represent graphically actions, conjectures, and preferences when the number of available actions is small. Let us focus our attention on the instructive case where player i's *opponent* $(-i)$ has *only two actions*, denoted ℓ ("left") and r ("right"). Consider, for instance, the function u_i represented by the payoff matrix in Figure 3.1 for player i.

$i \backslash -i$	ℓ	r
a	4	1
b	1	4
c	2	2
e	4	0
f	1	1

Figure 3.1: Matrix 1.

Given that $-i$ has only two actions, we can represent the conjecture

μ^i of i about $-i$ with a single number: $\mu^i(r)$, the subjective probability that i assigns to r. Each action a_i corresponds to a function that assigns to each value of $\mu^i(r)$ the expected payoff of a_i. Since the expected payoff function $(1-\mu^i(r))u_i(a_i,\ell)+\mu^i(r)u_i(a_i,r)$ is linear in the probabilities,[7] every action is represented by a line, such as aa, bb, cc etc., in Figure 3.2.

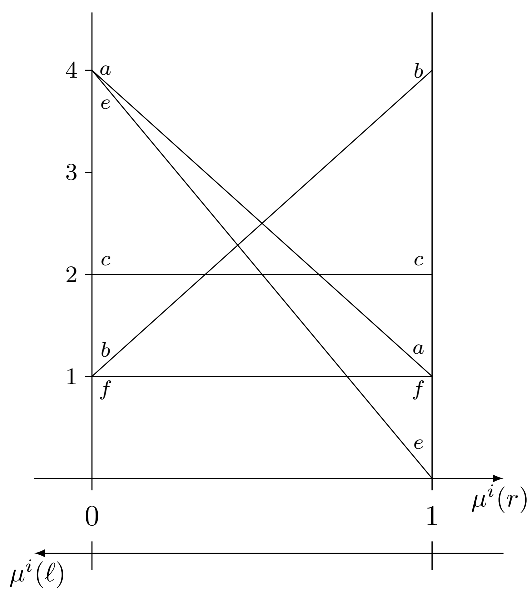


Figure 3.2: Expected payoff as a function of beliefs.

From Figure 3.2 it is apparent that only actions a, b, and e are "justifiable" by some conjecture. In particular, if $\mu^i(r)=0$, then a and e are both optimal; if $0<\mu^i(r)<\frac{1}{2}$, then a is the only optimal action (the line aa yields the maximum expected payoff). If $\mu^i(r)>\frac{1}{2}$, then b is the only optimal action (the line bb yields the maximum expected payoff). If $\mu^i(r)=\frac{1}{2}$, then there are two optimal actions: a and b.

[7]More precisely, the expected payoff function $u_i(a_i,\mu^i)$ is affine in μ^i, that is,

$$u_i(a_i, x\mu^i+(1-x)\nu^i)=xu_i(a_i,\mu^i)+(1-x)u_i(a_i,\nu^i)$$

for all $a_i\in A_i$, $\mu^i,\nu^i\in\Delta(A_{-i})$, and $x\in[0,1]$.

3.2.1 Mixed Actions

In principle, instead of choosing directly a certain action, a player could delegate his decision to a randomizing device, such as spinning a roulette wheel, or tossing a coin. In other words, a player could simply choose the *probability* of playing any given action.

Definition 4. *A random choice by player i, also called **mixed action**, is a probability measure $\alpha_i \in \Delta(A_i)$. An action $a_i \in A_i$ is also called **pure action**.*

Remark 2. *The set of pure actions can be regarded as a subset of the set of mixed actions, i.e., $A_i \subseteq \Delta(A_i)$ (cf. Remark 1).*

It is assumed that (according to i's beliefs) the random draw of an action of i is stochastically independent of the other players' actions. For example, the following situation is *excluded*: i chooses his action according to the (random) weather and he thinks his opponents are doing the same, so that there is correlation between a_i and a_{-i} even though there is no causal link between a_i and a_{-i} (this type of correlation will be discussed in section 6.2 of Chapter 6). More importantly, player i knows that moves are simultaneous and therefore by changing his actions he cannot cause any change in the probability distribution of opponents' actions. Hence, if player i has conjecture μ^i and chooses the mixed action α_i, where both have finite support, the subjective probability of each possible action profile (a_i, a_{-i}) is $\alpha_i(a_i)\mu^i(a_{-i})$ and i's expected payoff is

$$\begin{aligned} u_i(\alpha_i, \mu^i) &= \mathbb{E}_{\alpha_i \times \mu^i}(u_i) = \sum_{a_i \in \text{supp}\alpha_i} \sum_{a_{-i} \in \text{supp}\mu^i} \alpha_i(a_i)\mu^i(a_{-i})u_i(a_i, a_{-i}) \\ &= \sum_{a_i \in \text{supp}\alpha_i} \alpha_i(a_i)u_i(a_i, \mu^i). \end{aligned}$$

If the opponent has only two feasible actions, it is possible to use a graph to represent the lotteries corresponding to pure and mixed actions. For each action a_i, we consider a corresponding point in the Cartesian plane with coordinates given by the utilities that i obtains for each of the two actions of the opponent. If the actions of the opponents are ℓ and r, we denote such coordinates $x = u_i(\cdot, \ell)$ and $y = u_i(\cdot, r)$. Any *pure* action a_i corresponds to the vector $(x, y) = (u_i(a_i, \ell), u_i(a_i, r))$ (a row in the payoff matrix of

i). The same holds for the mixed actions: α_i corresponds to the vector $(x, y) = (u_i(\alpha_i, \ell), u_i(\alpha_i, r))$. The set of points (vectors) corresponding to the mixed actions is simply the *convex hull* of the points corresponding to the pure actions.[8] Figure 3.3 represents such a set for the matrix in Figure 3.1.

How can conjectures be represented in such a figure? It is quite simple. Every conjecture induces a preference relation on the space of payoff pairs. Such preferences can be represented through a map of iso-expected payoff curves (or indifference curves). Let (x, y) be the generic vector of (expected) payoffs corresponding to ℓ and r respectively. The expected payoff induced by conjecture μ^i is $\mu^i(\ell)x + \mu^i(r)y$. Therefore, the indifference curves are straight lines defined by the equation $y = \frac{\overline{u}}{\mu^i(r)} - \frac{\mu^i(\ell)}{\mu^i(r)}x$, where $\overline{u}$ denotes the constant expected payoff. Every conjecture μ^i corresponds to a set of parallel lines with negative slope (or, in the extreme cases, horizontal or vertical slope) determined by the orthogonal vector $(\mu^i(\ell), \mu^i(r))$. The direction of increasing expected payoff is given precisely by such orthogonal vector. The optimal actions (pure or mixed) can be graphically determined considering, for any conjecture μ^i, the highest line of iso-expected payoff among those that touch the set of *feasible* payoff vectors (the shaded area in Figure 3.3).

Allowing for mixed actions, the set of feasible (expected) payoffs vectors is a convex polyhedron (as in Figure 3.3). To see this, note that if α_i and β_i are mixed actions, then also $p \cdot \alpha_i + (1 - p) \cdot \beta_i$ (with $0 \leq p \leq 1$) is a mixed action, where $p \cdot \alpha_i + (1 - p) \cdot \beta_i$ is the function that assigns to each pure action a_i the weight $p\alpha_i(a_i) + (1 - p)\beta_i(a_i)$. Thus, all the convex combinations of feasible payoff vectors are feasible payoff vectors, once we allow for randomization. This geometrical intuition can be extended to all *finite* games. In general, each mixed action α_i corresponds to a vector of expected payoffs $(u_i(\alpha_i, a_{-i}))_{a_{-i} \in A_{-i}}$. Let

$$\mathbf{U} = \left\{ \mathbf{u} \in \mathbb{R}^{A_{-i}} : \exists \alpha_i \in \Delta(A_i), \mathbf{u} = (u_i(\alpha_i, a_{-i}))_{a_{-i} \in A_{-i}} \right\}$$

be the set of feasible payoff vectors when randomization is allowed; then $\mathbf{U}$ is a compact and convex polyhedron in $\mathbb{R}^{A_{-i}}$ with extreme points contained

[8] The **convex hull** of a set of points $X \subseteq \mathbb{R}^k$ is the smallest convex set containing X, that is, the intersection of all the convex sets containing X.

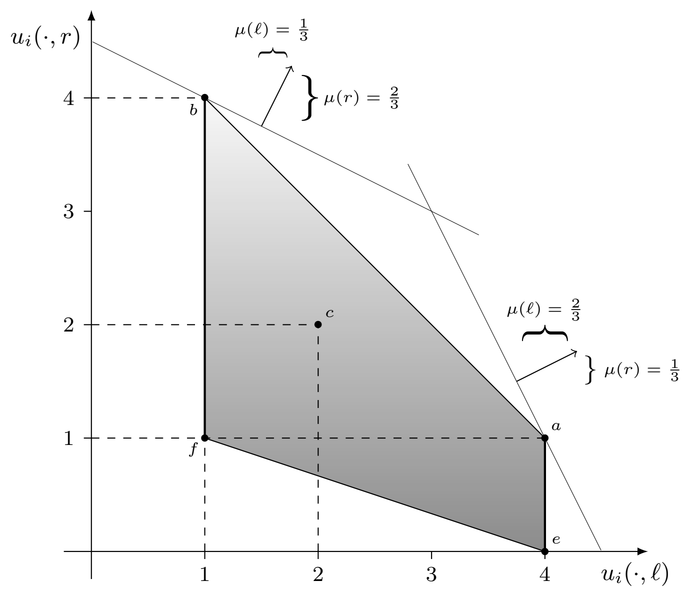


Figure 3.3: State-contingent representation of expected payoff.

in the set

$$\left\{\mathbf{u} \in \mathbb{R}^{A_{-i}} : \exists a_i \in A_i, \mathbf{u} = \left(u_i(a_i, a_{-i})\right)_{a_{-i} \in A_{-i}}\right\}$$

of payoff vectors induced by pure actions. For example, in Figure 3.3 $\mathbf{U}$ is the shaded area; the set of extreme points is

$$\left\{\left(u_i(a,\ell), u_i(a,r)\right), \left(u_i(b,\ell), u_i(b,r)\right), \left(u_i(e,\ell), u_i(e,r)\right), \left(u_i(f,\ell), u_i(f,r)\right)\right\},$$

which is a subset of $\left\{\left(u_i(a_i,\ell), u_i(a_i,r)\right)\right\}_{a_i \in A_i}$ because the vector

$$(u_i(c,\ell), u_i(c,r)) = (2,2) = \frac{1}{3}(1,1) + \frac{1}{3}(4,1) + \frac{1}{3}(1,4)$$

$$= \frac{1}{3}\left(u_i(f,\ell), u_i(f,r)\right) + \frac{1}{3}\left(u_i(a,\ell), u_i(a,r)\right) + \frac{1}{3}\left(u_i(b,\ell), u_i(b,r)\right)$$

lies in the interior of $\mathbf{U}$.

However, the idea that players use coins or roulette wheels to randomize their choices may seem weird and unrealistic. Furthermore, as illustrated by Figure 3.3, for any conjecture μ^i and any mixed action α_i, there is always a pure action a_i that yields the same or a higher expected payoff than α_i (check all possible slopes of the iso-expected payoff curves and verify how the set of optimal points looks like).[9] Hence, a player cannot be *strictly* better off by choosing a mixed action rather than a pure action.

The point of view we adopt in this textbook is that expected-utility maximizing players never randomize (although their choices might depend on extrinsic, payoff-irrelevant signals, as in Chapter 6, Section 6.2). Nonetheless, it will be shown that in order to assess the justifiability of a given pure action it is analytically convenient to introduce mixed actions. (In Chapter 5, we discuss interpretations of mixed actions that do not involve randomization.)

3.3 Best Replies and Undominated Actions

Definition 5. *A (mixed) action α_i^* is a **best reply** to conjecture μ^i if*

$$\forall \alpha_i \in \Delta(A_i),\ u_i(\alpha_i^*, \mu^i) \geq u_i(\alpha_i, \mu^i),$$

that is

$$\alpha_i^* \in \arg \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i).$$

The set of pure *actions that are best replies to conjecture μ^i is denoted by*[10]

$$r_i(\mu^i) = A_i \cap \arg \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i).$$

*The correspondence $r_i : \Delta(A_{-i}) \rightrightarrows A_i$ is called **best reply correspondence**.*[11] *An action a_i is called **justifiable** if there exists a conjecture $\mu^i \in \Delta(A_{-i})$ such that $a_i \in r_i(\mu^i)$.*

Note that, even if A_i is finite, one cannot conclude without proof that the set of pure best replies to a conjecture μ^i is nonempty, i.e., that $r_i(\mu^i) \neq \emptyset$. In principle, it could happen that in order to maximize

[9] A formal proof of this result will be provided in the next section.

[10] Recall: we regard A_i is a subset of $\Delta(A_i)$ (see Remark 2).

[11] A *correspondence* $\varphi : X \rightrightarrows Y$ is a multi-function that assigns to every element $x \in X$ a set of elements $\varphi(x) \subseteq Y$. A correspondence $\varphi : X \rightrightarrows Y$ can be equivalently expressed as a function with domain X and codomain 2^Y, the power set of Y.

expected payoff it is necessary to use mixed actions that assign positive probability to more than one pure action. However, we will show that $r_i(\mu^i) \neq \emptyset$ for every μ^i provided that A_i is finite (or, more generally, that A_i is compact and u_i continuous in a_i, two properties that trivially hold if A_i is finite).

The following result shows, as anticipated, that a rational player does not need to use mixed actions. Therefore, it can be assumed without loss of generality that his choice is restricted to the set of pure actions.

Lemma 1. *Consider a finite or compact-continuous game. Fix arbitrarily a player $i \in I$, a conjecture $\mu^i \in \Delta(A_{-i})$ and a mixed action α_i^*; then α_i^* is a best reply to μ^i if and only if every pure action in the support of α_i^* is a best reply to μ^i, that is*

$$\alpha_i^* \in \arg\max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i) \ \Leftrightarrow\ \mathrm{supp}\alpha_i^* \subseteq r_i(\mu^i).$$

Proof. We prove the result only for finite games. In the finite case the proof is quite simple. But since this is the first proof of this textbook, we will go over it rather slowly.

(Only if) We first show that if $\mathrm{supp}\alpha_i^*$ is not included in $r_i(\mu^i)$, then α_i^* is not a best reply to μ^i (this is the contrapositive of the "only if" implication).[12] Let a_i^* be a pure action such that $\alpha_i^*(a_i^*) > 0$ and assume that, for some α_i, $u_i(\alpha_i, \mu^i) > u_i(a_i^*, \mu^i)$, so that $a_i^* \in \mathrm{supp}\alpha_i^* \backslash r_i(\mu^i)$.[13] Since $u_i(\alpha_i, \mu^i)$ is a weighted average of the values $\left(u_i(a_i, \mu^i)\right)_{a_i \in A_i}$, there must be a pure action a_i' such that $u_i(a_i', \mu^i) > u_i(a_i^*, \mu^i)$. But then we can construct a mixed action α_i' that satisfies $u_i(\alpha_i', \mu^i) > u_i(\alpha_i^*, \mu^i)$ by "shifting probability weight" from a_i^* to a_i', which is possible because $\alpha_i^*(a_i^*) > 0$. Specifically, for every $a_i \in A_i$, let

$$\alpha_i'(a_i) = \begin{cases} 0, & \text{if } a_i = a_i^*, \\ \alpha_i^*(a_i') + \alpha_i^*(a_i^*), & \text{if } a_i = a_i', \\ \alpha_i^*(a_i), & \text{if } a_i \neq a_i^*, a_i'. \end{cases}$$

[12] The **contrapositive** of $p \Rightarrow q$ is $\neg q \Rightarrow \neg p$, where p and q are sentences and $\neg$ means "not." An implication $p \Rightarrow q$ holds if and only if its contrapositive $\neg q \Rightarrow \neg p$ holds. In this case, p says that $\alpha_i^* \in \arg\max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i)$, q says that $\mathrm{supp}\alpha_i^* \subseteq r_i(\mu^i)$, and $\neg q$ says that there is some $a_i^* \in \mathrm{supp}\alpha_i^*$ such that $a_i^* \notin r_i(\mu^i)$.

[13] Recall that $X \backslash Y$ denotes the set of elements of X that do not belong to Y: $X \backslash Y = \{x \in X : x \notin Y\}$.

α'_i is a mixed action since $\sum_{a_i \in A_i} \alpha'_i(a_i) = \sum_{a_i \in A_i} \alpha^*_i(a_i) = 1$. Moreover, it can be easily checked that

$$u_i(\alpha'_i, \mu^i) - u_i(\alpha^*_i, \mu^i) = \alpha^*_i(a^*_i)[u_i(a'_i, \mu^i) - u_i(a^*_i, \mu^i)] > 0,$$

where the inequality holds by assumption. Thus α^*_i is not a best reply to μ^i.

(If) Next we show that if each pure action in the support of α^*_i is a best reply, then α^*_i is also a best reply. It is convenient to introduce the following notation:

$$\hat{u}_i(\mu^i) = \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i).$$

By definition of $r_i(\mu^i)$,

$$\forall a_i \in r_i(\mu^i), \ \hat{u}_i(\mu^i) = u_i(a_i, \mu^i) \tag{3.3.1}$$
$$\forall a_i \in A_i, \ \hat{u}_i(\mu^i) \geq u_i(a_i, \mu^i). \tag{3.3.2}$$

Assume that $\text{supp}\alpha^*_i \subseteq r_i(\mu^i)$. Then, for every $\alpha_i \in \Delta(A_i)$

$$\begin{aligned} u_i(\alpha^*_i, \mu^i) &= \sum_{a_i \in \text{supp}\alpha^*_i} \alpha^*_i(a_i) u_i(a_i, \mu^i) = \sum_{a_i \in r_i(\mu^i)} \alpha^*_i(a_i) u_i(a_i, \mu^i) = \hat{u}_i(\mu^i) \\ &= \hat{u}_i(\mu^i) \sum_{a_i \in A_i} \alpha_i(a_i) = \sum_{a_i \in A_i} \alpha_i(a_i) \hat{u}_i(\mu^i) \geq \sum_{a_i \in A_i} \alpha_i(a_i) u_i(a_i, \mu^i). \end{aligned}$$

The first equality holds by definition, the second follows from $\text{supp}\alpha^*_i \subseteq r_i(\mu^i)$, the third holds by eq. (3.3.1), the fourth and fifth are obvious and the inequality follows from (3.3.2). ■

In the matrix of Figure 3.1, for example, any mixed action that assigns positive probability only to a and/or b is a best reply to the uniform conjecture $\mu^i = \frac{1}{2}\delta_\ell + \frac{1}{2}\delta_r$ that assigns probability $\frac{1}{2}$ to ℓ and r.[14] Clearly, the set of pure best replies to the uniform conjecture is $r_i\left(\frac{1}{2}\delta_\ell + \frac{1}{2}\delta_r\right) = \{a, b\}$.

Note that if at least one pure action is a best reply among all pure and mixed actions, then the maximum that can be attained by constraining

[14] Recall that δ_x denotes the Dirac probability measure that assigns probability 1 to point x.

the choice to pure actions is necessarily equal to what could be attained choosing among (pure and) mixed actions, i.e.,

$$r_i(\mu^i) \neq \emptyset \ \Rightarrow \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i) = \max_{a_i \in A_i} u_i(a_i, \mu^i).$$

This observation along with Lemma 1 yields the following:

Corollary 1. *Consider a finite or compact-continuous game. Fix arbitrarily a player $i \in I$ and a conjecture $\mu^i \in \Delta(A_{-i})$; then*

$$r_i(\mu^i) = \arg\max_{a_i \in A_i} u_i(a_i, \mu^i).$$

Hence, it is not necessary to use mixed actions to maximize expected payoff. Moreover, the best reply correspondence is non-empty-valued, that is, $r_i(\mu^i) \neq \emptyset$ for every $\mu^i \in \Delta(A_{-i})$.

Proof. For every conjecture μ^i, the expected payoff function $u_i(\cdot, \mu^i) : \Delta(A_i) \to \mathbb{R}$ is continuous in α_i and the domain $\Delta(A_i)$ is compact.[15] Hence, the function has at least one maximizer α_i^*. By Lemma 1, $\text{supp}\alpha_i^* \subseteq r_i(\mu^i)$, so that $r_i(\mu^i) \neq \emptyset$. As we have seen above, this implies that

$$\max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \mu^i) = \max_{a_i \in A_i} u_i(a_i, \mu^i)$$

and therefore $r_i(\mu^i) = \arg\max_{a_i \in A_i} u_i(a_i, \mu^i)$. ■

Recall that, for a given function $f : X \to Y$ and a given subset $C \subseteq X$,

$$f(C) = \{y \in Y : \exists x \in C, y = f(x)\}$$

denotes the set of images of elements of C. Analogously, for a given correspondence $\psi : X \rightrightarrows Y$,

$$\psi(C) = \{y \in Y : \exists x \in C, y \in \psi(x)\}$$

denotes the set of elements y that belong to the image $\psi(x)$ of some point $x \in C$. In particular, we use this notation for the best reply

[15] For the infinite case, see Appendix 3.4. If A_i is finite, $\Delta(A_i)$ is a compact subset of a Euclidean space.

correspondence. For example, $r_i\left(\Delta(A_{-i})\right)$ is the set of justifiable actions of player i (see Definition 5).

A question should spontaneously arise at this point: can we characterize the set of justifiable actions with no reference to conjectures and expected payoff maximization? In other words, can we conclude that an action will not be chosen by a rational player without checking directly that it is not a best reply to any conjecture? The answer comes from the concept of dominance. But, even if we are only concerned with the justifiability of pure actions, we will have to compare them with mixed actions. To anticipate: a (pure) action is justifiable if and only it is not dominated by any mixed action.

Definition 6. *A mixed action α_i* ***dominates*** *a (pure or) mixed action β_i if it yields a* strictly higher *expected payoff irrespective of the choices of the other players:*

$$\forall a_{-i} \in A_{-i},\ u_i(\alpha_i, a_{-i}) > u_i(\beta_i, a_{-i}).$$

The set of pure *actions of agent i that are* not dominated *by any mixed action is denoted by ND_i.*

The set of undominated actions can be formally written as follows:

$$\begin{aligned} ND_i &= \{a_i \in A_i : \forall \alpha_i \in \Delta(A_i), \exists a_{-i} \in A_{-i}, u_i(\alpha_i, a_{-i}) \leq u_i(a_i, a_{-i})\} \\ &= A_i \backslash \{a_i \in A_i : \exists \alpha_i \in \Delta(A_i), \forall a_{-i} \in A_{-i}, u_i(\alpha_i, a_{-i}) > u_i(a_i, a_{-i})\}. \end{aligned}$$

The proof of the following statement is left as an exercise for the reader.

Remark 3. *If a (pure) action a_i is dominated by a mixed action α_i then, for every conjecture $\mu^i \in \Delta(A_{-i})$, $u_i(a_i, \mu^i) < u_i(\alpha_i, \mu^i)$. Hence, a dominated action is not justifiable; equivalently, a justifiable action cannot be dominated.*

In the matrix of Figure 3.1, for instance, action f is dominated by c, which in turn is dominated by the mixed action $\alpha_i = \frac{1}{2}\delta_a + \frac{1}{2}\delta_b$ that assigns probability $\frac{1}{2}$ to a and b. Therefore, $ND_i \subseteq \{a, b, e\}$. Given that a, b, and e are best replies to at least one conjecture, we have that $\{a, b, e\} \subseteq ND_i$. Hence, $ND_i = \{a, b, e\}$.

The following lemma states that the converse of Remark 3 holds. Therefore, it provides a complete answer to our previous question about the characterization of the set of actions that a rational player would never choose.

Lemma 2. *(Wald [69], Pearce [54]) Fix arbitrarily a player $i \in I$ and an action $a_i^* \in A_i$ in a finite or compact-continuous game. There exists a conjecture $\mu^i \in \Delta(A_{-i})$ such that a_i^* is a best reply to μ^i* if and only if *a_i^* is not dominated by any (pure or) mixed action. In other words, the set of undominated (pure) actions and the set of justifiable actions coincide:*

$$ND_i = r_i\left(\Delta(A_{-i})\right).$$

Here we provide a proof for the *finite* case, which relies on an important result about linear algebra known as Farkas' lemma.[16] Intuitively, the payoff vector $\mathbf{u}^* = \left(u_i\left(a_i^*, a_{-i}\right)\right)_{a_{-i} \in A_{-i}}$ corresponding to an undominated action a_i^* must be on the "efficient frontier" of the convex set $\mathbf{U}$ of feasible payoff vectors (the North-East part of the boundary of the shaded area in Figure 3.3). Therefore we can find an hyperplane H (a line in Figure 3.3) going through $\mathbf{u}^*$ and separating space $\mathbb{R}^{A_{-i}}$ in two half-spaces (half-planes in Figure 3.3), so that one of them contains $\mathbf{U}$. Hyperplane H can be written as the set of vectors $\mathbf{y} \in \mathbb{R}^{A_{-i}}$ such that $\mathbf{y} \cdot \mathbf{x}^* = \mathbf{u}^* \cdot \mathbf{x}^*$, where $\mathbf{0} \neq \mathbf{x}^* \in \mathbb{R}^{A_{-i}}$. One can use Farkas' lemma to ensure that, since $\mathbf{y}^*$ is on the "efficient" part of the boundary, then $\mathbf{x}^*$ is a nonnegative vector, and therefore it can be normalized so that its elements sum to 1 and it becomes a conjecture $\mu^i \in \Delta\left(A_{-i}\right)$. By construction, this conjecture justifies action a_i^* as a best reply.

Lemma 3 (Farkas' Lemma)**.** *Let M be a $n \times m$ matrix and $\mathbf{c} \in \mathbb{R}^n$ be an n-dimensional vector. Then exactly one of the following statements is true:*
(1) there exists a vector $\mathbf{x} \in \mathbb{R}^m$ such that $M\mathbf{x} \geq \mathbf{c}$;
(2) there exists a vector $\mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{y} \geq \mathbf{0}$, $\mathbf{y}^{\mathrm{T}} M = \mathbf{0}^{\mathrm{T}}$ and $\mathbf{y}^{\mathrm{T}}\mathbf{c} > 0$.

Proof of Lemma 2. We use Farkas' Lemma to prove the non-obvious part of Lemma 2: if an action is undominated then it is justifiable, or—by contraposition—if it is not justifiable then it is dominated. Fix i and a_i^* arbitrarily. Since the game is finite, let $k = |A_i|$ and $m = |A_{-i}|$, and label

[16] In the statement of Farkas' Lemma and in the related proofs, we use boldface symbols to denote vectors (i.e., functions from $\{1, ..., n\}$ to $\mathbb{R}$). When vectors are represented as matrices, the default intrepretation is that they are columns. Thus, the inner product $\mathbf{y} \cdot \mathbf{x} = \sum_s y_s x_s$ in matrix algebra notation is written as $\mathbf{y}^{\mathrm{T}}\mathbf{x}$, where $\mathbf{y}$ and $\mathbf{x}$ are column vectors and $\mathbf{y}^{\mathrm{T}}$ is the row vector obtained from $\mathbf{y}$ by transposition.

elements in A_i as $\{1, 2, ..., k\}$ and elements in A_{-i} as $\{1, 2, ..., m\}$.[17] Then, construct a $k \times m$ matrix U in which the (w, z)-th coordinate is given by

$$U_{w,z} = u_i\left(a_i^*, z\right) - u_i\left(w, z\right).$$

With this, $\Delta\left(A_{-i}\right)$ is a subset of $\mathbb{R}^m$ and a_i^* is justifiable if and only if we can find $\mu^i \in \Delta\left(A_{-i}\right)$ such that $U\mu^i \geq \mathbf{0}$. This condition can be rewritten as follows: a_i^* is justifiable if and only if we can find $\mathbf{x} \in \mathbb{R}^m$ satisfying inequality:

$$M\mathbf{x} \geq \mathbf{c}, \tag{3.3.3}$$

where M is a $(k + m + 1) \times m$ matrix and $\mathbf{c}$ is a $(k + m + 1)$-dimensional vector defined as follows:

$$M = \begin{bmatrix} U \\ I_m \\ \mathbf{1}_m^{\mathrm{T}} \end{bmatrix}, \; \mathbf{c} = \begin{bmatrix} \mathbf{0}_k \\ \mathbf{0}_m \\ 1 \end{bmatrix}$$

(I_m denotes the m-dimensional identity matrix, that is an $m \times m$ matrix having 1s along the main diagonal and 0s everywhere else; $\mathbf{1}_\ell$ and $\mathbf{0}_\ell$ ($\ell = k$, or $\ell = m$) denote the ℓ-dimensional vectors having respectively 1s and 0s everywhere; from now on, the dimensionality indexes will be omitted). To see this, note that matrix inequality (3.3.3) can be written as the system of inequalities

$$\begin{cases} U\mathbf{x} \geq \mathbf{0} \\ x_1 \geq 0, ..., x_m \geq 0 \\ \sum_{z=1}^m x_z \geq 1 \end{cases};$$

if $\sum_{z=1}^m x_z > 1$ instead of $\sum_{z=1}^m x_z = 1$, then $\mathbf{x} \notin \Delta\left(A_{-i}\right)$; but this is immaterial because we can normalize by substituting $\mathbf{x} = (x_1, ..., x_m)$ with $\mathbf{x}' = \left(\frac{x_1}{\sum_{z=1}^m x_z}, ..., \frac{x_m}{\sum_{z=1}^m x_z}\right) \in \Delta\left(A_{-i}\right)$ and satisfy inequality (3.3.3), since $U\mathbf{x} \geq 0$ if and only if $U\mathbf{x}' \geq 0$.

Thus, if a_i^* is *not* justifiable, inequality (3.3.3) does *not* hold and Farkas' lemma implies that we can find $\mathbf{y} \in \mathbb{R}^{k+m+1}$ such that

$$\begin{cases} \mathbf{y} \geq \mathbf{0} \\ \mathbf{y}^{\mathrm{T}}\mathbf{c} > 0 \\ \mathbf{y}^{\mathrm{T}}M = \mathbf{0}^{\mathrm{T}} \end{cases}.$$

[17] Recall that $|X|$ denotes the cardinality of set X.

By construction, $\mathbf{y}^{\mathrm{T}}\mathbf{c} = y_{k+m+1}$; thus the second condition can be rewritten as $y_{k+m+1} > 0$. Furthermore, $\mathbf{y}^{\mathrm{T}}M = \mathbf{0}^{\mathrm{T}}$ implies that

$$\sum_{w=1}^{k} y_w \left[u_i\left(a_i^*, z\right) - u_i\left(w, z\right)\right] = -\left(y_{k+z} + y_{k+m+1}\right) < 0$$

for every $z \in A_{-i} = \{1, 2, ..., m\}$, where the inequality holds because $y_{k+m+1} > 0$ and $y_{k+z} \geq 0$. Since the left hand side is non-zero and the vector $(y_1, ..., y_k)$ is nonnegative, we must have $y_w > 0$ for some $w \in A_i$. Then, we can construct a probability vector

$$\bar{\mathbf{y}} = \left(\frac{y_1}{\sum_{w=1}^{k} y_w}, ..., \frac{y_k}{\sum_{w=1}^{k} y_w}\right) \in \Delta(A_i) \subseteq \mathbb{R}^k$$

such that $u_i\left(a_i^*, z\right) < \sum_{w=1}^{k} \bar{y}_w u_i\left(w, z\right)$ for every $z \in A_{-i}$. We conclude that a_i^* is dominated by mixed action $\bar{\mathbf{y}}$. ■

We can thus conclude that *a rational player always chooses undominated actions* and that *each undominated action is justifiable as a best reply to some conjecture.*

In some interesting situations of strategic interaction an action dominates all others. In those cases, a rational player should choose that action. This consideration motivates the following:

Definition 7. *An action a_i^* is* ***dominant*** *if it dominates every other action, i.e., if*

$$\forall a_i \in A_i \backslash \left\{a_i^*\right\}, \forall a_{-i} \in A_{-i},\ u_i(a_i^*, a_{-i}) > u_i(a_i, a_{-i}).$$

You should try to prove the following statement as an exercise:

Remark 4. *Fix an action a_i^* arbitrarily. The following conditions are equivalent:*
(0) action a_i^* is dominant;
(1) action a_i^* dominates every mixed *action $\alpha_i \neq a_i^*$;*
(2) action a_i^* is the unique best reply to every *conjecture.*

The following example illustrates the notion of dominant action. The example shows that, assuming that players are motivated by their material self-interest, individual rationality may lead to socially undesirable outcomes.

Example 3. (Linear Public Good Game). In a community composed of n individuals it is possible to produce a quantity g of a public good using an input x according to the production function $y = kx$. Both y and x are measured in monetary units (say in Euros). To make the example interesting, assume that the productivity parameter k satisfies $1 > k > \frac{1}{n}$. A generic individual i can freely choose how many Euros to contribute to the community for the production of the public good. The community members cannot sign a binding agreement on such contributions because no authority with coercive power can enforce the agreement. Let W_i be player i's wealth. The game can be represented as follows: $A_i = [0, W_i]$, $a_i \in A_i$ is the contribution of i; consequences are allocations of the public and private good $(y, W_1 - a_1, \ldots, W_n - a_n)$; agents are selfish,[18] hence their utility function can be written as

$$v_i(y, W_1 - a_1, \ldots, W_n - a_n) = y - a_i,$$

which yields the payoff function

$$u_i(a_1, \ldots, a_n) = k \sum_{j=1}^{n} a_j - a_i.$$

It can be easily checked that $a_i^* = 0$ is dominant for any player i: just rewrite u_i as

$$u_i(a_i, a_{-i}) = k \sum_{j \neq i} a_j - (1 - k) a_i$$

and recall that $k < 1$. The profile of dominant actions $(0, \ldots, 0)$ is Pareto-dominated by any symmetric profile of positive contributions $(\varepsilon, \ldots, \varepsilon)$ (with $\varepsilon > 0$). Indeed, $u_i(0, \ldots, 0) = 0 < (nk - 1)\varepsilon = u_i(\varepsilon, \ldots, \varepsilon)$, where the inequality holds because $k > \frac{1}{n}$. Let $S(a) = \sum_{i=1}^{n} u_i(a)$ be the social surplus; the surplus maximizing profile is $\hat{a}_i = W_i$ for each i. ▲

An action could be a best reply only to conjectures that assign zero probability to some action profiles of the other players. For instance, action e in Figure 3.1 is justifiable as a best reply only if i is certain that $-i$ does not choose r. Let us say that a player i is **cautious** if his conjecture does not rule out any $a_{-i} \in A_{-i}$. Formally, let

$$\Delta^o(A_{-i}) = \left\{ \mu^i \in \Delta(A_{-i}) : \mathrm{supp}\mu^i = A_{-i} \right\}$$

[18] And risk-neutral.

be the set of such *full-support* conjectures. In the *finite* case, we can write $\Delta^o(A_{-i})$ as follows:[19]

$$\Delta^o(A_{-i}) = \left\{\mu^i \in \Delta(A_{-i}) : \forall a_{-i} \in A_{-i}, \mu^i(a_{-i}) > 0\right\}.$$

A *rational and cautious* player i chooses actions in $r_i(\Delta^o(A_{-i}))$. These considerations motivate the following definition and results.

Definition 8. *A mixed action α_i **weakly dominates** another (pure or) mixed action β_i if it yields at least the same expected payoff for every action profile a_{-i} of the other players and strictly more for at least one $\bar{a}_{-i}$, that is,*

$$\begin{array}{rcl} \forall a_{-i} & \in & A_{-i},\ u_i(\alpha_i, a_{-i}) \geq u_i(\beta_i, a_{-i}), \\ \exists \bar{a}_{-i} & \in & A_{-i},\ u_i(\alpha_i, \bar{a}_{-i}) > u_i(\beta_i, \bar{a}_{-i}). \end{array}$$

The set of pure *actions that are* not weakly dominated *by any mixed action is denoted by NWD_i. Such actions are also called **admissible.***

The set NWD_i of admissible actions can be formally written as follows:[20]

$$\begin{aligned} & NWD_i \\ = & \left\{ a_i \in A_i : \forall \alpha_i \in \Delta(A_i), \begin{array}{l} (\exists \bar{a}_{-i} \in A_{-i}, u_i(\alpha_i, \bar{a}_{-i}) < u_i(a_i, \bar{a}_{-i})) \\ \vee \\ (\forall a_{-i} \in A_{-i}, u_i(\alpha_i, a_{-i}) \leq u_i(a_i, a_{-i})) \end{array} \right\} \\ = & \left\{ a_i \in A_i : \forall \alpha_i \in \Delta(A_i), \begin{array}{l} (\exists \bar{a}_{-i} \in A_{-i}, u_i(\alpha_i, \bar{a}_{-i}) < u_i(a_i, \bar{a}_{-i})) \\ \vee \\ (\forall a_{-i} \in A_{-i}, u_i(\alpha_i, a_{-i}) = u_i(a_i, a_{-i})) \end{array} \right\}, \end{aligned}$$

which means that a_i is not weakly dominated if, for every α_i, either (1) a_i yields a strictly higher payoff than α_i for at least one a_{-i}, or (2) a_i yields a weakly higher expected payoff than α_i for every a_{-i}; if case (1) does not hold, then $u_i(\alpha_i, a_{-i}) \geq u_i(a_i, a_{-i})$ for every a_{-i}, and then a_i is not weakly dominated by α_i if and only if a_i and α_i are payoff-equivalent.

The reader should try to prove the following remark as an exercise:

[19] In the finite case, we can interpret $\Delta^o(A_{-i})$ as the relative interior of $\Delta(A_{-i})$, that is, the intersection of $\Delta(A_{-i})$ with the open set $\mathbb{R}^{A_{-i}}_{++}$.

[20] Symbol $\vee$ represents the non-exclusive "or." For example, if c is the proposition "the weather is cold" and r is the proposition "the weather is rainy," $c \vee r$ means that the weather is either cold, or rainy, or both; thus, $c \vee r$ is false if the weather is both warm (not cold) and dry (not rainy).

Remark 5. *If an action a_i^* in a finite or compact-continuous game is the* unique *best reply to some conjecture μ^i, then it is admissible.*

Next we analyze the relationship between admissible actions and the actions that might be chosen by a rational and cautious player. The following lemma (proved in the appendix) says that a rational and cautious player would never choose a weakly dominated action.

Lemma 4. *Fix arbitrarily a player $i \in I$ and an action $a_i^* \in A_i$ in a finite or compact-continuous game. If there exists a full-support conjecture μ^i that justifies a_i^*, then a_i^* is admissible:*

$$r_i\left(\Delta^o(A_{-i})\right) \subseteq NWD_i.$$

The following example shows that, if A_i is infinite, the inclusion can be strict, that is, there can be admissible actions that are not best replies to any full-support conjecture.

Example 4. Let $A_1 = [0,1]$, $A_2 = \{0,1\}$,

$$\begin{array}{llll} u_1: & A_1 \times A_2 & \to & \mathbb{R}, \\ & (a_1, a_2) & \mapsto & a_2 a_1 + (1-a_2)\sqrt{1-(a_1)^2}, \end{array}$$

and u_2 is any continuous function in $\mathbb{R}^{A_1 \times A_2}$. Clearly, this is a compact-continuous game. Let $\mu \in [0,1]$ denote the probability of $a_2 = 1$. Then

$$u_1\left(a_1, \mu\right) = \mu a_1 + (1-\mu)\sqrt{1-(a_1)^2}.$$

This expected payoff function is strictly concave in a_1; hence, there is a unique best reply for each $\mu \in [0,1]$. The first-order condition yields the best reply function

$$r_1\left(\mu\right) = \frac{\mu}{\sqrt{1-2\mu+2\mu^2}} \in [0,1]\,.$$

Thus, $r_1(0) = 0$, $r_1(1) = 1$, and $0 < r_1(\mu) < 1$ if $0 < \mu < 1$, which implies that $0 \notin r_1\left(\Delta^o\left(A_2\right)\right)$ and $1 \notin r_1\left(\Delta^o\left(A_2\right)\right)$. Yet, neither 0 nor 1 is weakly dominated (see Remark 5). ▲

This example shows that the exact converse of Lemma 4 does not hold, because there are infinite compact-continuous games with admissible actions that cannot be justified as best replies to full-support conjectures. Yet, the following lemma (proved in the appendix) states that in all *finite* games admissibility is equivalent to justifiability by full-support conjectures.

Lemma 5. *(Cf. Wald [69], Pearce [54]) Let G be a finite game. Then, for every $i \in I$,*

$$NWD_i = r_i\left(\Delta^o(A_{-i})\right).$$

Next we consider an interesting special case:

Definition 9. *An action a_i^* is* ***weakly dominant*** *if it weakly dominates every other (pure) action, i.e., if for every other action $\widehat{a}_i \in A_i \backslash \{a_i^*\}$ the following conditions hold:*

$$\begin{aligned} \forall a_{-i} &\in A_{-i}, u_i(a_i^*, a_{-i}) \geq u_i(\widehat{a}_i, a_{-i}), \\ \exists \widehat{a}_{-i} &\in A_{-i}, u_i(a_i^*, \widehat{a}_{-i}) > u_i(\widehat{a}_i, \widehat{a}_{-i}). \end{aligned}$$

You should prove the following statement as an exercise:

Remark 6. *Fix an action a_i^* arbitrarily. The following conditions are equivalent:*
(0) a_i^* is weakly dominant;
(1) a_i^* weakly dominates every mixed *action $\alpha_i \neq a_i^*$;*
(2) a_i^* is the unique *best reply to every strictly positive conjecture $\mu^i \in \Delta^o(A_{-i})$;*
(3) a_i^* is the only *action with the property of being a best reply to* every *conjecture μ^i (note that if μ^i is not strictly positive, $\mu^i \in \Delta(A_{-i}) \backslash \Delta^o(A_{-i})$, then there may be other best replies to μ^i).*

If a rational and cautious player has a weakly dominant action, then he will choose such an action. There are interesting economic examples where individuals have weakly dominant actions.

Example 5. (Second Price Auction). An art merchant has to auction a work of art (e.g., a painting) at the highest possible price. However, he does not know how much such work of art is worth to the potential buyers. The buyers are collectors who buy the artwork with the only objective to keep it, i.e., they are not interested in what its future price might be. The authenticity of the work is not an issue. The potential buyer i is willing to spend at most $v_i > 0$ to buy it. Such valuation is completely subjective, meaning that if i were to know the other players' valuations, he would not change his own.[21] Following the advice of W. Vickrey, winner of Nobel Prize for Economics,[22] the merchant decides to adopt the following auction rule: the artwork will go to the player who submits the highest offer, but the *price paid* will be equal to the *second highest offer* (in case of a tie between the maximum offers, the work will be assigned by a random draw). This auction rule induces a game among the buyers $i = 1, \dots, n$ where $A_i = [0, +\infty)$, and

$$u_i(a_i, a_{-i}) = \begin{cases} v_i - \max a_{-i}, & \text{if } a_i > \max a_{-i}, \\ 0, & \text{if } a_i < \max a_{-i}, \\ \frac{1}{1+|\arg\max a_{-i}|}(v_i - \max a_{-i}), & \text{if } a_i = \max a_{-i}, \end{cases}$$

where $\max a_{-i} = \max_{j \neq i}(a_1, ..., a_{i-1}, a_{i+1}, \dots, a_n)$. It turns out that offering $a_i^* = v_i$ is the weakly dominant action (can you prove it?). Hence, if the potential buyer i is rational and cautious, he will offer exactly v_i. Doing so, he will expect to make some profits. In fact, being cautious, he will assign a positive probability to event $[\max a_{-i} < v_i]$. Since by offering v_i he will obtain the object only in the event that the price paid is lower than his valuation v_i, the expected payoff from this offer is strictly positive. ▲

3.3.1 Comparative Risk Aversion and Justifiability

In this section we study the impact of risk aversion on the set of justifiable actions. In general, attitudes toward risk are captured by players' von Neumann-Morgenstern utility functions $v_i : Y \to \mathbb{R}$ $(i \in I)$, where Y is the set of outcomes, or consequences. Assuming monetary consequences

[21] If a buyer were to take into account a potential future resale, things might be rather different. In fact, other potential buyers could hold some relevant information that affects the estimate of how much the artwork could be worth in the future. Similarly, if there were doubts regarding the authenticity of the work, it would be relevant to know the other buyers' valuations.

[22] See Vickrey [66].

and selfish preferences, $Y = \mathbb{R}^I$ and $v_i\left((y_j)_{j\in I}\right) = \upsilon_i(y_i)$.[23] In this case, risk aversion is captured by the concavity of υ_i. In comparative terms, the preferences over monetary lotteries represented by $\hat{\upsilon}_i$ exhibit more risk aversion than the preferences represented by υ_i if $\hat{\upsilon}_i$ is a concave and strictly increasing transformation of υ_i, that is, there is a concave and strictly increasing function φ such that $\hat{\upsilon}_i = \varphi \circ \upsilon_i$.

The same comparative criterion can be directly applied to the payoff functions u_i. Consider again the case of monetary outcomes and selfish preferences and let $g = (g_i)_{i\in I} : A \to \mathbb{R}^I$, where $g_i : A \to \mathbb{R}$ is the monetary outcome for player $i \in I$. Then the payoff function of i is

$$u_i = \upsilon_i \circ g_i : \begin{array}{ccc} A & \to & \mathbb{R}, \\ a & \mapsto & \upsilon_i\left(g_i(a)\right). \end{array}$$

Thus, $\hat{u}_i = \hat{\upsilon}_i \circ g_i$ exhibits more risk aversion than $u_i = \upsilon_i \circ g_i$ if there is a concave and strictly increasing function φ such that

$$\hat{u}_i = \hat{\upsilon}_i \circ g_i = (\varphi \circ \upsilon_i) \circ g_i = \varphi \circ (\upsilon_i \circ g_i) = \varphi \circ u_i.$$

This motivates the following:

Definition 10. *Fix two games in reduced form $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ and $\hat{G} = \langle I, (A_i, \hat{u}_i)_{i\in I}\rangle$ and a player $i \in I$, we say that i is* ***more risk averse in $\hat{G}$ than in G*** *if there exists a concave and strictly increasing function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\hat{u}_i = \varphi \circ u_i$.*

We show that, for players who rank actions according to subjective expected utility theory, higher risk aversion weakly expands, in the sense of set inclusion, the set of justifiable actions.[24] At first, the result may look counterintuitive: indeed, higher risk aversion of the decision maker, let us say player i, would increase the attractiveness of the "safer" actions, that is, actions $a_i \in A_i$ such that $u_i(a_i, a_{-i})$ is somewhat low for each profile $a_{-i} \in A_{-i}$, but does not change much with the choice of $a_{-i} \in A_{-i}$. On the other hand, "unsafe" actions a_i that are best replies to some deterministic conjectures may become instead *less* attractive. How can the set of justifiable actions change monotonically?

[23] We use the slightly different symbols v_i and υ_i because v_i is defined on $\mathbb{R}^I$ while υ_i is defined on $\mathbb{R}$.

[24] See Weinstein [73] and Battigalli et al. [17, Proposition 1].

The key to understand the result intuitively is to note that the set of justifiable actions expands with risk aversion if and only if, for every action a_i which is justifiable with low risk aversion, the nonempty set $\left\{\mu^i \in \Delta\left(A_{-i}\right) : a_i \in r_i\left(\mu_i\right)\right\}$ of conjectures justifying a_i does *not* become *empty* when risk aversion is higher. In particular, if an "unsafe" action is justified by a deterministic conjecture, an increase in risk aversion cannot make such actions unjustifiable, because the set of best replies to deterministic conjectures is invariant to strictly increasing transformations of the payoff function u_i. The following example illustrates: as risk aversion increases, safe actions become easier to justify, in the sense that the set of justifying conjectures becomes larger; in particular, the set can be empty with low risk aversion and nonempty with higher risk aversion; on the other hand, if an unsafe action is—despite being unsafe—justifiable, then the set of justifying conjectures shrinks as risk aversion increases, but it cannot become empty.

Example 6. Consider the following game form with monetary outcomes, where $I = \{1,2\}$, $A_1 = \{t,m,b\}$, $A_2 = \{\ell, r\}$, and outcome function $g : A \to \mathbb{R}^I$; player 1's component of the outcome function $g_1 : A \to \mathbb{R}$ is represented in Figure 3.4.

	ℓ	r
t	0	1
m	$1/3$	$1/3$
b	1	0

Figure 3.4: Function g_1 which represents monetary gains of player 1.

If player 1 is risk neutral, i.e., $v_1 : \mathbb{R}^I \to \mathbb{R}$ is defined by $v_1(x,y) = x$ and $u_1 = v_1 \circ g = g_1$, then action m is not justifiable. Indeed, for every belief $\mu^1 \in \Delta(A_2)$, we have

$$\mathbb{E}_{\mu^1}\left(u_1(m,\cdot)\right) = \frac{1}{3} < \frac{1}{2} \leq \max\left\{\mathbb{E}_{\mu^1}\left(u_1(t,\cdot)\right), \mathbb{E}_{\mu^1}\left(u_1(b,\cdot)\right)\right\}.$$

Then, let us suppose that the utility function v_1 is replaced by $v_{1,\theta} : \mathbb{R}^I \to \mathbb{R}$ defined by

$$v_{1,\theta}(x,y) = x^{1/\theta},$$

where $\theta \geq 1$ parametrizes risk aversion. Then, the payoff function is defined by $u_{1,\theta}(a) = (g_1(a))^{1/\theta}$ for all $a \in A$.

It is not difficult to check that the set of justifiable actions of player 1 is

$$r_{1,\theta}\left(\Delta(A_{-1})\right) = \begin{cases} \{t,b\}, & \text{if } \theta \in [1, \log_2 3), \\ \{t,m,b\}, & \text{if } \theta \in [\log_2 3, \infty). \end{cases}$$

Therefore, the collection of justifiable actions expands as θ increases. On the other hand, notice that the set of beliefs justifying actions b and t *shrinks* as θ increases above the threshold $\overline{\theta} = \log_2 3$, see Figure 3.5. ▲

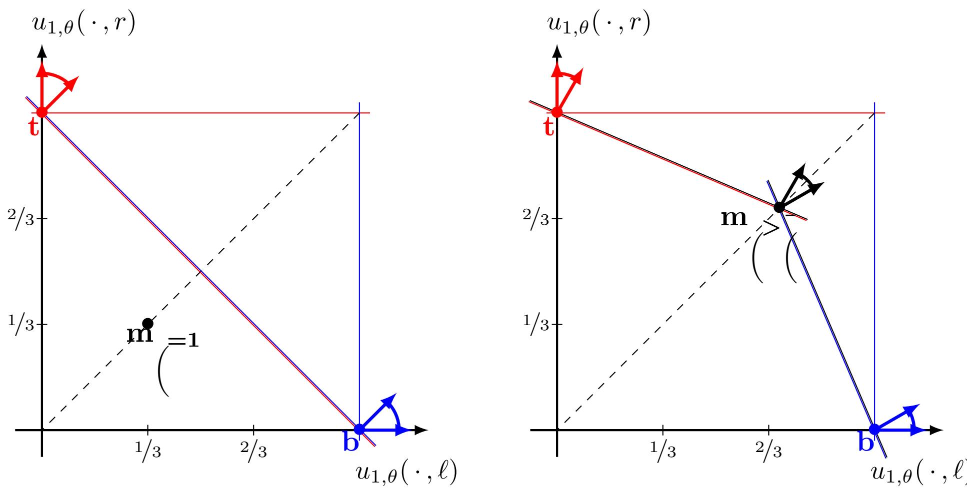


Figure 3.5: As θ increases, the set of beliefs justifying b and t shrinks.

Lemma 2 can be used to prove the following result:

Theorem 1. *Let $G = \langle I, (A_i, u_i)_{i \in I} \rangle$ and $\hat{G} = \langle I, (A_i, \hat{u}_i)_{i \in I} \rangle$ be two finite or compact-continuous games in reduced form, and suppose that player $i \in I$ is more risk averse in $\hat{G}$ than in G. Then*

$$r_i\left(\Delta(A_{-i})\right) \subseteq \hat{r}_i\left(\Delta(A_{-i})\right).$$

Proof. By assumption, there exists a concave and strictly increasing function $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\hat{u}_i = \varphi \circ u_i$. Let us fix any justifiable action of i in G, that is, $a_i \in r_i\left(\Delta(A_{-i})\right)$. We must show that $a_i \in \hat{r}_i\left(\Delta(A_{-i})\right)$, where $\hat{r}_i$ is the best reply correspondence associated to payoff function $\hat{u}_i$. By Lemma 2, a_i is not dominated in G by any mixed action $\alpha_i \in \Delta(A_i)$:

$$\forall \alpha_i \in \Delta(A_i), \exists a_{-i} \in A_{-i}, u_i(a_i, a_{-i}) \geq u_i(\alpha_i, a_{-i}).$$

Therefore, for every mixed action $\alpha_i \in \Delta(A_i)$, there exists a profile $a_{-i} \in A_{-i}$ such that

$$\begin{aligned}\hat{u}_i(a_i, a_{-i}) &= \varphi\left(u_i(a_i, a_{-i})\right) \geq \varphi\left(u_i(\alpha_i, a_{-i})\right) = \varphi\left(\mathbb{E}_{\alpha_i}(u_i(\,\cdot\,, a_{-i}))\right) \\ &\geq \mathbb{E}_{\alpha_i}(\varphi(u_i(\,\cdot\,, a_{-i}))) = \mathbb{E}_{\alpha_i}(\hat{u}_i(\,\cdot\,, a_{-i})) = \hat{u}_i(\alpha_i, a_{-i}),\end{aligned}$$

that is, a_i is not dominated in $\hat{G}$. Here, the first inequality follows from the monotonicity of φ and the second follows from Jensen's inequality.[25]

Hence, again by Lemma 2, action a_i belongs to $\hat{r}_i\left(\Delta(A_{-i})\right)$. ■

3.3.2 Best Replies and Undominated Actions in Nice Games

Many models used in applications of game theory to economics and other disciplines have the following features: players choose actions in compact (closed and bounded) intervals of the real line, their payoff functions are continuous, and the payoff function of each player is strictly concave, or—at least—strictly quasi-concave,[26] in his own action. Such games are called "nice;" see Moulin [48].

Definition 11. *A static game $G = \langle I, (A_i, u_i)_{i \in I}\rangle$ is **nice** if, for every player $i \in I$, A_i is a compact interval $[\underline{a}_i, \bar{a}_i] \subset \mathbb{R}$, the payoff function $u_i : A \to \mathbb{R}$ is continuous (that is, jointly continuous in all its arguments), and the function $u_i(\cdot, a_{-i}) : [\underline{a}_i, \bar{a}_i] \to \mathbb{R}$ is strictly quasi-concave for each $a_{-i} \in A_{-i}$.*

[25] The concave transformation of the expectation of a random variable is larger than the expectation of the concave transformation of that random variable, see Aliprantis and Border [3, Theorem 11.24].

[26] A function $f : X \to \mathbb{R}$ with convex domain is **strictly quasi-concave** if for every pair of *distinct* points $x', x'' \in X$ and every $t \in (0, 1)$,

$$f(tx' + (1-t)x'') > \min\{f(x'), f(x'')\}.$$

Remark 7. *Let* $A_i = [\underline{a}_i, \bar{a}_i] \subset \mathbb{R}$ *and fix* $a_{-i} \in A_{-i}$. *Function* $u_i(\cdot, a_{-i}) : [\underline{a}_i, \bar{a}_i] \to \mathbb{R}$ *is strictly quasi-concave if and only if the following two conditions hold: (1) there is a* unique *best reply* $r_i(a_{-i}) \in [\underline{a}_i, \bar{a}_i]$ *and (2)* $u_i(\cdot, a_{-i})$ *is strictly increasing on* $[\underline{a}_i, r_i(a_{-i})]$ *and strictly decreasing on* $[r_i(a_{-i}), \bar{a}_i]$.

Example 7. Consider Cournot's oligopoly game with capacity constraints: I is the set of firms, $a_i \in [0, \bar{a}_i]$ is the output of firm i and $\bar{a}_i$ is its capacity. The payoff function of firm i is

$$u_i(a) = a_i P\left(a_i + \sum_{j \neq i} a_j\right) - C_i(a_i),$$

where $P : [0, \sum_i \bar{a}_i] \to \mathbb{R}_+$ is the downward-sloping inverse demand function and $C_i : [0, \bar{a}_i] \to \mathbb{R}_+$ is i's cost function. The Cournot game is *nice* if $P(\cdot)$ and each cost function $C_i(\cdot)$ are continuous and if, for each i and a_{-i}, $P(a_i + \sum_{j \neq i} a_j)a_i - C_i(a_i)$ is strictly quasi-concave. The latter condition is easy to satisfy. For example, if each C_i is strictly increasing and convex, and $P(\sum_{j \in I} a_j) = \max\{0, \bar{p} - \beta \sum_{j \in I} a_j\}$, then strict-quasi concavity holds (prove this as an exercise).[27] ▲

Nice games allow an easy computation of the sets of justifiable actions and of dominated actions. We are going to compare best replies to deterministic conjectures, uncorrelated conjectures, correlated conjectures, pure actions not dominated by mixed actions, and pure actions not dominated by other pure actions. First note that, by definition, for every player i the following inclusions hold:[28]

$$\begin{aligned} r_i(A_{-i}) &\subseteq r_i(\mathrm{I}\Delta(A_{-i})) \subseteq r_i(\Delta(A_{-i})) \subseteq ND_i \\ &\subseteq \{a_i \in A_i : \forall b_i \in A_i, \exists a_{-i} \in A_{-i}, u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})\}, \end{aligned}$$

[27] But strict concavity does *not* hold. Can you see why? Consider the neighborhood of a point where $P(\sum_{j \in I} a_j)$ becomes zero.

[28] If you do not know how to deal with probability measures on infinite sets, just consider the simple ones, i.e., those that assign positive probability to a finite set of points. The set of simple probability measures on an infinite domain X is a subset of $\Delta(X)$, but under the assumptions considered here it is possible to restrict one's attention to this smaller set.

where

$$\mathrm{I}\Delta(A_{-i}) = \left\{ \mu^i \in \Delta(A_{-i}) : \exists (\mu^i_j)_{j\in I\setminus\{i\}} \in \bigtimes_{j\in I\setminus\{i\}} \Delta(A_j), \mu^i = \prod_{j\neq i} \mu^i_j \right\}$$

is the set of *product* probability measures on A_{-i}, that is, the set of beliefs that satisfy independence across opponents. In words, the set of best replies to deterministic conjectures in A_{-i} is contained in the set of best replies to probabilistic independent conjecture, which is contained in the set of best replies to probabilistic (possibly correlated) conjectures, which is contained in the set of actions not dominated by mixed actions, which is contained in the set of (pure) actions not dominated by other pure actions. In each case the inclusion may hold as an equality. The following technical results imply that—in nice games—all these (weak) inclusions indeed hold as equalities. We start by providing conditions under which the best reply correspondence is actually a continuous *function*.

Lemma 6. *Consider a compact-continuous game such that, for each $i \in I$, each A_i is convex and $u_i(\cdot, a_{-i})$ is strictly quasi-concave for each $a_{-i} \in A_{-i}$. Then each player i has a well-defined and continuous best reply function on the domain A_{-i}.*

Proof. By convexity of A_i and strict quasi-concavity of u_i, the set $\arg\max_{a_i\in A_i} u_i(a_i, a_{-i})$ is a singleton for each a_{-i}. To prove this, it is sufficient to show that, if $a'_i \neq a''_i$ and $u_i(a'_i, a_{-i}) = u_i(a''_i, a_{-i})$, then $u_i(a'_i, a_{-i}) < \max_{a_i\in A_i} u_i(a_i, a_{-i})$. To see this, let a'_i and a''_i be as above. By convexity of A_i, for each $t \in (0, 1)$, $ta'_i + (1-t)a''_i \in A_i$; by strict quasi-concavity of u_i, $u_i(ta'_i+(1-t)a''_i, a_{-i}) > \min\{u_i(a'_i, a_{-i}), u_i(a''_i, a_{-i})\}$. Since $u_i(a'_i, a_{-i}) = u_i(a''_i, a_{-i})$ the result follows.

Next we show that the function $r_i(a_{-i}) = \arg\max_{a_i\in A_i} u_i(a_i, a_{-i})$ is continuous, that is, for any sequence $\left(a^k_{-i}\right)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} a^k_{-i} = \bar{a}_{-i}$ the corresponding sequence of best replies $\left(r_i(a^k_{-i})\right)_{k\in\mathbb{N}}$ converges to $r_i(\bar{a}_{-i})$ $[\lim_{k\to\infty} r_i(a^k_{-i}) = r_i(\bar{a}_{-i})]$. Note first that, since the sequence $\left(r_i(a^k_{-i})\right)_{k\in\mathbb{N}}$ is contained in the compact set A_i, it must have at least one accumulation point $\bar{a}_i \in A_i$.[29] Therefore, there exists a subsequence $\left(a^{k_\ell}_{-i}\right)_{\ell\in\mathbb{N}}$ such that

[29]By the Bolzano–Weierstrass' theorem, every bounded sequence of real numbers has a convergent subsequence; see, e.g., Ok [52, Chapter A, p.52].

$\lim_{\ell\to\infty} r_i(a_{-i}^{k_\ell}) = \bar{a}_i$. The definition of best reply implies

$$\forall a_i \in A_i,\ \forall \ell \geq 1,\ u_i(r_i(a_{-i}^{k_\ell}), a_{-i}^{k_\ell}) \geq u_i(a_i, a_{-i}^{k_\ell}).$$

Taking the limit for $\ell \to \infty$ on both sides for any given a_i, the continuity of u_i yields

$$\forall a_i \in A_i,\ u_i(\bar{a}_i, \bar{a}_{-i}) \geq u_i(a_i, \bar{a}_{-i}).$$

Hence, $\bar{a}_i$ is a best reply to $\bar{a}_{-i}$. Since the best reply is unique, $\bar{a}_i = r_i(\bar{a}_{-i})$. This is true for every accumulation point. Therefore, the sequence $\left(r_i(a_{-i}^k)\right)_{k\in\mathbb{N}}$ has only one accumulation point, $r_i(\bar{a}_{-i})$, which is equivalent to say that $\lim_{k\to\infty} r_i(a_{-i}^k) = r_i(\bar{a}_{-i})$, as desired. ■

Lemma 6 implies that the set of best replies to deterministic conjectures is connected, which in turn yields the following convenient result:

Corollary 2. *In every* nice game, *the set of best replies of each player i to deterministic conjectures is a compact interval:*

$$r_i\left(A_{-i}\right) = \left[\min r_i\left(A_{-i}\right), \max r_i\left(A_{-i}\right)\right].$$

Proof. Since $r_i(\cdot)$ is a continuous function and A_{-i} is convex (hence, connected) and compact, $r_i\left(A_{-i}\right) \subseteq \mathbb{R}$ is convex and compact,[30] hence it is a compact interval. ■

We can now state and prove the main result of this section.

Lemma 7. *In every* nice game, *the set of best replies of each player i to deterministic conjectures coincides with the set of actions not dominated by other* pure *actions, and therefore it also coincides with the set of best replies to independent or correlated probabilistic conjectures, and with the set of actions not dominated by mixed actions:*

$$\begin{aligned} r_i(A_{-i}) &= r_i(\mathrm{I}\Delta(A_{-i})) = r_i(\Delta(A_{-i})) = ND_i \\ &= \{a_i \in A_i : \forall b_i \in A_i, \exists a_{-i} \in A_{-i}, u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})\}. \end{aligned}$$

Proof. Let $ND_{i,p} \subseteq A_i$ denote the set of player i's pure actions not dominated by other pure actions. We prove that $ND_{i,p} \subseteq r_i(A_{-i})$, that is, $A_i \backslash r_i(A_{-i}) \subseteq A_i \backslash ND_{i,p}$. Since we already noticed that

$$r_i(A_{-i}) \subseteq r_i(\mathrm{I}\Delta(A_{-i})) \subseteq r_i(\Delta(A_{-i})) \subseteq ND_{i,p},$$

[30] See Proposition 2 in Ok [52, D.2].

this implies the thesis. By Corollary 2,

$$r_i\left(A_{-i}\right) = \left[\min r_i\left(A_{-i}\right), \max r_i\left(A_{-i}\right)\right].$$

Therefore, it is enough to show that all the actions below $\min r_i\left(A_{-i}\right)$ or above $\max r_i\left(A_{-i}\right)$ are dominated. Fix any $a_i < \min r_i\left(A_{-i}\right)$. Thus, for each $a_{-i} \in A_{-i}$, $a_i < \min r_i\left(A_{-i}\right) \leq r_i\left(a_{-i}\right)$. By definition, $u_i(\cdot, a_{-i})$ attains its maximum at $r_i\left(a_{-i}\right)$; thus, by strict quasi-concavity, $u_i(\cdot, a_{-i})$ is strictly increasing on $[a_i, r_i\left(a_{-i}\right)]$. It follows that

$$\forall a_{-i} \in A_{-i},\ u_i(a_i, a_{-i}) < u_i(\min r_i\left(A_{-i}\right), a_{-i}) \leq u_i(r_i\left(a_{-i}\right), a_{-i}).$$

Therefore, every $a_i < \min r_i\left(A_{-i}\right)$ is strictly dominated by $\min r_i\left(A_{-i}\right)$. A similar argument shows that every $a_i > \max r_i\left(A_{-i}\right)$ is strictly dominated by $\max r_i\left(A_{-i}\right)$. ■

3.3.3 Supermodular Nice Games

Best reply correspondences in nice games are continuous real-valued functions, which in general may not be monotone. We now consider a class of nice games where best reply correspondences are increasing functions, the so-called supermodular nice games, in which players have "strategic complementarities."[31] A game exhibits strategic complementarities if each player's incentive to increase his action becomes stronger for higher conjectured actions of the co-players. For example, if firms i and j are competing on prices and firm i thinks that j is going to increase its price, then firm i has a stronger incentive to increase (or, equivalently, a weaker incentive to decrease) its own price as well; see also Example 8 below. The notion of strategic complementarities is captured by the "supermodularity" of the payoff functions, which we define below for the special case of nice games.

We first introduce some notation. Let J be a finite index set, e.g., $J = \{1, \ldots, n\}$, or $J = I$, the player set in a game. Since $\mathbb{R}^J$ (the set of all functions from J to $\mathbb{R}$) is isomorphic to the Euclidean space $\mathbb{R}^{|J|}$, we call "vectors" the elements of $\mathbb{R}^J$. For each $x, y \in \mathbb{R}^J$, we say that $x \leq y$ if and only if $x_i \leq y_i$ for each $i \in J$; this is the usual coordinatewise incomplete

[31] Here we restrict our attention to nice games for the sake of simplicity, but the theory of supermodular games is much more general. See Topkis [65].

order on Euclidean spaces. Accordingly, for each $x, y \in \mathbb{R}^J$, we define the **order box** $[x, y] \subseteq \mathbb{R}^J$ as

$$[x, y] = \{z \in \mathbb{R}^J : x \leq z \leq y\}.$$

Note that $[x, y]$ is non-empty if and only if $x \leq y$, and $[x, y]$ is the usual compact interval if $x, y \in \mathbb{R}$.

Given a non-empty order box $[x, y] \subseteq \mathbb{R}^J$, a function $f : [x, y] \to \mathbb{R}$ is said to be **supermodular** if, for all $a, b \in [x, y]$,

$$f(a) + f(b) \leq f(\max(a, b)) + f(\min(a, b)), \tag{3.3.4}$$

where $\max(\cdot, \cdot)$ and $\min(\cdot, \cdot)$ denote the following binary operations on vectors:[32]

$$\begin{aligned}(a, b) &\mapsto (\max\{a_i, b_i\})_{i \in J},\\ (a, b) &\mapsto (\min\{a_i, b_i\})_{i \in J}.\end{aligned}$$

Definition 12. *A* nice *game $G = \langle I, (A_i, u_i)_{i \in I}\rangle$ is **supermodular** if, for each $i \in I$, the payoff function $u_i : A \to \mathbb{R}$ is supermodular.*

Note that the above definition is meaningful, because in nice games A is indeed an order box in $\mathbb{R}^I$. The idea behind the notion of supermodularity can be better understood by rewriting inequality (3.3.4) as

$$f(\max(a, b)) - f(a) \geq f(b) - f(\min(a, b)), \tag{3.3.5}$$

or also

$$f(\max(a, b)) - f(b) \geq f(a) - f(\min(a, b)), \tag{3.3.6}$$

which suggests a property of "increasing differences." To see this, consider the following example represented in Figure 3.6. Let $a = (2, 1)$ and $b = (1, 2)$ be points in the order box $X = [(1, 1), (2, 2)]$, and fix a supermodular function $f : X \to \mathbb{R}$. Then, horizontal arrows represent inequality (3.3.5), while vertical arrows represent inequality (3.3.6).

[32] When $a, b \in \mathbb{R}$ then $\max(a, b) = \max\{a, b\}$. But in the general case $a, b \in \mathbb{R}^J$, $\max(a, b)$ is the maximal element of the binary set $\{a, b\}$ if and only if, either $a_i \geq b_i$ for all $i \in J$, or $b_i \geq a_i$ for all $i \in J$; otherwise, $\max(a, b)$ does not even belong to $\{a, b\}$. A similar observation holds for $\min(a, b)$. In the more general framework of partially ordered sets, these binary operations are denoted by $a \vee b$ (maximum) and $a \wedge b$ (minimum).

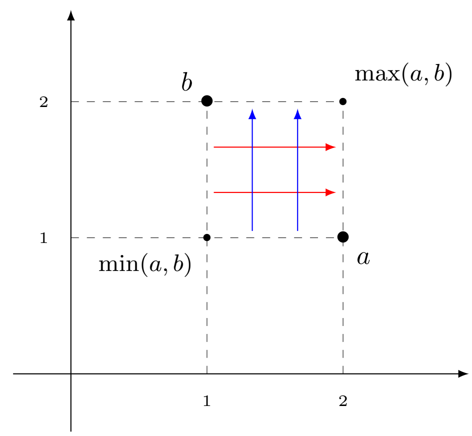


Figure 3.6: Increasing differences in a supermodular game.

Suppose that $X = A_1 \times A_2$ and let $f(a_1, a_2) = u_1(a_1, a_2)$ be the payoff function of player 1 in a two-person game. Then the "increasing differences" property says that player 1 gains more (or loses less) from increasing his action when the action of player 2 is higher.

We now make the connection with supermodularity more precise. Given non-empty order boxes $[x, y] \subseteq \mathbb{R}^{J_1}$ and $[a, b] \subseteq \mathbb{R}^{J_2}$, a function $f : [x, y] \times [a, b] \to \mathbb{R}$ is said to have **increasing differences in** $[a, b]$ if, for all $c \leq d$ in $[a, b]$, the function $\tilde{f} : [x, y] \to \mathbb{R}$ defined by

$$z \mapsto \tilde{f}(z) = f(z, d) - f(z, c)$$

is weakly increasing. Equivalently, $\tilde{f}(z) \leq \tilde{f}(w)$ for all $z, w \in [x, y]$ such that $z \leq w$.

The following result provides a useful criterion to check whether a function is supermodular:

Topkis' Characterization Theorem. (Topkis [65]) *Let* $[x, y] = \times_{i \in I}[x_i, y_i] \subseteq \mathbb{R}^I$ *be an order box. Then a function* $f : [x, y] \to \mathbb{R}$ *is supermodular if and only if* f *has increasing differences in* $\times_{j \in I \setminus \{i\}}[x_j, y_j]$ *for all* $i \in I$. *Additionally, if* f *is twice differentiable,*[33] *then* f *is*

[33] To avoid details involving boundary conditions, we stipulate that a function defined on an order box is differentiable if it has a differentiable extension defined on an open set containing the box.

supermodular if and only if the second-order partial derivatives $\partial^2 f/\partial x_i \partial x_j$ are nonnegative for all $1 \leq i < j \leq n$.

The second part of the Topkis' Characterization Theorem has an intuitive explanation. Indeed, if f is differentiable, then f has increasing differences if and only if $\partial f/\partial x_i$ is increasing in x_j for all $i \neq j$. It follows that, if f is also twice differentiable, then f has increasing differences if and only if $\partial^2 f/\partial x_i \partial x_j$ is a nonnegative function for all $i \neq j$.

Example 8. Consider a price-setting duopoly with differentiated products, so that $I = \{1, 2\}$ and $A_1 = A_2 = [0, \bar{p}]$, for some sufficiently large $\bar{p} > 0$. The joint demand function $D = (D_1, D_2) : A \to \mathbb{R}^2$ is defined by

$$(p_1, p_2) \mapsto (-d_1 p_1 + e_{12} p_2 + f_1, -d_2 p_2 + e_{21} p_1 + f_2)$$

for some collection of parameters $d_i, e_{ij}, f_i > 0$ $(i, j \in \{1, 2\},\ i \neq j)$. Such system of demand functions can be derived from quite standard preferences of consumers (see, for example, Motta [51]).

Each firm i has constant marginal cost $c_i < \bar{p}$. Since firms compete in prices, each firm i chooses p_i so as to maximize its payoff (profit)

$$u_i(p_i, p_j) = (p_i - c_i)\, D_i(p_i, p_j) = (p_i - c_i)\, (-d_i p_i + e_{ij} p_j + f_i)\,.$$

It follows from Topkis' Characterization Theorem that this is a supermodular nice game:

$$\frac{\partial u_i}{\partial p_i}(p_i, p_j) = -2 d_i p_i + d_i c_i + e_{ij} p_j + f_i,$$

thus

$$\frac{\partial^2 u_i}{\partial p_i \partial p_j}(p_1, p_2) = e_{ij} > 0.$$

This means that each firm has the incentive to increase the price of its own product if the other firm is doing so. In particular, we verify that the best reply functions are weakly increasing, and indeed strictly increasing when they attain an interior solution. From the first-order condition $\frac{\partial u_i}{\partial p_i} = 0$ for interior solutions, and taking into account that $\frac{\partial^2 u_i}{\partial p_i^2} = -2 d_i < 0$, we obtain the best reply function

$$r_i\,(p_j) = \min\left\{\frac{1}{2}\left(c_i + \frac{e_{ij} p_j + f_i}{d_i}\right), \bar{p}\right\},$$

that is, the best reply function is strictly increasing up to the point where it attains the upper bound $\bar{p}$. ▲

The following noteworthy lemma clarifies that the monotonicity result of Example 8 is a general property of supermodular nice games.

Lemma 8. *Let $G = \langle I, (A_i, u_i)_{i \in I} \rangle$ be a supermodular nice game. Then, for each $i \in I$, the best reply correspondences r_i are weakly increasing continuous functions. Moreover, if the payoff functions u_i are twice differentiable and $\frac{\partial^2 u_i}{\partial a_i^2} < 0$, $\frac{\partial^2 u_i}{\partial a_i \partial a_j} > 0$ for all $i, j \in I$, $j \neq i$, then the restriction of each r_i to the subset*

$$\{a_{-i} \in A_{-i} : \min A_i < r_i(a_{-i}) < \max A_i\}$$

where it attains interior values is a strictly increasing function.

Proof. We prove here the second part of the result, while the first part is proved in Appendix 3.4.2. Fix any $i \in I$ and note that, by Lemma 6, r_i is a continuous function. Pick any $a_{-i} \in A_{-i}$ such that $\min A_i < r_i(a_{-i}) < \max A_i$ (recall that $A_i = [\min A_i, \max A_i]$ because G is nice). Since $r_i(a_{-i})$ is an interior solution and u_i is twice differentiable, $r_i(a_{-i})$ satisfies the first and second-order conditions:

$$\begin{aligned} \frac{\partial u_i}{\partial a_i}(r_i(a_{-i}), a_{-i}) &= 0, \\ \frac{\partial^2 u_i}{\partial a_i^2}(r_i(a_{-i}), a_{-i}) &< 0, \end{aligned}$$

where the inequality is strict by assumption.[34] Hence, by the Implicit Function theorem, we obtain

$$\frac{\partial r_i}{\partial a_j}(a_{-i}) = -\frac{\frac{\partial}{\partial a_j}\left(\frac{\partial u_i}{\partial a_i}(r_i(a_{-i}), a_{-i})\right)}{\frac{\partial}{\partial a_i}\left(\frac{\partial u_i}{\partial a_i}(r_i(a_{-i}), a_{-i})\right)} = -\frac{\frac{\partial^2 u_i}{\partial a_i \partial a_j}(r_i(a_{-i}), a_{-i})}{\frac{\partial^2 u_i}{\partial a_i^2}(r_i(a_{-i}), a_{-i})} > 0$$

for each $j \neq i$. The last inequality follows from the hypothesis that $\frac{\partial^2 u_i}{\partial a_i \partial a_j} > 0$ for all $i \neq j$ and from the second-order condition. ■

[34] Strict quasi-concavity of u_i in a_i implies that $\frac{\partial^2 u_i}{\partial a_i^2} \leq 0$ at an interior maximum, but it does not imply that $\frac{\partial^2 u_i}{\partial a_i^2} < 0$. Consider, for example, a two-person game with $u_i(a_i, a_{-i}) = -(a_i - a_{-i})^4$. This payoff function is strictly concave (hence, strictly quasi-concave) in a_i, but $\frac{\partial^2 u_i}{\partial a_i^2} = 0$ at $a_i = a_{-i}$.

3.4 Appendix: Compact-Continuous Games

In this appendix, we add technical details about the analysis of compact-continuous games and we prove the general version of Lemma 2 (on best replies and dominance) and related results. The proof of Lemma 4 is similar. We first need some preliminaries about probability measures. We provide the necessary concepts to understand the analysis of mixed actions and probabilistic conjectures in the context of compact-continuous games. These preliminaries should be enough to allow the reader to understand the aforementioned proofs, but to fully master the mathematical concepts involved, the reader should consult appropriate primary sources, such as Aliprantis and Border [3]. In particular, the remarks stated in this Appendix are not proved; the reader should first try to prove them by herself or himself and then check on primary sources.

3.4.1 Probability Measures on Compact Sets and Expected Values

Sigma-Algebras and Probability Measures We expand on what we already explained in Section 3.1. For the reader's convenience, we repeat some concepts already introduced there. The following introductory discussion is heuristic.

To model uncertainty with an infinite state space X, we do not necessarily consider all the subsets of X as "events." We think of **events** as subsets of elements of X that verify some *statements*, which in turn can be obtained as negations, conjunctions, and disjunctions of simpler statements. For example, we may be uncertain about which point x is going to be selected, by an agent, or by a random device, in the compact interval $[0,1]$. Then, the open sub-interval (a,b) of $[0,1]$ corresponds to the statement "x is larger than a and smaller than b," which in turn is the conjunction of the two statements "x is larger than a" and "x is smaller than b." It is natural to assume that we can assign a probability to such elementary statements; with this, it is also natural to assume that we are able to assign a probability to their conjunction, i.e., to the open interval (a,b).

The negation of the conjunction "x is larger than a and smaller than b" is logically equivalent to the disjunction "either x is at most a, or x is at least b," which corresponds to the set $[0,a]\cup[b,1]=[0,1]\backslash(a,b)$. Of course, if we assign a probability p to a statement s, then we also assign probability $1-p$ to the negation of s. If statement s corresponds to set E, then the negation of s corresponds to the complement of E. Hence, if we can assign probability p to E, we also assign probability $1-p$ to E^c, the complement of E. In our example, we can assign probabilities to $(a, 1]$, $[0, b)$, and (a, b), therefore we also assign probabilities to $[0, a]$, $[b, 1]$, and $[0, a] \cup [b, 1]$ respectively.

Finally, suppose we can assign a probability to each statement s_n, corresponding to event E_n. It is then quite natural to assume that we can also assign a probability to the statement "s_n holds for at least one $n \in \mathbb{N}$." For example, since we can assign probability to each statement "x is at most $(n-1)/(2n)$," then we can also assign probability to the statement "for at least one $n \in \mathbb{N}$, x is at most $\frac{1}{2}\frac{n-1}{n}$." Each elementary statement corresponds to the compact interval $\left[0, \frac{1}{2}\frac{n-1}{n}\right]$ and "for at least one $n \in \mathbb{N}$, x is at most $\frac{1}{2}\frac{n-1}{n}$" corresponds to

$$\bigcup_{n=1}^{\infty} \left[0, \frac{1}{2}\frac{n-1}{n}\right] = \left[0, \frac{1}{2}\right).$$

With this in mind, we abstract from the particular meaning of the statements and their representation as subsets of X and we identify the essential properties of the collection of subsets of X to which we are willing to assign a probability. A **sigma-algebra** on a set X is a collection of subsets $\mathcal{X} \subseteq 2^X$ called **events** such that

1. $X \in \mathcal{X}$,
2. for each $E \in \mathcal{X}$, also the complement of E belongs to $\mathcal{X}$, i.e., $E^c = X \backslash E \in \mathcal{X}$,
3. for each sequence of subsets E_n in $\mathcal{X}$, i.e., $(E_n)_{n=1}^{\infty} \in \mathcal{X}^{\mathbb{N}}$, also their union belongs to $\mathcal{X}$, that is, $\bigcup\limits_{n=1}^{\infty} E_n \in \mathcal{X}$.

Remark 8. *Let $\mathcal{X}$ be a sigma-algebra on X. Properties 1 and 2 imply that $\emptyset \in \mathcal{X}$. For every sequence of subsets $(E_n)_{n=1}^{\infty}$,*

$$\left(\bigcup_{n=1}^{\infty} E_n\right)^c = \bigcap_{n=1}^{\infty} (E_n)^c;$$

if follows that, by properties 2 and 3, if $(E_n)_{n=1}^{\infty} \in \mathcal{X}^{\mathbb{N}}$ *then* $\bigcap_{n=1}^{\infty} E_n \in \mathcal{X}$*.*

Example 9. The collections $\{X, \emptyset\}$ and 2^X are sigma-algebras on X respectively called the **trivial** and the **discrete** sigma-algebra. ▲

Remark 9. *Let J be an arbitrary index set and consider the indexed family $\{\mathcal{X}_j\}_{j\in J}$ of sigma-algebras on X. Then also the intersection $\bigcap_{j\in J} \mathcal{X}_j$ is a sigma-algebra on X.*

A pair $(X, \mathcal{X})$ where $\mathcal{X}$ is a sigma-algebra on X is called a **measurable space**. A **probability measure** on a measurable space $(X, \mathcal{X})$ is a function $\mu : \mathcal{X} \to [0,1]$ such that

- $\mu(X) = 1$,
- for every sequence $(E_n)_{n=1}^{\infty} \in \mathcal{X}^{\mathbb{N}}$ of pairwise disjoint events

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

Remark 10. *For every probability measure μ on a measurable space $(X, \mathcal{X})$ the following holds:*
(1) $\mu(\emptyset) = 0$;
(2) if $(E_n)_{n=1}^{\infty} \in \mathcal{X}^{\mathbb{N}}$ is monotone increasing ($E_n \subseteq E_{n+1}$ for every $n \in \mathbb{N}$), then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n);$$

(3) if $(E_n)_{n=1}^{\infty} \in \mathcal{X}^{\mathbb{N}}$ is monotone decreasing ($E_{n+1} \subseteq E_n$ for every $n \in \mathbb{N}$), then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

Borel Sigma-Algebras and Probability Measures Let $X \subseteq \mathbb{R}^m$ be a nonempty compact subset.[35] In this case, the natural sigma-algebra of

[35] All the definitions and results stated in this appendix for compact subsets of Euclidean spaces also apply for compact subsets of metric spaces.

events to be considered is the so called **Borel sigma-algebra**, that is, the intersection of all the sigma-algebras containing all the closed subsets of X (or, equivalently, all the sets of the form $X \backslash C$ with C closed). Since 2^X is a sigma-algebra containing all the closed subsets, such intersection is well defined. We let $\mathcal{B}(X)$ denote the Borel sigma-algebra on any compact subset X of a Euclidean space. Whenever we consider a compact subset X of a Euclidean space, it is understood that the relevant measurable space is $(X, \mathcal{B}(X))$ and that the relevant probability measures are those with domain $\mathcal{B}(X)$, i.e., the Borel probability measures. The set of Borel probability measures on X is denoted $\Delta(X)$.

For every $\mu \in \Delta(X)$, with $X \subseteq \mathbb{R}^m$, we can define the corresponding cumulative distribution function (cdf) $F_\mu : \mathbb{R}^m \to [0,1]$ as follows:

$$\forall \hat{x} \in \mathbb{R}^m,\ F_\mu(\hat{x}) = \mu(\{x \in X : x_1 \leq \hat{x}_1, ..., x_m \leq \hat{x}_m\}).$$

Given F_μ, we can recover μ starting with the probability of events of the form $E = X \cap \bigtimes_{j=1}^m [\underline{x}_j, \overline{x}_j]$, that is

$$\mu(E) = \int_{\underline{x}_1}^{\overline{x}_1} \cdots \int_{\underline{x}_m}^{\overline{x}_m} \mathrm{d}F_\mu(x_1, \ldots, x_m);$$

then, intuitively, one can compute the probabilities of the unions, intersections, and complements of events whose probability has already been obtained.

Example 10. Let $X = [0,1]$. When the probability assigned to every interval is its length, so that—for example—$\mu((a,b]) = b - a$, then we obtain the uniform probability measure, or Lebesgue measure. If $\mu((a,b]) = F(b) - F(a)$ for some right-continuous non-decreasing function $F : \mathbb{R} \to [0,1]$, then $\mu = \mu_F$ is the probability measure generated by cdf F, and $F = F_\mu$ is the cdf generated by μ. If F is differentiable with integrable derivative f, then

$$\mu((a,b]) = \int_a^b f(x)\mathrm{d}x.$$

▲

Expected Values and Convergence of Sequences of Probability Measures For every continuous function $\varphi : X \to \mathbb{R}$ defined on a compact subset X of a Euclidean space and every measure $\mu \in \Delta(X)$, the **expected value** of φ given μ

$$\mathbb{E}_\mu(\varphi) = \int_X \varphi(x)\,\mu(\mathrm{d}x)$$

is the Riemann-Stieltjes integral

$$\mathbb{E}_\mu(\varphi) = \int \varphi(x)\,\mathrm{d}F_\mu.$$

Fix a sequence of measures $(\mu_n)_{n=1}^\infty \in [\Delta(X)]^{\mathbb{N}}$ and a measure $\bar{\mu}$; we say that μ_n **converges (weakly)** to $\bar{\mu}$, written $\mu_n \to \bar{\mu}$, if

$$\lim_{n\to\infty} \mathbb{E}_{\mu_n}(\varphi) = \mathbb{E}_{\bar{\mu}}(\varphi)$$

for every continuous (hence bounded) function $\varphi : X \to \mathbb{R}$.

Example 11. Let $X = [0,1]$. Then $\mu_n \to \bar{\mu}$ if and only if $\lim_{n\to\infty} F_{\mu_n}(x) = F_{\bar{\mu}}(x)$ for every x at which $F_{\bar{\mu}}$ is continuous. ▲

With this, we can define closed sets in $\Delta(X)$ and a notion of continuity of real-valued functions defined on $\Delta(X)$. First, we stipulate that a subset $C \subseteq \Delta(X)$ is **closed** if it contains the limits of all converging sequences of measures in C; that is, for all $(\mu_n)_{n\in\mathbb{N}} \in [\Delta(X)]^{\mathbb{N}}$ and $\bar{\mu} \in \Delta(X)$ such that $\mu_n \to \bar{\mu}$ and $\{\mu_n\}_{n\in\mathbb{N}} \subseteq C$, $\bar{\mu} \in C$. We say that $\psi : \Delta(X) \to \mathbb{R}$ is **continuous** if $\psi^{-1}(C)$ is closed for every closed subset $C \subseteq \mathbb{R}$. This is equivalent to requiring that $\lim_{n\to\infty} \psi(\mu_n) = \psi(\bar{\mu})$ for every converging sequence $\mu_n \to \bar{\mu}$. (See Aliprantis and Border [3, Theorem 15.11] and related material.) Note, when we consider the subset of all the Dirac measures $\{\mu \in \Delta(X) : \exists x \in X, \mu = \delta_x\}$, which can be identified with X, we obtain the usual notions of closed sets in X and continuity of real-valued functions defined on X.

A set $D \subseteq \Delta(X)$ is **open** if $X \backslash D$ is closed. We say that a collection of open sets $\mathcal{O}$ **covers** $\Delta(X)$ if the union of all such subsets contains $\Delta(X)$:

$$\Delta(X) \subseteq \bigcup \mathcal{O}.$$

$\Delta(X)$ endowed with its collection of open subsets is **compact** in the following sense: for every collection of open subsets $\mathcal{O}$ that covers $\Delta(X)$

there is a *finite* sub-collection $\mathcal{F} \subseteq \mathcal{O}$ that also covers $\Delta(X)$. Furthermore, $\Delta(X)$ is a **metrizable space**: this means that there exists a function $d : \Delta(X) \times \Delta(X) \to \mathbb{R}_+$, called **compatible metric** (or distance), such that

- for all $\mu', \mu'' \in \Delta(X)$, $d(\mu', \mu'') = 0$ if and only if $\mu' = \mu''$;
- for all $\mu', \mu'', \mu''' \in \Delta(X)$, $d(\mu', \mu''') \leq d(\mu', \mu'') + d(\mu'', \mu''')$;
- for all $(\mu_n)_{n=1}^{\infty} \in [\Delta(X)]^{\mathbb{N}}$ and $\bar{\mu} \in \Delta(X)$, $\mu_n \to \bar{\mu}$ if and only if for every $\varepsilon > 0$ there is some $n_\varepsilon \in \mathbb{N}$ such that $d(\mu_n, \bar{\mu}) < \varepsilon$ for all $n \geq n_\varepsilon$.[36]

Given nonempty compact subsets of Euclidean spaces X and Y, we can now define in a natural way notions of openness of subsets of $\Delta(X) \times \Delta(Y)$ and $X \times \Delta(Y)$, and of (joint) continuity for real-valued functions defined on $\Delta(X) \times \Delta(Y)$ and $X \times \Delta(Y)$. We omit the details.

3.4.2 Best Replies in Compact-Continuous Games

As explained in Section 3.4.1, for every compact subset X of a Euclidean space (more generally, for every compact subset of a metric space), the set of Borel probability measures $\Delta(X)$ is metrizable and compact; see [3, Theorem 15.11]. Every sequence in a compact metrizable space has a convergent subsequence. Also recall that, by definition of convergence in $\Delta(X)$, for every continuous function $f : X \to \mathbb{R}$, if $\mu_n \to \mu$ then $\mathbb{E}_{\mu_n}(f) \to \mathbb{E}_{\mu}(f)$. Therefore:

Lemma 9. *Let $X \subseteq \mathbb{R}^m$ be compact and let $f : X \to \mathbb{R}$ be continuous. For every sequence $(\mu_n)_{n=1}^{\infty} \in [\Delta(X)]^{\mathbb{N}}$ there is a convergent subsequence $(\mu_{n_k})_{k=1}^{\infty}$ and $\lim_{k\to\infty} \mathbb{E}_{\mu_{n_k}}(f) = \mathbb{E}_{\bar{\mu}}(f)$, where $\bar{\mu} = \lim_{k\to\infty} \mu_{n_k}$.*

Also recall from Section 3.4.1 that if X and Y are compact subsets of Euclidean spaces, then $X \times \Delta(Y)$, or—equivalently—$\Delta(X) \times Y$, is a compact metrizable space. With this, we prove an important property of the best reply correspondence.

[36] Compactness of the metrizable space $\Delta(X)$ implies that it is also complete (every Cauchy sequence converges) and separable ($\Delta(X)$ contains a countable subset D such that, for every $\mu \in \Delta(X)$ there is a sequence $(\mu_n)_{n=1}^{\infty} \in D^{\mathbb{N}}$ such that $\mu_n \to \mu$); see, e.g., Aliprantis and Border [3, Theorem 3.28].

Lemma 10. *Fix a compact-continuous game G. For every player $i \in I$, and every closed subset $C_{-i} \subseteq A_{-i}$, the restriction $r_i|_{\Delta(C_{-i})}$ of the best reply correspondence to $\Delta(C_{-i})$ is non-empty-valued, has a closed graph, and a closed image, that is,*

$$\forall \mu^i \in \Delta(C_{-i}),\ r_i\left(\mu^i\right) \neq \emptyset,$$

the set

$$\mathrm{Gr}\left(r_i|_{\Delta(C_{-i})}\right) = \left\{\left(\mu^i, a_i\right) \in \Delta(C_{-i}) \times A_i : a_i \in r_i\left(\mu^i\right)\right\}$$

is closed in the compact metrizable space $\Delta(A_{-i}) \times A_i$, and the set $r_i(\Delta(C_{-i})) \subseteq A_i$ is closed.

Proof. Fix $\mu^i \in \Delta(C_{-i})$ arbitrarily. Since u_i is continuous on A, $u_i\left(\cdot, \mu^i\right) : A_i \to \mathbb{R}$ is continuous. By compactness of A_i, $\arg\max_{a_i \in A_i} u_i\left(a_i, \mu^i\right) \neq \emptyset$. Since $r_i\left(\mu^i\right) = \arg\max_{a_i \in A_i} u_i\left(a_i, \mu^i\right)$ (Corollary 1), $r_i\left(\mu^i\right) \neq \emptyset$.[37] Hence r_i is non-empty-valued.

To prove that $\mathrm{Gr}\left(r_i|_{\Delta(C_{-i})}\right)$ is closed, we show that, for every convergent sequence contained in $\mathrm{Gr}\left(r_i|_{\Delta(C_{-i})}\right)$, the limit of the sequence belongs to $\mathrm{Gr}\left(r_i|_{\Delta(C_{-i})}\right)$: Suppose that $\left(\mu_n^i, a_{i,n}\right) \to \left(\bar{\mu}^i, \bar{a}_i\right)$, where $\mu_n^i \in \Delta(C_{-i})$ and $a_{i,n} \in r_i\left(\mu_n^i\right)$ for every $n \in \mathbb{N}$, so that the range of the sequence is included in $\mathrm{Gr}\left(r_i|_{\Delta(C_{-i})}\right)$. Then,

$$\forall a_i \in A_i, \forall n \in \mathbb{N},\ u_i\left(a_{i,n}, \mu_n^i\right) \geq u_i\left(a_i, \mu_n^i\right).$$

By continuity of u_i, taking the limit as $n \to \infty$ for each fixed a_i, we obtain

$$\forall a_i \in A_i,\ u_i\left(\bar{a}_i, \bar{\mu}^i\right) \geq u_i\left(a_i, \bar{\mu}^i\right).$$

Therefore $\bar{a}_i \in r_i\left(\bar{\mu}^i\right)$ and $\bar{\mu}^i \in \Delta(C_{-i})$ (since $\Delta(C_{-i})$ is a closed subset of $\Delta(A_{-i})$), that is, $\left(\bar{\mu}^i, \bar{a}_i\right) \in \mathrm{Gr}\left(r_i|_{\Delta(C_{-i})}\right)$. To see that $r_i(\Delta(C_{-i}))$ is closed, fix arbitrarily a convergent sequence $(a_{i,n})_{n=1}^{\infty}$ such that $a_{i,n} \to \bar{a}_i$ and $\{a_{i,n}\}_{n=1}^{\infty} \subseteq r_i(\Delta(C_{-i}))$. We must show that $\bar{a}_i \in r_i(\Delta(C_{-i}))$. For every $n \in \mathbb{N}$ there is some conjecture $\mu_n^i \in \Delta(C_{-i})$ that justifies $a_{i,n}$ as a best reply: $a_{i,n} \in r_i\left(\mu_n^i\right)$. Since C_{-i} is a closed subset of a compact set, C_{-i} is compact. Hence also $\Delta(C_{-i})$ is compact and the sequence $\left(\mu_n^i\right)_{n=1}^{\infty}$

[37] Recall that we defined $r_i\left(\mu^i\right)$ as the set of mixed actions α_i that maxmize $u_i\left(\cdot, \mu^i\right) : \Delta(A_i) \to \mathbb{R}$ and are also Dirac measures, hence belong to A_i as well.

has a convergent subsequence $\left(\mu^i_{n_k}\right)_{k=1}^{\infty}$, with $\lim_{k\to\infty} \mu^i_{n_k} = \bar{\mu}^i \in \Delta\left(C_{-i}\right)$ (Lemma 9). Since $\mathrm{Gr}\left(r_i\right)$ is closed, $\lim_{k\to\infty}\left(\mu^i_{n_k}, a_{n_k}\right) \in \mathrm{Gr}\left(r_i\right)$, that is,

$$\lim_{k\to\infty} a_{n_k} = \lim_{n\to\infty} a_n = \bar{a}_i \in r_i\left(\bar{\mu}^i\right).$$

Since $\bar{\mu}^i \in \Delta\left(C_{-i}\right)$, we conclude that $\bar{a}_i \in r_i\left(\Delta\left(C_{-i}\right)\right)$ as desired. ∎

3.4.3 Proof of the Duality Lemma of Wald and Pearce

The structure of the argument is as follows: We first observe that a mixed action $\bar{\alpha}_i$ is justifiable if and only if

$$0 \geq \min_{\mu^i\in\Delta(A_{-i})} \max_{\alpha_i\in\Delta(A_i)} \left[u_i\left(\alpha_i, \mu^i\right) - u_i\left(\bar{\alpha}_i, \mu^i\right)\right].$$

Then we invoke the maxmin theorem, which states that we can invert the min and max operations in the expression above without changing the final value. Next we note that in the *max-min* formula the second operation can be replaced with a minimization with respect to deterministic conjectures, that is, co-players' action profiles. This finally yields the equivalence with the fact that $\bar{\alpha}_i$ is not dominated by any mixed action.

Since finite games can be regarded as compact-continuous, the proof applies to all finite games. Considering the case of finite games, the reader can understand the core of the argument without the distraction of measure-theoretic considerations.

For any given conjecture μ^i, mixed action $\bar{\alpha}_i$, and alternative mixed action α_i, $\left[u_i\left(\alpha_i, \mu^i\right) - u_i\left(\bar{\alpha}_i, \mu^i\right)\right]$ represents the strength of the "incentive to deviate" from $\bar{\alpha}_i$ to α_i. Thus, $\bar{\alpha}_i$ is justifiable if and only if the maximal incentive to deviate is non-positive for at least one μ^i, that is, the minimum of the maximal incentive to deviate is not positive. The following result and proof make the previous argument more formal.

Lemma 11. *Let $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ be a compact-continuous game. Then, for every $i \in I$ and $\bar{\alpha}_i \in \Delta(A_i)$, the following are equivalent:*

(J) $\exists \bar{\mu}^i \in \Delta\left(A_{-i}\right), \forall \alpha_i \in \Delta\left(A_i\right), u_i\left(\bar{\alpha}_i, \bar{\mu}^i\right) \geq u_i\left(\alpha_i, \bar{\mu}^i\right)$,

(ND) $0 \geq \min_{\mu^i\in\Delta(A_{-i})} \max_{\alpha_i\in\Delta(A_i)} \left[u_i\left(\alpha_i, \mu^i\right) - u_i\left(\bar{\alpha}_i, \mu^i\right)\right]$.

Proof. *Preliminary observations:*[38] Let $X \subseteq \mathbb{R}^m$ be compact, and let the function $f : X \to \mathbb{R}$ be continuous. Then $\Delta(X)$ is compact metrizable and the map $\mu \mapsto \mathbb{E}_\mu(f)$ is continuous. Therefore

$$\sup_{x\in X} f(x) = \max_{x\in X} f(x), \quad \sup_{\mu\in\Delta(X)} \mathbb{E}_\mu(f) = \max_{\mu\in\Delta(X)} \mathbb{E}_\mu(f).$$

If X and Y are compact subsets of Euclidean spaces and $\varphi : \Delta(X) \times \Delta(Y) \to \mathbb{R}$ is continuous on $\Delta(X) \times \Delta(Y)$ (a product of compact metrizable spaces), then the maps $\mu \mapsto \max_{\nu\in\Delta(X)} \varphi(\nu,\mu)$ and $\nu \mapsto \min_{\mu\in\Delta(Y)} \varphi(\nu,\mu)$ are continuous.[39] Finally, recall that the payoff function u_i is extended to $\Delta(A_i) \times \Delta(A_{-i})$ as follows:

$$u_i(\alpha_i, \mu^i) = \mathbb{E}_{\alpha_i\times\mu^i}(u_i(\cdot,\cdot)).$$

Now, a probability measure $\mu^i \in \Delta(A_{-i})$ is such that $u_i(\alpha_i,\mu^i) \leq u_i(\bar{\alpha}_i,\mu^i)$ for all $\alpha_i \in \Delta(A_i)$, if and only if

$$0 \geq \sup_{\alpha_i\in\Delta(A_i)} \left[u_i(\alpha_i,\mu^i) - u_i(\bar{\alpha}_i,\mu^i)\right] = \max_{\alpha_i\in\Delta(A_i)} \left[u_i(\alpha_i,\mu^i) - u_i(\bar{\alpha}_i,\mu^i)\right].$$

Since the right-hand side is continuous in μ^i on the compact set $\Delta(A_{-i})$, there is a conjecture μ^i that makes it non-positive if and only if its minimum with respect to μ^i is non-positive. Hence, (J) holds if and only if

$$0 \geq \min_{\mu^i\in\Delta(A_{-i})} \max_{\alpha_i\in\Delta(A_i)} \left[u_i(\alpha_i,\mu^i) - u_i(\bar{\alpha}_i,\mu^i)\right],$$

that is, if and only if (ND) holds. ■

We wish to relate justifiability (condition J) to lack of dominance. To make this connection, we are going to use an important result that can be proved independently, the **maxmin theorem**:

Maxmin Theorem. (Sion [61]) *Let X and Y be compact subsets of Euclidean spaces, and let $f : X \times Y \to \mathbb{R}$ be continuous on $X \times Y$; then*

$$\min_{\mu\in\Delta(X)} \max_{\nu\in\Delta(Y)} \mathbb{E}_{\mu\times\nu}(f) = \max_{\nu\in\Delta(Y)} \min_{\mu\in\Delta(X)} \mathbb{E}_{\mu\times\nu}(f).$$

We can now prove the duality lemma of Wald and Pearce stating that an action is justifiable if and only if it is not dominated:

[38] To be skipped when considering finite games.
[39] See Aliprantis and Border [3, Theorem 17.31].

Lemma 12. (Wald [69], Pearce [54]) *Let G be a compact-continuous game. Then, for every $i \in I$ and $\bar{\alpha}_i \in \Delta(A_i)$, the following are equivalent:*

(J) $\exists \bar{\mu}^i \in \Delta(A_{-i}), \forall \alpha_i \in \Delta(A_i), u_i(\bar{\alpha}_i, \bar{\mu}^i) \geq u_i(\alpha_i, \bar{\mu}^i)$,

(MM) $\forall \alpha_i \in \Delta(A_i), \exists a_{-i} \in A_{-i}, 0 \geq u_i(\alpha_i, a_{-i}) - u_i(\bar{\alpha}_i, a_{-i})$.

Proof. By Lemma 11, (J) is equivalent to

$$0 \geq \min_{\mu^i \in \Delta(A_{-i})} \max_{\alpha_i \in \Delta(A_i)} \left[u_i(\alpha_i, \mu^i) - u_i(\bar{\alpha}_i, \mu^i)\right]. \tag{3.4.1}$$

By the Maxmin Theorem,

$$\min_{\mu^i \in \Delta(A_{-i})} \max_{\alpha_i \in \Delta(A_i)} \left[u_i(\alpha_i, \mu^i) - u_i(\bar{\alpha}_i, \mu^i)\right] = \max_{\alpha_i \in \Delta(A_i)} \min_{\mu^i \in \Delta(A_{-i})} \left[u_i(\alpha_i, \mu^i) - u_i(\bar{\alpha}_i, \mu^i)\right].$$

Therefore, (3.4.1) is equivalent to

$$\forall \alpha_i \in \Delta(A_i),\ 0 \geq \min_{\mu^i \in \Delta(A_{-i})} \left[u_i(\alpha_i, \mu^i) - u_i(\bar{\alpha}_i, \mu^i)\right]. \tag{3.4.2}$$

Since $\left[u_i(\alpha_i, \mu^i) - u_i(\bar{\alpha}_i, \mu^i)\right]$ is affine in μ^i,

$$\min_{\mu^i \in \Delta(A_{-i})} \left[u_i(\alpha_i, \mu^i) - u_i(\bar{\alpha}_i, \mu^i)\right] = \min_{a_{-i} \in A_{-i}} \left[u_i(\alpha_i, a_{-i}) - u_i(\bar{\alpha}_i, a_{-i})\right]$$

for every $\alpha_i \in \Delta(A_i)$. Therefore, (3.4.2) is equivalent to

$$\forall \alpha_i \in \Delta(A_i),\ 0 \geq \min_{a_{-i} \in A_{-i}} \left[u_i(\alpha_i, a_{-i}) - u_i(\bar{\alpha}_i, a_{-i})\right],$$

which is equivalent to condition (MM). The previous chain of equivalences proves the claim. ■

3.4.4 Cautiously Justifiable and Admissible Actions

Proof of Lemma 4. We must prove that, in a compact-continuous game, $r_i(\Delta^o(A_{-i})) \subseteq NWD_i$. By way of contraposition, we show that if a_i^* is weakly dominated, then it is not a best reply to any full-support conjecture, that is, $a_i^* \notin r_i\left(\Delta^o(A_{-i})\right)$. Thus, suppose that a_i^* is weakly dominated by some $\alpha_i \in \Delta(A_i)$, that is,

$$\begin{aligned} \forall a_{-i} &\in A_{-i},\ u_i(\alpha_i, a_{-i}) \geq u_i(a_i^*, a_{-i}), \\ \exists \bar{a}_{-i} &\in A_{-i},\ u_i(\alpha_i, \bar{a}_{-i}) > u_i(a_i^*, \bar{a}_{-i}). \end{aligned} \tag{3.4.3}$$

Let $\epsilon = u_i(\alpha_i, \bar{a}_{-i}) - u_i(a_i^*, \bar{a}_{-i}) > 0$. Continuity of u_i implies that the function

$$a_{-i} \mapsto u_i(\alpha_i, a_{-i}) - u_i(a_i^*, a_{-i}),$$

is continuous as well. Therefore, there exists $\eta > 0$ such that

$$\forall a_{-i} \in A_{-i},\ d\left(a_{-i}, \bar{a}_{-i}\right) < \eta \Rightarrow u_i(\alpha_i, a_{-i}) - u_i(a_i^*, \bar{a}_{-i}) > \frac{\epsilon}{2}$$

(d denotes the standard Euclidean distance). Note that the (relatively) open set $O_{-i} = \{a_{-i} \in A_{-i} : d\left(a_{-i}, \bar{a}_{-i}\right) < \eta\}$ is assigned strictly positive probability by every full-support probability measure μ^i.

So, fix any $\mu^i \in \Delta^o(A_{-i})$. Then, the definition of O_{-i} and eq.s (3.4.3) yield

$$\begin{aligned} u_i(\alpha_i, \mu^i) - u_i(a_i^*, \mu^i) &= \int_{A_{-i}} \left[u_i(\alpha_i, a_{-i}) - u_i(a_i^*, a_{-i})\right] \mu^i\left(\mathrm{d}a_{-i}\right) \\ &= \int_{O_{-i}} \left[u_i(\alpha_i, a_{-i}) - u_i(a_i^*, a_{-i})\right] \mu^i\left(\mathrm{d}a_{-i}\right) \\ &\quad + \int_{A_{-i} \setminus O_{-i}} \left[u_i(\alpha_i, a_{-i}) - u_i(a_i^*, a_{-i})\right] \mu^i\left(\mathrm{d}a_{-i}\right) \\ &\geq \int_{O_{-i}} \left[u_i(\alpha_i, a_{-i}) - u_i(a_i^*, a_{-i})\right] \mu^i\left(\mathrm{d}a_{-i}\right) \\ &> \frac{\epsilon}{2}\mu^i(O_{-i}) \\ &> 0, \end{aligned}$$

that is, $u_i(\alpha_i, \mu^i) > u_i(a_i^*, \mu^i)$. Since μ^i is arbitrary, we conclude that $a_i^* \notin r_i\left(\Delta^o(A_{-i})\right)$. ■

Proof of Lemma 5. The inclusion $r_i\left(\Delta^o(A_{-i})\right) \subseteq NWD_i$ holds by Lemma 4. We prove that if an action a_i^* of player i is not weakly dominated by any mixed action, then it is **cautiously justifiable**, i.e., justifiable by a full-support conjecture. With an argument similar to the proof of Lemma 2, we use Farkas' lemma to prove this statement by contraposition, that is, if a_i^* is not cautiously justifiable, then it is weakly dominated.

Fix i and a_i^* arbitrarily. Since the game is finite, we let $k = |A_i|$ and $m = |A_{-i}|$, and we label elements in A_i as $\{1, 2, ..., k\}$ and elements in A_{-i} as $\{1, 2, ..., m\}$. Then, we construct a $k \times m$ matrix U in which the (w, z)-th coordinate is given by

$$U_{w,z} = u_i\left(a_i^*, z\right) - u_i\left(w, z\right).$$

With this, $\Delta^o\left(A_{-i}\right)$ is a subset of $\mathbb{R}^m_{++}$, the strictly positive orthant of $\mathbb{R}^m$, and a_i^* is cautiously justifiable if and only if there exists $\mu^i \in \Delta^o\left(A_{-i}\right)$ such that $U\mu^i \geq \mathbf{0}$. This condition can be rewritten as follows: a_i^* is cautiously justifiable if and only if, for some $p > 0$, we can find $\mathbf{x} \in \mathbb{R}^m$ satisfying

$$M\mathbf{x} \geq \mathbf{c}, \tag{3.4.4}$$

where M is a $(k+m+1) \times m$ matrix and $\mathbf{c}$ is a $(k+m+1)$-dimensional vector[40] defined as follows:

$$M = \begin{bmatrix} U \\ I_m \\ \mathbf{1}_m^{\mathrm{T}} \end{bmatrix}, \mathbf{c} = \begin{bmatrix} \mathbf{0}_k \\ \mathbf{p}_m \\ 1 \end{bmatrix}.$$

Here, I_m denotes the m-dimensional identity matrix, that is an $m \times m$ matrix having 1s along the main diagonal and 0s everywhere else; $\mathbf{1}_\ell$ and $\mathbf{0}_\ell$ ($\ell = k$, or $\ell = m$) denote the ℓ-dimensional vectors having respectively 1s and 0s everywhere; $\mathbf{p}_m$ denotes the m-dimensional vector having $p > 0$ everywhere; from now on, the dimensionality indexes will be omitted.

To see why a_i^* is cautiously justifiable if and only there exists $\mathbf{x} \in \mathbb{R}^m$ such that (3.4.4) holds, note that matrix inequality (3.4.4) can be written as the system of inequalities

$$\begin{cases} U\mathbf{x} \geq \mathbf{0} \\ x_1 \geq p, ..., x_m \geq p \\ \sum_{z=1}^m x_z \geq 1 \end{cases}.$$

Note that, if $\sum_{z=1}^m x_z > 1$ instead of $\sum_{z=1}^m x_z = 1$, then $\mathbf{x} \notin \Delta^o\left(A_{-i}\right)$; but this is immaterial because we can normalize by substituting $\mathbf{x} =$

[40] Recall that, in matrix calculus, an n-dimensional vector is represented as an $n \times 1$ matrix (column vector), and by transposition we obtain a $1 \times n$ matrix (row vector).

$(x_1, ..., x_m)$ with $\mathbf{x}' = \left(\frac{x_1}{\sum_{z=1}^m x_z}, ..., \frac{x_m}{\sum_{z=1}^m x_z}\right) \in \Delta^o(A_{-i})$ and satisfy inequality (3.4.4), since $M\mathbf{x} \geq 0$ if and only if $M\mathbf{x}' \geq 0$.
Thus, if a_i^* is *not* cautiously justifiable, inequality (3.4.4) does *not* hold and Farkas' lemma implies the existence of some $\mathbf{y} \in \mathbb{R}^{k+m+1}$ such that

$$\begin{cases} \mathbf{y} \geq \mathbf{0} \\ \mathbf{y}^{\mathrm{T}}\mathbf{c} > 0 \\ \mathbf{y}^{\mathrm{T}}M = \mathbf{0}^{\mathrm{T}} \end{cases}.$$

By construction,

$$\mathbf{y}^{\mathrm{T}}\mathbf{c} = p\sum_{z=1}^{m} y_{k+z} + y_{k+m+1} > 0. \tag{3.4.5}$$

Furthermore, $\mathbf{y}^{\mathrm{T}}M = \mathbf{0}^{\mathrm{T}}$ implies that

$$\sum_{w=1}^{k} y_w \left[u_i\left(a_i^*, z\right) - u_i\left(w, z\right)\right] = -\left(y_{k+z} + y_{k+m+1}\right) \leq 0 \tag{3.4.6}$$

for every $z \in A_{-i} = \{1, 2, ..., m\}$, where the inequality holds because $y_{k+m+1} \geq 0$ and $y_{k+z} \geq 0$. Since $\mathbf{y} \geq \mathbf{0}$ and $p > 0$, (3.4.5) implies that there exists $z' \in \{1, 2, ..., m+1\}$ such that $y_{k+z'} > 0$. If $z' = m+1$, then the left-hand side of (3.4.6) is non-zero for every $z \in \{1, 2, ..., m\}$. If $z' \in \{1, 2, ..., m\}$, then the left-hand side of (3.4.6) is non-zero for $z = z'$. Since the vector $(y_1, ..., y_k)$ is nonnegative, we must have $y_w > 0$ for some $w \in A_i$. Then, we can construct a probability vector

$$\bar{\mathbf{y}} = \left(\frac{y_1}{\sum_{w=1}^k y_w}, ..., \frac{y_k}{\sum_{w=1}^k y_w}\right) \in \Delta(A_i) \subseteq \mathbb{R}^k$$

such that $u_i\left(a_i^*, z\right) < \sum_{w=1}^k \bar{y}_w u_i\left(w, z\right)$ for at least one $z \in A_{-i}$, and $u_i\left(a_i^*, z\right) \leq \sum_{w=1}^k \bar{y}_w u_i\left(w, z\right)$ for every $z \in A_{-i}$. We conclude that a_i^* is weakly dominated by mixed action $\bar{\mathbf{y}}$. ■

3.4.5 Increasing Best Reply Functions

We conclude this appendix by providing the proof of the first part of Lemma 8. The proof follows from an important result on supermodular functions:

Topkis' Maximum Theorem. (Topkis [65]) *Let $X \subseteq \mathbb{R}$ and $Y \subseteq \mathbb{R}^n$ be non-empty order boxes, and fix a* continuous supermodular *function $f : X \times Y \to \mathbb{R}$.*

Then, for each $y \in Y$, the set $\arg\max_{x \in X} f(x, y)$ has a greatest and a least element, which we denote by $M_f(y)$ and $m_f(y)$, respectively. In addition, the functions M_f and m_f are weakly increasing.

Proof. By assumption, X is an order box of $\mathbb{R}$, i.e., it is a (nonempty) compact interval. Let $\hat{f} : Y \rightrightarrows \mathbb{R}$ be the correspondence defined by

$$y \mapsto \hat{f}(y) = \arg\max_{x \in X} f(x, y).$$

Since f is continuous, then standard arguments (see the proof of Lemma 10) show that, for each $y \in Y$, the set $\hat{f}(y)$ is a closed (hence compact) subset of $X \subseteq \mathbb{R}$. Therefore $M_f(y) = \max \hat{f}(y)$ and $m_f(y) = \min \hat{f}(y)$ are well defined.

It remains to show that $M_f(a) \leq M_f(b)$ for all $a, b \in Y$ such that $a \leq b$. Define $\bar{a} = M_f(a)$ and $\bar{b} = M_f(b)$ so that, in particular, we have $f(\bar{a}, a) \geq f(\min(\bar{a}, \bar{b}), a)$. Since f is supermodular on $X \times Y$, we obtain

$$\begin{aligned} f(\bar{a}, a) + f(\bar{b}, b) &\leq f(\max(\big((\bar{a}, a), (\bar{b}, b)\big) + f(\min\big((\bar{a}, a), (\bar{b}, b)\big)) \\ &= f(\max(\bar{a}, \bar{b}), b) + f(\min(\bar{a}, \bar{b}), a). \end{aligned}$$

By the previous observation, it follows that $f(\bar{b}, b) \leq f(\max(\bar{a}, \bar{b}), b)$, so that $\max(\bar{a}, \bar{b}) \in \hat{f}(b)$. Therefore the conclusion that $\bar{a} \leq \bar{b}$ follows from the fact that $\bar{b}$ is the greatest element of $\hat{f}(b)$. The proof for m_f is similar. ■

Proof of the First Part of Lemma 8. Recall that, according to Remark 7, the best-reply correspondence r_i is actually a continuous function for each $i \in I$. The claim follows from Topkis' Maximum Theorem: it is enough to set $f = u_i$, $X = A_i$, and $Y = A_{-i}$, so that $r_i = M_f$, which is weakly increasing. ■

4

Rationalizability and Iterated Dominance

The analysis, so far, has been based on a set of minimal epistemic assumptions:[1] every player knows the sets of possible actions and his own payoff function. From this analysis we derived a basic, decision-theoretic principle of rationality: A rational player should not choose those actions that are dominated by mixed actions. This principle of rationality can be interpreted in a descriptive way (assuming that a rational player will not choose dominated actions), or in a normative way (a player should not choose dominated actions).

The concept of dominance is sufficient to obtain interesting results in some interactive situations: those where it is not necessary to guess the other players' actions in order to make a correct decision. Simple social dilemmas, like the Prisoners' Dilemma or contributing to a public good, have this feature. But when we analyze strategic reasoning, the decision-theoretic rationality principle is just a starting point: **thinking strategically** *means trying to anticipate the co-players' moves and plan a best response, taking into account that also the co-players are intelligent individuals trying to do just the same.*

Strategic thinking is based on knowledge of the rules of the game

[1] We call "epistemic" the assumptions about knowledge, conjectures, and beliefs of players.

and preferences, knowledge of the other players' knowledge of rules and preferences, and so on and so forth. As a starting point, we assume that players have *complete information*, i.e., common knowledge of the rules of the game and everybody's preferences. In Example 1 complete information implies that the functional forms and the parameters K, θ_1, θ_2 are common knowledge: both players know them, both players know that both players know them, and so on and so forth.

The complete information assumption is sufficiently realistic for some economic situations in which the consequences of players' actions are purely monetary outcomes and players are selfish and risk neutral.[2] The complete information assumption is also a useful simplification in that it allows to focus on other aspects of strategic thinking than players' knowledge of the game and of the knowledge of other players. Later on, we will introduce strategic thinking in games with *in*complete information and we will show that the techniques developed for games with complete information can be extended to the more general case.

4.1 Assumptions about Players' Beliefs

The analysis will proceed in a series of steps. After introducing the concept of rationality and its behavioral consequences, we will characterize which actions are consistent not only with rationality, but also with the belief that everybody is rational. Then we will characterize which actions are consistent not only with the previous assumptions, but also with the further assumption that each player believes that each other player believes that everybody is rational; and so on. Thus, at each further step we characterize more restrictive assumptions about players' rationality and beliefs.

Such assumptions can be formally represented as **events**, i.e., as subsets of a space Ω of "states of the world" where every state $\omega \in \Omega$ is a conceivable configuration of actions, conjectures, and beliefs concerning other players' beliefs. This formal representation provides an expressive and precise language that allows to make the analysis more rigorous

[2] Recall that what matters are players' preferences over *lotteries* of consequences, hence also their risk attitudes. But in some cases (for instance, some oligopolistic models) risk attitudes are irrelevant for the strategic analysis and risk neutrality can therefore be assumed with no loss of generality.

and transparent.[3] However, it calls for the preliminary introduction of some rather complex material, which is not strictly needed to understand the basics. Therefore we opt for a compromise. On the one hand, in order to make the presentation more transparent and to avoid verbose sentences, we use symbols to denote the assumptions concerning players' behavior and beliefs; we refer to these assumptions as "events" and denote a conjunction of assumptions by the symbol of intersection between events. On the other hand, rather than providing an explicit definition for such events, we only define mathematically the sets of actions (and conjectures) that are consistent with them. Of course, we can afford to do this precisely because we know how to represent assumptions about rationality and beliefs as events and how to derive characterizations of behavior based on increasingly sophisticated strategic thinking by means of rigorous mathematical methods. We ask you, the reader, to trust us on this derivation. Since the results are quite intuitive, we hope you will not object. (If you do object, perhaps you are ready for the complete mathematical analysis of strategic thinking: good news for us!)

To see what we mean, consider the rationality assumption. We denote by R_i the event "player i is rational," whereas $R = \bigcap_{i\in I} R_i$ denotes the event "everybody (in the game) is rational." In the previous chapter we saw that $\times_{i\in I} ND_i$ (where ND_i is the set of undominated actions of i) is the set of action profiles consistent with R (rationality).

The rationality assumption simply establishes a relation between conjectures (beliefs about the behavior of other players) and actions. This relation is represented by the best reply correspondence. Now we proceed introducing some assumptions concerning players' beliefs about each other beliefs, also called "**interactive beliefs**."

Denote by E a generic event about players' actions and beliefs (for example, $E = R$). We represent with the symbol $\mathrm{B}(E)$ the event "everybody believes E," meaning "everybody assigns probability one to E." Further, we write $\mathrm{B}(\mathrm{B}(E)) = \mathrm{B}^2(E)$ ("everybody believes that everybody believes E"), and more generally $\mathrm{B}^k(E) = \mathrm{B}(\mathrm{B}^{k-1}(E))$ for any integer $k > 1$.

Consider the conjunction of the following assumptions: R, $\mathrm{B}(R)$, $\mathrm{B}^2(R), \ldots, \mathrm{B}^k(R)$, etc. Is it possible to provide a characterization in terms

[3]See, for example, Battigalli and Bonanno [8], and Dekel and Siniscalchi [26].

of actions? In other words, what sets of action profiles are consistent with assumptions $R \cap \mathrm{B}(R)$, $R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R), \ldots, R \cap \left(\bigcap_{k=1}^{K} \mathrm{B}^k(R)\right)$, etc.? In order to answer this question, it is useful to introduce a mapping, which is derived from the best reply correspondence and assigns to any "cross-product" subset C of A a subset of action profiles that are "rationalized," or "justified" by C.

4.2 The Rationalization Operator

Fix some finite set X and a collection $\mathcal{C}$ of subsets of X, i.e., $\mathcal{C} \subseteq 2^X$ (we postpone the analysis of the case of an infinite X). In this textbook we call "operator" a map from $\mathcal{C}$ to itself, that is, a mapping that associates with each subset of X in $\mathcal{C}$ another subset of X in $\mathcal{C}$.[4]

As a matter of notation, for a fixed set X_{-i} (for instance, $X_{-i} = A_{-i}$) and a subset $Y_{-i} \subseteq X_{-i}$, we denote by $\Delta(Y_{-i})$ the subset of probability measures on X_{-i} that assign probability one to Y_{-i}:

$$\Delta(Y_{-i}) = \{\mu^i \in \Delta(X_{-i}) : \mu^i(Y_{-i}) = 1\}$$

(another slight abuse of notation).

In particular, we are going to consider the set $X = A = \times_{i \in I} A_i$ of action profiles, and the collection $\mathcal{C}$ of all "Cartesian" subsets of A, i.e., subsets with the cross-product form $C = \times_{i \in I} C_i$, where $C_i \subseteq A_i$ for every i. Then, if the game is finite,

$$\mathcal{C} = \{C \in 2^A : \exists\, (C_i)_{i \in I} \in \bigtimes_{i \in I} 2^{A_i}, C = \bigtimes_{i \in I} C_i\}.$$

Let $C_{-i} = \times_{j \in I \setminus \{i\}} C_j$. We define the following sets:

$$\rho_i(C_{-i}) = \left\{a_i \in A_i : \exists \mu^i \in \Delta(C_{-i}), a_i \in r_i(\mu^i)\right\} = r_i(\Delta(C_{-i})),$$

$$\rho(C) = \bigtimes_{i \in I} \rho_i(C_{-i}).$$

[4] In mathematics, the term "operator" is mostly used for maps from a space of functions to a space of functions (possibly the same). But sets can always be represented by functions (e.g., indicator functions). Therefore the present use of the term "operator" is consistent with standard mathematical language.

The interpretation is as follows: $\rho(C)$ *is the set of action profiles that could be chosen by rational players if every i is certain that the co-players choose in C_{-i}.* Therefore, we say that $\rho(C)$ is the set of action profiles "rationalized" (or "justified") by C and we call the mapping $\rho : \mathcal{C} \to \mathcal{C}$ "**rationalization operator**."[5] Note that $\rho(\emptyset) = \emptyset$.

	ℓ	c	r
T	$3, 2$	$0, 1$	$0, 0$
M	$0, 2$	$3, 1$	$0, 0$
B	$1, 1$	$1, 2$	$3, 0$

Figure 4.1: A 3×3 game.

Example 12. In a 3×3 game like the one in Figure 4.1, each player has $2^3 = 8$ subsets of actions (including the empty set). Hence, the collection of subsets $\mathcal{C}$ contains $8 \times 8 = 64$ elements, $7 \times 7 = 49$ of them are nonempty (since $\rho(\emptyset) = \emptyset$, we may ignore the empty subsets). It would be very tedious to determine $\rho(C)$ for each (nonempty) $C \in \mathcal{C}$. We give just a few examples:

$$
\begin{aligned}
\rho(\{T, M, B\} \times \{\ell, c, r\}) &= \{T, M, B\} \times \{\ell, c\}, \\
\rho(\{T, M, B\} \times \{\ell, c\}) &= \{T, M\} \times \{\ell, c\}, \\
\rho(\{T, M\} \times \{\ell, c\}) &= \{T, M\} \times \{\ell\}, \\
\rho(\{M, B\} \times \{\ell, c\}) &= \{T, M\} \times \{\ell, c\}, \\
\rho(\{T, M\} \times \{\ell\}) &= \{T\} \times \{\ell\}, \\
\rho(\{T, M\} \times \{r\}) &= \{B\} \times \{\ell\}, \\
\rho(\{T\} \times \{\ell\}) &= \{T\} \times \{\ell\}, \\
\rho(\{B\} \times \{r\}) &= \{B\} \times \{c\}.
\end{aligned}
$$

To see this, check the best replies to deterministic conjectures, note that r is strictly dominated by l and c, next note that B is strictly dominated by $\alpha_1 = \frac{1}{2}\delta_T + \frac{1}{2}\delta_M$ in the smaller game obtained by deleting the right

[5] The rationalization operator represents an example of **justification operator**, a concept which was first explicitly presented by Milgrom and Roberts [46].

column r. This shows that, in general, we may have $\rho(C) \subseteq C$, $C \subseteq \rho(C)$ (here this inclusion holds as an equality for $C = \{T\} \times \{l\}$), $\rho(C) \subsetneq C$ and $C \subsetneq \rho(C)$.[6] ▲

Remark 11. *Lemma 2 implies that $\rho_i(A_{-i})$ is the set of undominated actions for player i. Hence, $\rho(A) = \bigtimes_{i\in I} ND_i$.*

Remark 12. *The rationalization operator is* monotone*: for every pair of subsets $E, F \in \mathcal{C}$, if $E \subseteq F$ then $\rho(E) \subseteq \rho(F)$. To see this note that, for every i, $E_{-i} \subseteq F_{-i}$ implies $\Delta(E_{-i}) \subseteq \Delta(F_{-i})$, which in turn implies $\rho_i(E_{-i}) = r_i(\Delta(E_{-i})) \subseteq r_i(\Delta(F_{-i})) = \rho_i(F_{-i})$.*

Since the domain and codomain of ρ coincide, it makes sense to define the k-th iteration of ρ (the k-fold composition of ρ with itself) recursively as follows. For each $C \in \mathcal{C}$, define $\rho^0(C) = C$ for convenience; then for each $k \geq 1$,

$$\rho^k(C) = \rho(\rho^{k-1}(C)).$$

Note that we are using the standard definition of the iteration of a self-map $f : X \to X$, that is, $f^0 = \mathrm{Id}_X$ (the identity function on X) by convention, and $f^k = f \circ f^{k-1}$ for each $k \geq 1$. In this case $X = \mathcal{C}$ and $f = \rho$.

Rationalization in Infinite Games The analysis extends seamlessly from finite to compact-continuous games. To be completely general, we should define the rationalization operator for all the (Borel) measurable Cartesian subsets of A. But we can restrict our attention to the closed Cartesian subsets of A. In fact, compactness of A and continuity of the payoff functions imply that, for every $i \in I$ and every closed subset $C_{-i} \subseteq A_{-i}$, the set of conjectures $\Delta(C_{-i}) = \left\{\mu^i \in \Delta(A_{-i}) : \mu^i(C_{-i}) = 1\right\}$ is well defined[7] and that the set of best replies $r_i(\Delta(C_{-i}))$ is closed (see Lemma 14 in Appendix 4.7). Note that a Cartesian product $C = \bigtimes_{i\in I} C_i$ is closed if and only if each "factor" C_i is closed. Therefore, whenever $C = \bigtimes_{i\in I} C_i$ is closed, also $\rho(C) = \bigtimes_{i\in I} r_i(\Delta(C_{-i}))$ is closed, and the same holds for each $\rho^k(C)$ ($k \in \mathbb{N}$). For each $i \in I$, let $\mathcal{K}_{A_i} \subseteq 2^{A_i}$ denote the collection of *closed* (hence compact) subsets of A_i, including A_i itself.

[6] Symbol "$\subsetneq$" means "not weakly included in."

[7] Because the closed C_{-i} is necessarily measurable, hence the probability $\mu^i(C_{-i})$ is well defined.

With this, define the collection of closed Cartesian products

$$\mathcal{C} = \left\{ C \in 2^A : \exists \left(C_i\right)_{i\in I} \in \bigtimes_{i\in I} \mathcal{K}_{A_i}, C = \bigtimes_{i\in I} C_i \right\}.$$

Then, for each $C \in \mathcal{C}$, $\rho(C) = \bigtimes_{i\in I} r_i\left(\Delta\left(C_{-i}\right)\right)$ is well defined and $\rho(C) \in \mathcal{C}$. Therefore, $\rho : \mathcal{C} \to \mathcal{C}$ is a well-defined self-map.

4.3 Rationalizability: The Powers of ρ

The set of action profiles consistent with R (rationality of all the players) is $\rho(A)$. Therefore, if every player is rational and believes R, only those action profiles that are rationalized by $\rho(A)$ can be chosen. It follows that the set of action profiles consistent with $R \cap \mathrm{B}(R)$ is $\rho(\rho(A)) = \rho^2(A)$. Iterating the procedure one more step, it is relatively easy to see that the set of action profiles consistent with $R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$ is $\rho(\rho^2(A)) = \rho^3(A)$. The general relationship between events about rationality and beliefs and corresponding sets of action profiles is given by the Table 4.1.

Table 4.1: Behavioral implications of assumptions on rationality and beliefs.

Assumptions about rationality and beliefs	Behavioral implications
R	$\rho(A)$
$R \cap \mathrm{B}(R)$	$\rho^2(A)$
$R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$	$\rho^3(A)$
...	...
$R \cap \left(\bigcap_{k=1}^{K} \mathrm{B}^k(R)\right)$	$\rho^{K+1}(A)$
...	...

Note that, as one should expect, the sequence of subsets $(\rho^k(A))_{k=1}^{\infty}$ is weakly decreasing, i.e., $\rho^{k+1}(A) \subseteq \rho^k(A)$ $(k \in \mathbb{N})$. This fact can be easily derived from the monotonicity of the rationalization operator: by definition, $\rho^1(A) = \rho(A) \subseteq A = \rho^0(A)$; if $\rho^k(A) \subseteq \rho^{k-1}(A)$, the monotonicity of ρ implies $\rho^{k+1}(A) = \rho(\rho^k(A)) \subseteq \rho(\rho^{k-1}(A)) = \rho^k(A)$. By the induction principle $\rho^{k+1}(A) \subseteq \rho^k(A)$ for every k. Every monotonically decreasing sequence of subsets has a well-defined limit: the intersection of all the sets in the sequence. Therefore it makes sense to define:

$$\rho^\infty(A) = \bigcap_{k\geq 1} \rho^k(A).$$

Example 13. In the game of Example 12, iterating ρ starting from A we delete one action at each step and we stop after 4 steps with one action left for each player:

$$\begin{aligned}
\rho(A) &= \rho(\{T,M,B\}\times\{\ell,c,r\}) = \{T,M,B\}\times\{\ell,c\},\\
\rho^2(A) &= \rho(\{T,M,B\}\times\{\ell,c\}) = \{T,M\}\times\{\ell,c\},\\
\rho^3(A) &= \rho(\{T,M\}\times\{\ell,c\}) = \{T,M\}\times\{\ell\},\\
\rho^k(A) &= \rho(\{T,M\}\times\{\ell\}) = \{T\}\times\{\ell\},\ (\forall k\geq 4).
\end{aligned}$$

▲

It is easy to see that $\rho^\infty(A)$ may contain more than one element:

	b	s	c
B	$4,3$	$0,2$	$0,0$
S	$0,1$	$3,4$	$0,0$
C	$1,1$	$1,2$	$5,0$

Figure 4.2: A modified Battle of the Sexes.

Example 14. Here is a story for the payoff matrix in Figure 4.2: Rowena and Colin have to decide independently of each other whether to go to a Bach concert, a Stravinsky concert, or a Chopin concert. Rowena likes Bach more than Stravinsky, she loves Chopin, but she prefers to go to another concert with Colin rather than listening to Chopin alone. Colin *hates* Chopin, he likes Stravinsky more than Bach, but he prefers Bach with Rowena to Stravinsky alone. This is a variation on the "Battle of the Sexes." Iterating ρ from $A = \{B,S,C\}\times\{b,s,c\}$ we get

$$\begin{aligned}
\rho(A) &= \rho(\{B,S,C\}\times\{b,s,c\}) = \{B,S,C\}\times\{b,s\},\\
\rho^2(A) &= \rho(\{B,S,C\}\times\{b,s\}) = \{B,S\}\times\{b,s\},\\
\rho^k(A) &= \rho(\{B,S\}\times\{b,s\}) = \{B,S\}\times\{b,s\},\ (\forall k\geq 3).
\end{aligned}$$

Strategic thinking leads Rowena to avoid the Chopin concert, but it does not allow to solve the Bach-Stravinsky coordination problem. ▲

We argued informally that an action profile a is consistent with rationality and common belief in rationality if and only if $a \in \rho^{\infty}(A)$. Such action profiles are called "rationalizable":

Definition 13. *(Bernheim [21], Pearce [54]) An action profile $a \in A$ is **rationalizable** if $a \in \rho^{\infty}(A)$.*

In finite games, the set of rationalizable action profiles is obtained in a finite number of steps (the number depends on the game). In infinite games with compact action spaces and continuous payoff functions the set of rationalizable profile is obtained with countably many steps (i.e., finitely many, or a denumerable sequence of steps).

Theorem 2. *(a) If G is a finite game, then there exists a positive integer K such that $\rho^{K+1}(A) = \rho^{K}(A) = \rho^{\infty}(A) \neq \emptyset$. (b) If G is a compact-continuous game, then the set of rationalizable action profiles, $\rho^{\infty}(A)$, is nonempty, compact, and satisfies $\rho^{\infty}(A) = \rho\left(\rho^{\infty}(A)\right)$.*

Proof. For the proof of part (b), which contains measure-theoretic arguments, see Appendix 4.7. Here we prove only part (a). If A_i is finite, for every conjecture μ^i the set of best replies $r_i(\mu^i)$ is nonempty. Then for each nonempty Cartesian set $C \subseteq A$, $\rho(C)$ is nonempty (for every i, there exists at least a conjecture $\mu^i \in \Delta(C_{-i})$ with $\emptyset \neq r_i(\mu^i) \subseteq \rho_i(C_{-i})$). Thus, $\rho^k(A) \neq \emptyset$ implies $\rho^{k+1}(A) \neq \emptyset$. Since $A \neq \emptyset$, it follows by induction that $\rho^k(A) \neq \emptyset$ for every k.

The sequence of subsets $(\rho^k(A))_{k=1}^{\infty}$ is weakly decreasing. Also, if $\rho^k(A) = \rho^{k+1}(A)$, then $\rho^k(A) = \rho^{\ell}(A)$ for every $\ell \geq k$. Given that A is finite, the inclusion $\rho^{k+1}(A) \subseteq \rho^k(A)$ can be strict only for a finite number of steps K (in particular, $K \leq \sum_{i \in I}(|A_i| - 1)$, where $|A_i|$ denotes the number of elements of A_i: when $\rho^{k+1}(A) \subset \rho^k(A)$, at least one action for at least one player i is eliminated, but at least one action for each player is never eliminated). All the above implies that $\rho^{K+1}(A) = \rho^{K}(A) = \rho^{\infty}(A) \neq \emptyset$. ■

It is possible to give an alternative and useful characterization of rationalizable actions. Consider the following:

Definition 14. *A set $C \in \mathcal{C}$ has the **best reply property** if $C \subseteq \rho(C)$.*

Remark 13. *By Theorem 2, the set $\rho^\infty(A)$ of rationalizable action profiles of a finite or compact-continuous game has the best reply property.*

We leave the proof of the following as an exercise:

Remark 14. *For every pair of subsets $E, F \in \mathcal{C}$, if both E and F have the best reply property then also $C = \times_{i\in I}(E_i \cup F_i)$ has the best reply property.*

Theorem 3. *In every finite or compact-continuous game, an action profile $a \in A$ is rationalizable if and only if $a \in C$ for some subset $C \in \mathcal{C}$ with the best reply property.*

Proof. (Only if) If a is rationalizable then $a \in \rho^\infty(A) = \rho(\rho^\infty(A))$, where the equality follows from Theorem 2. Hence, a belongs to a set with the best reply property, viz., $\rho^\infty(A)$.

(If) Let $a \in C \subseteq \rho(C)$ for some $C \in \mathcal{C}$. We will prove by induction that $C \subseteq \rho^k(C) \subseteq \rho^k(A)$ for each k. Therefore $a \in C \subseteq \bigcap_{k\in\mathbb{N}} \rho^k(A) = \rho^\infty(A)$, i.e., a is rationalizable.

Basis step. Since $C \subseteq \rho(C)$, $C \subseteq A$ and ρ is monotone, it follows that $C \subseteq \rho^1(C) \subseteq \rho^1(A)$.

Inductive step. Suppose that $C \subseteq \rho^k(C) \subseteq \rho^k(A)$. By monotonicity

$$\rho(C) \subseteq \rho(\rho^k(C)) \subseteq \rho(\rho^k(A)).$$

Since $C \subseteq \rho(C)$ and $\rho(\rho^k(\cdot)) = \rho^{k+1}(\cdot)$, we obtain

$$C \subseteq \rho^{k+1}(C) \subseteq \rho^{k+1}(A).$$

■

Remark 15. *The proof of Theorem 3 clarifies that, in every finite or compact-continuous game, for every $C \in \mathcal{C}$ with the best reply property, $C \subseteq \rho^\infty(A)$. Since Theorem 2 implies that $\rho^\infty(A)$ has the best reply property, then $\rho^\infty(A)$ must be the largest set with the best reply property.*[8]

We gave a definition of rationalizability based on iterations of the rationalization operator ρ because it is the most intuitive. An alternative

[8] More precisely, given the partially ordered collection $(\mathcal{C}, \subset)$, $\rho^\infty(A)$ is the unique maximal set with the best reply property. In other words, $\rho^\infty(A)$ is the unique set $C \in \mathcal{C}$ such that $C \subseteq \rho(C)$ and, for every $D \in \mathcal{C}$, $C \subset D$ implies $D \nsubseteq \rho(D)$.

definition of rationalizability is suggested by Theorem 3: a is rationalizable if it belongs to some set $C \in \mathcal{C}$ with the best reply property. This second definition is equivalent to the first one for all games where $\rho^{\infty}(A) = \rho(\rho^{\infty}(A))$. But there are some "badly behaved" infinite games where $\rho(\rho^{\infty}(A)) \subset \rho^{\infty}(A)$ (where $\subset$ means *strict* inclusion). The second definition of rationalizability is valid for those games too, while the first one only gives a necessary condition. For the purposes of these lectures, this "technicality" can be neglected. We used the first definition because we find it more intuitive.

4.4 Comparative Risk Aversion and Rationalizability

We have shown in Theorem 1 of Chapter 3 that higher risk aversion implies a larger set of justifiable actions. An inductive argument extends such monotonicity result to the set of rationalizable action profiles: a weak increase in risk aversion for all players weakly expands the set of rationalizable actions of each player.[9] The interesting novelty is that the set of rationalizable actions of a player may expand just because the risk aversion of *another* player increases as shown by the following example:

Example 15. Consider the following two-person game form with monetary consequences:

$g_1, g_2:$	b'	b''
a'	0, 1	1, 0
a''	$\frac{1}{3}$,0	$\frac{1}{3}$,1
a'''	1, 1	0, 0

Let $u_{\theta,1} = g_1^{\frac{1}{\theta}}$, where $\theta \geq 1$ parametrizes risk aversion, and let $u_2 = v_2 \circ g_2$ for any continuous and strictly increasing v_2 (the risk attitudes of player 2 are immaterial). We consider the rationalizability correspondence $\theta \mapsto \rho_{\theta}^{\infty}(A)$. Note that every action of player 2 is a best response to some belief, and the set of rationalizable actions of player 2 is $\{b', b''\}$ if a'' is

[9] See Weinstein [73], who also consider other solution concepts. For an analogous result about "ambiguity aversion" see the working paper version of Battigalli et al. [17, Proposition 1].

justifiable for player 1, and $\{b'\}$ if a'' is unjustifiable. In the latter case, the only rationalizable action of player 1 is a''', the best reply to b'. With this, the calculations of Example 6 imply

$$\rho_\theta^\infty(A) = \rho_\theta^3(A) = \begin{cases} \{(a''', b')\}, & \text{if } 1 \le \theta < \log_2 3, \\ A_1 \times A_2, & \text{if } \theta \ge \log_2 3. \end{cases}$$

▲

Theorem 4. *Let $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ and $\hat{G} = \langle I, (A_i, \hat{u}_i)_{i\in I}\rangle$ be two compact-continuous games, and suppose every player $i \in I$ is more risk averse in $\hat{G}$ than in G. Then the set of rationalizable profiles of G is included in the set of rationalizable profiles of $\hat{G}$.*

Proof. Let A_i^k and $\hat{A}_i^k$ respectively denote the sets of k-step rationalizable actions of i in G and $\hat{G}$. We prove by induction that $A_i^k \subseteq \hat{A}_i^k$ for every $i \in I$ and $k \in \mathbb{N}$.

Basis step. By Theorem 1, $A_i^1 = r_i(\Delta(A_{-i})) \subseteq \hat{r}_i(\Delta(A_{-i})) = \hat{A}_i^1$ for every $i \in I$, where $\hat{r}_i$ is i's best reply correspondence in game $\hat{G}$, that is, according to payoff function $\hat{u}_i$.

Inductive Step. Suppose, by way of induction, that $A_j^k \subseteq \hat{A}_j^k$ for every $j \in I$ and fix any $i \in I$, then $A_{-i}^k \subseteq \hat{A}_{-i}^k$. Since $A_i^{k+1} = r_i\left(\Delta\left(A_{-i}^k\right)\right)$ and $\hat{A}_i^{k+1} = \hat{r}_i\left(\Delta\left(\hat{A}_{-i}^k\right)\right)$, we must show that $r_i\left(\Delta\left(A_{-i}^k\right)\right) \subseteq \hat{r}_i\left(\Delta\left(\hat{A}_{-i}^k\right)\right)$. By Theorem 1, $r_i\left(\Delta\left(A_{-i}^k\right)\right) \subseteq \hat{r}_i\left(\Delta\left(A_{-i}^k\right)\right)$. Since $A_{-i}^k \subseteq \hat{A}_{-i}^k$ by the inductive hypothesis, then also $\Delta\left(A_{-i}^k\right) \subseteq \Delta\left(\hat{A}_{-i}^k\right)$ and $\hat{r}_i\left(\Delta\left(A_{-i}^k\right)\right) \subseteq \hat{r}_i\left(\Delta\left(\hat{A}_{-i}^k\right)\right)$. Therefore $r_i\left(\Delta\left(A_{-i}^k\right)\right) \subseteq \hat{r}_i\left(\Delta\left(A_{-i}^k\right)\right) \subseteq \hat{r}_i\left(\Delta\left(\hat{A}_{-i}^k\right)\right)$, which implies $r_i\left(\Delta\left(A_{-i}^k\right)\right) \subseteq \hat{r}_i\left(\Delta\left(\hat{A}_{-i}^k\right)\right)$. ■

4.5 Iterated Dominance

The equivalence between best replies and undominated actions allows us to characterize the actions that survive the iterated dominance procedure. In order to give the precise definition of this procedure, we first define the concept of dominance within a subset of action profiles.

Definition 15. *Fix a nonempty Cartesian subset C of A: $\emptyset \neq C \in \mathcal{C}$. An action a_i is **dominated in** C if $a_i \in C_i$ and there is a mixed action $\alpha_i \in \Delta(C_i)$ such that*

$$\forall a_{-i} \in C_{-i},\ u_i(a_i, a_{-i}) < u_i(\alpha_i, a_{-i}).$$

We denote by $\mathrm{ND}_i(C) \subseteq A_i$ the set of actions in C_i that are not dominated *in C and let $\mathrm{ND}(C) = \bigtimes_{i \in I} \mathrm{ND}_i(C) \subseteq A$ denote the set of undominated action profiles in C.*

Remark 16. *Operator* ND *is not monotone (can you explain why?).*

Using the standard notation for iterations (k-fold composition) of self-maps, we can represent the iterated dominance procedure through the following sequence of sets $\mathrm{ND}(A)$, $\mathrm{ND}(\mathrm{ND}(A)) = \mathrm{ND}^2(A)$, ..., $\mathrm{ND}^k(A)$, ... Essentially, the idea is to first eliminate all dominated actions, thus obtaining $\mathrm{ND}(A)$. Then one moves on to eliminate all dominated actions in the *restricted game* with set of action profiles $\mathrm{ND}(A)$, thus obtaining $\mathrm{ND}^2(A)$, then eliminate all the dominated actions from the restricted game with set of action profiles $\mathrm{ND}^2(A)$, thus obtaining $\mathrm{ND}^3(A)$, and so on.

Definition 16. *$a \in A$ is a profile of **iteratively undominated actions** if $a \in \mathrm{ND}^\infty(A) = \bigcap_{k \geq 1} \mathrm{ND}^k(A)$.*

Theorem 5. *(Pearce [54]) Fix a finite or compact-continuous game; for every $k = 1, 2, \ldots$, $\rho^k(A) = \mathrm{ND}^k(A)$. Therefore, an action profile is rationalizable if and only if it is iteratively undominated.*

We first prove a quite simple preliminary result about the sequence of subsets $\left(\rho^k(A)\right)_{k=1}^{\infty}$. For any nonempty subset $C_i \subseteq A_i$ define the **constrained best reply** correspondence $r_i(\cdot|C_i) : \Delta(A_{-i}) \rightrightarrows C_i$ as

$$r_i(\mu^i|C_i) = \arg\max_{a_i \in C_i} u_i(a_i, \mu^i).$$

Recall that $r_i(\mu^i) = \arg\max_{a_i \in A_i} u_i(a_i, \mu^i)$. Therefore, for each $\mu^i \in \Delta(A_{-i})$,

$$\emptyset \neq r_i(\mu^i) \subseteq C_i \Rightarrow r_i(\mu^i) = r_i(\mu^i|C_i); \tag{4.5.1}$$

in words, if there is at least one best reply (as is the case in all finite, or compact-continuous games) and all the best replies satisfy the constraint of being in C_i, then constrained and unconstrained best replies must coincide (check you can prove this; more generally, prove that $r_i(\mu^i) \cap C_i \neq \emptyset \Rightarrow r_i(\mu^i|C_i) \subseteq r_i(\mu^i)$).[10] For each $C \in \mathcal{C}$ and $i \in I$, let

$$\overline{\rho}_i(C) = r_i(\Delta(C_{-i})|C_i) = \bigcup_{\mu^i \in \Delta(C_{-i})} \arg\max_{a_i \in C_i} u_i(a_i, \mu^i);$$

$$\overline{\rho}(C) = \bigtimes_{i \in I} \overline{\rho}_i(C).$$

Since $\overline{\rho}_i(C)$ is obtained by *constrained* maximization, it follows that $\overline{\rho}(C) \subseteq C$ for every C. The rationalization operator ρ, instead, does not satisfy this property (if C, for instance, is the set of dominated action profiles, then $\rho(C) \cap C = \emptyset$). Also, like ND, operator $\overline{\rho}$ is not monotone, whereas ρ is monotone. Yet, ρ and $\bar{\rho}$ coincide whenever the constraint "does not bite":

Remark 17. *Suppose that $r_i(\mu^i) \neq \emptyset$ for each $i \in I$ and $\mu^i \in \Delta(A_{-i})$. Then,*

$$\forall C \in \mathcal{C},\ \rho(C) \subseteq C \Rightarrow \rho(C) = \bar{\rho}(C). \tag{4.5.2}$$

Proof. Suppose that $\rho(C) \subseteq C$. Since $\rho(C) = \bigtimes_{i \in I} r_i(\Delta(C_{-i}))$, $r_i(\mu^i) \subseteq r_i(\Delta(C_{-i})) \subseteq C_i$ for each i and $\mu^i \in \Delta(C_{-i})$ (the first inclusion holds by definition, the second by assumption). Under the stated assumptions $r_i(\mu^i) \neq \emptyset$ for every i and μ^i. Therefore, by (4.5.1), $r_i(\mu^i) = r_i(\mu^i|C_i)$ for every i and $\mu^i \in \Delta(C_{-i})$, which implies $\rho_i(C_{-i}) = r_i(\Delta(C_{-i})) = r_i(\Delta(C_{-i})|C_i) = \bar{\rho}_i(C)$ for each i, that is, $\rho(C) = \bar{\rho}(C)$. ■

This implies that if ρ and $\bar{\rho}$ are iterated starting from the set A of *all* action profiles, the same sequence of subsets obtains:

Lemma 13. *In a finite or compact-continuous game, for every $k = 1, 2, ...$, $\rho^k(A) = \overline{\rho}^k(A)$.*[11]

Proof. By definition $\rho^0(A) = \overline{\rho}^0(A)$. Suppose by way of induction that $\rho^k(A) = \overline{\rho}^k(A)$. As already shown, the monotonicity of ρ implies

[10] In some infinite games, for some i, μ^i, and C_i, $r_i(\mu^i) = \emptyset$ and $r_i(\mu^i|C_i) \neq \emptyset$. See Example 16.

[11] As should be clear from the proof, the lemma holds for all games where the best reply correspondences are non-empty-valued.

$\rho(\rho^k(A)) = \rho^{k+1}(A) \subseteq \rho^k(A)$; hence $\rho(\rho^k(A)) = \rho(\bar{\rho}^k(A)) \subseteq \bar{\rho}^k(A)$. Under the compactness-continuity assumption, $r_i(\mu^i) \neq \emptyset$ for each $i \in I$ and $\mu^i \in \Delta(A_i)$. Therefore eq. (4.5.2) applies and yields $\rho(\bar{\rho}^k(A)) = \bar{\rho}(\bar{\rho}^k(A))$. Thus, $\rho^{k+1}(A) = \rho(\rho^k(A)) = \bar{\rho}(\bar{\rho}^k(A)) = \bar{\rho}^{k+1}(A)$. ■

The lemma can be reformulated as follows. Say that an action a_i is **iteratively justifiable** if (1) it is justifiable (2) it is justifiable in the reduced game G^1 obtained by elimination of all the non justifiable actions, (3) it is justifiable in the reduced game G^2 obtained by the elimination of all non justifiable actions in G^1, and so on. The actions that are iteratively justifiable are exactly those obtained by iterating the operator $\bar{\rho}$. Hence, Lemma 13 states that, for every player, the set of rationalizable actions coincides with the set of iteratively justifiable actions.

Example 16. Here is an example that violates the compactness-continuity hypothesis of Theorem 2 and where, as a consequence, $r_i(\mu^i)$ may be empty and the thesis of Lemma 13 does not hold: $I = \{1,2\}$, $A_1 = \{0,1\}$, $A_2 = \{1 - \frac{1}{k} : k = 1,2,...\} \cup \{1\}$, the payoff functions are given by the infinite bi-matrix in Figure 4.3.

$1\backslash 2$	0	...	$1-\frac{1}{k}$	...	1
0	$1,\ 0$	...	$1,\ 1-\frac{1}{k}$	...	$1,\ 0$
1	$0,\ 0$	...	$0,\ 0$	...	$0,\ 1$

Figure 4.3: A countably infinite game.

Thus, $u_2\left(1-\frac{1}{k},\mu^2\right) = \mu^2(0)\left(1-\frac{1}{k}\right)$, $u_2(1,\mu^2) = \mu^2(1) = 1-\mu^2(0)$. If $\mu^2(1) \geq \frac{1}{2}$ the best reply is $a_2 = 1$. If $\mu^2(1) < \frac{1}{2}$, then $u_2(1-\frac{1}{k},\mu^2) > u_2(1,\mu^2)$ for k sufficiently large, but there is no best reply because $u_2(1-\frac{1}{k},\mu^2)$ is strictly increasing in k. Note also that $a_1 = 0$ strictly dominates $a_1 = 1$. Clearly $\rho(A) = \bar{\rho}(A) = \{0\} \times \{1\}$ and $\rho_2(\{0\}) = \emptyset$, $\bar{\rho}_2(\{0\} \times \{1\}) = \{1\}$. Therefore $\rho^2(A) = \emptyset \neq \{0\} \times \{1\} = \bar{\rho}^2(A)$. How is compactness-continuity violated? A_1 is finite, hence compact; A_2 is a closed and bounded subset of the real line (A_2 contains the limit of the sequence $(1-\frac{1}{k})_{k=1}^{\infty}$), hence it is also compact; but u_2 is discontinuous at

$(0,1)$: $\lim_{k\to\infty} u_2\left(0, 1-\frac{1}{k}\right) = 1 \neq 0 = u_2(0,1)$.[12] ▲

Theorem 5 follows quite easily from Lemma 13: indeed Lemma 2 implies that the iteratively undominated actions coincide with the iteratively justifiable ones, and by Lemma 13 the iteratively justifiable actions coincide with the rationalizable actions. The details are as follows:

Proof of Theorem 5. *Basis step.* Lemma 2 implies that $\rho(A) = \mathrm{ND}(A)$.

Inductive step. Assume that $\rho^{k-1}(A) = \mathrm{ND}^{k-1}(A)$ (inductive hypothesis) and consider the game G^{k-1} where the set of actions of each player i is $\rho_i^{k-1}(A)$, and the payoff functions are obtained from the original game by restricting their domain to $\rho^{k-1}(A)$. The inductive hypothesis implies that the set of undominated action profiles in G^{k-1} is $\mathrm{ND}(\rho^{k-1}(A)) = \mathrm{ND}(\mathrm{ND}^{k-1}(A)) = \mathrm{ND}^k(A)$. By Lemma 2, $\mathrm{ND}(\mathrm{ND}^{k-1}(A)) = \overline{\rho}(\mathrm{ND}^{k-1}(A))$. The inductive hypothesis and Lemma 13 yield $\overline{\rho}(\mathrm{ND}^{k-1}(A)) = \overline{\rho}(\rho^{k-1}(A)) = \overline{\rho}(\overline{\rho}^{k-1}(A)) = \rho^k(A)$. Hence $\mathrm{ND}^k(A) = \rho^k(A)$. ■

So far, we have considered a procedure of iterated elimination which is **maximal**, in the sense that at any step *all* the dominated actions of all players are eliminated (where dominance holds in the restricted game that resulted from previous iterated eliminations). However, one can show that to compute the set of action profiles that are iteratively undominated (and therefore rationalizable) it is sufficient to iteratively eliminate *some* actions which are dominated for *some* player until in the restricted game so obtained there are no dominated actions left. To simplify the exposition, we restrict the analysis of non-maximal iterated dominance procedures to *finite* games.[13]

Definition 17. *An* ***iterated dominance procedure*** *is a sequence* $(C^k)_{k=0}^K \in \mathcal{C}^{K+1}$ *of nonempty Cartesian subsets of* A *such that (i)* $C^0 = A$, *(ii) for each* $k = 1, \ldots, K$, $\mathrm{ND}(C^{k-1}) \subseteq C^k \subset C^{k-1}$ *(*$\subset$ *is the* strict *inclusion), and (iii)* $\mathrm{ND}(C^K) = C^K$.

[12] For the mathematically savy: it is easy to change the example so that A_2 is not compact and u_2 is trivially continuous, just endow A_2 with the discrete topology.

[13] For an extension to infinite, compact-continuous games see Dufwenberg and Stegeman [29], who also provide counterexamples when compactness-continuity fails.

In words, an iterated dominance procedure is a sequence of steps starting from the full set of action profiles (condition *i*) and such that at each step k, for at least one player, at least one of the actions that are dominated given the previous steps is eliminated (condition *ii*); the elimination procedure can stop only when no further elimination is possible (condition *iii*).

Theorem 6. *Fix a finite game. For every iterated dominance procedure* $(C^k)_{k=0}^K$, C^K *is the set of rationalizable action profiles.*

Proof. Fix an iterated dominance procedure, i.e., a sequence of subsets $(C^k)_{k=0}^K$ that satisfies *(i)*-*(ii)*-*(iii)* of Definition 17.

Claim. *For each* $k = 0, 1, \ldots, K$,

$$\rho^{k+1}(A) \subseteq \rho(C^k) = \mathrm{ND}(C^k).$$

The proof of the claim is by induction:

Basis step. By Lemma 2 $\rho(A) = \mathrm{ND}(A)$; by *(i)* $C^0 = A$. Therefore $\rho^1(A) \subseteq \rho(C^0) = \mathrm{ND}(C^0)$ (the weak inclusion actually holds as an equality).

Inductive step. Suppose that $\rho^{k+1}(A) \subseteq \rho(C^k) = \mathrm{ND}(C^k)$. By *(ii)*, $\mathrm{ND}(C^k) \subseteq C^{k+1} \subset C^k$. By monotonicity of ρ and the inductive hypothesis,

$$\rho^{k+2}(A) \subseteq \rho(\mathrm{ND}(C^k)) \subseteq \rho(C^{k+1}),$$

and

$$\rho(C^{k+1}) \subseteq \rho(C^k) = \mathrm{ND}(C^k) \subseteq C^{k+1}.$$

Since the game is finite, there is at least one best reply to every conjecture and eq. (4.5.2) holds; thus,

$$\rho(C^{k+1}) = \bar{\rho}(C^{k+1}) = \mathrm{ND}(C^{k+1}),$$

where the second equality follows from Lemma 2. Collecting inclusions and equalities, $\rho^{k+2}(A) \subseteq \rho(C^{k+1}) = \mathrm{ND}(C^{k+1})$. □

The claim implies that $\rho^{K+1}(A) \subseteq \rho(C^K) = \mathrm{ND}(C^K)$. By *(iii)*, $C^K = \mathrm{ND}(C^K)$. Therefore $\rho^{\infty}(A) \subseteq \rho(C^K) = C^K$, that is, every rationalizable profile belongs to C^K, and C^K has the best reply property. By Theorem 3, C^K must be the set of rationalizable profiles. ■

4.6 Rationalizability and Iterated Dominance in Nice Games

Recall that in a **nice game** the action set of each player is a compact interval in the real line, payoff functions are continuous and each u_i is strictly quasi-concave in the own action a_i (see Definition 11). Therefore, nice games are compact-continuous. It turns out that the analysis of rationalizability and iterated dominance in nice games is nice indeed! The analysis, however, requires some technicalities.

Recall that, for the games under considerations, $\mathcal{C}$ denotes the collection of *closed* subsets of A with a cross-product form $C = \times_{i \in I} C_i$. Define the following operators on $\mathcal{C}$: for every $C \in \mathcal{C}$,

$$\begin{aligned}
\mathrm{r}(C) &= \bigtimes_{i \in I} r_i(C_{-i}), \;\; \iota\rho(C) = \bigtimes_{i \in I} r_i(\mathrm{I}\Delta(C_{-i})), \\
\mathrm{ND}_p(C) &= \bigtimes_{i \in I} \{a_i \in C_i : \forall b_i \in C_i, \exists a_{-i} \in C_{-i}, u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})\},
\end{aligned}$$

where $\mathrm{I}\Delta(C_{-i})$ is the set of product probability measures on C_{-i}, or independent (uncorrelated) conjectures (see Section 3.3.2). Note that r and $\iota\rho$ are self-maps on $\mathcal{C}$, because $r_i(C_{-i})$ and $r_i(\mathrm{I}\Delta(C_{-i}))$ are closed for every closed set C_{-i} (see the Appendix 4.7). Similarly, for every $C \in \mathcal{C}$, $\mathrm{ND}_p(C)$ is closed; hence, also ND_p is a self-map on $\mathcal{C}$. In fact, for every $i \in I$ and $b_i \in C_i$, since u_i is continuous, the set

$$\{a_i \in C_i : \exists a_{-i} \in C_{-i}, u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})\}$$

is closed. Therefore, also the intersection of such sets

$$\bigcap_{b_i \in C_i} \{a_i \in C_i : \exists a_{-i} \in C_{-i}, u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})\}$$

$$= \{a_i \in C_i : \forall b_i \in C_i, \exists a_{-i} \in C_{-i}, u_i(a_i, a_{-i}) \geq u_i(b_i, a_{-i})\}$$

is closed.

Iterating $\iota\rho$ from A yields a solution procedure called **independent rationalizability**,[14] which is what obtains if (a) players are rational, (b)

[14]This was the solution originally called "rationalizability" by Bernheim [21] and Pearce [54] before the epistemic analysis proved that the cleaner and more basic concept is the "correlated" version $\rho^{\infty}(A)$.

they hold independent (i.e., non correlated) conjectures, and (c) there is common belief of (a) and (b).

Iterating r from A yields a solution procedure called **point-rationalizability**, see Bernheim [21]. This procedure has no autonomous conceptual interest,[15] because there is no good reason to assume at the outset that players hold deterministic conjectures. Point rationalizability is just an ancillary solution concept that turns out to be convenient in the analysis of special games, including nice games.

Iterating ND_p from A yields the **iterated deletion of (pure) actions strictly dominated by (pure) actions**. In general games with compact action sets and continuous payoff functions, $\mathrm{ND}^k(A) \subseteq (\mathrm{ND}_p)^k(A)$ for all k.[16] Clearly $((\mathrm{ND}_p)^k(A))_{k=1}^{\infty}$ is a much simpler procedure than $(\mathrm{ND}^k(A))_{k=1}^{\infty}$, but it is weaker and it does not have a satisfactory conceptual foundation. So, in general, it should be regarded as an ancillary algorithm that allows to simplify the analysis of compact-continuous games.

As an exercise, use monotonicity to prove the following statements:

Remark 18. *For each k, $\mathrm{r}^k(A) \subseteq \iota\rho^k(A) \subseteq \rho^k(A)$. Hence $\mathrm{r}^{\infty}(A) \subseteq \iota\rho^{\infty}(A) \subseteq \rho^{\infty}(A)$.*

Remark 19. *For each $C \in \mathcal{C}$, if $C \subseteq \mathrm{r}(C)$ ($C \subseteq \iota\rho(C)$), then $C \subseteq \mathrm{r}^{\infty}(A)$ ($C \subseteq \iota\rho^{\infty}(A)$).*

Thus, if $\mathrm{r}^{\infty}(A) \neq \emptyset$, then $\iota\rho^{\infty}(A) \neq \emptyset$ and $\rho^{\infty}(A) \neq \emptyset$. Also, suppose that there is some $C \in \mathcal{C}$ that contains at least two elements and has the "deterministic" best reply property $C \subseteq \mathrm{r}(C)$. Then, the results above imply that there is a multiplicity of rationalizable (and independent rationalizable) action profiles. These simple results show why point rationalizability is a useful solution concept even though it does not have an interesting conceptual foundation. One nice feature of nice games is that Lemma 7 yields

[15] Point rationalizability obtains if (a) players are rational, (b) they hold deterministic conjectures, and (c) there is common belief of (a) and (b).

[16] The result holds more generally for games with compact actions sets, where the payoff functions are jointly continuous in the opponents' actions and upper semi-continuous in the own action.

Theorem 7. *In every nice game, for every $k \in \mathbb{N}$,*

$$\mathrm{r}^k(A) = \iota\rho^k(A) = \rho^k(A) = \mathrm{ND}^k(A) = (\mathrm{ND}_p)^k(A)$$

and each set is a product of closed intervals; hence

$$\mathrm{r}^\infty(A) = \iota\rho^\infty(A) = \rho^\infty(A) = \mathrm{ND}^\infty(A) = (\mathrm{ND}_p)^\infty(A)$$

and, again, each set is a product of closed intervals.

Proof. The second statement follows easily from the first, which is trivially true for $k = 0$. Suppose, by way of induction, that the first statement holds for a given k. By the inductive hypothesis the set $C = \mathrm{r}^k(A)$ is a product of closed intervals. Then Lemma 7 and Corollary 2 (applied to the restricted action sets C_i, $i \in I$) yield

$$\mathrm{r}(\mathrm{r}^k(A)) = \iota\rho(\mathrm{r}^k(A)) = \rho(\mathrm{r}^k(A)) = \mathrm{ND}(\mathrm{r}^k(A)) = \mathrm{ND}_p(\mathrm{r}^k(A)).$$

By the inductive hypothesis

$$\mathrm{r}^{k+1}(A) = \iota\rho^{k+1}(A) = \rho^{k+1}(A) = \mathrm{ND}^{k+1}(A) = (\mathrm{ND}_p)^{k+1}(A).$$

■

Example 17. As an application, consider the market for a crop analyzed in Chapter 1 (Section 1.6.1), but this time suppose that there is a finite number $n \geq 2$ of firms who decide how much to produce of a crop to be sold on the market 6 months later at the highest price that allows demand to absorb total output; thus, they compete *à la* Cournot.[17] Each firm has cost function

$$C(q_i) = \frac{1}{2m}(q_i)^2,\ q_i \in \left[0, \frac{4\alpha}{\beta}\right]$$

($\frac{4\alpha}{\beta}$ is a capacity constraint given by the available land).[18] *Market demand* is given by the function $D(p; n) = \max\{0, n(\alpha - \beta p)\}$ with $0 < \beta < 2$; therefore price as a function of *average quantity* is

$$P\left(\frac{1}{n}\sum_{i=1}^{n} q_i\right) = \max\left\{0, \frac{1}{\beta}\left(\alpha - \frac{1}{n}\sum_{i=1}^{n} q_i\right)\right\}.$$

[17]The following analysis is adapted from Boergers and Janssen [25].

[18]The capacity constraint is loose enough to simplify the calculations.

It is interesting to relate the analysis of rationalizable oligopolistic behavior to the standard competitive equilibrium for these market fundamentals.

(1) First, verify that the game is nice. The payoff function of each firm i is the profit function

$$\pi_i(q_i, q_{-i}) = \begin{cases} \frac{q_i}{\beta}\left(\alpha - \frac{q_i}{n} - \frac{1}{n}\sum_{j\neq i} q_j\right) - \frac{1}{2m}(q_i)^2, & \text{if } n\alpha > q_i + \sum_{j\neq i} q_j, \\ -\frac{1}{2m}(q_i)^2, & \text{if } n\alpha \leq q_i + \sum_{j\neq i} q_j. \end{cases}$$

This function is clearly continuous.[19] Now, fix q_{-i} arbitrarily. If $\sum_{j\neq i} q_j \geq n\alpha$ then $\pi_i(q_i, q_{-i}) = -\frac{1}{2m}(q_i)^2$, which is strictly decreasing in q_i. If $\sum_{j\neq i} q_j < n\alpha$, then π_i is initially increasing in q_i up to the point $r_i(q_{-i})$ where marginal revenue equals marginal cost (see below), and then it becomes strictly decreasing.[20]

(2) The best reply function is easily obtained from the first-order conditions when $\alpha > \frac{1}{n}\sum_{j\neq i} q_j$, in the other case the best reply is the corner solution $q_i = 0$:

$$r_i(q_{-i}) = \max\left\{0, \frac{m(\alpha - \frac{1}{n}\sum_{j\neq i} q_j)}{\beta + \frac{2m}{n}}\right\}.$$

(3) The monopolistic output is

$$q^{n,M} := r_i(0, \ldots, 0) = \frac{m\alpha}{\beta + \frac{2m}{n}}.$$

---

[19] Note that $\frac{q_i}{\beta}\left(\alpha - \frac{q_i}{n} - \frac{1}{n}\sum_{j\neq i} q_j\right) - \frac{1}{2m}(q_i)^2 = -\frac{1}{2m}(q_i)^2$ if $n\alpha = q_i + \sum_{j\neq i} q_j$.

[20] There is a kink at $\hat{q}_i = n\alpha - \sum_{j\neq i} q_j$, but this is immaterial. Note also that π_i is *not* concave in q_i: in the left neighborhood of the kink point $\hat{q}_i$ the derivative is

$$\begin{aligned} \frac{\partial \pi_i}{\partial q_i}(q_i, q_{-i}) &= P\left(\frac{q_i}{n} + \frac{1}{n}\sum_{j\neq i} q_j\right) + \frac{q_i}{n}P'\left(\frac{q_i}{n} + \frac{1}{n}\sum_{j\neq i} q_j\right) - \frac{q_i}{m} \\ &\approx \frac{\hat{q}_i}{n}P'\left(\frac{\hat{q}_i}{n} + \frac{1}{n}\sum_{j\neq i} q_j\right) - \frac{\hat{q}_i}{m} < -\frac{\hat{q}_i}{m} \end{aligned}$$

where the approximation holds because q_i is close to $\hat{q}_i$ and $P\left(\frac{q_i}{n} + \frac{1}{n}\sum_{j\neq i} q_j\right) \approx 0$, and the inequality holds because $P' < 0$. In a right neighborhood of $\hat{q}_i$, $\frac{\partial \pi_i}{\partial q_i}(q_i, q_{-i}) = -\frac{q_i}{m} \approx -\frac{\hat{q}_i}{m}$. Therefore, for $\varepsilon > 0$ small enough,

$$\frac{\partial \pi_i}{\partial q_i}\pi_i(\hat{q}_i - \varepsilon, q_{-i}) < \frac{\partial \pi_i}{\partial q_i}\pi_i(\hat{q}_i + \varepsilon, q_{-i}),$$

which implies that π_i is *not* concave.

If every competitor produces at maximum capacity, $\frac{1}{n}\sum_{j\neq i} q_j > \alpha$ (because $\beta < 2$ and capacity is $4\alpha/\beta$), so the best reply is

$$r_i\left(\frac{4\alpha}{\beta}, \dots, \frac{4\alpha}{\beta}\right) = 0.$$

Since the game is nice, $\rho(A) = \left[0, q^{n,M}\right]^n$. If this set has the best reply property, this is also the set of rationalizable profiles. To check this, it is enough to verify whether the best reply to the most pessimistic conjecture consistent with rationality is zero, that is $r_i(q^{n,M}, ..., q^{n,M}) = 0$; in the affirmative case $\left[0, q^{n,M}\right]^n$ has the best reply property. The condition for this is $\frac{n-1}{n}\frac{m\alpha}{\beta+\frac{2m}{n}} \geq \alpha$, or

$$\beta \leq m\left(\frac{n-3}{n}\right).$$

Therefore, if $\beta < m$ there is $\bar{n}$ large enough so that, for each $n > \bar{n}$, $\left[0, q^{n,M}\right]^n$ has the best reply property, hence there is a huge multiplicity of rationalizable outcomes. Conversely, if $\beta > m$, then

$$\beta > m\left(\frac{n-3}{n}\right)$$

for every n; hence, in this case the set $\left[0, q^{n,M}\right]^n$ does *not* have the best reply property whatever the number n of firms. Furthermore, it can be shown that if $\beta > m$ there is a unique rationalizable outcome.[21] By symmetry, each firm has the same rationalizable output $q^{n,*}$ which must solve $q = \frac{m(\alpha - \frac{n-1}{n}q)}{\beta+\frac{2m}{n}}$, because the singleton $\{(q^{n,*}, \dots, q^{n,*})\}$ has the best reply property. Therefore

$$q^{n,*} = \frac{m\alpha}{\beta + m\frac{n+1}{n}}$$

with corresponding rationalizable price

$$p^{n,*} := P\left(\frac{m\alpha}{\beta + m\frac{n+1}{n}}\right) = \frac{\alpha + \frac{m}{n}\frac{\alpha}{\beta}}{\beta + m\frac{n+1}{n}}$$

[21] If $\beta > m$, the joint best reply function

$$(q_i)_{i=1}^n \mapsto (r_i(q_{-i}))_{i=1}^n$$

is a contraction, which implies that the sequence $\boldsymbol{\theta}\rho^k(A))_{k\in\mathbb{N}}$ shrinks to a point.

(of course, this is the Cournot-Nash equilibrium).
(4) Finally, the competitive equilibrium price with these market fundamentals is

$$p^* := \frac{\alpha}{\beta + m},$$

with average output

$$q^* := \frac{m\alpha}{\beta + m}.$$

As pointed out in Section 1.6.1 of the Chapter 1, $\beta > m$ is the "cobweb stability" condition. Under this condition, the unique rationalizable output is $q^{n,*} = \frac{m\alpha}{\beta + m\frac{n+1}{n}}$. Of course,

$$\begin{aligned} \lim_{n\to\infty} q^{n,*} &= \lim_{n\to\infty} \frac{m\alpha}{\beta + m\frac{n+1}{n}} = \frac{m\alpha}{\beta + m} = q^*, \\ \lim_{n\to\infty} p^{n,*} &= \lim_{n\to\infty} \frac{\alpha + \frac{m}{n}\frac{\alpha}{\beta}}{\beta + m\frac{n+1}{n}} = \frac{\alpha}{\beta + m} = q^*. \end{aligned}$$

To sum up, under the "cobweb stability" condition $\beta > m$, the rationalizable outcome is unique and approximates the so called "rational expectations equilibrium." ▲

4.7 Appendix: Compact-Continuous Games

4.7.1 Preliminaries on the Rationalization Operator

Recall that, for the case of compact-continuous games, we defined $\mathcal{C}$ as the collection of *closed* Cartesian subsets of A. In light of the results presented in Section 3.4.2, we can show that ρ is a self map from $\mathcal{C}$ to $\mathcal{C}$.

Lemma 14. *For every $C \in \mathcal{C}$, $\rho(C) \in \mathcal{C}$; furthermore, if $C \neq \emptyset$ then $\rho(C) \neq \emptyset$.*

Proof. Let $C = \times_{i\in I} C_i \in \mathcal{C}$. Since C is closed and $\rho(C) = \times_{i\in I} r_i(\Delta(C_{-i}))$, Lemma 10 implies that the Cartesian set $\rho(C)$ is also closed; hence $\rho(C) \in \mathcal{C}$; furthermore, if $C \neq \emptyset$, Lemma 10 implies that $\rho(C) \neq \emptyset$. ■

For future reference, we record an important property of compact sets:

Lemma 15. (Finite-intersection property of compact sets) *Let X be compact and let $\{C_j\}_{j\in J}$ be an indexed family of* closed *subsets of X such that, for every finite $F \subseteq J$, the finite indexed subfamily $\{C_j\}_{j\in F}$ has nonempty intersection: $\bigcap_{j\in F} C_j \neq \emptyset$. Then also $\{C_j\}_{j\in J}$ has nonempty intersection: $\bigcap_{j\in J} C_j \neq \emptyset$.*

4.7.2 Proof of Theorem 2 (b)

We must prove that $\rho^{\infty}(A)$ is nonempty, compact, and satisfies $\rho^{\infty}(A) = \rho(\rho^{\infty}(A))$. Recall that we defined $\rho^{0}(A) = A$ for convenience. We first prove by induction on $k \in \mathbb{N}_0$ (where $\mathbb{N}_0$ is the set of nonnegative integers) that $\rho^{k}(A) \neq \emptyset$ for every $k \geq 0$. The basis step is trivial: $\rho^{0}(A) = A$ where $A \neq \emptyset$ by assumption. Now assume by way of induction that $\rho^{k}(A) \in \mathcal{C}$ is nonempty for some k. Then Lemma 14 implies that $\rho^{k+1}(A) = \rho\left(\rho^{k}(A)\right) \neq \emptyset$. Next we show that $\rho^{\infty}(A) \neq \emptyset$. Since the sequence $\left(\rho^{k}(A)\right)_{k=1}^{\infty}$ is weakly decreasing,

$$\forall \ell \in \mathbb{N},\ \bigcap_{k=1}^{\ell} \rho^{k}(A) = \rho^{\ell}(A) \neq \emptyset.$$

Then, the finite-intersection property of compact sets applied to the indexed family $\{\rho^{\ell}(A)\}_{\ell\in\mathbb{N}}$ (Lemma 15) implies that

$$\rho^{\infty}(A) = \bigcap_{\ell=1}^{\infty} \rho^{\ell}(A) \neq \emptyset.$$

Since $\rho^{\infty}(A)$ is an intersection of closed set, it is closed. Since A is compact, the closed subset $\rho^{\infty}(A)$ is compact.

Finally, we prove that $\rho^{\infty}(A) = \rho(\rho^{\infty}(A))$. The inclusion $\rho(\rho^{\infty}(A)) \subseteq \rho^{\infty}(A)$ follows from the monotonicity of ρ: For every $k \in \mathbb{N}_0$, $\rho^{\infty}(A) \subseteq \rho^{k}(A)$. Since ρ is monotone,

$$\rho(\rho^{\infty}(A)) \subseteq \bigcap_{k=0}^{\infty} \rho\left(\rho^{k}(A)\right) = \bigcap_{\ell=1}^{\infty} \rho^{\ell}(A) = \rho^{\infty}(A).$$

Next we show that $\rho^{\infty}(A) \subseteq \rho(\rho^{\infty}(A))$. Pick any profile $(a_i)_{i\in I} \in \rho^{\infty}(A)$. For every k, let A_i^k denote the set of "k-rationalizable" actions of player

i, so that $\rho^k(A) = \bigtimes_{i\in I} A_i^k$ and $\rho^\infty(A) = \bigcap_{k=1}^{\infty} \bigtimes_{i\in I} A_i^k$. With this, for each $i \in I$, we can find a sequence of conjectures $\left(\mu_n^i\right)_{n=0}^{\infty}$ such that $\mu_n^i \in \Delta\left(A_{-i}^n\right) \subseteq \Delta\left(A_{-i}\right)$ and $a_i \in r_i\left(\mu_n^i\right)$ for every $n \in \mathbb{N}_0$. By Lemma 9, $\left(\mu_n^i\right)_{n=0}^{\infty}$ has a convergent subsequence $\left(\mu_{n_k}^i\right)_{k=1}^{\infty}$ with limit $\mu_{n_k}^i \to \bar{\mu}^i \in \Delta\left(A_{-i}\right)$. By Lemma 10, $a_i \in r_i\left(\bar{\mu}^i\right)$ for each $i \in I$. To prove that $\left(a_i\right)_{i\in I} \in \rho\left(\rho^\infty(A)\right)$, we only have to show that $\bar{\mu}^i\left(\bigcap_{n=1}^{\infty} A_{-i}^n\right) = 1$ for each $i \in I$. Note that, for every n, there is some k such that $n_k \geq n$, so that $\mu_{n_k}^i\left(A_{-i}^n\right) = 1$ because $A_{-i}^{n_k} \subseteq A_{-i}^n$. Therefore, by the *portmanteau theorem* [3, Theorem 15.3], for every A_{-i}^n (a closed set)

$$\bar{\mu}^i\left(A_{-i}^n\right) \geq \limsup_{k\to\infty} \mu_{n_k}^i\left(A_{-i}^n\right) = 1,$$

that is, $\bar{\mu}^i\left(A_{-i}^n\right) = 1$. Since $\bar{\mu}^i$ is countably additive, it must be continuous.[22] Therefore,

$$\bar{\mu}^i\left(\bigcap_{n=1}^{\infty} A_{-i}^n\right) = \lim_{n\to\infty} \bar{\mu}^i\left(A_{-i}^n\right) = 1.$$

■

[22]This means that $\bar{\mu}^i\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \bar{\mu}^i\left(E_k\right)$ for every weakly *decreasing* sequence of measurable sets $\left(E_k\right)_{k=1}^{\infty}$, and $\bar{\mu}^i\left(\bigcup_{k=1}^{\infty} F_k\right) = \lim_{k\to\infty} \bar{\mu}^i\left(F_k\right)$ for every weakly *increasing* sequence of measurable subsets $\left(F_k\right)_{k=1}^{\infty}$.

5

Pure Equilibrium

In Chapter 4, we analyzed the behavioral implications of rationality and common belief in rationality when there is complete information, that is, common knowledge of the rules of the game and of players' preferences. There are interesting strategic situations where these assumptions about rationality, belief and knowledge imply that each player can correctly predict the opponents' behavior, e.g., many oligopoly games. But this is not the case in general: whenever a game has multiple rationalizable profiles there is at least one player i such that many conjectures μ^i are consistent with common belief in rationality, but at most one of them can be correct. In this chapter, we analyze situations where players' predictions are correct, and we discuss why predictions should (or should not) be correct. Let us give a preview.

We start with *Nash*'s classic definition of equilibrium (in pure actions): An action profile $a^* = (a_i^*)_{i \in I}$ is an **equilibrium** if each action a_i^* is a best reply to the actions of the other players a_{-i}^*.

Nash equilibrium play follows from the assumptions that players are rational and hold correct conjectures about the behavior of other players. But why should this be the case? Can we find scenarios that make the assumption of correct conjectures compelling, or at least plausible? We will discuss two types of scenarios: (1) those that make Nash equilibrium an "obvious way to play the game," and (2) those that make Nash equilibrium a steady state of an adaptive process. In each case, we will point out that Nash equilibrium is justified under rather restrictive assumptions, and that weakening such assumptions suggests interesting generalizations of the

Nash equilibrium concept whereby players form probabilistic conjectures about the behavior of co-players. We then analyze these generalizations in Chapter 6 on "probabilistic equilibria."

More specifically, Nash equilibrium may be the "obvious way to play the game" under complete information either as the outcome of sophisticated strategic reasoning, or as the result of a non-binding pre-play agreement. If it were obvious how the game should be played, then each player i could predict the obvious way to play of the co-players, viz. a^*_{-i}, and play the best reply a^*_i. Yet, if sophisticated strategic reasoning is given by rationality and common belief in rationality, then Nash equilibrium is obtained only in those games where rationalizable actions and equilibrium actions coincide. This coincidence is implied by some assumptions about action sets and payoff functions that are satisfied in several economic applications, but in general the relevant concept under this scenario is rationalizability, not Nash equilibrium.

What about pre-play agreements? True, a non-binding agreement to play a particular action profile a^* has to be a Nash equilibrium, otherwise the agreement would be self-defeating. But it will be shown in Chapter 6 that more general "probabilistic agreements" that make behavior depend on some extraneous random variables may be more efficient. Such probabilistic agreements are called "**correlated equilibria**" and generalize the Nash equilibrium concept.

Next we turn to adaptive processes. The general scenario is that a given game is played recurrently by agents that each time are drawn at random from large populations corresponding to different players/roles in the game (e.g., seller and buyer, or male and female). The agents drawn from each population are matched, play among themselves, and then are separated. If each player's conjecture about the opponents' behavior in the current play is based on his observations of the opponents' actions in the past and this player best responds, we obtain a dynamic process in which at each point in time some agent switches from one action to another. *If* the process converges so that each agent playing in a given role ends up choosing the *same* action over and over, the steady state must look like a Nash equilibrium: indeed the experience of those playing in role i is that the opponents keep playing some profile a^*_{-i} and thus they keep choosing a best reply, say, a^*_i.[1] Since this is true for each i, $(a^*_i)_{i\in I}$ must be a Nash

[1] Suppose for simplicity that there is only one best reply.

equilibrium.

But it is possible that the process converges to an heterogeneous situation: a fraction q_i^1 of the agents in the population i choose some action a_i^1, another fraction q_i^2 of agents choose action a_i^2 and so on. If the process has stabilized, $(q_j^1, q_j^2, \ldots)$ will (approximately) represent the observed frequencies of the actions of those playing in role j, and i will conjecture that a_j^1, $a_j^2, \ldots$ are played with probabilities q_j^1, $q_j^2, \ldots$. If each action a_i^1, a_i^2, ... is a best reply to such conjecture, given a bit of inertia, i will keep choosing whatever he was choosing before and the heterogeneous behavior within each population will persist. In this case, the steady state is described by *mixed* actions whose support is made of best responses to the mixed actions of the opponents. This is called "**mixed (Nash) equilibrium**." It turns out that a mixed equilibrium formally is a Nash equilibrium of the extended game where each player i chooses in the set $\Delta(A_i)$ of mixed actions and payoffs are computed by taking expected values (under the assumption of independence across players). Indeed, this is the standard definition of mixed equilibrium.

The analysis above relies on the assumption that, when a game is played recurrently, each agent can observe the frequency of play of different actions. But often the information feedback is poorer. For example, one may observe only a variable determined by the players' actions, such as the number of customers of a firm in a price-setting oligopoly. In such a context, it may happen that players' conjectures are wrong and yet they are confirmed by players' information feedback, so that they keep best responding to such wrong conjectures and the system is in a steady state which is not a Nash (or mixed Nash) equilibrium. A steady state whereby players best respond to confirmed (non contradicted) conjectures is called "conjectural equilibrium," or "**self-confirming equilibrium**." Since correct conjectures (conjectures that correspond to the actual behavior of the other players) are necessarily confirmed, every Nash equilibrium is necessarily self-confirming.

To sum up, a serious effort to provide reasonable justifications for the Nash equilibrium concept leads us to think about scenarios where strategic interaction would plausibly lead to players best responding to correct conjectures about the opponents. But for each such scenario we need additional assumptions (on the payoff functions, or on information feedback) to justify Nash play. Without such additional assumptions we are left with interesting generalizations of the Nash equilibrium concept.

The rest of the chapter is organized as follows. In Section 5.1, we define (pure) Nash equilibrium and provide sufficient conditions for its existence. In Section 5.2 we focus on symmetric equilibria of nice games and provide a simple existence proof. In Section 5.3 we analyze nice games with strategic complementarities and show in these games there is a tight relationship between Nash equilibrium and rationalizability. Finally, in Section 5.4 we go back to the interpretation of the Nash equilibrium concept, paving the way for the generalizations analyzed in Chaper 6.

5.1 Nash Equilibrium

A Nash[2] equilibrium is a situation in which every player is rational and holds *correct* conjectures about the actions of the other players.

Definition 18. *An action profile $a^* = (a_i^*)_{i\in I}$ is a **Nash equilibrium** if, for every $i \in I$, $a_i^* \in r_i(a_{-i}^*)$.*

Remark 20. *An action profile a^* is a Nash equilibrium if and only if the singleton $\{a^*\}$ has the best reply property. Hence, by Theorem 3, every Nash equilibrium is rationalizable; if there is a unique rationalizable action profile, then it is necessarily the unique Nash equilibrium.*

By inspection of simple games, such as "Matching Pennies," it is obvious that not all games have Nash equilibria. The following classic theorem provides *sufficient* conditions for the existence of at least one Nash equilibrium.

Theorem 8. *Consider a* compact-continuous *game $G = \langle I, (A_i, u_i)_{i\in I}\rangle$. If, for every player $i \in I$, A_i is* convex *and, the function $u_i(\cdot, a_{-i}) : A_i \to \mathbb{R}$ is* quasi-concave *for every $a_{-i} \in A_{-i}$, then G has a Nash equilibrium.*

[2] John Nash, who has been awarded the Nobel Prize for Economics (joint with John Harsanyi and Reinhard Selten) in 1994, was the first to give a general definition of this equilibrium concept. Nash analyzed the equilibrium concept for the *mixed* extension of a game, and proved the existence of mixed equilibria for all games with a finite number of pure actions (see Definition 12 and Theorem 7).

Almost all others equilibrium concepts used in non-cooperative game theory can be considered as generalizations or "refinements" of the equilibrium proposed by Nash. Perhaps, this is the reason why the other equilibrium concepts that have been proposed do not take their name after one of the researchers who introduced them in the literature.

The proof of this Theorem 8 follows from preliminary results of independent interest. The first one states that the graph of the best-reply correspondence is closed:

Closed-Graph Lemma *If A is compact and $u_i : A \to \mathbb{R}$ is continuous, then the restriction of the best reply correspondence to subdomain A_{-i}, $r_i|_{A_{-i}} : A_{-i} \rightrightarrows A_i$, is nonempty valued and its graph*

$$\mathrm{Gr}\left(r_i|_{A_{-i}}\right) = \{(a_{-i}, a_i) \in A_{-i} \times A_i : a_i \in r_i\left(a_{-i}\right)\}$$

is closed.

The closed-graph property is a special case of Lemma 10 in Appendix 4.7, which considers best replies to probabilistic conjectures. Since the proof of Lemma 10 involves measure-theoretic arguments, it is useful to provide a separate proof for this simpler case (see also the proof of Lemma 6).

Proof of the Closed-Graph Lemma. Since u_i is continuous, for each $a_{-i} \in A_{-i}$, the section $u_i\left(\cdot, a_{-i}\right) : A_i \to \mathbb{R}$ is also continuous, therefore, it attains a maximum on the compact set A_i. Thus, $r_i(a_{-i}) \neq \emptyset$ for each $a_{-i} \in A_{-i}$.

To show that $\mathrm{Gr}\left(r_i|_{A_{-i}}\right)$ is closed, we prove that the limit of every convergent sequence in $\mathrm{Gr}\left(r_i|_{A_{-i}}\right)$ also belongs to $\mathrm{Gr}\left(r_i|_{A_{-i}}\right)$. Let $\left(a_{-i}^n, a_i^n\right)_{n=1}^{\infty} \in \mathrm{Gr}\left(r_i|_{A_{-i}}\right)^{\mathbb{N}}$ be such that $\lim_{n\to\infty}\left(a_{-i}^n, a_i^n\right) = (\bar{a}_{-i}, \bar{a}_i)$. By definition of $\mathrm{Gr}\left(r_i|_{A_{-i}}\right)$, a_i^n is a best reply to a_{-i}^n for each n, thus,

$$\forall a_i \in A_i, \forall n \in \mathbb{N},\ u_i\left(a_i^n, a_{-i}^n\right) \geq u_i\left(a_i, a_{-i}^n\right). \tag{5.1.1}$$

Since u_i is continuous, taking the limit of each side of (5.1.1) as $n \to \infty$ for each a_i, we obtain

$$\forall a_i \in A_i,\ u_i\left(\bar{a}_i, \bar{a}_{-i}\right) \geq u_i\left(a_i, \bar{a}_{-i}\right),$$

that is, $\bar{a}_i \in r_i\left(\bar{a}_{-i}\right)$. Therefore $(\bar{a}_{-i}, \bar{a}_i) \in \mathrm{Gr}\left(r_i|_{A_{-i}}\right)$. ■

Next we report without proof an important fixed-point theorem:[3]

[3] For a proof, see Ok [52], pp. 331.

Fixed-Point Theorem of Kakutani *Let X be a nonempty, convex, and compact subset of $\mathbb{R}^n$ and let $\varphi : X \rightrightarrows X$ be a correspondence such that the graph $\mathrm{Gr}(\varphi) = \{(x, y) \in X \times X : y \in \varphi(x)\}$ is closed and, for each $x \in X$, $\varphi(x)$ is nonempty and convex; then φ has a fixed point, i.e., there is some $x^* \in X$ such that $x^* \in \varphi(x^*)$.*

Proof of Theorem 8. To ease notation, we use the symbol r_i also for the restriction of the best-reply correspondence to sub-domain A_{-i}. By the Closed-Graph Lemma, each r_i is nonempty valued and has a closed graph. Fix a_{-i} arbitrarily and let $u_i^*(a_{-i}) = \max_{a_i \in A_i} u_i(a_i, a_{-i})$ denote the maximum payoff of i given a_{-i}. By quasi-concavity of u_i, for every $a_{-i} \in A_{-i}$ and every real number $y \in \mathbb{R}$, the set $\{a_i \in A_i : u_i(a_i, a_{-i}) \geq y\}$ is convex. In particular, the set $\{a_i \in A_i : u_i(a_i, a_{-i}) \geq u_i^*(a_{-i})\}$ is convex. Since

$$r_i(a_{-i}) = \{a_i \in A_i : u_i(a_i, a_{-i}) \geq u_i^*(a_{-i})\}$$

(where the weak inequality holds as an equality), it follows that $r_i(a_{-i})$ is convex. Now define the joint best reply correspondence $r : A \rightrightarrows A$ as follows:

$$\forall a \in A,\ r(a) = \bigtimes_{i \in I} r_i(a_{-i}).$$

Since $r_i : A_{-i} \rightrightarrows A_i$ is nonempty and convex valued with a closed graph for each $i \in I$, the same holds also for the joint best reply correspondence $r : A \rightrightarrows A$. Then, by the Fixed-Point Theorem of Kakutani, there exists $a^* \in A$ such that $a^* \in r(a^*)$, that is, $a_i^* \in r_i(a_{-i}^*)$ for each $i \in I$. Such fixed point is a Nash equilibrium. ■

5.2 Equilibrium in Nice Symmetric Games

As one can see from the proof of Theorem 8, rather advanced mathematical tools are necessary to prove a relatively general existence result. Here, we analyze the more specific case in which all players are in a symmetric position and have unidimensional action sets. This allows us to use more elementary tools to show that, under convenient simplifying assumptions, there exists an equilibrium in which all players choose the same action. The fact that all players choose the same action is not part of the thesis of Theorem 8, which did not assume symmetry. Therefore the next result is not a special case of Theorem 8.[4] Before we state the theorem, we must formally spell out the definition of symmetric game and symmetric equilibrium.

Definition 19. *A static* game $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ *is* ***symmetric*** *if the players have the same set of actions, denoted by $\hat{A}$ (hence, for every $i \in I$, $A_i = \hat{A}$) and*

$$u_i\left(a\right) = u_{\pi(i)}\left(a \circ \pi^{-1}\right)$$

for every player $i \in I$, every action profile $a : I \to \hat{A}$, and every bijection (permutation) $\pi : I \to I$. A Nash equilibrium *a^* of a symmetric game G is* symmetric *if $a^* : I \to \hat{A}$ is constant, that is, if $a^*_i = a^*_j$ for all $i, j \in I$.*

Note that, if $a = (a_i)_{i\in I}$ then $a' = a \circ \pi^{-1} = \left(a_{\pi^{-1}(i)}\right)_{i\in I}$. In words, if $j = \pi(i)$, then the action a'_j played by j in the permuted profile a' is the action $a_i = a_{\pi^{-1}(j)}$ played by i in the original profile a. Hence, symmetry requires

$$u_i\left(\left(a_j\right)_{j\in I}\right) = u_{\pi(i)}\left(\left(a_{\pi^{-1}(j)}\right)_{j\in I}\right)$$

so that player $\pi(i)$ plays the *same* action in profile $a' = a \circ \pi^{-1}$ as player i in profile a.

Example 18. The Cournot oligopoly game of Example 7 is *symmetric* if all firms have the same capacity constraint and total cost function: for some $k > 0$, and $C : [0, k] \to \mathbb{R}$, $\bar{a}_i = k$ and $C_i(\cdot) = C(\cdot)$ for all $i \in I$. To see this, fix $a = (a_j)_{j\in I} \in [0, k]^I = \hat{A}^I$ arbitrarily and let $\pi : I \to I$ be a bijection, or permutation, then

$$\begin{aligned} u_i\left(a\right) &= a_i P\left(\sum_{j\in I} a_j\right) - C_i(a_i) \\ &= a_{\pi^{-1}(\pi(i))} P\left(\sum_{j\in I} a_{\pi^{-1}(j)}\right) - C_{\pi(i)}(a_{\pi^{-1}(\pi(i))}) \\ &= u_{\pi(i)}\left(\left(a_{\pi^{-1}(j)}\right)_{j\in I}\right) = u_{\pi(i)}\left(a \circ \pi^{-1}\right), \end{aligned}$$

where the second equality holds because $\pi^{-1}\left(\pi(i)\right) = i$, summations are commutative, and $C_i = C = C_{\pi(i)}$ by assumption. ▲

[4] However, by adding symmetry to the hypotheses of Theorem 8, one can prove the existence of a symmetric Nash equilibrium.

Recall that a game is **nice** if the action set of each player is a compact interval in the real line, payoff functions are continuous, and each player's payoff function is strictly quasi-concave in his own action.

Theorem 9. *Every* symmetric nice *game has a symmetric Nash equilibrium.*

Proof. To ease notation, we assume without loss of generality that the common action set is the interval $\hat{A} = [0, 1]$, which is just a normalization. Since u_i is continuous and strictly quasi-concave in a_i, there exists one and only one best reply to every (deterministic) conjecture a_{-i}. Let $\hat{r}_i(x) \in [0, 1]$ denote the unique best reply to the symmetric (deterministic) conjecture a_{-i} such that $a_j = x$ for each $j \in I \backslash \{i\}$. Since the game is symmetric, $\hat{r}_i(x)$ must be independent of i. Thus, we let $\hat{r} : [0, 1] \to [0, 1]$ denote the common **best reply function for symmetric conjectures**. By Lemma 6, each $r_i(\cdot)$ is a continuous function; therefore, $\hat{r}(\cdot)$ must be *continuous*.

Now, let us introduce an auxiliary function $f : [0, 1] \to \mathbb{R}$ as follows: $f(x) = x - \hat{r}(x)$. Function f is continuous[5] with $f(0) \leq 0$ and $f(1) \geq 0$. It follows from the Intermediate Value Theorem[6] that there exists a point $x^* \in [0, 1]$ such that $f(x^*) = 0$, that is, $x^* = \hat{r}(x^*)$. For each player $i \in I$, action $a_i^* = x^*$ is a best reply to the symmetric conjecture a_{-i}^* such that $a_j^* = x^*$ for each $j \neq i$. This means that the profile $(a_i^*)_{i \in I}$ with $a_i^* = x^*$ for each $i \in I$ is a symmetric Nash equilibrium. ■

The reason why we considered symmetric nice games is that they are often used in applied theory and we could use elementary results in real analysis to prove the existence of a symmetric equilibrium. A key step in the proof is that every continuous function $\hat{r} : [0, 1] \to [0, 1]$ has a **fixed point** $x^* = \hat{r}(x^*)$. This is an elementary special case of the Fixed-Point Theorem of Kakutani, because $\hat{r} : [0, 1] \to [0, 1]$ is continuous if and only if the singleton-valued correspondence $x \mapsto \{\hat{r}(x)\}$ (which is trivially non-empty and convex-valued) has a closed graph.

[5] If $\varphi, \psi : [0, 1] \to \mathbb{R}$ are continuous functions, then, for every $k \in \mathbb{R}$, $(\varphi + k\psi) : [0, 1] \to \mathbb{R}$ ($x \mapsto \varphi(x) + k\psi(x)$) is also continuous.

[6] See, for example, Binmore [22], p. 88.

5.3 Rationalizability and Equilibrium in Supermodular Nice Games

In Section 3.3.3, we have seen that best reply correspondences in supermodular nice games are weakly increasing functions (Lemma 8). This has an important consequence: the set of Nash equilibria has a largest and a smallest element, and they characterize the set of rationalizable actions profiles. To articulate and prove this result, we first record in the following lemma a structural property of the set $\rho^{\infty}(A)$ of rationalizable action profiles in a supermodular nice game, namely, $\rho^{\infty}(A)$ is an order box in $\mathbb{R}^I$ such that the lowest (respectively, largest) rationalizable action of each player is the best reply to the lowest (respectively, largest) rationalizable actions of the co-players.

Lemma 16. *Let $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ be a supermodular nice game, and—for each $i \in I$—let $\underline{a}_i^0$ and $\bar{a}_i^0$ respectively denote the smallest and largest action of player i, that is, $A_i = [\underline{a}_i^0, \bar{a}_i^0]$. Then, for each $n \in \mathbb{N}$,*

$$\rho^n(A) = \bigtimes_{i\in I}[\underline{a}_i^n, \bar{a}_i^n] = \bigtimes_{i\in I}\left[r_i\left(\underline{a}_{-i}^{n-1}\right), r_i\left(\bar{a}_{-i}^{n-1}\right)\right].$$

Therefore,

$$\rho^{\infty}(A) = \bigtimes_{i\in I}[\underline{a}_i^*, \bar{a}_i^*] = \bigtimes_{i\in I}\left[r_i\left(\underline{a}_{-i}^*\right), r_i\left(\bar{a}_{-i}^*\right)\right],$$

where, for each $i \in I$, $\underline{a}_i^* = \lim_{n\to\infty} \underline{a}_i^n$ and $\bar{a}_i^* = \lim_{n\to\infty} \bar{a}_i^n$.

Proof. Since G is a nice game, Theorem 7 implies that, for each $n \in \mathbb{N}$, $\rho^n(A)$ is an order box of best replies to deterministic conjectures, that is,

$$\rho^n(A) = \bigtimes_{i\in I}\left[\min r_i\left([\underline{a}_{-i}^{n-1}, \bar{a}_{-i}^{n-1}]\right), \max r_i\left([\underline{a}_{-i}^{n-1}, \bar{a}_{-i}^{n-1}]\right)\right],$$

where, for each $i \in I$, $[\underline{a}_{-i}^{n-1}, \bar{a}_{-i}^{n-1}]$ is the order box in $\mathbb{R}^{I\setminus\{i\}}$ of the $(n-1)$-rationalizable action profiles of the co-players. By Lemma 8, each r_i is a weakly increasing function; therefore

$$\begin{aligned}
\min r_i\left([\underline{a}_{-i}^{n-1}, \bar{a}_{-i}^{n-1}]\right) &= r_i\left(\underline{a}_{-i}^{n-1}\right),\\
\max r_i\left([\underline{a}_{-i}^{n-1}, \bar{a}_{-i}^{n-1}]\right) &= r_i\left(\bar{a}_{-i}^{n-1}\right).
\end{aligned}$$

Therefore, $\rho^n(A) = \bigtimes_{i\in I}[\underline{a}_i^n, \bar{a}_i^n] = \bigtimes_{i\in I}\left[r_i\left(\underline{a}_{-i}^{n-1}\right), r_i\left(\bar{a}_{-i}^{n-1}\right)\right]$.

Since $(\rho^n(A))_{n=1}^{\infty} = \left(\bigtimes_{i\in I}[\underline{a}_i^n, \bar{a}_i^n]\right)_{n=1}^{\infty}$ is a weakly decreasing sequence of Cartesian subsets, for each $i \in I$, $(\underline{a}_i^n)_{n=1}^{\infty}$ is weakly increasing and $(\bar{a}_i^n)_{n=1}^{\infty}$ is weakly decreasing, therefore these sequences have limits $\underline{a}_i^* = \lim_{n\to\infty}\underline{a}_i^n$ and $\bar{a}_i^* = \lim_{n\to\infty}\bar{a}_i^n$. Thus,

$$\rho^{\infty}(A) = \bigcap_{n=1}^{\infty}\rho^n(A) = \bigcap_{n=1}^{\infty}\bigtimes_{i\in I}[\underline{a}_i^n, \bar{a}_i^n] = \bigtimes_{i\in I}[\underline{a}_i^*, \bar{a}_i^*].$$

By Theorem 2 and 7,

$$\rho^{\infty}(A) = \rho(\rho^{\infty}(A)) = \mathrm{r}\left(\bigtimes_{i\in I}[\underline{a}_i^*, \bar{a}_i^*]\right) = \bigtimes_{i\in I} r_i\left([\underline{a}_{-i}^*, \bar{a}_{-i}^*]\right) = \bigtimes_{i\in I}[\underline{a}_i^*, \bar{a}_i^*].$$

Again, Lemma 8 implies that

$$\begin{aligned} \min r_i\left([\underline{a}_{-i}^*, \bar{a}_{-i}^*]\right) &= r_i\left(\underline{a}_{-i}^*\right), \\ \max r_i\left([\underline{a}_{-i}^*, \bar{a}_{-i}^*]\right) &= r_i\left(\bar{a}_{-i}^*\right) \end{aligned}$$

for each $i \in I$. Therefore, $\rho^{\infty}(A) = \bigtimes_{i\in I}[\underline{a}_i^*, \bar{a}_i^*] = \bigtimes_{i\in I}\left[r_i\left(\underline{a}_{-i}^*\right), r_i\left(\bar{a}_{-i}^*\right)\right]$. ■

Lemma 16 implies the main result of this section:

Theorem 10. *For every supermodular nice game G the set of rationalizable action profiles is an order box $\rho^{\infty}(A) = [\underline{a}^*, \bar{a}^*]$ where $\underline{a}^*$ is the smallest Nash equilibrium and $\bar{a}^*$ is the largest Nash equilibrium, that is, $\underline{a}^* \leq a^* \leq \bar{a}^*$ for every Nash equilibrium a^*.*

Proof. By Lemma 16,

$$[\underline{a}^*, \bar{a}^*] = \bigtimes_{i\in I}[\underline{a}_i^*, \bar{a}_i^*] = \bigtimes_{i\in I}\left[r_i\left(\underline{a}_{-i}^*\right), r_i\left(\bar{a}_{-i}^*\right)\right].$$

Therefore $\underline{a}_i^* = r_i\left(\underline{a}_{-i}^*\right)$ and $\bar{a}_i^* = r_i\left(\bar{a}_{-i}^*\right)$ for each $i \in I$, which implies that $\underline{a}^*$ and $\bar{a}^*$ are Nash equilibria. Fix any Nash equilibrium a^*. Since a^* is rationalizable (see Remark 20), it follows that $a^* \in [\underline{a}^*, \bar{a}^*]$, that is, $\underline{a}^* \leq a^* \leq \bar{a}^*$. ■

Corollary 3. *Let $G = \langle I, (A_i, u_i)_{i \in I} \rangle$ be a supermodular nice game with a unique Nash equilibrium. Then there is a unique rationalizable action profile, the Nash equilibrium.*

Proof. By Theorem 10, the set of rationalizable action profiles is an order box $[\underline{a}^*, \bar{a}^*]$ where $\underline{a}^*$ and $\bar{a}^*$ are Nash equilibria. By uniqueness of the Nash equilibrium, $\underline{a}^* = \bar{a}^*$; therefore $[\underline{a}^*, \bar{a}^*]$ is a singleton. ∎

For instance, suppose that the (increasing) best reply functions of a two-player supermodular nice game are represented in Figure 5.1. Then, there are three Nash equilibria x, y, and z. It follows from Theorem 10 that the set of rationalizable action profiles is the order box $[x, z]$, which is represented by the thick dashed line.

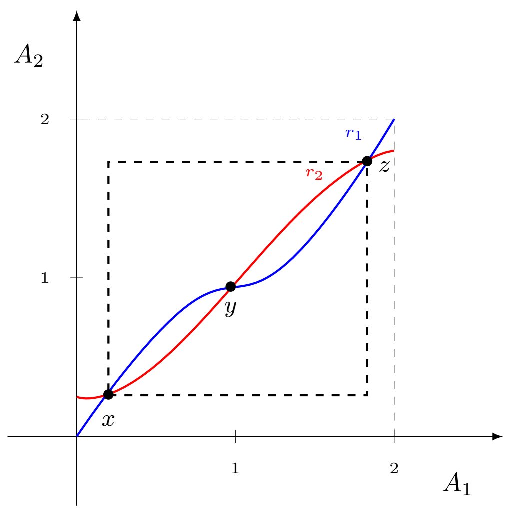


Figure 5.1: A supermodular nice game with $A_1 = A_2 = [0, 2]$.

Finally, we report an additional result which follows from Theorem 10.

Theorem 11. *In every* symmetric *supermodular nice game the set of rationalizable action profiles has a smallest element $\underline{a}^*$ and a largest element $\bar{a}^*$ and they are* symmetric *Nash equilibria.*

Part of the proof of Theorem 11 relies on the following lemma, which shows the symmetry of the equilibrium set in symmetric games.[7]

Lemma 17. *The set of Nash equilibria of a symmetric game is symmetric.*

Proof. If the set of Nash equilibria is empty, then the statement vacuously holds. So, in what follows, we assume that this set is non-empty. Let a^* be a Nash equilibrium of a symmetric game. We must show that, for any permutation $\pi : I \to I$, $a^* \circ \pi$ is also a Nash equilibrium. Let $\hat{A}$ denote the common action set. Since a^* is an equilibrium

$$\forall i \in I, \forall b \in \hat{A},\ u_i\left(a^*\right) \geq u_i\left(b, a^*_{-i}\right).$$

By symmetry

$$\begin{aligned}\forall i \in I, \forall b \in \hat{A},\ u_i\left(a^* \circ \pi\right) &= u_{\pi(i)}\left(\left(a^* \circ \pi\right) \circ \pi^{-1}\right) = u_{\pi(i)}\left(a^*\right)\\ &\geq u_{\pi(i)}\left(b, a^*_{-\pi(i)}\right) = u_i\left(b, \left(a^* \circ \pi\right)_{-i}\right).\end{aligned}$$

Therefore $a^* \circ \pi$ is an equilibrium. ■

Proof of Theorem 11. By Theorem 10, $\rho^\infty(A) = [\underline{a}^*, \bar{a}^*]$, where $\underline{a}^*$ and $\bar{a}^*$ are, respectively, the smallest and largest Nash equilibrium. Lemma 17 shows that the set of Nash equilibria of a symmetric game is symmetric, that is, for every Nash equilibrium $a^* \in \hat{A}^I$—where $\hat{A}$ is the common action set of each player—and every permutation (bijection) $\pi : I \to I$, $a^* \circ \pi$ is also a Nash equilibrium. Now, let $\hat{A} \subseteq \mathbb{R}$ (as is the case in a symmetric nice game), fix $a = (a_i)_{i\in I} \in \hat{A}^I$ arbitrarily, and note that $a \circ \pi \leq a$ for every permutation only if a is symmetric, that is, $a_i = a_j$ for all $i, j \in I$. To see this, pick any two players i and j so that $a_i \leq a_j$ and let π be any permutation such that $j = \pi(i)$; since $a \circ \pi \leq a$, then $a_j = a_{\pi(i)} \leq a_i$. Since $a_i \leq a_j$ and $a_j \leq a_i$, then $a_i = a_j$. By a similar argument, if $a \leq a \circ \pi$ for every permutation π, then a is symmetric. In particular, by symmetry of the set of Nash equilibria, $\underline{a}^* \circ \pi$ is a Nash equilibrium and $\underline{a}^* \leq \underline{a}^* \circ \pi$, $\bar{a}^* \circ \pi$ is a Nash equilibrium and $\bar{a}^* \circ \pi \leq \bar{a}^*$; thus, $\underline{a}^*$ and $\bar{a}^*$ are symmetric. ■

[7] Note that also the set of rationalizable profiles of a symmetric game is symmetric. Since it is a Cartesian product, it must have the form $\hat{C}^I$ for some subset $\hat{C}$ of the common action set $\hat{A}$. We could use the symmetry of the rationalizable set, rather than the symmetry of the equilibrium set, to prove Theorem 11.

5.4 Interpretations of the Nash Equilibrium Concept

Nash equilibrium is the most well known and applied equilibrium concept in economic theory, besides the competitive equilibrium. Indeed, we have argued in the Introduction that, in principle, any economic situation (and more generally any social interaction) can be represented as a non-cooperative game. The property according to which every action is a best reply to the other players' actions seems to be essential in order to have an equilibrium in a non-cooperative game.

Nonetheless, we should refrain from accepting this conclusion without further reflection. Why does the Nash equilibrium represent an interesting theoretical concept? When should we expect that the actions simultaneously chosen by the players form an equilibrium? Why should players hold correct conjectures regarding each other's behavior?

We propose a few different interpretations of the Nash equilibrium concept, each addressing the questions above. Such interpretations can be classified in two subgroups: (1) a Nash equilibrium represents "an obvious way to play," (2) a Nash equilibrium represents a stationary (stable) state of an adaptive process. In some cases, we will also introduce a corresponding generalization of the equilibrium concept which is appropriate under the given interpretation and will be analyzed in Chapter 6.

(1) Equilibrium as an "obvious way to play": Assume *complete information*, i.e., common knowledge of the interactive situation represented by G (recall from Chapter 1, G is a mathematical structure reperesenting the interactive situation; yet, to ease language, in this discussion we also let symbol G denote the situation itself).[8] It could be the case that from such common knowledge and shared assumptions about behavior and beliefs, or from some prior events that occurred before the game, the players could positively conclude that a specific action profile a^* represents an "obvious way to play" G. If a^* represents an obvious way to play, every player i expects that everybody else chooses his action in a^*_{-i}. Moreover, if a^* is to be played by rational players, it must be the case that no player i has incentives to choose an action different from a^*_i:

[8] Complete information is sufficient to justify the considerations that follow, but it is not strictly necessary.

if i had an incentive to choose $a_i \neq a_i^*$, not only a_i^* would not be chosen, but also the other players would have no reason to believe that a^* is the obvious way to play. Hence, a^* must be a Nash equilibrium. Indeed, in his PhD thesis, John Nash motivates his solution concept as follows (Nash [50], p. 23):

> "We proceed by investigating the question: what would be a 'rational' prediction of the behavior to be expected of rational playing the game in question? By using the principles that a rational prediction should be unique, that the players should be able to deduce and make use of it, and that such knowledge on the part of each player of what to expect the others to do should not lead him to act out of conformity with the prediction, one is led to the concept of a solution defined before."

What can make a^* an obvious way to play?

(1.a) Deductive interpretation: The players analyze G as an interactive decision problem and look for a rational solution. If such solution exists and is unique, then it corresponds to an obvious way to play. As an example of solution of a game, consider the case where G has a unique rationalizable profile a^* (this is the case for several models in economic theory, such as Example 17 in Section 4.6). Then, if all players are rational and their conjectures are derived from the assumption that there is common belief in rationality, they will choose exactly the actions in a^*. Moreover, a^* is necessarily the unique Nash equilibrium of the game (Remark 20).

Rationalizability captures deductive strategic thinking in a compelling way, and it justifies Nash equilibrium as an obvious way to play when there is a unique rationalizable outcome. But what about games where Nash equilibrium action profiles are a strict subset of rationalizable action profiles? If we insist in the deductive interpretation, perhaps we should just stick to rationalizability.

(1.b) Non-binding agreement: Suppose that the players are able to communicate before playing the game and that they reach an agreement to choose the actions specified in the action profile a^*. Suppose also that such agreement is not legally binding for the parties, it is simply a "gentlemen agreement" based on players honoring their words. Further, players attach little value to honoring their words: if i believes that a different action yields him a higher utility, he will not choose a_i^*. All players are perfectly

aware of this. Therefore, the agreement is credible, or "self-enforcing" only if no player has an incentive to deviate from the agreement, i.e., only if a^* is a Nash equilibrium.

Is it really necessary that players agree on a specific action profile? Perhaps they could agree on making the action profile actually played (for instance, the way they coordinate in a "Battle of the Sexes" game) depend on some exogenous random variable, such as the weather. In some cases, this would allow to reach, in expectation, fairer outcomes. We will go back to this point in the next chapter.

(2) Equilibrium as a stationary state of an adaptive process.[9] When introductory economic textbooks explain why in a competitive market the price should reach the equilibrium level that equates demand and supply, they almost inevitably rely on informal dynamic justifications. They argue, for instance, that if there is an excess demand the sellers will realize that they are able to sell their goods for a higher price; conversely, if there is an excess supply the sellers will lower their prices to be able to sell the residual unsold goods. Essentially, such arguments rely more or less explicitly on the existence of a feedback effect that pushes the prices towards the equilibrium level. These arguments cannot be formalized in the standard competitive equilibrium model, where market prices are taken as parametrically given by the economic agents and then determined by the demand=supply conditions. They nonetheless provide an intuitive support to the theory.

Similar arguments can be used to explain why the actions played in a game should eventually reach the equilibrium position. In some sense, the conceptual framework provided by game theory is better suited to formalize this kind of arguments. As explained in the Introduction, in a game model every variable that the analyst tries to explain, i.e., any "endogenous" variable, is directly or indirectly determined by the players' actions, according to precise rules that are part of the game (unlike prices in a competitive market model, where the price-formation mechanism is not specified). Assuming that the given game represents an interactive situation that players face recurrently, one can formulate assumptions regarding how players modify their actions taking into account the

[9] To go deeper on this topic the reader can start with the survey by Battigalli *et al.* [15]. In the already mentioned paper by Milgrom and Roberts [46], the connections between the concept of rationalizability and adaptive processes are explored.

outcomes of previous interactions, thus representing formally the feedback process.

One can distinguish two different types of adaptive dynamics: learning and evolutionary dynamics. Here, we will present only a general and brief description.[10]

(2.a) Learning dynamics. Assume that a given game G is played recurrently and that players are interested in maximizing their current expected payoff (a reason for this may be that they do not value the future or that they believe that their current actions do not affect in any way future payoffs). Players have access to information on previous outcomes. Based on such information, they modify their conjectures about their opponents' behavior in the current period. Let a^* be a Nash equilibrium. If in period t every player i expects his opponents to choose a^*_{-i}, then a^*_i is one of his best replies. It is therefore possible that i chooses a^*_i. If this happens, what players observe at the end of period t will confirm their conjectures, which then will remain unchanged for the following period. So even a small inertia (that is a preference, *coeteris paribus*, to repeat the previous action) will induce players to repeat in period $t+1$ the previous actions a^*. Analogously, a^* will be played also in period $t+2$ and so on. Hence, the equilibrium a^* is a stationary state of the process.

We just argued that every Nash equilibrium is a stationary state of plausible learning processes (we presented the argument in an informal way, but a formalization is possible). (i) Do such processes always converge to a steady state? (ii) Is it true that, for any plausible learning process, every stationary state is a Nash equilibrium? It is not possible, in general, to give affirmative answers. First, one has to be more specific about the dynamics of the recurrent interaction. For instance, one has to specify exactly what players are able to observe about the outcomes of previous interactions: is it the action profile of the other players? Is it some variables that depend on such actions? Is it only their own payoffs? Another relevant issue is whether game G is played always by the same agents, or in every period agents are randomly matched with "strangers." An exact specification of these assumptions shows that the process does not always converge to a stationary state. Moreover, as we will see in the next section, there may be stationary states that do not satisfy the Nash equilibrium condition

[10] In Chapter 7, we present an elementary analysis of learning dynamics relating their long-run behavior to equilibrium concepts and rationalizability.

of Definition 18. There are two reasons for this: (a) Players' conjectures can be confirmed by observed outcomes even if they are not correct. (b) If players are randomly drawn from large populations, then the state variable of the dynamic process is given by the fractions of agents in the population that choose each action. But then it is possible that in a stationary state two different agents, playing in the same role, choose different actions. Even though no agent actually randomizes, such situations look like "mixed equilibria," in which players choose randomly among some of the best replies to their probabilistic conjectures, and such conjectures happen to be correct. We analyze notions of equilibrium corresponding to situations (a) and (b) in the next chapter.

(2.b) Evolutionary dynamics. In the analysis of adaptive processes in games, the analogy with evolutionary biology has often been exploited. Consider, for instance, a symmetric two-person game. In every period two agents are drawn from a large population. Then they meet, they interact, they obtain some payoff and finally they split. Individuals are compared to animals, or plants, whose behavior is determined by a genetic code transmitted to their offsprings. If action a is more successful than b, then the agents programmed to play a reproduce themselves faster than those programmed to play b, and therefore the ratio between the fraction of agents playing a and the fraction of agents playing b increases.

In evolutionary biology this theoretical approach based on game theory has been highly successful and has allowed to explain phenomena that appeared as paradoxical within a more naive evolutionary framework.[11] In the social sciences, evolutionary dynamics are used as a metaphor to represent, at an aggregate level, learning phenomena such as imitation, in which agents modify the way they play not on the basis of their direct personal experiences alone, but also observing the behavior of others in the same circumstances. Behaviors that turn out to be more successful are imitated and spread more rapidly.

As for learning dynamics, although Nash equilibria represent stationary states of evolutionary dynamics, the process need not always converge to a stationary state. Furthermore, there may be stationary states that do not satisfy the best reply property of Definition 18. Indeed, in this context a state is represented by the fractions in the population that use different actions. It may be the case that in a stationary state distinct agents of the

[11] See the monograph by Maynard Smith [45].

same population do different things.[12]

These considerations motivate the definition of more general equilibrium concepts whereby players may hold probabilistic, rather than deterministic conjectures about the opponents' behavior. For this reason we refer to these more general concepts—analyzed in Chapter 6—as "probabilistic equilibria."

[12] See the monographs by Weibull [71] and Sandholm [60].

6

Probabilistic Equilibria

Up to now we considered mixed actions only with reference to dominance relations: if a (pure) action is dominated by a mixed action, then a rational player should not choose it. We already noticed (Lemma 1), though, that an expected utility maximizer has no strict incentive to choose a mixed action. If there were a small cost for tossing a coin or spinning a roulette wheel, then no expected utility maximizer would choose a mixed action.

We adopt the point of view that (rational) players do not actually choose mixed actions. Nevertheless, it is possible to reconcile this point of view with an interpretation of mixed actions that makes them relevant independently of the equivalence results stated in Lemma 2 and Theorem 5. Specifically, in Section 6.1 we adopt the so-called "mass-action interpretation" of mixed actions put forward by John Nash in his thesis (Nash 1950): the individuals playing a given game G are drawn at random from large populations of agents and matched to play G; mixed actions represent *statistical distributions* of actions in populations; random matching implies that the share of population i playing a particular action a_j determines the probability that a_j is played in G; in a "mixed (Nash) equilibrium," any other player i correctly estimates this probability. In Section 6.2, we go back to the interpretation of an equilibrium as a non-binding self-enforcing agreement and notice that players may want to agree on probabilistic decision rules that link their behavior to the realization of extraneous random variables; such probabilistic agreements typically induce a spurious correlation between the actions of different players and are thus called "correlated equilibria." Finally, in Section 6.3 we interpret equilibria as the steady states of adaptive process whereby the game is played recurrently and players learn from their personal experience; this requires a specification of players' information feedback; in a steady state, called "self-confirming equilibrium," each agent holds a probabilistic conjecture confirmed by (i.e., consistent with) the feedback he receives and chooses a best reply; whether a steady state is a (mixed) Nash equilibrium depends on the properties of feedback.

6.1 Mixed Equilibrium

Consider the following game between home *owners* and *thieves*. Each owner has an apartment where he stores goods worth a total value of V. Owners have the option to keep an alarm system, which costs $c < V$.[1] The alarm system is not detectable by thieves. Each thief can decide whether to attempt a theft (burglary) or not. If there is no alarm system, the thieves successfully seize the goods and resell them to some dealer, making a profit of $V/2$. If there is an alarm system, the attempted burglary is detected and the police is automatically alerted. The thieves in such event need to leave all goods in place and try to escape. The probability they get caught is $\frac{1}{2}$ and in this case they are sanctioned with a monetary fine P and then released.[2]

If the situation just described were a simple game between one owner and one thief, it could be represented using the matrix below:

O\T	$Burglary$	No
$Alarm$	$V-c, -\frac{P}{2}$	$V-c, 0$
No	$0, \frac{V}{2}$	$V, 0$

Figure 6.1: Matrix 2

It is easy to check that Matrix 2 has no equilibria according to Definition 18.

[1] We assume for simplicity that c represents the cost of installing an alarm system as well as the cost of keeping it active. Furthermore, we neglect the possibility to insure against theft.

[2] Prisons do not exist. If thieves cannot pay the amount P, they are inflicted an equivalent punishment.

However, the game between owners and thieves is more complex. There are two large populations, and we assume for simplicity that they have the same size: the population of owners (each one with an apartment) and the population of thieves. Thieves randomly distribute themselves across apartments. For any given owner, the probability of an attempted burglary is equal to the fraction of thieves that decide to attempt a burglary. From the thieves' perspective, the probability that an apartment has an alarm system is given by the overall fraction of owners keeping an alarm system.

Assume that the game is played recurrently. The fractions of agents that choose the different actions evolve according to some adaptive process with the following features. At the end of each period it is possible to access (reading them on the newspapers) the statistics of the numbers of successful and unsuccessful burglaries. Players are fairly inert in that they tend to replicate the actions chosen in the previous period. However, they occasionally decide to revise their choices on the basis of the previous period statistics. Since in each period only a few agents revise their choices, such statistics change slowly.

An owner not equipped with an alarm system decides to install it if and only if the expected benefit is larger than the cost, i.e., if the proportion of attempted burglaries is larger than c/V. Conversely, an owner equipped with an alarm system will decide to get rid of it if and only if the proportion of attempted burglaries is lower than c/V. The proportion of attempted burglaries changes only slowly and the owners equate the probability of being robbed today with the one of the previous period. The probability that makes an owner indifferent between his two actions is c/V. When indifferent, an owner sticks to the previous period action.

Analogously, a thief that was active in the previous period decides not to attempt a burglary in the current period if and only if the fraction of apartments equipped with an alarm system (which is also the fraction of unsuccessful burglaries) is larger than $V/(V + P)$. A thief that was not active in the previous period attempts a burglary in the current one if and only if the fraction is lower than $V/(V + P)$.

The state variables of this process are the fraction α of installed alarm systems and the fraction β of attempted burglaries. It is not hard to see that α grows (resp., decreases) if and only if $\beta > \frac{c}{V}$ (resp. $\beta < c/V$). Similarly, β grows (resp., decreases) if and only if $\alpha < V/(V + P)$ (resp., $\alpha > V/(V + P)$). If $\alpha = V/(V + P)$ and $\beta = c/V$, then the state of the system does not change, i.e., $(\alpha, \beta) = (V/(V+P), c/V)$ is a stationary state, or rest point, of the dynamic process. Whether this rest point is stable depends on details that we have not specified. The mixed action pair $(\alpha_1, \alpha_2) = (V/(V+P), c/V)$ is said to be a *mixed* (Nash) *equilibrium* of the matrix game above.

In this example, we have interpreted the mixed action of j, say α_j, both as a *statistical distribution* of actions in population j and as a *conjecture* of the agents of population i about the action of the opponent. In a stable environment every conjecture about j, α_j, is *correct* in the sense that it corresponds to the statistical distribution of actions in population j. Furthermore, α_j is such that every agent in population i is *indifferent* among the actions that are chosen by a positive fraction of agents.

Next, we present a general definition of mixed equilibrium and then show that it is characterized by the aforementioned properties. To simplify the analysis, we restrict our attention to *finite* games, but the following concepts and results can be extended to compact-continuous games.

Definition 20. *The* ***mixed extension*** *of a finite game* $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ *is a game* $\overline{G} = \langle I, (\Delta(A_i), \overline{u}_i)_{i\in I}\rangle$ *where*

$$\overline{u}_i(\alpha) = \sum_{a=(a_j)_{j\in I}\in A} u_i(a) \prod_{j\in I} \alpha_j(a_j),$$

for all i *and* $\alpha = (\alpha_i)_{i\in I} \in \bigtimes_{i\in I} \Delta(A_i)$.

Note, the payoff functions of the mixed extension are obtained calculating the expected payoff corresponding to a vector of mixed actions under the assumption that the actions of different players are *statistically independent.*

Definition 21. *(Nash [49]) Fix a finite game* G*. A mixed action profile* α *is a* ***mixed equilibrium*** *of game* G *if* α *is a Nash equilibrium of the mixed extension of* G*.*

Note that every (pure) Nash equilibrium $(a_i^*)_{i\in I}$ corresponds to a mixed equilibrium $(\alpha_i^*)_{i\in I}$ such that each α_i^* assigns probability one to a_i^*: mixed equilibrium is a generalization of (pure) Nash equilibrium.

Theorem 12. *Every finite game* G *has at least one mixed equilibrium.*[3]

[3] The same holds for any compact-continuous game.

Proof. It is easy to verify that the mixed extension $\overline{G}$ of a finite game G satisfies all the assumptions of Theorem 8: for each i, $\Delta(A_i)$ is a nonempty, convex and compact subset of $\mathbb{R}^{A_i}$, $\overline{u}_i$ is a multi-linear function and therefore it is both continuous in α and weakly concave, indeed linear,[4] in α_i. ■

When we introduced mixed actions in Chapter 3, we said that, even if we do not assume that players randomize, mixed actions still play an important technical role. The space of mixed actions is convex and expected payoff is multi-linear in the mixed actions of different players. In Chapter 3, such convexity-linearity structure was exploited to characterize the set of justifiable (pure) actions (Lemma 2). In Theorem 12, the convexification-linearization allowed by the mixed extension of a game yields existence of an equilibrium.

The following result introduces an alternative characterization of mixed equilibria, which is not based on the mixed extension of the game. With a small abuse of notation, we let $r_i(\alpha_{-i})$ denote the set of best replies to the conjecture $\mu^{\alpha_{-i}} \in \Delta(A_{-i})$ obtained as a product of the marginal measures α_j $(j \neq i)$, that is, $r_i(\alpha_{-i}) = r_i(\mu^{\alpha_{-i}})$ where

$$\forall \alpha_{-i} \in \bigtimes_{j \neq i} \Delta(A_j), \forall a_{-i} \in A_{-i},\ \mu^{\alpha_{-i}}(a_{-i}) = \prod_{j \neq i} \alpha_j(a_j).$$

Theorem 13. *A mixed action profile $(\alpha_i)_{i \in I}$ is a mixed equilibrium if and only if, for every $i \in I$, $\mathrm{supp}\alpha_i \subseteq r_i(\alpha_{-i})$ (in the finite case, every action played with positive probability is a best reply to the mixed action profile played by the opponents).*

Proof. The statement follows directly from the definition of mixed equilibrium and from Lemma 1. ■

Corollary 4. *All actions played with positive probability in a mixed equilibrium are rationalizable.*

[4] More precisely, for each α_{-i}, the section of $\bar{u}_i$ at α_{-i}, $\bar{u}_{i,\alpha_{-i}} : \Delta(A_i) \to \mathbb{R}$, is *affine*: for all $\alpha_i, \beta_i \in \Delta(A_i)$ and $\lambda \in [0,1]$,

$$\bar{u}_{i,\alpha_{-i}}(\lambda\alpha_i + (1-\lambda)\beta_i) = \lambda\bar{u}_{i,\alpha_{-i}}(\alpha_i) + (1-\lambda)\bar{u}_{i,\alpha_{-i}}(\beta_i).$$

Proof. By Theorem 13, the set $C = \times_{i\in I}\text{supp}\alpha_i$ has the best reply property. Hence, Theorem 3 implies that every action played with positive probability is rationalizable. ■

The reader should verify by example that the converse of Corollary 4 does not hold: there are (finite) games with rationalizable actions that are not played with positive probability in any mixed equilibrium.

Theorem 12 and Corollary 4 provide a linear programming algorithm to compute all mixed equilibria of finite two-person games. In such games the payoff function of player i can be represented by a matrix with generic entry $u_i^{k\ell}$. Player 1 chooses the rows (indexed by k) and player 2 the columns (indexed by ℓ).

Step 1: Eliminate all iteratively dominated actions (by Theorem 5 and Corollary 4 such actions are played with zero probability in equilibrium). The order of elimination is irrelevant (Theorem 6).

Step 2: For any pair of nonempty subsets $A_1^* \subseteq A_1$ and $A_2^* \subseteq A_2$, compute the set of mixed equilibria, (α_1, α_2), such that $\text{supp}\alpha_1 = A_1^*$ and $\text{supp}\alpha_2 = A_2^*$. This set, which could be empty, is computed as follows (we consider the non trivial case in which each set contains at least two actions). To simplify the notation, assume that $A_1^* = \{1, ..., K\}$ and $A_2^* = \{1, ..., L\}$ and denote by α_1^k (respectively, α_2^ℓ) the probability of action k (resp. ℓ) of player 1 (resp. 2). Solve the following systems of linear equations and inequalities with unknown $\alpha_1 = (\alpha_1^1, ..., \alpha_1^K) \in \Delta(A_1^*)$ and $\alpha_2 = (\alpha_2, ..., \alpha_2^L) \in \Delta(A_2^*)$:

$$\begin{aligned} \sum_{k=1}^{K} u_2^{k\ell}\alpha_1^k &= \sum_{k=1}^{K} u_2^{k1}\alpha_1^k,\ \ell = 2, ..., L, \\ \sum_{k=1}^{K} u_2^{k\ell}\alpha_1^k &\leq \sum_{k=1}^{K} u_2^{k1}\alpha_1^k,\ \ell = L+1, ..., |A_2|; \end{aligned} \tag{6.1.1}$$

$$\begin{aligned} \sum_{\ell=1}^{L} u_1^{k\ell}\alpha_2^\ell &= \sum_{\ell=1}^{L} u_1^{1\ell}\alpha_2^\ell,\ k = 2, ..., K, \\ \sum_{\ell=1}^{L} u_1^{k\ell}\alpha_2^\ell &\leq \sum_{\ell=1}^{L} u_1^{1\ell}\alpha_2^\ell,\ k = K+1, ..., |A_1|. \end{aligned} \tag{6.1.2}$$

The equations in (6.1.1) determines the set of mixed actions of player 1 that make player 2 indifferent between the actions in subset A_2^*. The inequalities determines the set of mixed actions of player 1 that make $a_2 = 1$ (and so *all the actions in* A_2^*) weakly preferred to the actions that do not belong to A_2^*. For any α_1 that satisfies (6.1.1), player 2 has no incentive to "deviate" from any mixed action with support A_2^*. Similar considerations hold for system (6.1.2). Such system determines the set of mixed actions of player 2 that make player 1 indifferent among all the actions in A_1^* and at the same time make such actions weakly preferred to all the others. Therefore, *the indifference conditions for player 1 determine the equilibrium randomization(s) of player 2, the indifference conditions for player 2 determine the equilibrium randomization(s) of player 1.*

In the previous example (Owners and Thieves) the equilibrium is determined as follows: First, note that the best reply to a deterministic conjecture is unique. However, no pair of pure actions is an equilibrium. Hence, the equilibrium is necessarily mixed (existence follows from Theorem 12). An equilibrium with support $A_1^* = \{Alarm, No\}$, $A_2^* = \{Burglary, No\}$ is given by the following (we are writing $\alpha = \alpha_1(A)$ and $\beta = \alpha_2(B)$):

Indifference condition for $i = 1$ *(Owner):*

$$V - c = V(1 - \beta).$$

Indifference condition for $i = 2$ *(Thief):*

$$\frac{V}{2}(1 - \alpha) - \frac{P}{2}\alpha = 0.$$

Solving the system, $\alpha = \frac{V}{V+P}$, $\beta = \frac{c}{V}$, which is the stationary state identified by the (informal) analysis of a plausible dynamic process.

6.2 Correlated Equilibrium

Mixed equilibria can be interpreted as statistical distributions over action profiles that are in some sense stationary. Such distributions are obtained as a product of the marginal distributions corresponding to the equilibrium mixed actions: if $(\alpha_i^*)_{i\in I}$ is a mixed equilibrium, then the probability of any profile $(a_i)_{i\in I}$ is $\prod_{i\in I} \alpha_i^*(a_i)$; this means that actions are statistically independent across players. Statistical independence is justified by the interpretation of mixed equilibrium as a stationary vector of statistical distributions of actions in $|I|$ populations whose agents are randomly matched with the agents of the other populations to play the game G.

Now we present a different concept of probabilistic equilibrium that, formally, generalizes the concept of mixed equilibrium by allowing *correlated* distributions over players' action profiles. The interpretation of this equilibrium concept given here is rather different from the one proposed for the mixed equilibrium.[5]

Let us consider once again the case in which players are able to communicate and sign a *non-binding* agreement before the game is played. If this agreement simply prescribes to play a certain profile a^*, then—as noted earlier—a^* must be a Nash equilibrium, otherwise it would not be self-enforcing. However, players could reach more sophisticated self-enforcing agreements, in which the chosen actions depend on the realizations of extraneous random variables that do not directly affect the payoffs. We first illustrate this idea with a simple example.

	B	S
B	$3,1$	$0,0$
S	$0,0$	$1,3$

Figure 6.2: Matrix 3.

Matrix 3 represents the classic *Battle of the Sexes* (BoS): Rowena (row player) and Colin (column player) would like to coordinate and go to the same concert, either Bach or Stravinsky, but Rowena prefers Bach and Colin prefers Stravinsky. Suppose that Rowena and Colin have to agree on how to play the BoS the next day, when no further communication between them will be possible. There exists two simple self-enforcing agreements, (B, B) and (S, S). However, the first favors Rowena, the second favors Colin, and neither player wants to give up. How to sort this out? Colin can make the following proposal that would ensure in expected values a fair distributions of the gains from coordination: "If tomorrow's weather is $b$ad then both of us choose Bach, if instead it is $s$unny then we both choose Stravinsky." Notice that the weather forecasts are uncertain: there is a

[5]Proposition 19 of Chapter 7 offers an adaptive interpretation.

50% probability of bad weather and 50% of sunny weather. The agreement generates an expected payoff of 2 for both players. Rowena understands that the idea is smart. Indeed, the agreement is self-enforcing as both players have an incentive to respect it if they expect the other to do the same. For instance, if Rowena expects that Colin sticks to the agreement and waking up she observes that the weather is $b$ad, then she expects that Colin will go to the Bach concert and she wants to play B. Similarly, if she observes that the weather is $s$unny, she expects that Colin will go to the Stravinsky concert and she wants to play S.

In other words, a sophisticated agreement can use exogenous and not directly relevant random variables to coordinate players beliefs and behavior. In such an agreement the conjecture of a player about the behavior of others (and so their best replies) depends on the observed realization of such random variables and the actions of different players are correlated, albeit spuriously.

Clearly, it is not always possible to condition the choice on some commonly observable random variable. Furthermore, even if this were possible, players could still find it more convenient to base their respective choices on random variables that are only partially correlated. Consider, for example, the following bi-matrix game:

	a	b
a	$6,6$	$2,7$
b	$7,2$	$0,0$

Figure 6.3: Matrix 4.

If Rowena and Colin rely on a commonly observed random variable to design a self-enforcing agreement, they can only achieve a probability distribution over pure Nash equilibria. Actually this is true in every game: suppose that the agreement says "if x is (commonly) observed, each $j \in I$ must take action $\sigma_j(x)$," then when i observes x he infers that each co-player j will take action $\sigma_j(x)$, and—in order to make the agreement self-enforcing—$\sigma_i(x)$ must be a best reply to $\sigma_{-i}(x)$, thus $(\sigma_i(x))_{i\in I}$ must be a Nash equilibrium.[6] The two pure equilibria of the

[6]One can generalize and obtain distributions over mixed equilibria if players choose mixed actions according to the observed realization x.

game in Matrix 4 yield the payoff pairs $(7, 2)$ and $(2, 7)$. So, the sum of the expected payoffs attainable with self-enforcing agreements that rely on a commonly observed random variable is $\mu \times (2+7) + (1-\mu) \times (7+2) = 9$, where μ is the probability of observing a realization x that yields (a, b) ($\mu = \mathbb{P}(\{x : \sigma_1(x) = a, \sigma_2(x) = b\})$).

If instead Rowena and Colin rely on different (but correlated) random variables, they can do better. For example, suppose that Rowena observes whether $X = \mathbf{a}$ or $X = \mathbf{b}$, and Colin observes whether $Y = \mathbf{a}$ or $Y = \mathbf{b}$, where (X, Y) is a pair of random variables with the following joint distribution:

$X\backslash Y$	**a**	**b**
a	1/3	1/3
b	1/3	0

They agree that each one should play a when she or he observes **a**, and b when she or he observes **b**. It can be checked that this agreement is self-enforcing. Suppose that Rowena observes $X = \mathbf{a}$, then she assigns the same (conditional) probability, $\frac{1}{2} = \frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{3}}$, to actions a and b by Colin; thus, the expected payoff of choosing a is 4, while the expected payoff of choosing b is only 3.5. If she observes $X = \mathbf{b}$, then she is sure that Colin observes $Y = \mathbf{a}$ and plays a; so she best responds with b. A symmetric argument applies to Colin. Therefore no player wants to deviate, whatever she or he observes. The sum of the expected payoffs in this agreement is $\frac{1}{3}(6+6) + \frac{1}{3}(2+7) + \frac{1}{3}(7+2) = 10$.

This rather informal discussion motivates the following definitions and results. As we did for mixed equilibria, for the sake of simplicity we restrict the analysis to *finite* games.

Definition 22. *(see Aumann [5] and [6]) A probabilistic self-enforcing agreement, or* ***correlated equilibrium****, is a structure $\langle \Omega, p, (T_i, \tau_i, \sigma_i)_{i \in I} \rangle$, where (Ω, p) is a finite probability space ($p \in \Delta(\Omega)$), the sets T_i are finite*[7]

[7] Finiteness is just a simplification without substantial loss of generality when A is finite. If Ω and T_i are infinite, they must be endowed with sigma-algebras of events, probability measure p is defined on the sigma-algebra of Ω, and all the functions τ_i and σ_i must be measurable.

and the functions $\tau_i : \Omega \to T_i$ *and* $\sigma_i : T_i \to A_i$ *are such that,*

$$\forall i \in I, \forall t_i \in T_i, \forall a_i \in A_i, \text{ if } p\left(\left\{\omega' : \tau_i(\omega') = t_i\right\}\right) > 0 \text{ then}$$

$$\sum_{\omega} p(\omega|t_i) u_i \left(\sigma_i(t_i), \sigma_{-i}(\tau_{-i}(\omega))\right) \geq \sum_{\omega} p(\omega|t_i) u_i \left(a_i, \sigma_{-i}(\tau_{-i}(\omega))\right).$$

In the above definition τ_i represents the random variable, also called **signal**, observed by player i, whereas σ_i represents the **strategy,** or **decision function,** imposed by the agreement on player i;

$$p(\omega|t_i) = \begin{cases} p(\omega)/p\left(\{\omega' : \tau_i(\omega') = t_i\}\right), & \text{if } \tau_i(\omega) = t_i, \\ 0, & \text{if } \tau_i(\omega) \neq t_i, \end{cases}$$

is the probability of state ω conditional on observing t_i (which is well defined if the probability of t_i is positive)[8] and

$$\sigma_{-i}(\tau_{-i}(\omega)) = \left(\sigma_j(\tau_j(\omega))\right)_{j \neq i}$$

is the action profile chosen by the other players (according to the agreement) if state ω occurs.

The definition requires that players have no incentive to deviate from the behavior prescribed by the agreement, whatever the realization of their signals. Thus, *conditional on every possible observation, the expected payoff of a deviation cannot be higher than the expected payoff obtained by sticking to the agreement.* The inequalities expressing these conditions are sometimes called *incentive compatibility constraints.*

Clearly, a correlated equilibrium induces a probability measure over players' action profiles, $\mu \in \Delta(A)$, where

$$\forall a \in A,\ \mu(a) = p\left(\{\omega : \forall i \in I, \sigma_i(\tau_i(\omega)) = a_i\}\right). \tag{6.2.1}$$

Formally, measure $\mu \in \Delta(A)$ is the image of measure $p \in \Delta(\Omega)$ through the pushforward given by the composite function $\sigma \circ \tau : \Omega \to A$, that is,

[8] This is the special case of the standard definition of conditional probability:

$$p(F) > 0 \Rightarrow p(E|F) = \frac{p(E \cap F)}{p(F)}.$$

If $E = \{\omega\}$ (a singleton) and $F = \{\omega' : \tau_i(\omega') = t_i\}$, then the formula in the main text obtains.

$\mu = p \circ (\sigma \circ \tau)^{-1}$, where $\tau = (\tau_i)_{i \in I} : \Omega \to T$ and $\sigma = (\sigma_i)_{i \in I} : T \to A$; more explicitly,

$$\mu(E) = p((\sigma \circ \tau)^{-1}(E)) = \sum_{\omega : \sigma(\tau(\omega)) \in E} p(\omega)$$

for each subset of action profiles $E \subseteq A$. To analyze the properties of such induced probability measures we introduce the notion of "canonical" correlated equilibrium:

Definition 23. *A self-enforcing probabilistic agreement* $\langle \Omega', p', (\tau_i', \sigma_i')_{i \in I} \rangle$ *where*
(1) $\Omega' = A$,
(2) $p' \in \Delta(A)$,
(3) $\forall i \in I, \forall a = (a_i)_{i \in I} \in A$, $\tau_i'(a) = a_i$ ($\tau_i' = \mathrm{proj}_{A_i}$ *is the projection function from* A *onto* A_i*),*
(4) $\forall i \in I$, $\forall a_i \in A_i$, $\sigma_i'(a_i) = a_i$ ($\sigma_i' = \mathrm{Id}_{A_i}$ *is the identity function on* A_i*),*
is called ***canonical*** *correlated equilibrium.*

Although, formally, it is the whole structure $\langle \Omega', p', (\tau_i', \sigma_i')_{i \in I} \rangle$ that forms a canonical correlated equilibrium, it is standard to call "canonical correlated equilibrium" just the distribution $p' \in \Delta(A)$, as all the other elements of the structure are trivially determined by the given game G.

Theorem 14. *If* $\langle \Omega, p, (T_i, \tau_i, \sigma_i)_{i \in I} \rangle$ *is a correlated equilibrium, then* $p' = \mu$*, where* μ *satisfies (6.2.1), is a canonical correlated equilibrium distribution.*

A canonical correlated equilibrium with distribution μ can be interpreted as the following mechanism. An action profile is chosen at random according to the "agreed upon" probability measure $\mu \in \Delta(A)$. Then a "mediator" observes the realized profile $a = (a_i)_{i \in I}$ and *privately suggests* to each player i to play action a_i. Then player i chooses freely any action $a_i' \in A_i$, but he has no incentive to deviate from the suggested action a_i.

Proof of Theorem 14. Let $\langle \Omega, p, (T_i, \tau_i, \sigma_i)_{i \in I} \rangle$ be a correlated equilibrium; fix arbitrarily a player i and an action a_i with strictly positive marginal probability:

$$\mu(a_i) := (\text{marg}_{A_i}\mu)(a_i) = \sum_{a_{-i}\in A_{-i}} \mu(a_i, a_{-i}) = p\left(\{\omega : \sigma_i(\tau_i(\omega)) = a_i\}\right) > 0,$$

where the second equality follows from the definition of marginal probability, and the third one from eq. (6.2.1). We must show that i has no incentive to deviate from the "mediator's suggestion" a_i, that is

$$\forall a_i', \sum_{a_{-i}\in A_{-i}} \mu(a_{-i}|a_i)\left[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})\right] \geq 0,$$

where

$$\mu(a_{-i}|a_i) := (\text{marg}_{A_{-i}}\mu(\cdot|a_i))(a_{-i}) = \frac{p\left(\{\omega : \sigma(\tau(\omega)) = (a_i, a_{-i})\}\right)}{p\left(\{\omega : \sigma_i(\tau_i(\omega)) = a_i\}\right)}.$$

Since $\mu(a_i) > 0$, the previous system of inequalities is equivalent to

$$\forall a_i',\ \mu(a_i) \sum_{a_{-i}\in A_{-i}} \mu(a_{-i}|a_i)\left[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})\right] \geq 0.$$

Since $\mu(a_i, a_{-i}) = \mu(a_{-i}|a_i)\mu(a_i)$, the incentive compatibility constraints for μ can be expressed as

$$\forall a_i', \sum_{a_{-i}\in A_{-i}} \mu(a_i, a_{-i})\left[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})\right] \geq 0. \tag{6.2.2}$$

Hence it is sufficient to show that (6.2.2) holds. This system of inequalities will be derived from the incentive compatibility constraints of the correlated equilibrium $\langle \Omega, p, (T_i, \tau_i, \sigma_i)_{i\in I}\rangle$ that induces μ, thus proving the result.

For all $t_i \in \sigma_i^{-1}(a_i)$ such that $p\left(\{\omega : \tau_i(\omega) = t_i\}\right) > 0$ (there must be at least one t_i like this), the incentive compatibility constraints yield

$$\forall a_i', \sum_{\omega:\tau_i(\omega)=t_i} p(\omega|t_i)[u_i\left(a_i, \sigma_{-i}(\tau_{-i}(\omega))\right) - u_i\left(a_i', \sigma_{-i}(\tau_{-i}(\omega))\right)] \geq 0.$$

Multiplying each inequality by $p\left(\{\omega : \tau_i(\omega) = t_i\}\right)$, taking into account that $p(\omega|t_i)p\left(\{\omega : \tau_i(\omega) = t_i\}\right) = p(\omega)$ for each $\omega \in \tau_i^{-1}(t_i)$, and taking the summation w.r.t. the t_i's in $\sigma_i^{-1}(a_i)$ yields

$$\forall a_i', \sum_{\omega:\sigma_i(\tau_i(\omega))=a_i} p(\omega)[u_i\left(a_i, \sigma_{-i}(\tau_{-i}(\omega))\right) - u_i\left(a_i', \sigma_{-i}(\tau_{-i}(\omega))\right)] \geq 0. \tag{6.2.3}$$

Since u_i depends only on actions, the terms in (6.2.3) can be regrouped to obtain

$$\begin{aligned} \forall a_i', \qquad & \sum_{a_{-i}} \sum_{\omega \in (\sigma\circ\tau)^{-1}(a_i,a_{-i})} p(\omega)[u_i\left(a_i, a_{-i}\right) - u_i\left(a_i', a_{-i}\right)] \qquad (6.2.4) \\ = \quad & \sum_{a_{-i}} [u_i\left(a_i, a_{-i}\right) - u_i\left(a_i', a_{-i}\right)] \sum_{\omega \in (\sigma\circ\tau)^{-1}(a_i,a_{-i})} p(\omega) \geq 0, \end{aligned}$$

where $\sigma = (\sigma_j)_{j\in I} : T \to A$, $\tau = (\tau_j)_{j\in I} : \Omega \to T$, $\sigma \circ \tau : \Omega \to A$ is the composition of σ and τ, and $(\sigma \circ \tau)^{-1}(a_i, a_{-i})$ is the set of pre-images in Ω of action profile (a_i, a_{-i}).

By definition of μ, $\sum_{\omega\in(\sigma\circ\tau)^{-1}(a_i,a_{-i})} p(\omega) = \mu(a_i, a_{-i})$. Therefore (6.2.4) yields (6.2.2) as desired. ∎

The first part of the proof yields the following:

Remark 21. *A distribution $\mu \in \Delta(A)$ is induced by a correlated equilibrium (hence, by Theorem 14, it is a canonical correlated equilibrium) if and only if it satisfies the following system of linear inequalities in the probabilities $(\mu(a))_{a\in A}$:*

$$\forall i \in I, \forall a_i, a_i' \in A_i, \sum_{a_{-i}\in A_{-i}} \mu(a_i, a_{-i}) \left[u_i(a_i, a_{-i}) - u_i(a_i', a_{-i})\right] \geq 0.$$

Hence, the set of measures $\mu \in \Delta(A)$ induced by a correlated equilibrium (the set of canonical correlated equilibria) is a convex and compact polytope.[9] *The same holds for the set of expected payoff profiles induced by correlated equilibria.*

The following observation, the proof of which is elementary, clarifies that the correlated equilibrium concept generalizes the mixed equilibrium concept:

Remark 22. *Fix a mixed action profile α and let $\mu^\alpha \in \Delta(A)$ be the measure defined by $\mu^\alpha(a) = \prod_{i\in I} \alpha_i(a_i)$ for every $a \in A$. Then α is a mixed equilibrium if and only if μ^α is a canonical correlated equilibrium.*[10]

[9] A polytope in $\mathbb{R}^2$ is just a polygon, a polytope in $\mathbb{R}^3$ is a polyhedron. "Polytope" is the generalization of these geometrical objects to $\mathbb{R}^n$.

[10] The word "correlated" here may be a bit misleading, as we are considering the case where players' actions are mutually independent. Nonetheless, this is a special case of the general definition of correlated equilibrium.

Like pure and mixed Nash equilibrium, also correlated equilibrium is a refinement of rationalizability:

Theorem 15. *If μ is a canonical correlated equilibrium, then every action to which μ assigns a positive marginal probability is rationalizable.*

Proof. For every i, let $C_i = \left\{a_i : \sum_{a_{-i}} \mu(a_i, a_{-i}) > 0\right\}$ be the set of i's actions to which μ assigns a positive marginal probability. It will be shown that $C = \times_{i\in I} C_i$ is a set with the best reply property. By Theorem 3, this implies that every action $a_i \in C_i$ is rationalizable, as desired. We write marginal and conditional probabilities using an obvious and natural notation: $\mu(a_i)$, $\mu(a_{-i}|a_i)$, $\mu(a_j|a_i)$.

For every i and a_i, if $a_i \in C_i$, then (by definition of C_i) $\mu(a_i) > 0$, and the conditional probability $\mu(\cdot|a_i) \in \Delta(A_{-i})$ is well defined. Also, for every $j \neq i$ and a_j, $\mu(a_j|a_i) > 0$ only if $\mu(a_j) > 0$. Hence, $\mathrm{supp}\mu(\cdot|a_i) \subseteq C_{-i}$. Finally, given that μ is a canonical correlated equilibrium it must be the case that

$$\forall a_i' \in A_i, \sum_{a_{-i}\in A_{-i}} \mu(a_{-i}|a_i)u_i(a_i, a_{-i}) \geq \sum_{a_{-i}\in A_{-i}} \mu(a_{-i}|a_i)u_i(a_i', a_{-i}),$$

that is, $a_i \in r_i(\mu(\cdot|a_i))$. Since this is true for each player i, it follows that C has the best reply property. ■

The concept of correlated equilibrium can be generalized assuming that different players can assign different probabilities to the states $\omega \in \Omega$. This generalization is called **subjective correlated equilibrium** (Brandenburger and Dekel [24]). It can be shown that an action a_i is rationalizable *if and only if* there exists a subjective correlated equilibrium in which a_i is chosen in at least one state ω (see Section 8.5.5 of Chapter 8).

6.3 Self-Confirming Equilibrium

We conclude this section on probabilistic equilibria introducing another generalization of the Nash equilibrium concept. In Section 5.4 we mentioned the possibility of interpreting an equilibrium as the stationary state of a learning process and we noted that such stationary states need not satisfy the Nash property. Consider the following simple example.

$1\backslash 2$	ℓ	r
t	$2,0$	$2,1$
b	$0,0$	$3,1$

Figure 6.4: Matrix 5.

The game in Matrix 5 has a unique rationalizable outcome, and so a unique Nash equilibrium (and a unique degenerate mixed equilibrium that coincides with the pure strategy equilibrium). If Rowena (player 1) knew the payoff function of Colin (player 2) and were to believe that Colin is rational, then she would expect him to play action r and would play the best reply b.

Instead, we assume the following: (i) information may be *incomplete*, each player knows his payoff function, but does not necessarily know the opponent's payoff function; (ii) the game is played recurrently and, at the end of each round, each player observes his realized payoff, but he does *not* observe directly the opponent's action; (iii) in every period players form probabilistic conjectures on the opponents' actions that are revised according to previous observations; consistently with Bayesian updating, when players observe something that was expected with probability one they do not revise their conjectures. Also, we assume for simplicity that (iv) players maximize the current expected payoff without worrying about the impact of current choices on future payoffs.

The situation of Colin is very simple: he will always chooses his dominant action r. Conversely, the situation of Rowena is not so simple. If in any given period t she expects ℓ with probability larger than $\frac{1}{3}$, then she best responds with t. In this case Rowena's payoff is equal to 2 *independently* of the choice made by Colin, hence she is not able to infer Colin's choice from observing the realization $u_1 = 2$, rather she observes something that she was expecting with certainty (to obtain a payoff of 2) and thus she does not revise her probabilistic conjecture. Hence also in period $t+1$ she repeats the same choice t and according to the same reasoning her conjecture remains unaffected. This shows that if the starting conjecture of Rowena assigns a sufficiently high probability to ℓ the process is stuck in (t, r), which is therefore a stationary state, even if it is not an equilibrium in the sense of Definitions 18 (Nash equilibrium) or 21 (mixed equilibrium).[11]

The example just described illustrates a situation in which players best respond to their conjectures and the information obtained *ex post* does not induce them to change their conjectures, even if they are incorrect. Situations of this kind are known as "**self-confirming equilibria**," or "conjectural equilibria."[12]

6.3.1 Pure Self-Confirming Equilibrium

As it is apparent from the previous example, in order to verify whether a certain situation is a self-confirming equilibrium, it is necessary to specify what players are able to observe *ex post*. We represent this information with a **feedback** function, $f_i : A \to M_i$, where M_i is a set of "messages" that i could receive at the end of each period. Let $f_{i,a_i} : A_{-i} \to M_i$ denote the section of f_i at a_i. Assuming that i remembers his choice, if i receives message m_i after he has chosen action a_i, he infers that the opponents must have played an action profile from the set $f_{i,a_i}^{-1}(m_i) = \{a_{-i} : f_i(a_i, a_{-i}) = m_i\}$ (In the previous example, $f_1(.) = u_1(.)$, $M_1 = \{0, 2, 3\}$, m_i means "You got m_i euros," $\{a_2 : f_1(t, a_2) = 2\} = \{\ell, r\}$, $\{a_2 : f_1(b, a_2) = 0\} = \{\ell\}$, $\{a_2 : f_1(b, a_2) = 3\} = \{r\}$). Suppose that i's conjecture is μ^i, i chooses the best reply $a_i^* \in r_i(\mu^i)$ and then observes m_i; suppose also that μ^i assigns probability one to the set $f_{i,a_i^*}^{-1}(m_i) = \{a_{-i} : f_i(a_i^*, a_{-i}) = m_i\}$; then the conjecture of i is confirmed (he observes what he expected with probability one), he sticks to it and keeps choosing a_i^*.

This preliminary discussion shows that in order to ascertain the stability of behavior under learning from personal experience, we must

[11] One could object to that if Rowena is patient, then she will want to *experiment* with action b so as to be able to observe (indirectly) Colin's behavior, even if that implies an expected loss in the current period. The objection is only partially valid. It can be shown that for any "degree of patience"(discount factor) there exists a set of initial conjectures that induce Rowena to choose always the safe action a rather than experimenting with b. It is true, however, that the more Rowena values future payoffs, the less it is plausible that the process will be stuck in (t, r).

[12] Both terms ("conjectural" and "self-confirming") have been used in the literature with reference to the same idea (see, for example, the literature review in [16]). For an analysis of how this concept has originated and its relevance for the analysis of adaptive processes in a repeated-iteration context see the survey by Battigalli *et al.* [15] and Chapter 7.

enrich the standard definition of "game" with an *ex post* information structure that specifies what players can learn after any round of a recurrent game has been played. We do so by adding to the static game (in reduced form) $G = \langle I, (A_i, u_i)_{i\in I}\rangle$ a profile of feedback functions $f = (f_i : A \to M_i)_{i\in I}$. We call the pair (G, f) **game with feedback**. Differently from the analysis of mixed and correlated equilibria, here we consider both finite and infinite games.

Definition 24. *A game with feedback (G, f) is **compact-continuous** if G is compact-continuous and, for each player i, the message space M_i is a compact subset of a Euclidean space $\mathbb{R}^{\ell_i}$ and the feedback function f_i is continuous.*

With this, we can define the self-confirming equilibrium concept (also called "conjectural equilibrium"), which captures with a static definition the possible steady states of learning dynamics. We start with equilibria in pure actions, which capture stability in a game recurrently played by the same set of individuals. Then we will generalize to distributions of pure actions.

Definition 25. *Fix a game with feedback (G, f). A profile of actions and conjectures $(a_i^*, \mu^i)_{i\in I} \in \bigtimes_{i\in I}(A_i \times \Delta(A_{-i}))$ is a **self-confirming equilibrium** of (G, f) if, for every player $i \in I$, the following conditions hold:*
*(1) (*rationality*) $a_i^* \in r_i(\mu^i)$,*
*(2) (*confirmed conjectures*) $\mu^i\left(f_{i,a_i^*}^{-1}(f_i(a^*))\right) = 1$.*
$(a_i^*)_{i\in I}$ is a self-confirming equilibrium action profile of (G, f) if there is some profile of conjectures $(\mu^i)_{i\in I}$ such that $(a_i^*, \mu^i)_{i\in I}$ is a self-confirming equilibrium of (G, f).

Note that the feedback obtained with non-justifiable actions is irrelevant for the set of self-confirming equilibria. Indeed, the rationality condition (1) implies that self-confirming equilibrium actions are necessarily justifiable; hence, the confirmed conjecture condition (2) actually applies only to justifiable actions. On the other hand, unlike Nash equilibria, in the analysis of self-confirming equilibria one cannot ignore the unjustifiable actions, because a self-confirming equilibrium action can be justified by a confirmed conjecture that (wrongly) assigns positive probability to an unjustifiable action of a co-player.

For example, in the game of Matrix 5 the action profile (t, r) is a self-confirming equilibrium pair if each player observes only his own payoff ($f_i = u_i$, $i = 1, 2$): Pick any (μ^1, μ^2) with $\mu^1(\ell) \geq \frac{1}{3}$ and note that $t \in r_1(\mu^1)$, $\{a_2 : u_1(t, a_2) = u_1(t, r)\} = A_2$, $\{a_1 : u_2(a_1, r) = u_2(t, r)\} = A_1$. If instead $\mu^1(\ell) < \frac{1}{3}$, then $r_1(\mu^1) = \{b\}$. Therefore, t is justified by a confirmed conjecture μ^1 only if μ^1 assigns positive probability to the unjustifiable action ℓ of player 2.

Remark 23. *Every Nash equilibrium a^* is a self-confirming equilibrium action profile for every profile of feedback functions f (let $\mu^i(a^*_{-i}) = 1$ for each i). If, for every player $i \in I$ and every justifiable action $a_i \in r_i(\Delta(A_{-i}))$, the section $f_{i,a_i} : A_{-i} \to M_i$ is one-to-one (injective), then the sets of Nash equilibrium and self-confirming equilibrium profiles of actions coincide.*

6.3.2 Self-Confirming Equilibrium with Anonymous Interaction

The Nash equilibrium concept has been generalized to allow for randomization, thus obtaining the mixed equilibrium concept. We motivated mixed equilibrium as a characterization of stationary patterns of behavior in situations of recurrent anonymous interaction, i.e., situations where a game is played recurrently and every role (e.g., the role of the row player) is played by agents drawn at random from large populations. A similar generalization can be applied to the self-confirming equilibrium concept. Again, we will interpret a mixed action $\alpha_i \in \Delta(A_i)$ as a representation of the fractions of agents in population i playing the different actions $a_i \in A_i$. Since different agents in population i may hold different conjectures, the extension should allow for some heterogeneity: in a stationary state different agents of the same population may choose different actions that are justified by different conjectures. This stands in stark contrast with the mixed Nash equilibrium concept, which relies on the assumption that conjectures are correct, hence common to all agents in any given population i. We will assume for the sake of simplicity that all the agents playing the same action hold the same conjecture. This restriction is without loss of generality.

Consider an agent from population i who holds conjecture μ^i and keeps playing a best reply $a_i \in r_i(\mu^i)$. In the case of anonymous interaction, what an agent playing in role i observes at the end of the game depends on the behavior of the agents playing in the other roles, but these agents are drawn at random; therefore, each message m_i will be observed, in the long run, with a frequency determined by the fraction of agents in population(s) $-i$ choosing the actions that (together with a_i) yield message m_i. Conjecture μ^i is confirmed if the subjective probabilities that μ^i assigns to each message given a_i, that is $(\mu^i(f_{i,a_i}^{-1}(m_i))_{m_i \in M_i}$, coincide with the observed long-run frequencies of these messages.

The following example clarifies this point:

$1\backslash 2$	ℓ	r
t	$2,3$	$2,1$
b	$0,0$	$3,1$

Figure 6.5: Matrix 6.

Example 19. Consider the game in Matrix 6 and assume that feedback and payoff functions coincide. Suppose that 25% of the column agents choose ℓ and 75% choose r. Then the probability that a row agent receives the message "You got zero euros" when he chooses b is $\frac{1}{4}$. Row agents do not know this fraction and could, for instance, believe that ℓ occurs with probability $\frac{1}{2}$, which would induce them to choose t; alternatively, a probability of $\frac{1}{5}$ would induce them to play b. It may happen, for instance, that half of the row agents believe that $\mathbb{P}(\ell) = \frac{1}{2}$ and the other half believe $\mathbb{P}(\ell) = \frac{1}{5}$. The former ones choose t, the latter ones b. If the fractions of column agents playing ℓ and r remain 25% and 75% respectively, then those who play b will observe "Zero euros" 25% of the times and "Three euros" 75% of the times. They will then realize that their beliefs were not correct and will keep on revising them until eventually they will coincide with the objective proportions, 25%:75%. These row agents will continue to play b. The other half of the row agents, those that believe that $\mathbb{P}(\ell) = \frac{1}{2}$ and choose t, do not observe anything new and therefore keep on believing and doing the same things. But is it possible that the fractions of column agents playing ℓ and r stay constant? Indeed it is. Suppose that 25% of them believe that $\mathbb{P}(t) = \frac{2}{5}$ and the rest believe that $\mathbb{P}(t) = \frac{1}{5}$. The former ones will choose ℓ (expected payoff: $\frac{6}{5} > 1$), the latter ones will choose r (as $1 > \frac{3}{5}$). Those choosing r do not receive any new information and keep on doing the same thing. Those choosing ℓ half of the times observe "Three euros," and the other half "Zero Euros." Their conjectures are not confirmed, but they will keep revising upward the probability of t and best replying with ℓ, until their conjectures converge to the long-run frequencies 50%-50%, i.e., the actual proportions of row agents choosing t and b. Then there is a stable situation characterized by the following fractions, or mixed actions: $\alpha_1(t) = \frac{1}{2}$, $\alpha_2(\ell) = \frac{1}{4}$. These fractions do not form a mixed Nash equilibrium, because the indifference conditions for a mixed Nash equilibrium yield $\alpha_1^*(t) = \frac{1}{3}$ and $\alpha_2^*(\ell) = \frac{1}{3}$. ▲

It should be clear from the discussion that here, as in our interpretation of the mixed equilibrium concept, mixed actions are interpreted as statistical distributions of pure actions in populations of agents playing in the same role.

Let us move to a general definition. Denote by $\mathbb{P}^{f_i}_{a_i,\mu^i}(m_i)$ the probability of receiving message m_i determined by action a_i and conjecture μ^i, given the feedback function f_i: in the finite case,

$$\mathbb{P}^{f_i}_{a_i,\mu^i}(m_i) = \sum_{a_{-i}:f_i(a_i,a_{-i})=m_i} \mu^i(a_{-i}).$$

Similarly, $\mathbb{P}^{f_i}_{a_i,\alpha_{-i}}(m_i)$ is the probability of m_i determined by a_i and the mixed action profile α_{-i}, given f_i:

$$\mathbb{P}^{f_i}_{a_i,\alpha_{-i}}(m_i) = \sum_{a_{-i}:f_i(a_i,a_{-i})=m_i} \prod_{j\neq i} \alpha_j(a_j).$$

In general, $\mathbb{P}^{f_i}_{a_i,\mu^i}(\cdot) \in \Delta(M_i)$ and $\mathbb{P}^{f_i}_{a_i,\alpha_{-i}}(\cdot) \in \Delta(M_i)$ are, respectively, the probability measures on M_i obtained as pushforward of μ^i and $\mu^{\alpha_{-i}}$ through $f_{i,a_i}: A_{-i} \to M_i$, the section of i's feedback function at a_i:[13]

$$\mathbb{P}^{f_i}_{a_i,\mu^i}(\cdot) = \mu^i \circ f^{-1}_{i,a_i},\ \mathbb{P}^{f_i}_{a_i,\alpha_{-i}}(\cdot) = \mu^{\alpha_{-i}} \circ f^{-1}_{i,a_i}.$$

[13]If A_{-i} is infinite, it is endowed with a sigma-algebra of subsets $\mathcal{A}_{-i} \subseteq 2^{A_{-i}}$ containing all the singletons $\{a_{-i}\}$ $(a_{-i} \in A_{-i})$, conjectures are sigma-additive functions $\mu^i : \mathcal{A}_{-i} \to [0,1]$, and each section f_{i,a_i} of the feedback function is assumed to be $\mathcal{A}_{-i}$-measurable, so that—in particular— $f^{-1}_{i,a_i}(m_i) \in \mathcal{A}_{-i}$ for each message m_i. These assumptions are satisfied in compact-continuous games with feedback.

Definition 26. *Fix a game with feedback (G, f). A profile of mixed actions and conjectures $(\alpha_i, (\mu^i_{a_i})_{a_i \in \mathrm{supp}\alpha_i})_{i \in I}$ is an* ***anonymous self-confirming equilibrium*** *of (G, f) if for every role $i \in I$ and action $a_i \in \mathrm{supp}\alpha_i$ the following conditions hold:*
*(1) (*rationality*) $a_i \in r_i(\mu^i_{a_i})$,*
*(2) (*confirmed conjectures*) $\mathbb{P}^{f_i}_{a_i, \mu^i_{a_i}}(\cdot) = \mathbb{P}^{f_i}_{a_i, \alpha_{-i}}(\cdot)$.*
$\alpha = (\alpha_i)_{i \in I}$ is an anonymous self-confirming equilibrium profile if there is a profile of conjectures $((\mu^i_{a_i})_{a_i \in \mathrm{supp}\alpha_i})_{i \in I}$ such that $(\alpha_i, (\mu^i_{a_i})_{a_i \in \mathrm{supp}\alpha_i})_{i \in I}$ is an anonymous self-confirming equilibrium.

Remark 24. *Every mixed Nash equilibrium of G is an anonymous self-confirming equilibrium profile for every f.*

Remark 25. *Every pure self-confirming equilibrium $(a^*_i)_{i \in I}$ is a degenerate anonymous self-confirming equilibrium, which can be interpreted as a* symmetric *equilibrium where all agents of the same population i play the same action a^*_i.*

Since the anonymous version of the self-confirming equilibrium concept is more general, from now on we refer to it simply as "self-confirming equilibrium" without further qualifications.

Example 20. It can be checked as an exercise that the game in Matrix 6, assuming that the players observe ex post only their own payoff ($f_i = u_i$), admits three types of anonymous self-confirming equilibrium:

1. t is chosen by all the agents in population 1: α_2 can take any value since the conjecture of 1 is necessarily "confirmed" (1, choosing t, does not receive new information); $\alpha_1(t) = 1$, $\mu^1_t(\ell) \geq \frac{1}{3}$, $0 \leq \alpha_2(\ell) \leq 1$, $\mu^2_\ell(t) = 1$, $\mu^2_r(t) \leq \frac{1}{3}$;

2. r is chosen by all agents in population 2: α_1 can take any value since the conjecture of 2 is necessarily "confirmed" (2, choosing r, does not receive new information); $0 \leq \alpha_1(t) \leq 1$, $\mu^1_t(\ell) \geq \frac{1}{3}$, $\mu^1_b(\ell) = 0$, $\alpha_2(\ell) = 0$, $\mu^2_r(t) \leq \frac{1}{3}$;

3. all actions are chosen by a positive fractions of agents: the beliefs of those who choose "informative" actions (b for $i = 1$ and ℓ for $i = 2$) are correct; $\frac{1}{3} \leq \alpha_1(t) \leq 1$, $0 < \alpha_2(\ell) \leq \frac{1}{3}$, $\mu^1_t(\ell) \geq \frac{1}{3}$, $\mu^1_b(\ell) = \alpha_2(\ell)$, $\mu^2_\ell(t) = \alpha_1(t)$, $\mu^2_r(t) \leq \frac{1}{3}$.

Notice that the equilibria of type 1 include the Nash equilibrium (t, ℓ), those of type 2 include the Nash equilibrium (b, r), and those of type 3 include the mixed equilibrium $\alpha_1^*(t) = \frac{1}{3}$, $\alpha_2^*(\ell) = \frac{1}{3}$. ▲

Beliefs about Distributions

In the definition of anonymous self-confirming equilibrium we check whether actions played by a positive fraction of agents are justified by confirmed conjectures. This definition is adequate because a subjective expected utility maximizer ultimately cares only about the probabilities of the actions of his co-players. Yet, what an agent is uncertain about in a population game is the actual *distribution* of actions of the co-players, that is—in a two-population game—, the true $\alpha_{-i} \in \Delta(A_{-i})$.[14] Thus, an agent in population i has a belief $\nu^i \in \Delta(\Delta(A_{-i}))$. Suppose that an agent with such belief ν^i keeps playing a_i and that the true distribution of actions in the other population is α_{-i}, then his belief is confirmed if ν^i assigns probability one to the set of α'_{-i} that yield the same frequency distribution of messages as α_{-i} given a_i. Using the "pushforward" notation, the confirmation condition is

$$\nu^i\left(\left\{\alpha'_{-i} \in \Delta(A_{-i}) : \alpha'_{-i} \circ f_{i,a_i}^{-1} = \alpha_{-i} \circ f_{i,a_i}^{-1}\right\}\right) = 1.$$

Belief $\nu^i \in \Delta(\Delta(A_{-i}))$ yields a corresponding conjecture $\mu^i \in \Delta(A_{-i})$ by taking the subjective averages, according to ν^i, of the objective probabilities $\alpha_{-i}(a_{-i})$:[15]

$$\mu^i(a_{-i}) = \int_{\Delta(A_{-i})} \alpha_{-i}(a_{-i}) \nu^i(\mathrm{d}\alpha_{-i}).$$

With this, the definition of anonymous self-confirming equilibrium given above is equivalent to requiring that each action played by a positive fraction of agents be justified by the conjecture derived from a confirmed belief over distributions of opponents actions. However, such equivalence holds only under the assumption of subjective expected utility

[14] With more than one co-player drawn at random from different populations, one has to take into account that draws are independent across populations, which makes the analysis slightly more complex.

[15] These subjective averages are also called the "predictive probabilities" implied by belief ν^i.

maximization. With more general rationality criteria one has to use the conceptually more precise definition involving beliefs about distributions.[16]

6.3.3 Properties of Feedback, Nash and Self-Confirming Equilibria

In simultaneous-moves games, if it is possible to perfectly observe ex post the opponents' actions, then a self-confirming equilibrium is necessarily a Nash equilibrium. Indeed, in this case, conjectures are confirmed if and only if they are correct (Remark 23).[17] However, simultaneous-moves games are also used as reduced forms to analyze the equilibria of corresponding dynamic games. Specifically, it will be shown that each dynamic game admits a "strategic-form" representation whereby it is assumed that players choose in advance and simultaneously the contingent plans (strategies) that will determine their actions in every possible situation that may arise as the play unfolds (see Section 1.4.2 of the Introduction, and Section 9.3 of Chapter 9). In dynamic games, even if, at the end of the game, it is possible to observe all the *actions* that have *actually* been *taken*, it is not possible to observe *how the opponents would have played in circumstances that did not occur*. Hence, it is impossible, at least for some player, to observe the strategies of other players. For this reason, in dynamic games it is easier to find self-confirming equilibria that do not correspond to Nash equilibria of the strategic form.

Next we provide conditions on feedback that imply the equivalence of self-confirming equilibria and mixed equilibria. First we define the ex post information partition of A_{-i} corresponding to a given action a_i. Recall that $f_{i,a_i} = f_i(a_i, \cdot) : A_{-i} \to M_i$ is the section at a_i of the feedback function f_i. The section f_{i,a_i} shows how the feedback received by player i depends on the opponents' actions given a_i. For each $i \in I$ and $a_i \in A_i$, we let

$$\mathcal{F}_{-i}(a_i) = \{C_{-i} \in 2^{A_{-i}} : \exists m_i \in M_i, C_{-i} = f_{i,a_i}^{-1}(m_i)\}$$

[16] See Battigalli *et al.* [16]. Furthermore, if we model explicitly the learning process of agents who experiments rationally, the relevant beliefs are those about distributions (see, e.g., Fudenberg and Levine [33]).

[17] In the case of anonymous interaction, we can consider at least two scenarios: (a) the individuals observe the statistical distribution of the actions in previous periods, (b) the individuals observe the long-run frequencies of the opponents' actions. In both cases we can say that a conjecture is correct if it corresponds to the observed frequencies.

denote the partition of A_{-i} given by the pre-images of function f_{i,a_i}. For example, the game in Matrix 6 with $f_i = u_i$ $(i = 1, 2)$ has the following ex post partitions, for each player and action:

$$\begin{aligned} \mathcal{F}_{-1}(t) &= \{A_2\},\ \mathcal{F}_{-1}(b) = \{\{\ell\}, \{r\}\}, \\ \mathcal{F}_{-2}(\ell) &= \{\{t\}, \{b\}\},\ \mathcal{F}_{-2}(r) = \{A_1\}. \end{aligned}$$

Definition 27. *Feedback function f_i satisfies **own-action independence of feedback about others** relative to payoff function u_i if all the* justifiable *actions induce the same partition of A_{-i}, that is,*

$$\forall a_i', a_i'' \in r_i\left(\Delta\left(A_{-i}\right)\right),\ \mathcal{F}_{-i}(a_i') = \mathcal{F}_{-i}(a_i''). \tag{6.3.1}$$

*Game with feedback (G, f) satisfies **own-action independence** if condition (6.3.1) holds for each player $i \in I$.*

Note that condition (6.3.1) applies only to justifiable actions because, as explained above, only the feedback about such actions matters for the determination of self-confirming equilibria. It is easily checked that the game in Matrix 6 with feedback $f_i = u_i$ $(i = 1, 2)$ does *not* satisfies own-action independence of feedback about others.

Definition 28. *Feedback function f_i satisfies **observable payoffs** relative to payoff function u_i if for every action of i the feedback received by i determines the payoff of i, that is,*

$$\forall a_i \in A_i, \forall a_{-i}', a_{-i}'' \in A_{-i},\ f_i(a_i, a_{-i}') = f_i(a_i, a_{-i}'') \Rightarrow u_i(a_i, a_{-i}') = u_i(a_i, a_{-i}''). \tag{6.3.2}$$

*Game with feedback (G, f) satisfies **observable payoffs** if condition (6.3.2) holds for each player $i \in I$.*

Note that the same condition can be expressed as

$$u_i(a_i, a_{-i}') \neq u_i(a_i, a_{-i}'') \Rightarrow f_i(a_i, a_{-i}') \neq f_i(a_i, a_{-i}'').$$

In other words, the observable payoff condition says that for each player $i \in I$ there is a function $\pi_i : A_i \times M_i \to \mathbb{R}$ such that $u_i\left(a_i, a_{-i}\right) = \pi_i\left(a_i, f_i\left(a_i, a_{-i}\right)\right)$ for every profile $\left(a_i, a_{-i}\right)$.[18] The following example illustrates both properties of feedback.

[18]In a more compact and abstract form, $u_{i,a_i} = \pi_{i,a_i} \circ f_{i,a_i}$ for every a_i.

Example 21. Let G be a Cournot oligopoly. Suppose each firm $i \in I$ knows its own total cost function $C_i(\cdot)$ and the inverse market demand function $P(\cdot)$, and that it observes ex post only the market price:

$$f_i(a_i, a_{-i}) = P\left(a_i + \sum_{j \neq i} a_j\right),$$

where a_j is the output of firm j. The utility of firm i is its profit:[19]

$$u_i(a_i, a_{-i}) = \pi_i(a_i, f_i(a_i, a_{-i})) = a_i P\left(a_i + \sum_{j \neq i} a_j\right) - C(a_i).$$

Of course, the observable payoffs condition holds for this game with feedback. Furthermore, if $P(\cdot)$ is invertible, also own-action independence holds. Indeed, each firm i observes ex post the total output of the competitors independently of its own output choice, because

$$\sum_{j \neq i} a_j = P^{-1}(p) - a_i$$

for each a_i and observed price p. Next suppose instead that each firm i only observes the revenue

$$f_i(a_i, a_{-i}) = R_i(a_i, a_{-i}) = a_i P\left(\sum_{j \in I} a_j\right)$$

fetched for its output a_i, but it does not observe the unit price. In this case, the observable payoffs property still holds, because

$$u_i(a_i, a_{-i}) = \pi_i(a_i, f_i(a_i, a_{-i})) = R_i(a_i, a_{-i}) - C(a_i).$$

But own-action independence holds *if and only if* producing no output ($a_i = 0$) is *not justifiable*. Indeed, if $a_i = 0$, then i's revenue is zero for every output profile of the competitors and i cannot recover $\sum_{j \neq i} a_j$. If instead $a_i > 0$, then i can back out from its revenue the unit price

$$p = P\left(\sum_{j \in I} a_j\right) = R_i(a_i, a_{-i}) / a_i$$

[19] If firm i is owned by a possibly risk averse agent, then $u_i = v_i \circ \pi_i$, where v_i is concave and strictly increasing.

and the total output of the competitors:

$$\sum_{j \neq i} a_j = P^{-1}\left(\frac{R_i\left(a_i, a_{-i}\right)}{a_i}\right) - a_i.$$

Whether $a_i = 0$ is justifiable or not depends on details of the oligopoly model that we did not specify here. For example, if $C_i\left(a_i\right) = ca_i$ and there is some competitors' output profile $\hat{a}_{-i}$ such that $P\left(\sum_{j \neq i} \hat{a}_j\right) < c$, then $a_i = 0$ is justifiable as a best response to $\hat{a}_{-i}$ and own-action independence does *not* hold. ▲

Comment on Own-Action Independence The example clarifies that own-action independence of feedback *about others* may hold even if, for some or all i, feedback function f_i depends on a_i: market price in a quantity-setting oligopoly depends on own output, but the observation of market price gives feedback about the output of others that does not depend on own output.

Comment on Payoff Observability As the example suggests, observability of payoffs holds in many applications: what we call "payoff" in game theory is a player's "utility" expressed as a function of all players' actions, and it seems obvious that a player should observe how much utility he gets. Yet, payoff observability is not a tautological assumption. Suppose, for example, that players have other-regarding preferences. This means that they do not only care about the consequences of interaction for themselves, but also for (some) other players. In particular, they may care about the consumption of other players. If player i cares about his own income y_i and the incomes of other players y_{-i}, then his utility is a function of the form $v_i(y_i, y_{-i})$. The outcome function is $y = (y_j)_{j \in I} = (g_j(a))_{j \in I} = g(a)$. Assume that i observes only his own income; then $f_i(a) = g_i(a)$ and $u_i(a) = v_i(f_i(a), g_{-i}(a))$. Hence, we may have $f_i(a') = f_i(a'')$ and yet $u_i(a') \neq u_i(a'')$. Intuitively, $u_i(a)$ is not necessarily a utility "experienced"—hence observed—by i, it is a way of ranking action profiles according to the preferences over the induced income allocations. But player i observes only his own component of the income allocation; hence, he cannot know if the realized allocation is one he ranks highly or not.

The following lemma says that under own-action independence and payoff observability, an action can be justified by a confirmed conjecture if and only if it is a best reply to the true distribution of actions of the co-players.

Lemma 18. *Let (G, f) be a finite or compact-continuous game with feedback; fix any $i \in I$ and suppose that f_i satisfies own-action independence of feedback about others and observable payoffs relative to u_i. Then, for all $a_i^* \in A_i$ and $\alpha_{-i} \in \times_{j \neq i} \Delta(A_j)$ the following are equivalent:*
(1) there is a conjecture $\mu^i \in \Delta(A_{-i})$ such that $a_i^* \in r_i(\mu^i)$ and $\mathbb{P}^{f_i}_{a_i^*,\mu^i}(\cdot) = \mathbb{P}^{f_i}_{a_i^*,\alpha_{-i}}(\cdot)$,
(2) $a_i^* \in r_i(\alpha_{-i})$.

Proof. We prove the result for *finite* games.[20] Fix a_i^* and α_{-i}. It is obvious that (2) implies (1): let $\mu^i(a_{-i}) = \prod_{j \neq i} \alpha_j(a_j)$ for all $a_{-i} \in A_{-i}$, then (1) holds.

To prove that (1) implies (2), first note that *observability of payoffs* implies that, for each action $a_i \in A_i$, the section of the payoff function of i at a_i, $u_{i,a_i} : A_{-i} \to \mathbb{R}$, is constant on each cell of the ex post information partition $\mathcal{F}_{-i}(a_i)$. Indeed, fix any $C_{-i} \in \mathcal{F}_{-i}(a_i)$ and $a'_{-i}, a''_{-i} \in C_{-i}$. By definition of $\mathcal{F}_{-i}(a_i)$, a'_{-i} and a''_{-i} yield the same message given a_i: $f_i(a_i, a'_{-i}) = f_i(a_i, a''_{-i})$; hence, observability of payoffs implies $u_i(a_i, a'_{-i}) = u_i(a_i, a''_{-i})$. With this, we write $u_i(a_i, C_{-i})$ for the common value of $u_i(a_i, \cdot)$ on cell $C_{-i} \in \mathcal{F}_{-i}(a_i)$. Similarly, we write $f_i(a_i, C_{-i})$ for the common feedback message on cell $C_{-i} \in \mathcal{F}_{-i}(a_i)$ given a_i.[21] Thus,

$$\forall a_i \in A_i,\ \bar{u}_i(a_i, \alpha_{-i}) = \sum_{C_{-i} \in \mathcal{F}_{-i}(a_i)} u_i(a_i, C_{-i}) \sum_{a_{-i} \in C_{-i}} \prod_{j \neq i} \alpha_j(a_j). \quad (6.3.3)$$

Let (1) hold for $\mu^i \in \Delta(A_{-i})$. Then, in particular, a_i^* is justifiable, that is, $a_i^* \in r_i(\Delta(A_{-i}))$. The set of "objective best replies" $r_i(\alpha_{-i})$ contains at least one action a_i^o, which is necessarily justifiable. We are going to

[20] The proof for compact-continuous games is left as an exercise for the mathematically savvy reader. Consider that, in this (more general) case, feedback functions are measurable because they are continuous, and best reply correspondences are non-empty valued. The proof shows that the result can be further generalized (cf. Battigalli *et al.* [18]).
[21] The message is constant on $C_{-i} \in \mathcal{F}_{-i}(a_i)$ by definition of $\mathcal{F}_{-i}(a_i)$, whether or not payoffs are observable.

prove that $\bar{u}_i(a_i^*, \alpha_{-i}) \geq \bar{u}_i(a_i^o, \alpha_{-i})$. Since $a_i^o \in r_i(\alpha_{-i})$, this implies (2): $a_i^* \in r_i(\alpha_{-i})$.

Since both a_i^* and a_i^o are justifiable, own-action independence implies the common-feedback condition $\mathcal{F}_{-i}(a_i^o) = \mathcal{F}_{-i}(a_i^*)$. Common feedback and eq. (6.3.3) yield

$$\bar{u}_i(a_i^o, \alpha_{-i}) = \sum_{C_{-i} \in \mathcal{F}_{-i}(a_i^*)} u_i(a_i^o, C_{-i}) \sum_{a_{-i} \in C_{-i}} \prod_{j \neq i} \alpha_j(a_j). \tag{6.3.4}$$

The confirmed-conjecture condition stated in (1) implies that the objective and subjective probabilities of each message $f_i(a_i^*, C_{-i})$ coincide:

$$\forall C_{-i} \in \mathcal{F}_{-i}(a_i^*), \sum_{a_{-i} \in C_{-i}} \prod_{j \neq i} \alpha_j(a_j) = \sum_{a_{-i} \in C_{-i}} \mu^i(a_{-i}). \tag{6.3.5}$$

Eqs. (6.3.4)-(6.3.5) yield

$$\bar{u}_i(a_i^o, \alpha_{-i}) = \sum_{C_{-i} \in \mathcal{F}_{-i}(a_i^*)} u_i(a_i^o, C_{-i}) \sum_{a_{-i} \in C_{-i}} \mu^i(a_{-i}) = u_i(a_i^o, \mu^i).$$

Eqs. (6.3.3)-(6.3.5) yield

$$\bar{u}_i(a_i^*, \alpha_{-i}) = \sum_{C_{-i} \in \mathcal{F}_{-i}(a_i^*)} u_i(a_i^*, C_{-i}) \sum_{a_{-i} \in C_{-i}} \mu^i(a_{-i}) = u_i(a_i^*, \mu^i).$$

Since $a_i^* \in r_i\left(\mu^i\right)$, the foregoing equalities imply

$$\bar{u}_i(a_i^*, \alpha_{-i}) = u_i(a_i^*, \mu^i) \geq u_i(a_i^o, \mu^i) = \bar{u}_i(a_i^o, \alpha_{-i}).$$

Hence, $a_i^* \in r_i(\alpha_{-i})$. ■

Theorem 16. *Let (G, f) be a compact-continuous game with feedback that satisfies own-action independence and observable payoffs. Then the sets of mixed Nash equilibria and self-confirming equilibrium profiles of mixed actions coincide.*

Proof. Every mixed Nash equilibrium is also self-confirming (see Remark 24). To prove the converse, fix any selfconfirming equilibrium $(\alpha_i, (\mu^i_{a_i})_{a_i \in \text{supp}\alpha_i})_{i \in I}$. By definition, for each player $i \in I$ and action

$a_i \in \text{supp}\alpha_i$, we have $a_i \in r_i\left(\mu^i_{a_i}\right)$ and $\mathbb{P}^{f_i}_{a_i,\mu^i_{a_i}}(\cdot) = \mathbb{P}^{f_i}_{a_i,\alpha_{-i}}(\cdot)$. Thus, Lemma 18 implies that $\text{supp}\alpha_i \subseteq r_i(\alpha_{-i})$ for each $i \in I$. Hence, by Theorem 13, $(\alpha_i)_{i\in I}$ is a mixed Nash equilibrium. ■

Although observability of payoffs is not tautological, it is common in economic applications. Therefore the above theorem says that the key property for the comparison between (anonymous) self-confirming and (mixed) Nash equilibrium is own-action dependence/independence of feedback.

6.3.4 Comparative Statics for Self-Confirming Equilibrium

Here we analyze how the set of self-confirming equilibria is affected by changes in the informativeness of feedback and in risk aversion. We show that coarser feedback and higher risk aversion expand the set of self-confirming equilibria.[22] This means that the long-run outcome of learning is less predictable and less likely to coincide with Nash equilibrium when feedback is less informative or players are more risk averse.

Coarseness of Feedback By inspection of the definition of self-confirming equilibrium it is easily checked that if the feedback functions f_{i,a_i} are constant for all i and a_i, that is, if there is no feedback about co-players' actions, then every profile $(\alpha_i)_{i\in I}$ that assigns positive weight only to justifiable actions is a self-confirming equilibrium profile of mixed actions. This suggests that poor feedback yields a large set of self-confirming equilibria. We can make this intuition precise. First recall that a partition $\bar{\mathcal{F}}$ is coarser than a partition $\mathcal{F}$ if each cell of $\bar{\mathcal{F}}$ is a union of cells of $\mathcal{F}$.

Definition 29. *We say that game with feedback $\left(G, \bar{f}\right)$ has* ***coarser feedback*** *than (G, f) if, for every $i \in I$ and $a_i \in A_i$, $\bar{\mathcal{F}}_i(a_i)$ is coarser than $\mathcal{F}_i(a_i)$.*

The following result says that making feedback coarser expands the set of self-confirming equilibria.

[22] In the case of risk aversion, the result only refers to the set of pure, or non-anonymous self-confirming equilibria of games with obvervable payoffs.

Theorem 17. *Fix two compact-continuous games with feedback* (G, f) *and* $(G, \bar{f})$*. If* $(G, \bar{f})$ *has* ***coarser feedback*** *than* (G, f)*, then every self-confirming equilibrium of* (G, f) *is also a selfconfirming equilibrium of* $(G, \bar{f})$*.*

Proof. We prove the result for *finite* games.[23] Fix a self-confirming equilibrium $(\alpha_i, (\mu^i_{a_i})_{a_i \in \text{supp}\alpha_i})_{i\in I}$ of (G, f). Since (G, f) and $(G, \bar{f})$ differ only with respect to the feedback functions, the given profile satisfies the best response property (1) of Definition 26 in both games with feedback; therefore, we only have to show that the confirmed conjectures property (2) is satisfied also for $(G, \bar{f})$. Fix $i \in I$ and $a_i \in$suppα_i arbitrarily. Since the justifying conjecture $\mu^i_{a_i}$ is confirmed in (G, f), then

$$\forall C_{-i} \in \mathcal{F}_i(a_i),\ \mu^i_{a_i}(C_{-i}) = \sum_{a_{-i}\in C_{-i}} \prod_{j\neq i} \alpha_j(a_{-i}).$$

Since $\bar{\mathcal{F}}_i(a_i)$ is coarser than $\mathcal{F}_i(a_i)$,

$$\forall \bar{C}_{-i} \in \bar{\mathcal{F}}_i(a_i),\ \bar{C}_{-i} = \bigcup_{C_{-i}\in\mathcal{F}_i(a_i): C_{-i}\subseteq \bar{C}_{-i}} C_{-i}.$$

By additivity of $\mu^i_{a_i}$,

$$\forall \bar{C}_{-i} \in \bar{\mathcal{F}}_i(a_i),\ \mu^i_{a_i}\left(\bar{C}_{-i}\right) = \sum_{C_{-i}\in\mathcal{F}_i(a_i): C_{-i}\subseteq \bar{C}_{-i}} \mu^i_{a_i}(C_{-i}) = \sum_{a_{-i}\in \bar{C}_{-i}} \prod_{j\neq i} \alpha_j(a_{-i}).$$

Thus $\mu^i_{a_i}$ is confirmed in $(G, \bar{f})$ as well. ■

Risk Aversion Using Jensen's inequality and Lemma 2 (the duality result of Wald and Pearce), we have shown that an increase in risk aversion expands the set of rationalizable actions (Theorems 1 and 4). A simpler argument relying only on Jensen's inequality shows that an increase in risk aversion expands the set of non-anonymous (i.e., pure) self-confirming equilibria of games with observable payoffs.[24] The intuition is quite

[23] Again, the proof for compact-continuous games is left as an exercise for the mathematically savvy reader. Furthermore, the result holds for more general games (cf. Battigalli *et al.* [18]).

[24] See Battigalli *et al.* [16, Theorem 1] for a similar result about mixed self-confirming equilibria and "ambiguity aversion". See Weinstein [73] for results about Nash equilibrium and risk aversion.

simple: under payoff observability, in a pure self-confirming equilibrium each player is certain about the payoff of his equilibrium action, but he may be uncertain about the payoff of deviations; keeping conjectures fixed as we increase risk aversion, deviations with uncertain payoffs become less attractive. Therefore, the same profile of actions and conjectures is a self-confirming equilibrium also with higher risk aversion. On the other hand, higher risk aversion may make "safe" actions more attractive and thus yield new self-confirming equilibria. The following example illustrates.

Example 22. Consider again the game form with monetary payoffs of Example 15 and assume that feedback functions coincide with monetary payoff functions:

$g_1 = f_1, g_2 = f_2:$	b'	b''
a'	0, 1	1, 0
a''	$\frac{1}{3}$,0	$\frac{1}{3}$,1
a'''	1, 1	0, 0

The ex post information partitions are

$$\begin{aligned}\mathcal{F}_1\left(a'\right) &= \mathcal{F}_1\left(a'''\right) = \left\{\{b'\},\{b''\}\right\}, \mathcal{F}_1\left(a''\right) = \{A_2\}, \\ \mathcal{F}_2(b') &= \mathcal{F}_2(b'') = \left\{\{a',a'''\},\{a''\}\right\}.\end{aligned}$$

In particular, action a'' does not allow to observe ex post whether player 2 chose b' or b''. Under risk neutrality a'' is not justifiable, hence it cannot be part of a self-confirming equilibrium. If instead risk aversion is sufficiently high, the uniform conjecture on A_2 justifies a'' and is confirmed, which implies that (a'', b'') is a pure self-confirming equilibrium action pair. Assume $v_1(m_1) = (m_1)^{1/\theta}$ (the risk attitudes of player 2 are immaterial). The pure (non-anonymous) equilibrium correspondence is

$$pSCE_\theta = \begin{cases} \{(a''', b')\}, & \text{if } 1 \leq \theta < \log_2 3, \\ \{(a''', b'), (a'', b'')\}, & \text{if } \theta \geq \log_2 3. \end{cases}$$

(Compare with Examples 6 and 15.) ▲

Theorem 18. *Suppose game with feedback (G, f) satisfies observable payoffs and let players in $\hat{G}$ be more risk averse than in G, then every pure self-confirming equilibrium of (G, f) is also a pure self-confirming equilibrium of $\left(\hat{G}, f\right)$.*

Proof. Fix arbitrarily a self-confirming equilibrium $\left(a_i^*, \mu^i\right)_{i \in I}$ of (G, f) and a player $i \in I$. Then $a_i^* \in r_i\left(\mu^i\right)$ and μ^i is confirmed (given profile a^*); since payoffs are observable, $u_i\left(a_i^*, \mu^i\right) = u_i\left(a_i^*, a_{-i}^*\right)$; therefore,

$$\forall a_i \in A_i, u_i\left(a_i^*, \mu^i\right) = u_i\left(a_i^*, a_{-i}^*\right) \geq u_i\left(a_i, \mu^i\right) = \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right), \qquad (6.3.6)$$

where $\mathbb{E}_{\mu^i}(u_{i,a_i})$ is just the explicit expression of $u_i\left(a_i, \mu^i\right)$ as an expectation. We must show that $\left(a_i^*, \mu^i\right)_{i \in I}$ is a self-confirming equilibrium of $\left(\hat{G}, f\right)$. Since we keep the feedback functions fixed, conjecture μ^i is confirmed in $\left(\hat{G}, f\right)$ as well; therefore, we only have to show that $a_i^* \in \hat{r}_i\left(\mu^i\right)$, where $\hat{r}_i$ is the best reply correspondence of i in $\hat{G}$, that is, given payoff function $\hat{u}_i$. Payoff observability holds for $\left(\hat{G}, f\right)$ as well: Let φ be the concave and strictly increasing function such that $\hat{u}_i = \varphi \circ u_i$, then

$$\begin{aligned} f_i(a_i, a'_{-i}) &= f_i(a_i, a''_{-i}) \Rightarrow u_i(a_i, a'_{-i}) = u_i(a_i, a''_{-i}) \\ &\Rightarrow \varphi\left(u_i(a_i, a'_{-i})\right) = \varphi\left(u_i(a_i, a''_{-i})\right) \Rightarrow \hat{u}_i(a_i, a'_{-i}) = \hat{u}_i(a_i, a''_{-i}). \end{aligned}$$

Therefore,

$$\hat{u}_i\left(a_i^*, \mu^i\right) = \hat{u}_i\left(a_i^*, a_{-i}^*\right).$$

With this,

$$\forall a_i \in A_i,\ \hat{u}_i\left(a_i^*, \mu^i\right) = \hat{u}_i\left(a_i^*, a_{-i}^*\right) = \varphi\left(u_{i,a_i^*}\left(a_{-i}^*\right)\right)$$

$$\geq \varphi\left(\mathbb{E}_{\mu^i}\left(u_{i,a_i}\right)\right) \geq \mathbb{E}_{\mu^i}\left(\varphi \circ u_{i,a_i}\right) = \mathbb{E}_{\mu^i}\left(\hat{u}_{i,a_i}\right) = \hat{u}_i\left(a_i, \mu^i\right)$$

where the first inequality follows from eq. (6.3.6) and the monotonicity of φ, and the second is Jensen's inequality, which follows from the concavity of φ. Since i was fixed arbitrarily, this proves that the given profile satisfies the best-reply and confirmed-conjectures properties in $\left(\hat{G}, f\right)$ and is therefore a self-confirming equilibrium of the latter. ■

As is clear from the proof, the result can be generalized as follows: whenever a pure or mixed action profile yields sure payoffs to the players and is a self-confirming equilibrium, then it still is a self-confirming equilibrium with more risk averse players. The following example shows that when equilibrium payoffs are uncertain, or risky, an equilibrium may disappear when risk aversion increases.

Example 23. Consider the game form with monetary outcomes in the following matrix:[25]

g_1, g_2	ℓ	r
t	6, 0	0, 1
m	4, 1	2, 1
b	3, 1	3, 0

(I) First assume that each player *observes his monetary payoff*, that is, $f_i = g_i$ $(i = 1, 2)$. It can be checked that if player 1 is *risk neutral* then $\left(m, \frac{1}{2}\delta_\ell + \frac{1}{2}\delta_r\right)$ is a mixed (Nash and self-confirming) equilibrium profile where player 1 gets 3 dollars on average. Note that player 1 is indifferent among his actions in this equilibrium, despite the fact that m is chosen with probability 1. If player 1 is risk averse this profile cannot be an equilibrium, because the safe action b becomes more attractive given the former equilibrium conjecture. (II) Next assume that player 2 observes his monetary payoff, but *player 1 observes nothing*. Under risk neutrality $\left(\left(m, \mu^1(\ell) = \frac{1}{2}\right), \left(r, \mu^2(m) + \mu^2(t) = 1\right)\right)$ gives a set of self-confirming equilibria (parameterized by $\mu^2(m) \in [0, 1]$), but these equilibria disappear under risk aversion, which makes b more attractive given conjecture $\mu^1(\ell) = \frac{1}{2}$. ▲

[25] The risk attitudes of player 2 are immaterial.

7

Learning and Solution Concepts

So far we avoided the details of learning dynamics. The analysis of such dynamics requires the use of mathematical tools whose knowledge we do not take for granted in this textbook (differential equations, difference equations, stochastic processes). Nonetheless, it is possible to state some elementary results about learning dynamics by addressing the following question: When is it the case that a *trajectory*, that is, an infinite sequence of action profiles $(a^t)_{t=1}^{\infty}$ (with $a^t = (a_i^t)_{i \in I} \in A$), is *consistent with adaptive learning*? We start with a qualitative answer.

Consider a *finite game* G that is played recurrently. Assume that all actions are observable and consider the point of view of a player i that observes that a given profile a_{-i} has *not* been played for a very long time, say for at least T periods. Then it is reasonable to assume that i assigns to a_{-i} a very small probability. If T is sufficiently large, and a_{-i} was not played in the periods $\hat{t}, \hat{t}+1, \ldots, \hat{t}+T$, then the probability of a_{-i} in $t > \hat{t}+T$ will be so small that the best reply to i's conjecture in t will *also* be the best reply to a conjecture that assigns probability *zero* to a_{-i}. In other words, i will choose in $t > \hat{t}+T$ only those actions that are best replies to conjectures μ^i such that $\mathrm{supp}\mu^i \subseteq \left\{a_{-i}^{\tau} : \hat{t} \le \tau < t\right\}$, that is, only actions in the set $r_i\left(\Delta\left(\left\{a_{-i}^{\tau} : \hat{t} \le \tau < t\right\}\right)\right)$.[1] Notice that this argument

[1] A_i is finite and $u_i(a_i, \mu^i)$ is continuous (in fact linear) with respect to probabilities $(\mu^i(a_{-i}))_{a_{-i} \in A_{-i}}$. It follows that if a_i^* is a best reply to a conjecture that assigns a "sufficiently small" probability to a_{-i}, then a_i^* is a best reply also to a slightly different

assumes only that i is able to compute best replies to conjectures. The argument is therefore consistent with a high degree of incompleteness of information.

7.1 Perfectly Observable Actions

Building on this intuition, one can use the rationalization operator ρ to define the consistency of a trajectory $(a^t)_{t=1}^{\infty}$ with adaptive learning under the assumption that the opponents' actions are perfectly observable.[2] Recall that, for any set $C = \times_{i\in I} C_i \subseteq A$ we let $\rho_i(C_{-i}) = r_i(\Delta(C_{-i}))$, which gives the set of profiles "justified" or "rationalized" by C: $\rho(C) = \times_{i\in I} \rho_i(C_{-i})$. Now, the set of action profiles chosen in a given interval of time is not, in general, a Cartesian product. For this reason, it is useful to generalize the definition of ρ as follows. For $C_{-i} \subseteq A_{-i}$ (not necessarily a product set) let $\rho_i(C_{-i}) = r_i(\Delta(C_{-i}))$. Then, let $C \subseteq A$, and define

$$\rho(C) = \bigtimes_{i\in I} \rho_i(\mathrm{proj}_{A_{-i}} C),$$

where $\mathrm{proj}_{A_{-i}} C = \{a_{-i} : \exists a_i \in A_i, (a_i, a_{-i}) \in C\}$ is the set of the other players' action profiles that are consistent with C. (If C is a product set, the usual definition obtains.) Note that, even with this more general definition, the ρ operator is *monotone*: $E \subseteq F$ implies $\rho(E) \subseteq \rho(F)$.

According to the reasoning developed above, suppose that players base their conjectures on past observations and they maximize their expected payoffs in each period. Then, if for a "very long"time only action profiles in C are observed, the current profile of best replies must be in $\rho(C)$. This explains the following definition.

Definition 30. *A trajectory $(a^t)_{t=1}^{\infty}$ is **consistent with adaptive learning** (when opponents' actions are perfectly observable) if for every $\hat{t}$ there exists a T such that, for all $t > \hat{t} + T$, $a^t \in \rho\left(\left\{a^\tau : \hat{t} \leq \tau < t\right\}\right)$.*

Next we define when a trajectory $(a^t)_{t=1}^{\infty}$ generates a "limit distribution"$\mu \in \Delta(A)$.

---

conjecture that assigns probability *zero* to a_{-i}.

[2]The following results are adapted from Milgrom and Roberts [46].

Definition 31. *$\mu \in \Delta(A)$ is the **limit distribution** of $(a^t)_{t=1}^{\infty}$ if for every $a \in A$*

$$\text{(i)} \lim_{t\to\infty} \frac{|\{\tau : 1 \le \tau < t, a^\tau = a\}|}{t} = \mu(a)$$

$$\text{(ii) if } \mu(a) = 0, \text{ then } \exists \hat{t} : \forall t > \hat{t},\ a^t \neq a.$$

This definition requires that the "long-run frequency" of every a is $\mu(a)$, and furthermore that if $\mu(a) = 0$ there exists a time starting from which the profile a is no longer chosen. Convergence of a trajectory to an action profile a^* is a special case: let $\mu(a^*) = 1$ in the definition, then $(a^t)_{t=1}^{\infty}$ converges to a^*, written $a^t \to a^*$, if there exists a time $\hat{t}$ such that, for every $t > \hat{t}$, $a^t = a^*$ (this is the standard notion of convergence for discrete spaces, that are formed by isolated points). Also, notice that not all trajectories admit a limit distribution.

The following lemma points out some properties necessarily satisfied by limit distributions.

Lemma 19. *If μ is the limit distribution of $(a^t)_{t=1}^{\infty}$, then for every $\hat{t}$, there exists T such that, for every $t > \hat{t} + T$,*

$$a^t \in \operatorname{supp}\mu \subseteq \{a^\tau : \hat{t} \le \tau < t\}.$$

Proof. Let μ be the limit distribution of $(a^t)_{t=1}^{\infty}$ and fix any given time $\hat{t}$ and profile a. If a is never chosen from $\hat{t}$ onwards, that is if $a \notin \left\{a^\tau : \hat{t} \le \tau < t\right\}$ for every $t > \hat{t}$, then, by Definition 31 (i),

$$\mu(a) = \lim_{t\to\infty} \frac{|\{\tau : 1 \le \tau < t, a^\tau = a\}|}{t} \le \lim_{t\to\infty} \frac{\hat{t}}{t} = 0,$$

i.e., $a \notin \operatorname{supp}\mu$. Therefore, for every $a \in \operatorname{supp}\mu$ there exists a time T'_a such that, for every $t > \hat{t} + T'_a$, $a \in \left\{a^\tau : \hat{t} \le \tau < t\right\}$. Let $T' = \max_{a\in\operatorname{supp}\mu} T'_a$ (this is well defined because A is finite). Then, for every $t > \hat{t} + T'$, $\operatorname{supp}\mu \subseteq \left\{a^\tau : \hat{t} \le \tau < t\right\}$. Moreover, if $\mu(a) = 0$ there exists a T''_a such that $\forall t > \hat{t} + T''_a$, $a^t \neq a$ (Definition 31 (ii)). Let $T'' = \max_{a\in A\setminus\operatorname{supp}\mu} T''_a$, then $\forall t > \hat{t} + T''$, $a^t \in \operatorname{supp}\mu$. Let $T = \max\{T', T''\}$, then

$$\forall t > \hat{t} + T,\ a^t \in \operatorname{supp}\mu \subseteq \left\{a^\tau : \hat{t} \le \tau < t\right\}.$$

■

Note that the membership relation means that, from a certain t onwards, only action profiles in the support of the limit distribution occur. The inclusion states that *all* action profiles in the support occur infinitely often.

It is now possible to present a few elementary results that relate adaptive learning with the solution and equilibrium concepts introduced earlier. The first result identifies a sufficient condition for consistency with adaptive learning.[3]

Theorem 19. *Let $(a^t)_{t=1}^{\infty}$ be a trajectory. If the limit distribution of $(a^t)_{t=1}^{\infty}$ (provided that it exists) is a canonical correlated equilibrium, then $(a^t)_{t=1}^{\infty}$ is consistent with adaptive learning.*

Proof. Let μ be the limit distribution of $(a^t)_{t=1}^{\infty}$ and fix an arbitrary $\hat{t}$. It follows from Lemma 19 and the monotonicity of ρ that there exists T such that, for every $t > \hat{t} + T$, $a^t \in \operatorname{supp}\mu$ and

$$\rho(\operatorname{supp}\mu) \subseteq \rho(\{a^\tau : \hat{t} \le \tau < t\}).$$

Suppose that μ is a canonical correlated equilibrium. Then $\operatorname{supp}\mu \subseteq \rho(\operatorname{supp}\mu)$ (the proof is almost identical to the one of Theorem 15). The claim follows. ∎

The next two results show that, in the long run, adaptive learning induces behavior consistent with the "complete-information" solution concepts of earlier chapters, even if consistency with adaptive learning does not require complete information.

Theorem 20. *Let $(a^t)_{t=1}^{\infty}$ be a trajectory consistent with adaptive learning. Then (1)*

$$\forall k \ge 0, \exists t_k, \forall t \ge t_k,\ a^t \in \rho^k(A), \tag{7.1.1}$$

so that only rationalizable actions are chosen in the long run; (2) if $a^t \to a^*$, then a^* is a Nash equilibrium.

Proof. (1) First recall that since A is finite there exists a K such that $\rho^k(A) = \rho^\infty(A)$ for every $k \ge K$. Then, (7.1.1) implies that from some time t_K onwards only rationalizable actions are chosen. The proof of (7.1.1) is by induction. The statement trivially holds for $k = 0$. Suppose

[3] See Definition 22 and Remark 14.

by way of induction that the statement is true for a given k. By consistency with adaptive learning, there exists a T_k such that

$$\forall t > t_k + T_k,\ a^t \in \rho(\{a^\tau : t_k \leq \tau < t\}).$$

By the inductive assumption, $\tau \geq t_k$ implies $a^\tau \in \rho^k(A)$. Hence,

$$\{a^\tau : t_k \leq \tau < t\} \subseteq \rho^k(A).$$

By monotonicity of ρ,

$$\rho(\{a^\tau : t_k \leq \tau < t\}) \subseteq \rho(\rho^k(A)) = \rho^{k+1}(A).$$

Taking $t_{k+1} = t_k + T_k + 1$, the above inclusions yield

$$\forall t \geq t_{k+1},\ a^t \in \rho(\{a^\tau : t_k \leq \tau < t\}) \subseteq \rho^{k+1}(A)$$

as desired.

(2) If $a^t \to a^*$, then there exists a $\hat{t}$ such that $a^t = a^*$ for every $t > \hat{t}$. Consistency of $(a^t)_{t=1}^\infty$ with adaptive learning implies that there exists T such that

$$\forall t > \hat{t} + T,\ a^* = a^t \in \rho(\{a^\tau : \hat{t} \leq \tau < t\}) = \rho(\{a^*\}).$$

Since $a^* \in \rho(\{a^*\})$, a^* is a Nash equilibrium. ■

7.2 Imperfectly Observable Actions

The definition of consistency with adaptive learning needs to be modified when players cannot perfectly observe the actions previously chosen by their opponents. Assume that if an action profile $a = (a_j)_{j \in I}$ is chosen player i only observes a signal (or message) $m_i = f_i(a)$. As in the section on self-confirming equilibria (Section 6.3), the profile of functions $f = (f_i : A \to M_i)_{i \in I}$ is a primitive element of the analysis.

Take a trajectory $(a^t)_{t=1}^\infty$. The set of signals observed by player i from time $\hat{t}$ (included) to time t (excluded) is $\{m_i : \exists \tau, \hat{t} \leq \tau < t, m_i = f_i(a^\tau)\}$. From i's perspective, the set of other players' action profiles that could have been chosen in the same time interval is

$$\left\{a_{-i} : \exists \tau, \hat{t} \leq \tau < t, f_i(a^\tau) = f_i(a_i^\tau, a_{-i})\right\}.$$

This motivates the following definition:

Definition 32. *A trajectory* $(a^t)_{t=1}^{\infty}$ *is* f-consistent with adaptive learning *if for every* $\hat{t}$ *there exists a* T *such that, for every* $t > \hat{t} + T$ *and* $i \in I$, $a_i^t \in \rho_i\left(\left\{a_{-i} : \exists \tau, \hat{t} \leq \tau < t, f_i(a^\tau) = f_i(a_i^\tau, a_{-i})\right\}\right)$.

Remark 26. *Imperfect observability implies that, in general, two players* i *and* j *obtain different information about the past actions of a third player* k*. This is the reason why the rationalization operator* ρ *could not be used to define the property of* f*-consistency with adaptive learning. If each section of feedback function* f_{i,a_i} *is injective (*$i \in I$*,* $a_i \in A_i$*), then there is perfect observability of other players' previous actions and the definition of the previous section (Definition 30) is obtained.*

It is now possible to generalize the results that relate consistency with adaptive learning and equilibrium. Given the discussion of Section 6.3, it should not be surprising that the relevant equilibrium concept for this scenario is the self-confirming equilibrium.[4]

Theorem 21. *Fix a game with feedback* (G, f) *and a trajectory* $(a^t)_{t=1}^{\infty}$*. If* $(a^t)_{t=1}^{\infty}$ *has a limit distribution which is part of an anonymous self-confirming equilibrium of* (G, f)*, then* $(a^t)_{t=1}^{\infty}$ *is* f*-consistent with adaptive learning.*

Proof. Fix an anonymous self-confirming equilibrium $(\alpha_i, (\nu_{a_i}^i)_{a_i \in \mathrm{supp}\alpha_i})_{i \in I}$. By the confirmed conjectures condition, for every $i \in I$ and $a_i \in \mathrm{supp}\alpha_i$, $\hat{f}_{i,a_i}(\nu_{a_i}^i) = \hat{f}_{i,a_i}(\alpha_{-i})$, where $\hat{f}_{i,a_i} : \Delta(A_{-i}) \to \Delta(M_i)$ is the pushforward induced by the section $f_{i,a_i} : A_{-i} \to M_i$ of i's feedback function at a_i. This implies $f_{i,a_i}(\mathrm{supp}\nu_{a_i}^i) = f_{i,a_i}(\mathrm{supp}\alpha_{-i})$ and therefore $\mathrm{supp}\nu_{a_i}^i \subseteq f_{i,a_i}^{-1}(f_{i,a_i}(\mathrm{supp}\alpha_{-i}))$, which can be rewritten more explicitly as

$$\mathrm{supp}\nu_{a_i}^i \subseteq \{a_{-i} \in A_{-i} : \exists \hat{a}_{-i} \in \mathrm{supp}\alpha_{-i},\ f_i(a_i, a_{-i}) = f_i(a_i, \hat{a}_{-i})\}. \tag{7.2.1}$$

In words, the conjecture that justifies a given action in a self-confirming equilibrium deems possible only (but possibly not all) the action profiles of the opponents that are consistent with the feedback generated by some of the profiles that actually occur with positive probability. This fact will be used in the proof of the proposition, which we are about to undertake.

[4]The following results are adapted from Gilli [38].

Before we delve into the mathematical details, we can provide the intuition behind the result. Recall that, starting from a certain period, say t, the only profiles of actions that occur are those in the support of the limit distribution—provided that it exists—and all such profiles occur infinitely often. If such limit distribution is an anonymous self-confirming equilibrium, then every agent, from t onwards, plays only actions in the support of the equilibrium distribution and keeps playing each of them infinitely often. Thus, the set of messages that every player receives starting from t coincides with the set of messages that are possible according to a justifying conjecture, and all such messages occur infinitely often. This implies that the given trajectory is f-consistent with adaptive learning.

Assume that product measure $\prod_{i\in I}\alpha_i$ of the self-confirming equilibrium is the limit distribution of trajectory $(a^t)_{t=1}^{\infty}$, which implies that α_i is the limit distribution of $(a_i^t)_{t=1}^{\infty}$ for each $i \in I$. Throughout, let $\hat{t}$ be fixed but arbitrary.

Fix $i \in I$. For every $a_i \in \mathrm{supp}\alpha_i$, by the rationality condition, we have $a_i \in r_i(\nu^i_{a_i})$. If we denote by $C^{a_i}_{-i}$ the set on the right hand side of (7.2.1), we have $\nu^i_{a_i} \in \Delta(C^{a_i}_{-i})$, and thus

$$a_i \in r_i(\Delta(C^{a_i}_{-i})) = \rho_i(C^{a_i}_{-i}).$$

Next, we show that there exists T_i' such that, for every $a_i \in \mathrm{supp}\alpha_i$ and $t > \hat{t} + T_i'$,

$$\rho_i(C^{a_i}_{-i}) \subseteq \rho_i\big(\{\hat{a}_{-i} \in A_{-i} : \exists\tau,\, \hat{t} \le \tau < t,\, f_i(a^\tau) = f_i(a_i^\tau, \hat{a}_{-i})\}\big). \quad (7.2.2)$$

By monotonicity of ρ_i, it is sufficient to show

$$C^{a_i}_{-i} \subseteq \{\hat{a}_{-i} \in A_{-i} : \exists\tau,\, \hat{t} \le \tau < t,\, f_i(a^\tau) = f_i(a_i^\tau, \hat{a}_{-i})\}. \quad (7.2.3)$$

Assume $\hat{a}_{-i} \in C^{a_i}_{-i}$, that is, there exists $a_{-i} \in \mathrm{supp}\alpha_{-i}$ such that

$$f_i(a_i, \hat{a}_{-i}) = f_i(a_i, a_{-i}) \quad (7.2.4)$$

(i.e., $\hat{a}_{-i}$ and a_{-i} generate the same feedback given a_i). Since both a_i and a_{-i} are in the support of the *actual* mixed strategies in the limit (product) distribution, the profile $a = (a_i, a_{-i})$ is also played with positive probability, that is, $a \in \mathrm{supp}\alpha$. Thus, a occurs infinitely often in the trajectory $(a^t)_{t=1}^{\infty}$ and there exists $\tau^* > \hat{t}$ such that $a^{\tau^*} = a$, which also implies $f_i(a^{\tau^*}) = f_i(a) = f_i(a_i, \hat{a}_{-i})$. Taking $T'_{i,a_i} > \tau^* - \hat{t}$, it is immediately verified that $\hat{a}_{-i}$ belongs to the right hand side of (7.2.3) for every $t > \hat{t} + T'_{i,a_i}$ (τ^* is the integer τ whose existence is required in the specification of the set). Thus, $\max\{T'_{i,a_i}\}_{a_i \in \mathrm{supp}\alpha_i}$ is the desired T'_i. It follows that, for every $t > \hat{t} + T'_i$

$$\mathrm{supp}\alpha_i \subseteq \bigcup_{a_i \in \mathrm{supp}\alpha_i} \rho_i(C^{a_i}_{-i}) \subseteq \rho_i\big(\{\hat{a}_{-i} \in A_{-i} : \exists\tau,\, \hat{t} \leq \tau < t,\, f_i(a^\tau) = f_i(a_i^\tau, \hat{a}_{-i})\}\big).$$

By Lemma 19, there exists T''_i such that, for every $t > \hat{t} + T''_i$, $a_i^t \in \mathrm{supp}\alpha_i$. Choosing $T_i > \max\{T'_i, T''_i\}$ and $T > \max\{T_i\}_{i \in I}$, we get:

$$\forall t > \hat{t}+T, \forall i \in I,\, a_i^t \in \rho_i\big(\{\hat{a}_{-i} \in A_{-i} : \exists\tau,\, \hat{t} \leq \tau < t,\, f_i(a^\tau) = f_i(a_i^\tau, \hat{a}_{-i})\}\big).$$

Since $\hat{t}$ is arbitrary, this complete the proof. ■

The following result is a partial converse:

Theorem 22. *Fix a game with feedback (G, f) and a trajectory $(a^t)_{t=1}^{\infty}$. If $(a^t)_{t=1}^{\infty}$ is f-consistent with adaptive learning and $a^t \to a^*$ then a^* is a self-confirming equilibrium profile.*

Proof. If $a^t \to a^*$, then there exists $\hat{t}$ such that $a^t = a^*$ for every $t > \hat{t}$. By f-consistency of $(a^t)_{t=1}^{\infty}$ with adaptive learning, it follows that there exists T such that, for every $t > \hat{t} + T$ and for every $i \in I$,

$$\begin{aligned} a_i^* &= a_i^t \in \rho_i\left(\left\{a_{-i} : \exists\tau, \hat{t} \leq \tau < t, f_i(a^\tau) = f_i(a_i^\tau, a_{-i})\right\}\right) \\ &= \rho_i\left(\left\{a_{-i} : f_i(a_i^*, a_{-i}^*) = f_i(a_i^*, a_{-i})\right\}\right). \end{aligned}$$

Hence, for every player i there exists a belief μ^i such that: (1) $a_i^* \in r_i(\mu^i)$, (2) $\mu^i\left(\left\{a_{-i} : f_i(a_i^*, a_{-i}^*) = f_i(a_i^*, a_{-i})\right\}\right) = 1$, which means that $(a_i^*, \mu^i)_{i \in I}$ is a self-confirming equilibrium. ■

8

Games with Incomplete Information

Let $G = \langle I, (A_i, u_i)_{i \in I} \rangle$ be a static game where the payoff functions $u_i : A \to \mathbb{R}$ are derived from an outcome function $g : A \to Y$ and utility functions $v_i : Y \to \mathbb{R}$, that is, $u_i = v_i \circ g$, $i \in I$. To make sense of the strategic analysis of the game, it is typically assumed that G is common knowledge, or—in other words—that there is *complete information*. The reason is apparent for "deductive" solutions concepts, such as rationalizability, that characterize the set of action profiles consistent with common knowledge of the game, rationality and common belief in rationality. Also the interpretation of the Nash equilibrium concept (or more generally of correlated equilibrium) as a "self-enforcing agreement" makes more sense under the complete information assumption. To see this, suppose that before the game is played the players agreed to play some Nash equilibrium profile a^*. Then, nobody has an incentive to deviate *if he expects that the others comply with the agreement*. But why should the players expect that the agreement is complied with? If the payoff functions of others are unknown, a player could suspect that some other player has an incentive to deviate and hence he may want to deviate too.

In some games the incentive to deviate when there is a small chance that the opponent might not honor the agreement is very strong. For example, consider the Pareto efficient agreement (a, c) of the following payoff matrix:

	c	d
a	$100, 100$	$0, 99$
b	$99, 0$	$99, 99$

If Rowena does not know Colin's payoff function, she may suspect—for example—that d is dominant for Colin, and hence switch to her safe action b provided she assigns to d a probability above 1%. A similar argument applies to Colin. Now suppose that payoff functions are mutually known, but Rowena is not sure that Colin knows her payoff function. Then she assigns a positive probability to Colin playing his safe action d instead of the agreed upon action c, and again she would go for the safe action, b, provided she assigns to d a probability above 1%. Continuing like this it can be shown that even a high degree of mutual knowledge of the game does not ensure that the agreement will be honored. On the other hand, common knowledge of the game makes the agreement self-enforcing.[1]

8.1 Games with Payoff Uncertainty

How can we represent a game with incomplete information? The problem can be viewed as follows: the outcome function $g : A \to Y$ and the utility functions $v_i : Y \to \mathbb{R}$ depend on a vector of parameters θ which is not commonly known. Hence, the payoff functions $u_i = v_i \circ g$ depend on a parameters vector which is not commonly known.[2] To represent this uncertainty, payoff functions can be written in a parameterized form

$$u_i : \Theta \times A \to \mathbb{R},$$

where Θ is the set of values of θ that are not excluded by what is commonly known among the players.

In order to represent what a player knows about θ, we assume for simplicity that θ is decomposable into subvectors θ_j, with $j \in \{0\} \cup I$, and that each player $i \in I$ knows the true value of θ_i, whereas θ_0 represents

[1] Vice versa, recall that the interpretation of a Nash equilibrium (pure or mixed) as a stationary state of an adaptive process does *not* require common knowledge, nor mutual knowledge of the game.

[2] The set of the opponents' actions could be unknown as well. We omit this source of uncertainty for the sake of simplicity.

residual uncertainty that would persist even if the players could credibly share their private information. To further simplify the analysis, we will sometimes assume that there is no such residual uncertainty (in many cases, this assumption is rather innocuous). In this case Θ_0 is a singleton and it makes sense to omit it from the notation, writing $\theta = (\theta_i)_{i \in I}$.

Example 24. Consider again the team of two players producing a public good (Example 1). Now we express the production function and cost-of-effort functions in a slightly different parametric form. The parameterized production function is

$$Q = \theta_0 (a_1)^{\theta_1^q} (a_2)^{\theta_2^q}.$$

The parameterized cost of the effort of player i measured in terms of output is

$$C_i = \frac{1}{2\theta_i^c}(a_i)^2.$$

The parameterized consequence function specifies a triple (Q, C_1, C_2) as a function of (θ, a_1, a_2):

$$g(\theta, a_1, a_2) = \left(\theta_0 (a_1)^{\theta_1^q} (a_2)^{\theta_2^q}, \frac{1}{2\theta_1^c}(a_1)^2, \frac{1}{2\theta_2^c}(a_2)^2 \right).$$

The *commonly known utility function* of player i is $v_i(Q, C_1, C_2) = Q - C_i$. The *parameterized payoff function* is then

$$u_i(\theta, a_1, a_2) = v_i(g(\theta, a_1, a_2)) = \theta_0 (a_1)^{\theta_1^q} (a_2)^{\theta_2^q} - \frac{1}{2\theta_i^c}(a_i)^2.$$

It is assumed to be common knowledge that $\theta_0 \in \Theta_0 \subseteq \mathbb{R}_+$ and $\theta_i = (\theta_i^q, \theta_i^c) \in \Theta_i \subseteq \mathbb{R}_+^2$ where the sets Θ_0, Θ_1 and Θ_2 are specified by the given model. Each player i knows his efficiency parameter $\theta_i = (\theta_i^q, \theta_i^c)$ and nobody knows θ_0, unless Θ_0 is a singleton. ▲

In Example 24 the set $\Theta = \Theta_0 \times \Theta_1 \times \Theta_2$ only includes parameters that directly affect the payoff functions of players and, for this reason, we call them **payoff-relevant** (θ_1^q and θ_2^q are output elasticities with respect to agents' actions, θ_1^c and θ_2^c are coefficients of the cost functions, and θ_0 is a "total factors productivity" parameter). However, parameter vector θ_i can also include **payoff-irrelevant** information, namely information that does not *directly* affect payoffs. As we are going to show in this chapter, even payoff-irrelevant parameters may affect the strategic analysis. For instance, a player may believe that some payoff-irrelevant information that he or other players observe is correlated with parameters affecting payoffs.

Unlike θ_i $(i \in I)$, the parameter θ_0 that captures residual uncertainty may be assumed to affect the payoff of at least one player: since it is common knowledge that players ignore θ_0 and cannot condition their behavior on θ_0, there is no reason to include payoff-irrelevant components in θ_0. Formally, it may be assumed without any loss of generality that, whenever $\theta_0' \neq \theta_0''$, there must be some player j, information profile $(\theta_i)_{i \in I}$, and action profile $a = (a_i)_{i \in I}$ such that $u_j(\theta_0', (\theta_i)_{i \in I}, a) \neq u_j(\theta_0'', (\theta_i)_{i \in I}, a)$.

Example 25. Consider the following Cournot oligopoly game with incomplete information: The market inverse demand function is

$$P(\theta_0, Q) = \max\{0, \theta_0 - Q\}, \text{ with } Q = \sum_{i=1}^{n} q_i.$$

The cost function of each firm i is given by

$$C_i(\theta_i, q_i) = \frac{1}{2\theta_i^c}(q_i)^2.$$

Then

$$u_i(\theta, q_1, ..., q_n) = q_i \max\left\{0, \theta_0 - \sum_{i=1}^{n} q_i\right\} - \frac{1}{2\theta_i^c}(q_i)^2. \tag{8.1.1}$$

Thus, θ_0 is an (unknown) shift parameter that captures market conditions, while θ_i^c parametrizes the cost function of firm i. Obviously, θ_0 and $(\theta_i^c)_{i \in I}$ are payoff relevant. Further, assume that each firm has a marketing department which performs a market analysis and issues a score to summarize its report; in particular, for every $i \in I$, let $\theta_i^m = \theta_0 + \varepsilon_i$, where ε_i is normally distributed with mean 0 and precision τ_i and $\varepsilon_i, \varepsilon_j$ are independent for every i, j with $i \neq j$, conditional on θ_0.[3] The payoff function of each firm i, given by eq. (8.1.1), is independent of $(\theta_j^m)_{j \in I}$; thus, each θ_i^m $(i \in I)$ is payoff-irrelevant. Nevertheless, θ_i^m may play a role in the strategic analysis: (1) firm i can condition its behavior on θ_i^m, (2)

[3] Assume, for simplicity, that the marketing department operational cost is equal to 0.

firm $j \neq i$ may believe that i will condition its behavior on θ_i^m and try to exploit the correlation between θ_j^m, θ_0 and θ_i^m to make infererence about the behavior of i, and so on. ▲

In the previous narrative, θ is mostly interpreted as a fixed parameter vector imperfectly and asymmetrically known by the players. To stress this interpretation we sometimes call θ **state of nature**. However, in many models, either the whole θ, or some of its components, are intepreted as the realization of an exogenous random variable. To illustrate, in Example 25 the shift parameter θ_0 may represent the realization of a demand shock. The specific interpretation of θ (or some of its components) as a fixed parameter or the realization of an exogenous random variable is more important for extensions of the Nash and selfconfirming equilibrium concepts to enviroments with incomplete and asymmetric information. Thus, we will consider it again as we discuss these extensions in Sections 8.6 and 8.7.

8.1.1 Private and Interdependent Values

In general, player i's payoff function, u_i, may depend on the whole profile $\theta = (\theta_j)_{j \in I}$, and not just on θ_i (see Example 24). However, the analysis is simpler when the consequence function g is commonly known and uncertainty only concerns the utility functions $v_i : Y \to \mathbb{R}$. In this case, each player i knows his own payoff function $u_i = v_i \circ g$ (given that g is commonly known and v_i represents i's preferences over lotteries of consequences). To represent other players' ignorance of i's utility (and payoff) function, such a function can be written in parameterized form $v_i(\theta_i, y)$, where θ_i is known only to player i. This yields parameterized payoff functions $u_i(\theta_i, a) = v_i(\theta_i, g(a))$, where i knows θ_i. Whenever u_i varies only with θ_i (and a) the game is said to have **private values**.

Suppose now that the consequence function g is *not* common knowledge. Uncertainty about the consequence function can be expressed by representing g in the parameterized form $g(\theta, a)$. Then, θ_i identifies not only i's preferences, but also i's private information about the consequence function. Payoff functions have the general parameterized form $u_i(\theta, a) = v_i(\theta_i, g(\theta, a))$. Since u_i depends on the whole parameter vector θ, the game is said to have **interdependent values**, meaning that u_i may vary with θ_j $(j \neq i)$.

Finally, we say that there is distributed knowledge of θ if parameter θ can be identified by pooling the information of all the players. More formally, let $\theta = (\theta_0, (\theta_i)_{i\in I})$ and let Θ_j denote the set of possible values of θ_j ($j \in \{0\} \cup I$). Then there is **distributed knowledge of** θ when Θ_0 is a *singleton*.

Example 26. Consider again Example 25. Then, there is distributed knowledge of θ if the (inverse) demand function is commonly known, i.e., $\Theta_0 = \{\bar{\theta}_0\}$. In this case, the model also features private values. ▲

Example 27. In the game of Example 24 there is distributed knowledge of θ if and only if there is only one possible value, say $\bar{\theta}_0$, of the total productivity parameter, i.e., $\Theta_0 = \{\bar{\theta}_0\}$. There are private values if and only if there is distributed knowledge of θ, hence common knowledge of total productivity $\bar{\theta}_0$, and moreover there is common knowledge of the output elasticities with respect to players' actions, $\bar{\theta}_1^q$ and $\bar{\theta}_2^q$, that is, $\Theta_i = \{\bar{\theta}_i^q\} \times \Theta_i^c$ ($i = 1, 2$). Otherwise there are interdependent values. ▲

To summarize, a strategic environment with incomplete information can be simply described by the following mathematical structure

$$\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i : \Theta \times A \to \mathbb{R})_{i\in I} \rangle ,$$

where all the sets Θ_j ($j \in I \cup \{0\}$) and A_i ($i \in I$) are nonempty, Θ_0 is the space of residual uncertainty, Θ_i the set of parameter values that i's opponents consider possible, and—as usual—$\Theta = \times_{i\in I\cup\{0\}} \Theta_i$, $A = \times_{i\in I} A_i$. We call θ_i, i's private information about parameter θ, the **information-type** of i (the term "type" will be used later for a more complex object). We call $\hat{G}$ "**game with payoff uncertainty**" rather than "game with incomplete information;" the latter terminology, although perfectly legitimate, is often used to indicate a game endowed with a description of the possible beliefs of players about θ — see Section 8.4. We assume that all the sets and functions specified by structure $\hat{G}$ are commonly known. Otherwise, it would mean that some aspects of players' uncertainty have been omitted from the formal representation.

Note, *the "size" of Θ_j measures the ignorance of each player $i \neq j$ about θ_j*. It is important to specify how ignorant each player is. Even if the modeler knew the true value of θ, say $\bar{\theta}$, he should still put the sets

Θ_j in the formal model. This is true even in the private value case. Why? Because in the strategic analysis of the game each player i puts himself in the shoes of each other player j and considers how j would behave for each information-type θ_j that i deems possible.

We can extend to games with payoff uncertainty some properties defined earlier.

Definition 33. *A game with payoff uncertainty $\hat{G}$ is **compact-continuous** if all the sets Θ_j $(j \in I \cup \{0\})$ and A_i $(i \in I)$ are compact subsets of Euclidean spaces, and every payoff function u_i is continuous $(i \in I)$.*

Definition 34. *A game with payoff uncertainty $\hat{G}$ is **nice** if (1) it is compact-continuous, (2) all the sets Θ_j $(j \in I \cup \{0\})$ are convex subsets of Euclidean spaces,*[4] *and, (3) for every player $i \in I$, A_i is a compact interval in $\mathbb{R}$, and the section of his payoff function $u_{i,\theta,a_{-i}} : A_i \to \mathbb{R}$ is strictly quasi-concave for all $(\theta, a_{-i}) \in \Theta \times A_{-i}$.*

The games of Examples 24 and 25 are nice if all the parameter sets and action sets are convex and compact.

8.2 Rationalizability and Payoff Uncertainty

The specification of a game with payoff uncertainty is sufficient to obtain some results about the outcomes of strategic interaction that are consistent with rationality and standard assumptions about players' beliefs (for instance, some degrees of mutual belief in rationality). Consider the following examples:

Example 28. There are two possible payoff functions for each player i, $u_i(\theta^a, \cdot)$ and $u_i(\theta^b, \cdot)$. Player 1 (Rowena) knows the true payoff functions while player 2 (Colin) does not know them. The situation can be represented with two matrices, corresponding to the two possible values of

[4] It is enough to assume that each Θ_j is connected, i.e., it is not contained in a union of disjoint open sets. Every convex set is connected. A subset of $\mathbb{R}$ is connected if and only if it is convex, but a connected subset of $\mathbb{R}^k$ may be nonconvex. What matters for our analysis is that, for every continuous function $f : X \to \mathbb{R}$ $(X \subseteq \mathbb{R}^k)$, if X is convex, or connected, then $f(X)$ is an interval.

the parameter θ, assuming that Rowena knows the true payoff matrix.[5] Action a dominates b in payoff matrix θ^a, while b is justifiable (not dominated) in matrix θ^b.

$\hat{G}^1$:

θ^a	c	d
a	$4, 0$	$2, 1$
b	$3, 1$	$1, 0$

θ^b	c	d
a	$2, 0$	$0, 1$
b	$0, 1$	$1, 2$

Consider the following epistemic assumptions: R_1, R_2, $\mathrm{B}_1(R_2)$, $\mathrm{B}_2(R_1)$, $\mathrm{B}_1(\mathrm{B}_2(R_1))$.[6] Obviously the implications of such hypotheses for Rowena's choice depend on the true value of θ (which is known only to her). As theorists, we do not make any assumption about θ, but rather we want to explain what may happen (consistently with the above mentioned assumptions) for each possible value $\theta \in \Theta$.
R_1 implies that Rowena chooses a if $\theta = \theta^a$ (a is dominant in matrix θ^a).
$R_2 \cap \mathrm{B}_2(R_1)$ implies that Colin chooses d. Indeed, $\mathrm{B}_2(R_1)$ implies that Colin is certain that Rowena would choose a if $\theta = \theta^a$. Hence, Colin assigns probability zero to the pair (θ^a, b). Also, notice that d is "dominant" when the set of possible pairs is restricted to $\left\{(\theta^a, a), (\theta^b, a), (\theta^b, b)\right\}$.
$R_1 \cap \mathrm{B}_1(R_2) \cap \mathrm{B}_1(\mathrm{B}_2(R_1))$ implies that Rowena expects d and as a consequence chooses b if $\theta = \theta^b$.
To sum up, the aforementioned assumptions imply that the outcome is (a, d) if $\theta = \theta^a$ and (b, d) if $\theta = \theta^b$. Note, the set of i's actions consistent with rationality and these epistemic assumptions can depend only on what i knows. Then, in this case, the set of possible actions of Colin (the singleton $\{d\}$) is independent of θ. ▲

Example 29. Players 1 and 2 receive an envelope with a money prize of k thousands Euros, with $k = 1, ..., K$. It is possible that both players receive the same prize. Every player knows the content of her own envelope and can either offer to exchange her envelope with that of the other player (action OE=*Offer to Exchange*), or not (action N=*do Not offer to*

[5] In this case, we can drop indexes from θ as only player 1 has private information and therefore there is a one-to-one correspondence between Θ and Θ_1. Thus, $\Theta_1 = \left\{\theta_1^a, \theta_1^b\right\}$, $\Theta_2 = \left\{\bar{\theta}_2\right\}$, $\Theta = \left\{\theta^a, \theta^b\right\}$ with $\theta^a = (\theta_1^a, \bar{\theta}_2)$ and $\theta^b = (\theta_1^b, \bar{\theta}_2)$.

[6] The meaning of these symbols is explained in Chapter 4.

exchange). The OE/N actions are taken simultaneously and the exchange takes place only if it is offered by both. In order to offer an exchange a player has to pay a small transaction cost ε. The players' payoff is given by the amount of money they end up with at the end of the game. For this example, $\Theta_i = \{1, ..., K\}$ and $u_i(\theta, a)$ is given by the following table:

$a_i \backslash a_j$	OE	N
OE	$\theta_j - \varepsilon$	$\theta_i - \varepsilon$
N	θ_i	θ_i

Therefore, a necessary condition for i to offer to exchange is that he assigns a positive probability to event $[\theta_j > \theta_i] \cap [a_j = \text{OE}]$. It can be shown that the assumptions of rationality and common belief in rationality imply that i keeps his envelope, whatever the content. Let us consider, for simplicity, only the case $K = 3$ (the general result can be obtained by induction):
R_i implies that i keeps the envelope if $\theta_i = 3$, since by offering to exchange he could obtain at most $3 - \varepsilon$.
$R_i \cap \mathrm{B}_i(R_j)$ implies that i keeps the envelope if $\theta_i \geq 2$. Indeed, i is certain that j would not offer to exchange if $\theta_j = 3$; hence i is certain that by offering to exchange he would obtain at most $2 - \varepsilon$.
$R_i \cap \mathrm{B}_i(R_j) \cap \mathrm{B}_i(\mathrm{B}_j(R_i))$ implies that i keeps the envelope whatever the value of θ_i. Indeed i is certain that j could offer to exchange only if $\theta_j = 1$; hence i is certain that by offering to exchange he could obtain at most $1 - \varepsilon$. ▲

Intuitive, step-by-step solutions such as those of the previous examples can be obtained by a simple extension of the rationalizability solution concept from games with complete information to games with payoff uncertainty. In general, we deem "rationalizable" any outcome consistent with the assumptions of *rationality and common belief in rationality*, which we denote using the symbol $R \cap \mathrm{CB}(R)$,[7] given the *common background knowledge*. Rationalizability in games of complete information, for instance, identifies the choices that are consistent with assumptions $R \cap \mathrm{CB}(R)$ given that the game is common knowledge.

[7] Recall that $\mathrm{CB}(R) = \bigcap_{k \geq 1} \mathrm{B}^k(R)$ denotes common probability-one belief in rationality.

In the analysis of strategic interaction with incomplete information, it is interesting to address the following question: What behavior is consistent with $R \cap \mathrm{CB}(R)$ given that the features of strategic interaction encoded by structure $\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i)_{i \in I} \rangle$ are common knowledge? The intuitive solutions of the previous examples suggest how to obtain the answer: eliminate, step by step, the pairs (θ_i, a_i) such that a_i is not a best reply to any probabilistic conjecture $\mu^i \in \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i})$ that assigns probability zero to every (θ_{-i}, a_{-i}) eliminated in the previous steps. By Lemma 2, in a given step of the elimination procedure one has to delete the pairs (θ_i, a_i) such that, for θ_i, a_i is dominated by some mixed action α_i given the previous steps of elimination.

More formally, this solution procedure can be defined *via* a generalization of the rationalization operator. Consider a finite or compact-continuous $\hat{G}$. First we extend the definition of best reply correspondence to the payoff-uncertainty setup: for every $\mu^i \in \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i})$ and $\theta_i \in \Theta_i$, let

$$r_i(\mu^i, \theta_i) = \arg \max_{a_i \in A_i} \mathbb{E}_{\mu^i}\left(u_{i,\theta_i,a_i}\right), \tag{8.2.1}$$

where $u_{i,\theta_i,a_i} : \Theta_0 \times \Theta_{-i} \times A_{-i} \to \mathbb{R}$ is the section of i's parameterized payoff function at (θ_i, a_i). When conjecture μ^i has a finite support, the expected utility formula is

$$\mathbb{E}_{\mu^i}\left(u_{i,\theta_i,a_i}\right) = \sum_{(\theta_0,\theta_{-i},a_{-i}) \in \mathrm{supp}\mu^i} u_i\left(\theta_0, \theta_i, \theta_{-i}, a_i, a_{-i}\right) \mu^i\left(\theta_0, \theta_{-i}, a_{-i}\right).$$

As usual, we slightly abuse notation and write $\Delta(\Theta_0 \times C_{-i})$ for the set of probability measures on $\Theta_0 \times (\bigtimes_{j \neq i} \Theta_j \times A_j)$ that assign probability one to $\Theta_0 \times C_{-i}$. The following lemma says that the best reply correspondence $(\mu^i, \theta_i) \mapsto r_i(\mu^i, \theta_i)$ defined above is "well behaved," that is, it is nonempty valued and it has a closed graph:[8]

Lemma 20. *If $\hat{G}$ is a compact-continuous game with payoff uncertainty, for every $i \in I$, and every nonempty closed subset $C_{0,-i} \subseteq \Theta_0 \times \Theta_{-i} \times A_{-i}$,*

$$\mathrm{Gr}\left(r_i|_{\Delta(C_{0,-i}) \times \Theta_i}\right) = \left\{\left(\mu^i, \theta_i, a_i\right) \in \Delta\left(C_{0,-i}\right) \times \Theta_i \times A_i : a_i \in r_i\left(\mu^i, \theta_i\right)\right\}$$

[8] The mathematical meaning of "closed" in the lemma is clear when $\Theta_0 \times \Theta_{-i} \times A_{-i}$ is finite, which implies that $\Delta\left(\Theta_0 \times \Theta_{-i} \times A_{-i}\right)$ is a compact (and convex) subset of $\mathbb{R}^{\Theta_0 \times \Theta_{-i} \times A_{-i}}$. For the more general compact-continuous case see Section 4.7 and, in particular, Lemma 10.

is closed, and $r_i\left(\mu^i, \theta_i\right) \neq \emptyset$ *for all* $\left(\mu^i, \theta_i\right) \in \Delta\left(C_{0,-i}\right) \times \Theta_i$*.*

Let $\mathcal{C}$ denote the collection of closed Cartesian subsets of the form

$$C = \Theta_0 \times \left(\bigtimes_{i \in I} C_i \right) \subseteq \Theta \times A,$$

where $C_i \subseteq \Theta_i \times A_i$ for each $i \in I$, and let

$$C_{0,-i} = \Theta_0 \times \left(\bigtimes_{j \neq i} C_j \right).$$

With this, define the **rationalization operator** $\rho : \mathcal{C} \to \mathcal{C}$ as follows: for each $C \in \mathcal{C}$,

$$\begin{aligned} \rho_i(C_{0,-i}) &= \left\{ (\theta_i, a_i) \in \Theta_i \times A_i : \exists \mu^i \in \Delta(C_{0,-i}), a_i \in r_i(\mu^i, \theta_i) \right\}, \\ \rho(C) &= \Theta_0 \times \left(\bigtimes_{i \in I} \rho_i(C_{0,-i}) \right). \end{aligned}$$

Intuitively, we keep the whole residual uncertainty space Θ_0 as is, because strategic thinking can rule out some pairs (θ_i, a_i) for some player i, but cannot rule out any possible θ_0. Formally, one can reinterpret θ_0 as the information type of an *inactive* player 0, so that A_0 is a singleton.

Remark 27. *ρ is* monotone*, that is, $E \subseteq F$ implies $\rho(E) \subseteq \rho(F)$ for all $E, F \in \mathcal{C}$; therefore, the sequence $\left(\rho^k(\Theta \times A)\right)_{k=1}^{\infty}$ is (weakly) decreasing.*

With this, we can extend the definition of rationalizability from games with complete information to games with payoff uncertainty.[9]

Definition 35. *A profile $\left(\theta_0, (\theta_i, a_i)_{i \in I}\right)$ is* ***rationalizable*** *if*

$$\left(\theta_0, (\theta_i, a_i)_{i \in I}\right) \in \rho^{\infty}(\Theta \times A) = \bigcap_{k \in \mathbb{N}} \rho^k(\Theta \times A);$$

an action a_i is ***rationalizable for information-type*** *θ_i if the pair (θ_i, a_i) is part of a rationalizable profile, that is,*

$$(\theta_i, a_i) \in \mathrm{proj}_{\Theta_i \times A_i} \rho^{\infty}(\Theta \times A).$$

[9] See Battigalli [7].

A proper formalization of events (like rationality, R) and belief operators (like $\mathrm{B}_i(\cdot)$ and $\mathrm{B}(\cdot)$) would allow to derive the following table, where it is implicitly assumed that there is common knowledge of the game with payoff uncertainty $\hat{G}$.

Assumptions about rationality and beliefs	Behavioral implications
R	$\rho(\Theta \times A)$
$R \cap \mathrm{B}(R)$	$\rho^2(\Theta \times A)$
$R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$	$\rho^3(\Theta \times A)$
...	...
$R \cap \left(\bigcap_{k=1}^{K} \mathrm{B}^k(R)\right)$	$\rho^{K+1}(\Theta \times A)$
...	...

In Example 28 $\rho^3(\Theta \times A) = \left\{(\theta^a, a), (\theta^b, b)\right\} \times \{d\}$. In Example 29 $\rho^K(\Theta \times A) = (\Theta_1 \times \{N\}) \times (\Theta_2 \times \{N\})$ (after K steps, where K is the number of possible prizes, only the pairs (θ_i, N) survive).

The results about rationalizability in games with complete information can be extended to games with payoff uncertainty:

Theorem 23. *If $\hat{G}$ is a finite or compact-continuous game with payoff uncertainty, then*

$$\mathrm{proj}_\Theta \rho^k (\Theta \times A) = \Theta$$

for all $k \in \mathbb{N} \cup \{\infty\}$, and

$$\rho\left(\rho^\infty(\Theta \times A)\right) = \rho^\infty(\Theta \times A)\,;$$

furthermore, for every $C \in \mathcal{C}$,

$$C \subseteq \rho(C) \Rightarrow C \subseteq \rho^\infty(\Theta \times A)\,.$$

The first part follows from Lemma 20 and says that we cannot eliminate any information-type θ_i, we can only eliminate pairs (θ_i, a_i) such that a_i is not rationalizable given θ_i; this implies that *the set of rationalizable actions is nonempty for every information-type of every player.* The second part says that restarting the iterations from the set of rationalizable profiles cannot further reduce it. The third part says that every set $C \in \mathcal{C}$ with the best reply property is contained in the set of rationalizable profiles. The proof is a rather straightforward extension of the proofs of Theorems 2 and 3.

Furthermore, also the equivalence between rationalizability and iterated dominance can be extended to the payoff-uncertainty setup. For any $C \in \mathcal{C}$, let $C_{i,\theta_i} = \{a_i \in A_i : (\theta_i, a_i) \in C_i\}$ denote the θ_i-section of C_i.

Definition 36. *Fix $C \in \mathcal{C}$, $i \in I$, $a_i \in A_i$, $\alpha_i \in \Delta(A_i)$ in a game with payoff uncertainty $\hat{G}$; mixed action α_i* ***dominates*** *a_i* ***given*** *θ_i* ***within*** *C, written $\alpha_i \gg_{(\theta_i, C)} a_i$, if* $\text{supp}\alpha_i \subseteq C_{i,\theta_i}$ *and*

$$\forall(\theta_0, \theta_{-i}, a_{-i}) \in \Theta_0 \times C_{-i},\ \mathbb{E}_{\alpha_i}(u_{i,\theta_0,\theta_i,\theta_{-i},a_{-i}}) > u_i(\theta_0, \theta_i, \theta_{-i}, a_i, a_{-i}),$$

where $u_{i,\theta,a_{-i}} : A_i \to \mathbb{R}$ is the section of i's parameterized payoff function at $\theta = (\theta_0, \theta_i, \theta_{-i}, a_{-i})$.

When mixed action α_i has a finite support, the expected utility formula is

$$\mathbb{E}_{\alpha_i}(u_{i,\theta_0,\theta_i,\theta_{-i},a_{-i}}) = \sum_{a'_i \in \text{supp}\alpha_i} u_i\left(\theta_0, \theta_i, \theta_{-i}, a'_i, a_{-i}\right) \alpha_i\left(a'_i\right).$$

This is like the standard notion of dominance by a mixed action (within a restricted set of possible profiles), but it is extended to take into account i's private information and i's uncertainty about (θ_0, θ_{-i}). Next define the set of undominated profiles given $C \in \mathcal{C}$, that is $\text{ND}(C)$, as follows:

$$\begin{aligned}
\text{ND}_i(C) &= C_i \backslash \left\{(\theta_i, a_i) \in C_i : \exists \alpha_i \in \Delta(C_{i,\theta_i}), \alpha_i \gg_{(\theta_i, C)} a_i\right\}, \\
\text{ND}(C) &= \Theta_0 \times \left(\bigtimes_{i \in I} \text{ND}_i(C)\right).
\end{aligned}$$

As in the complete-information case, $\text{ND} : \mathcal{C} \to \mathcal{C}$ is *not monotone.* Lemma 2 can be restated as follows:

Lemma 21. *Let $\hat{G}$ be a finite or compact-continuous game with payoff uncertainty. For each $C \in \mathcal{C}$, $i \in I$ and $(\theta_i, a_i^*) \in C_i$, the following statements are equivalent:*
(1) there is no *mixed action α_i such that $\alpha_i \gg_{(\theta_i, C)} a_i^*$,*
(2) there is a conjecture $\mu^i \in \Delta(\Theta_0 \times C_{-i})$ such that $a_i^* \in$ $\arg\max_{a_i \in C_{i,\theta_i}} \mathbb{E}_{\mu^i}(u_{i,\theta_i,a_i})$.

By a straightforward extension of the methods of Chapter 4, one can use Lemma 21 to prove the following result:

Theorem 24. *If $\hat{G}$ is a finite or compact-continuous game with payoff uncertainty, then $\rho^k(\Theta \times A) = \mathrm{ND}^k(\Theta \times A)$ for all $k \in \mathbb{N} \cup \{\infty\}$.*

Therefore a profile of information-types and actions $((\theta_i, a_i))_{i \in I}$ is rationalizable if and only if it survives the iterated elimination of those pairs (θ_i, a_i) such that a_i is dominated given θ_i within the set of profiles that survived the previous rounds of elimination.

Finally, in the case of *nice* games with payoff uncertainty, rationalizability is characterized by the iterated deletion of pairs (θ_i, a_i) such that either a_i is not a best reply to any deterministic conjecture, or a_i is dominated by another pure action. We can formally state this result by extending the definition of operators r and ND_p from games with complete information (cf. Section 4.6) to games with payoff uncertainty. For each $C \in \mathcal{C}$, let $\mathrm{r}_i(C_{0,-i})$ denote the set of pairs (θ_i, a_i) such that a_i is a best reply for θ_i to a deterministic conjecture in $C_{0,-i}$, that is,

$$\mathrm{r}_i(C_{0,-i}) = \bigcup_{\theta_i \in \Theta_i} r_i\left(C_{0,-i} \times \{\theta_i\}\right)$$

$$= \left\{(\theta_i, a_i) \in \Theta_i \times A_i : \exists \left(\theta_0, (\theta_j, a_j)_{j \neq i}\right) \in C_{0,-i}, a_i \in r_i\left(\left(\theta_0, (\theta_j, a_j)_{j \neq i}\right), \theta_i\right)\right\},$$

where (with the usual abuse of notation) we identify $C_{0,-i} \subseteq \Theta_0 \times \Theta_{-i} \times A_{-i}$ with the set of Dirac measures on $\Theta_0 \times \Theta_{-i} \times A_{-i}$ supported by points in $C_{0,-i}$, and let

$$\mathrm{r}\left(C\right) = \Theta_0 \times \left(\bigtimes_{i \in I} \mathrm{r}_i\left(C_{0,-i}\right)\right).$$

The iterations of operator (self-map) r give rise to an extension of the **point rationalizability** concept to games with payoff uncertainty. For each $C \in \mathcal{C}$, let

$$\mathrm{ND}_{p,i}(C) = C_i \backslash \left\{(\theta_i, a_i) \in C_i : \exists a_i' \in C_{i,\theta_i}, a_i' \gg_{(\theta_i, C)} a_i\right\}$$

and

$$\mathrm{ND}_p(C) = \Theta_0 \times \left(\bigtimes_{i \in I} \mathrm{ND}_{p,i}(C)\right).$$

The iterations of operator ND_p (self-map) yield a simple notion of iterated dominance by pure actions. With these definitions of the operators r and ND_p, we have:

Theorem 25. *If $\hat{G}$ is a nice game with payoff uncertainty, then*

$$\mathrm{r}^k(\Theta \times A) = \rho^k(\Theta \times A) = \mathrm{ND}_p^k(\Theta \times A)$$

for all $k \in \mathbb{N} \cup \{\infty\}$.

The proof is based on a straightforward extension of Lemma 7.

8.3 Directed Rationalizability

According to the previous analysis of rationalizability, restrictions on players' conjectures are only derived from strategic thinking, while restrictions on players' beliefs about the parameter profile θ do not play any role. To see this more formally, let $\rho_j^\infty(\Theta \times A) = \mathrm{proj}_{\Theta_j \times A_j}\rho^\infty(\Theta \times A)$ denote the set of rationalizable pairs (θ_j, a_j) of player j. By Theorem 23, an action a_i is rationalizable for information-type θ_i if and only if it is a best reply to some conjecture μ^i that assigns probability 1 to $\Theta_0 \times \left(\bigtimes_{j \neq i} \rho_j^\infty(\Theta \times A)\right)$, where $\mathrm{proj}_{\Theta_j}\rho_j^\infty(\Theta \times A) = \Theta_j$ for each co-player j (for each type, the set of rationalizable actions is nonempty). Therefore, the marginal beliefs about residual uncertainty and co-players' information-types are not in any way restricted by rationalizability, the restrictions concern only the relationship between information-types and actions implied by rationality and common belief in rationality.

Yet, when we analyze strategic interaction in environments with incomplete information, it may be natural to assume that some contextual restrictions on players beliefs about θ hold, and that it is common belief that they hold. For example, when two sellers compete they may have statistical information about the distribution of valuations among buyers. If this statistical information is common, we obtain a common restriction on sellers' beliefs about buyers' valuations.[10]

To simplify our informal terminology, let us say that some event E is **transparent** if E is true and there is common belief of E, in symbols, if

[10] If buyers are modeled as non-strategic agents rather than players, this is a common restriction on beliefs about θ_0.

$E \cap \mathrm{CB}(E)$ is the case. We are interested in characterizing the behavioral implications of rationality and common belief in rationality under the assumption that some features of players' conjectures are transparent. Specifically, we present an extension of the notion of rationalizability for games with payoff uncertainty whereby some restrictions on players' conjectures are posited and assumed to be transparent, thereby "directing" the result of the solution procedure toward a subset of outcomes. For this reason, we call this more flexible approach "directed rationalizability." To gain intuition on what we mean and how this procedure works, we first look at an example.

Example 30. Consider a modification of the game in Example 28. There are two possible payoff functions for each player i, $u_i(\theta^a, \cdot)$ and $u_i(\theta^b, \cdot)$. Player 1 (Rowena) knows the true payoff functions, while player 2 (Colin) does not know them. We represent this situation with two matrices, corresponding to the two possible values of the parameter θ.

θ^a	c	d
a	4,0	2,1
b	3,1	1,0

θ^b	c	d
a	2,1	0,0
b	0,0	1,1

In this game,

$$\rho^{\infty}(\Theta \times A) = \rho^{1}(\Theta \times A) = \left\{(\theta^a, a), (\theta^b, a), (\theta^b, b)\right\} \times \{c, d\}\,.$$

Indeed, action a is dominant for information-type θ^a, while both a and b are justifiable for θ^b. Moreover, both actions for Colin are justifiable. The procedure ends at the first step because, as it can be verified, action c (resp. d) is optimal for Colin if he thinks that θ^b (resp. θ^a) is more likely than θ^a (resp. θ^b).

Let $p = \mu^2\left(\{\theta^a\} \times A_1\right)$ denote the probability that Colin assigns to information-type θ^a. Assume now that it is transparent that $p > \frac{1}{2}$. Given such restriction, consider the following epistemic assumptions (we are not making the transparency of $p > \frac{1}{2}$ explicit in the notation): R, $R \cap \mathrm{B}(R)$ and $R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$.
R implies that Rowena chooses a if $\theta = \theta^a$ (a is dominant in matrix θ^a).
$R \cap \mathrm{B}(R)$ implies that Colin chooses d. Indeed, $\mathrm{B}_2(R_1)$ implies that Colin is certain that Rowena would choose a if $\theta = \theta^a$. Hence, Colin assigns probability zero to (θ^a, b), which implies $\mu^2\left(\theta^a, a\right) = p$. Since $p > \frac{1}{2}$,

action d yields a strictly higher expected payoff than c. To see this, let $q = \mu^2\left(\theta^b, b\right)$ denote the probability that Colin assigns to the pair (θ^b, b). The expected payoff of choosing d is

$$\mu^2\left(\theta^a, a\right) + \mu^2\left(\theta^b, b\right) = p + q \geq p > \frac{1}{2},$$

and the expected payoff of choosing c is

$$\mu^2\left(\theta^b, a\right) \leq 1 - p < \frac{1}{2}.$$

Since it is transparent that $p > \frac{1}{2}$, $R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$ implies that Rowena expects d and she chooses b if $\theta = \theta^b$.
Summing up, the aforementioned assumptions imply that the outcome is (a, d) if $\theta = \theta^a$ and (b, d) if $\theta = \theta^b$. ▲

The example has two features that make it very simple. First, there is no residual uncertainty and only one player is uninformed. Second, we only posited a restriction on beliefs about the co-player's information-type. It follows that there is no room for a possible dependence of the posited belief restriction on the own information-type of a player: in Example 30 there is only one possible information-type of Colin. Yet, there are reasons to make the analysis more flexible. For example, in many economic applications a player's information-type represents private information about an unknown statistical distribution. To consider a more specific example, when a couple splits and they have to agree on who is going to keep their house and pay a compensation to the other, it may be common belief who values the house more among the two of them, but not exactly by how much. This can be modeled as a type-dependent restriction on beliefs about θ: let i be the high-valuation player, if i's type/valuation is θ_i, then player i is certain that $\theta_{-i} < \theta_i$.

It may also be reasonable to posit joint restrictions concerning the subjective probabilities assigned to information-types *and* actions of co-players, such as (without residual uncertainty) the independence restriction

$$\mu^i\left(\theta_{-i}, a_{-i}\right) = \prod_{j \neq i} \mu^i\left(\theta_j, a_j\right).$$

Next we present an example of a different kind. It may be plausible to assume that the following is transparent in a two-person game with no residual uncertanity about θ: player i assigns strictly positive probability to every type of j, and that if a certain action a_j^* is *weakly* dominant for a type θ_j^*, then player i assigns strictly positive probability to a_j^* given θ_j^*. Thus, it is also transparent that $\mu^i\left(\theta_j^*, a_j^*\right) = \mu^i\left(a_j^*|\theta_j^*\right)\mu^i\left(\theta_j^*\right) > 0$. This seemingly innocuous assumption may have a significant impact.

Example 31. Consider the Envelope Game of Example 29 *without transaction costs*, that is, with $\varepsilon = 0$. Suppose that the only transparent features of players' conjectures are as follows: (a) if i's prize (information-type) is θ_i, then player i assigns strictly positive probability to every prize $\theta_{-i} \leq \theta_i$ of the co-player, and (b) if the co-player had the lowest prize, absent transaction costs, he would offer to exchange (weakly dominant action) with strictly positive probability. Then, an "unraveling" argument quite similar to the one offered in Example 29 shows that no type above the lowest offers to exchange. Note that the assumed restriction is plausible and consistent with strategic reasoning, but it does not *derive* from strategic reasoning. Indeed, if the posited restrictions are transparent, rationality and common belief in rationality imply the following: if a player has the lowest prize, then he is certain that—whatever he does—he ends up with the same prize; hence, he is indifferent. As a consequence, offering to exchange would be rationalizable for every type, because a player could assign probability 1 to the event that the co-player does not exchange. ▲

With this in mind, for each information-type θ_i of each player i, we posit a subset of conjectures in $\Delta(\Theta_0 \times \Theta_{-i} \times A_{-i})$. To simplify the analysis, we assume that all the sets Θ_j $(j \in I \cup \{0\})$ and A_i $(i \in I)$ are *finite*.[11] For each $i \in I$ and $\theta_i \in \Theta_i$, we let $\Delta_{i,\theta_i} \subseteq \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i})$ denote the restricted set of conjectures for information-type θ_i of player i. We maintain throughout this section that $\Delta_{i,\theta_i} \neq \emptyset$ for each i and θ_i. To illustrate, in Example 30 player 2 has no private information and

$$\Delta_2 = \left\{\mu^2 \in \Delta\left(\Theta_1 \times A_1\right) : \mu^2\left(\{\theta^a\} \times A_1\right) > \frac{1}{2}\right\},$$

[11] This simplification is relatively innocuous. The analysis can be extended to compact-continuous games with a modicum of additional technicalities.

whereas $\Delta_{1,\theta^a} = \Delta_{1,\theta^b} = \Delta(\{c,d\})$ (no restrictions for any information-type of player 1); in Example 31

$$\Delta_{i,\theta_i} = \left\{\mu^i : \begin{array}{cc} \forall\theta_{-i} \in \{1,...,\theta_i\}, & \mu^i(\{\theta_{-i}\} \times A_{-i}) > 0, \\ & \mu^i(1, \mathrm{OE}) > 0 \end{array}\right\}.$$

For a given profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$, define the **$\Delta$-rationalization operator** $\rho_\Delta : \mathcal{C} \to \mathcal{C}$ as follows: for each $C \in \mathcal{C}$,

$$\begin{aligned} \rho_{i,\Delta}(C_{0,-i}) &= \left\{(\theta_i, a_i) \in \Theta_i \times A_i : \exists\mu^i \in \Delta(C_{0,-i}) \cap \Delta_{i,\theta_i}, a_i \in r_i(\mu^i, \theta_i)\right\}, \\ \rho_\Delta(C) &= \Theta_0 \times \left(\bigtimes_{i\in I} \rho_{i,\Delta}(C_{0,-i})\right). \end{aligned}$$

It can be verified that, for any fixed Δ, the operator ρ_Δ is *monotone*, that is, $E \subseteq F$ implies $\rho_\Delta(E) \subseteq \rho_\Delta(F)$ for all $E, F \in \mathcal{C}$. There is also another noteworthy form of monotonicity. Consider two profiles $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ and $\Delta' = \left(\Delta'_{i,\theta_i}\right)_{i\in I,\theta_i\in\Theta_i}$, and write $\Delta \subseteq \Delta'$ if $\Delta_{i,\theta_i} \subseteq \Delta'_{i,\theta_i}$ for each $i \in I$ and $\theta_i \in \Theta_i$. The proof of the following statement is left as an exercise for the reader.

Remark 28. *Fix two profiles of restrictions, Δ and Δ'. If $\Delta \subseteq \Delta'$, then $\rho_\Delta(C) \subseteq \rho_{\Delta'}(C)$ for each $C \in \mathcal{C}$.*

For a fixed profile Δ, we define the k-th iteration of ρ_Δ recursively. For each $C \in \mathcal{C}$, let $\rho^0_\Delta(C) = C$, and, for each $k \geq 1$, let $\rho^k_\Delta(C) = \rho_\Delta(\rho^{k-1}_\Delta(C))$.

Definition 37. *Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of restricted sets of conjectures. A profile $\left(\theta_0, (\theta_i, a_i)_{i\in I}\right)$ is Δ-**rationalizable** if*

$$\left(\theta_0, (\theta_i, a_i)_{i\in I}\right) \in \rho^\infty_\Delta(\Theta \times A) = \bigcap_{k\in\mathbb{N}} \rho^k_\Delta(\Theta \times A);$$

*an action a_i is Δ-**rationalizable for information-type** θ_i if the pair (θ_i, a_i) is part of a Δ-rationalizable profile, that is,*

$$(\theta_i, a_i) \in \mathrm{proj}_{\Theta_i\times A_i}\rho^\infty_\Delta(\Theta \times A).$$

First note that we obtain the previous definition of rationalizability for games with payoff uncertainty when we let $\Delta_{i,\theta_i} = \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i})$ for every i and θ_i, that is, when no restrictions are assumed. Second, the set $\rho_\Delta^\infty(\Theta \times A)$ of Δ-rationalizable profiles depends on the specific profile Δ of restricted sets; different specifications of Δ may yield different solution sets. In Example 30, if it is transparent that Colin deems θ^a strictly more likely than θ^b, the solution is

$$\rho_\Delta^\infty(\Theta \times A) = \rho_\Delta^3(\Theta \times A) = \left\{(\theta^a, \mathrm{a}), \left(\theta^b, \mathrm{b}\right)\right\} \times \{\mathrm{d}\};$$

if instead Colin may *also* deem θ^b at least as likely as θ^a (that is, if there are no restrictions for Colin), the solution is

$$\rho_\Delta^\infty(\Theta \times A) = \rho_\Delta^1(\Theta \times A) = \left\{(\theta^a, \mathrm{a}), \left(\theta^b, \mathrm{a}\right), \left(\theta^b, \mathrm{b}\right)\right\} \times \{\mathrm{c}, \mathrm{d}\}.$$

This inclusion result (with respect to the restrictions represented by Δ) is not a coincidence: the reader can use Remark 28 to prove by induction the following statement.

Remark 29. *Fix two profiles of restrictions, Δ and Δ'. If $\Delta \subseteq \Delta'$, then $\rho_\Delta^\infty(\Theta \times A) \subseteq \rho_{\Delta'}^\infty(\Theta \times A)$.*

We call "directed rationalizability" the map $\Delta \mapsto \rho_\Delta^\infty(\Theta \times A)$ when we refer to the final result, and the map $\Delta \mapsto \left(\rho_\Delta^k(\Theta \times A)\right)_{k=1}^\infty$ when we refer to the iterated elimination procedure.

It can be shown that, for each profile Δ, $\rho_\Delta^\infty(\Theta \times A)$ charaterizes the behavioral implications of rationality and common belief in rationality when it is transparent that players' conjectures satisfy the restrictions represented by Δ. Of course, since ρ_Δ is monotone, the sequence $\left(\rho_\Delta^k(\Theta \times A)\right)_{k=1}^\infty$ is (weakly) decreasing. Thus, the abstract analysis of directed rationalizability is similar to the analysis of the simpler "undirected" version. But there is a conceptual difference: if the (transparent) restrictions represented by Δ also concern conjectures about co-players' *behavior*, then they may turn out to be inconsistent with implications of strategic thinking, that is, of the assumptions R, $R \cap \mathrm{B}(R)$, $R \cap \mathrm{B}(R) \cap \mathrm{B}^2(R)$,...; then, the set of Δ-rationalizable actions is empty. This is the case in the Envelope game of Example 29 (i.e., with positive transaction costs) if it is transparent that the lowest type offers to exchange with strictly positive probability, because rationality and common belief in rationality imply that *no* type offers to exchange. We now show that if, instead, the Δ-restrictions only concern beliefs about θ, then the solution is nonempty. This requires some preliminary notations and definitions.

Recall that, given a probability measure μ on a space $X \times Y$, we let $\text{marg}_X \mu$ denote its marginal on X, that is, $\text{marg}_X \mu(E) = \mu(E \times Y)$ for every $E \subseteq X$. We say that $\Delta = (\Delta_{i,\theta_i})_{i \in I, \theta_i \in \Theta_i}$ is a profile of **restrictions on exogenous beliefs**[12] if, for each $i \in I$ and each $\theta_i \in \Theta_i$, there is a nonempty set $\bar{\Delta}_{i,\theta_i} \subseteq \Delta(\Theta_0 \times \Theta_{-i})$ such that

$$\Delta_{i,\theta_i} = \left\{ \mu^i \in \Delta\left(\Theta_0 \times \Theta_{-i} \times A_{-i}\right) : \text{marg}_{\Theta_0 \times \Theta_{-i}} \mu^i \in \bar{\Delta}_{i,\theta_i} \right\}.$$

With this, we obtain the following extension of Theorem 23.

Theorem 26. *Let $\hat{G}$ be a finite game with payoff uncertainty. Consider a profile $\Delta = (\Delta_{i,\theta_i})_{i \in I, \theta_i \in \Theta_i}$ of restrictions on* exogenous *beliefs. Then*

$$\text{proj}_\Theta \rho_\Delta^k \left(\Theta \times A\right) = \Theta$$

for all $k \in \mathbb{N} \cup \{\infty\}$, and

$$\rho_\Delta \left(\rho_\Delta^\infty \left(\Theta \times A\right)\right) = \rho_\Delta^\infty \left(\Theta \times A\right);$$

furthermore, for every $C \in \mathcal{C}$,

$$C \subseteq \rho_\Delta \left(C\right) \Rightarrow C \subseteq \rho_\Delta^\infty \left(\Theta \times A\right).$$

Like the proof of Theorem 23, also the proof of Theorem 26 is left to the reader as an exercise. The key additional step is to show (by induction) that, for each step k and for all i and θ_i there is at least one conjecture $\mu^i \in \Delta_{i,\theta_i}$ such that $\mu^i\left(\Theta_0 \times \left(\bigtimes_{j \neq i} \text{proj}_{\Theta_j \times A_j} \rho_\Delta^{k-1}\left(\Theta \times A\right)\right)\right) = 1$. This is so because the latter condition only concerns conjectures about the behavior of co-players' *given* their information-types, while the former only restricts the probabilities of their information-types.

[12] Recall that we call "exogenous" something that we are not trying to explain as game theorists: we are not trying to explain (θ_0, θ_{-i}) nor—more generally—any beliefs about exogenous things.

8.4 Equilibrium and Beliefs: The Simplest Case

The directed rationalizability approach *allows for* the introduction of some restrictions on exogenous beliefs, i.e., beliefs about parameters or exogenous variables. In this section and the next we argue that an extension of the traditional Nash equilibrium approach *requires* the precise specification of exogenous beliefs.

In examples 28 and 29 the repeated elimination of dominated actions selects a unique outcome for each θ. In the complete information case we proved that if there is a unique rationalizable profile, this must be the unique equilibrium. Also in the present incomplete-information context we can interpret the rationalizable mapping from information-types to actions, *if* it is *unique*, as a unique equilibrium of the following kind. For every player i and every information-type θ_i there is a corresponding choice, which in general can be denoted by $\sigma_i(\theta_i)$—in Example 28 $\sigma_1(\theta^a) = a$, $\sigma_1(\theta^b) = b$, and $\sigma_2 = d$; in Example 29 $\sigma_i(\theta_i) = \text{N}$. The unique profile of rationalizable decision functions $\sigma = (\sigma_i : \Theta_i \to A_i)_{i \in I}$ has the following property: there exists no player i, no information-type θ_i and no belief over $\Theta_0 \times \Theta_{-i}$, such that i has an incentive to deviate from $\sigma_i(\theta_i)$ *provided that i expects every other player j to act according to the decision function* $\sigma_j(\cdot)$; furthermore, σ is the unique profile of decision functions with this property.

A decision function of player j can be interpreted as the conjecture of player i about how j would behave as a function of his information-type θ_j. Then, one can tentatively define an equilibrium in this context as a profile of decision functions such that every player, given his information-type, chooses a best reply to his conjecture about other players, and furthermore this conjecture is correct, i.e., it corresponds to how each co-player chooses as a function of his own information-type.

The problem of this tentative definition is that it is silent on the belief of each player i over $\Theta_0 \times \Theta_{-i}$. But in many games it is not possible to ascertain whether a profile of decision functions $(\sigma_i)_{i \in I}$ has the property that no player i has incentives to unilaterally deviate from σ_i (i.e., what the others expect him to choose) without further assumptions about players' beliefs about θ. This is shown by the following example.

Example 32. Consider the following variation of game $\hat{G}^1$ (Rowena knows the true payoff matrix, Colin does not):

$\hat{G}^2$:

θ^a	c	d
a	$4, 0$	$2, 1$
b	$3, 1$	$1, 0$

θ^b	c	d
a	$1, 1$	$0, 0$
b	$0, 1$	$2, 0$

How can we determine whether d is a best reply to the decision function $(\sigma_1(\theta^a), \sigma_1(\theta^b)) = (a,b)$? It all depends on the probability that player 2 (Colin) assigns to θ^a and θ^b. If Colin thinks that θ^a is more likely than θ^b, $\mathbb{P}_2(\theta^a) \geq \frac{1}{2}$, then d is a best reply, otherwise it is not. Suppose that $\mathbb{P}_2(\theta^a) < \frac{1}{2}$. In an "equilibrium" (σ_1, σ_2), we must have $\sigma_1(\theta^a) = a$, because a is dominant for information-type θ^a of Rowena.[13] In matrix θ^b Colin's payoff is independent of Rowena's action; thus, if $\mathbb{P}_2(\theta^a) < \frac{1}{2}$, whatever the value of $\sigma_1(\theta^b)$, Colin's best reply to the decision function σ_1 is c (indeed, c yields expected payoff $1 - \mathbb{P}_2(\theta^a) > \frac{1}{2}$, while d yields expected payoff $\mathbb{P}_2(\theta^a) < \frac{1}{2}$). Rowena's best response to c in matrix θ^b is a. Then, if $\mathbb{P}_2(\theta^a) < \frac{1}{2}$, the equilibrium profile should be $\sigma_1(\theta^a) = \sigma_1(\theta^b) = a$, $\sigma_2 = c$. ▲

The example shows that equilibrium analysis requires a specification of beliefs about other players' information-types. In general, *in order to define equilibrium behavior, it is necessary to enrich the mathematical structure $\hat{G}$ with the probabilistic beliefs of every player i about the residual uncertainty parameter θ_0 and his co-players' information-types* $\theta_{-i} = (\theta_j)_{j \in I \setminus \{i\}}$. We denote such beliefs by $p^i \in \Delta(\Theta_0 \times \Theta_{-i})$. It is then natural to ask: How is p^i determined? What does player j ($j \neq i$) know or believe about p^i? If p^i is just a subjective probability, then it does not seem very plausible to assume that j "knows" p^i, and it is even less plausible to assume that the belief profile $(p^i)_{i \in I}$ is "common knowledge." Then, even though by fixing a belief profile $(p^i)_{i \in I}$ we can determine whether each action $\sigma_i(\theta_i)$ ($\theta_i \in \Theta_i$) is a best reply to the profile of decision functions $\sigma_{-i} = (\sigma_j)_{j \neq i}$, it is not clear whether this best reply property is sufficient to ensure that there is no incentive to deviate. How can i be confident that j will follow his prescribed choice $\sigma_j(\theta_j)$ if i does not know p^j and therefore cannot know whether each $\sigma_j(\theta_j)$ ($\theta_j \in \Theta_j$) is in turn a best reply to σ_{-j}?

Actually, if we are to use a rigorous language, the word **"knowledge"** should only be used for a true belief determined by direct observation or

[13] Of course, "equilibrium" has yet to be defined in this context. The formal definition will follow.

logical deduction. As long as we assume that players cannot peer into other players' minds, we have to exclude that they can *know* anything about the beliefs of others. However, a player may know facts that, according to some psychological hypotheses, imply that the co-players' beliefs have some features, or maybe completely pin down the co-players' beliefs about unknown parameters, beliefs about such beliefs, and so on. If the psychological hypotheses are correct, this player's beliefs about the beliefs of others are also correct. In general, as we did in Section 8.3, we say that some event E is **transparent** if E is true and there is common belief of E. Then, the correctly posed question is whether the belief profile $(p^i)_{i\in I}$ is transparent. In the affirmative case, we can carry out a relatively simple equilibrium analysis with incomplete information.

Next, we consider a scenario where it is plausible to assume that $(p^i)_{i\in I}$ is transparent.[14] For each player/role $i \in I$ there is a large population of agents that can play in that role. Each one of them is characterized by an information-type θ_i (for instance, his preferences, his ability, his strength). Let $q_i \in \Delta(\Theta_i)$ denote the statistical distribution of θ_i within population i: $q_i(\theta_i)$ is the fraction of agents in the population i with information-type θ_i. Furthermore, assume that θ_0 is determined through a random experiment whose probabilities are fixed and transparent (for instance, assume that θ_0 depends on the color of a ball extracted from an urn whose composition has been publicly announced in front of all players at the same time); let the probability of θ_0 according to this experiment be $q_0(\theta_0)$. If players are randomly chosen from the corresponding populations, then the probability of meeting opponents characterized by the profile of information-types $\theta_{-i} = (\theta_j)_{j\neq i}$ is $\prod_{j\in I\setminus\{i\}} q_j(\theta_j)$.[15] If these statistical distributions are commonly known, then it is reasonable to assume that, for all i and θ_{-i}, $p^i(\theta_0, \theta_{-i}) = \prod_{j\in\{0\}\cup I\setminus\{i\}} q_j(\theta_j)$ and that the profile of "objective" beliefs $(p^i)_{i\in I}$ is transparent. Thus, we obtain a structure

$$BG^s = \left\langle I, \Theta_0, (\Theta_i, A_i, u_i, p^i)_{i\in I}\right\rangle$$

called **simple Bayesian game with type-independent beliefs** (in the following section we consider a more general definition of Bayesian game).

[14] This is what Harsanyi (1967-68) calls the "prior lottery model".

[15] As usual, we are assuming that every Θ_j is finite. The generalization to the infinite case, though conceptually trivial, requires the use of measure theoretic concepts.

Assuming that all the features of the interactive situation described by BG^s are transparent, we can meaningfully define as **equilibrium** a profile of decision functions $(\sigma_i)_{i\in I} \in \bigtimes_{i\in I} A_i^{\Theta_i}$ such that, for each i and θ_i,

$$\sigma_i(\theta_i) \in \arg\max_{a_i\in A_i} \sum_{\theta_0,\theta_{-i}} p^i(\theta_0,\theta_{-i})u_i(\theta_0,\theta_i,\theta_{-i},a_i,\sigma_{-i}(\theta_{-i})), \tag{8.4.1}$$

where $\sigma_{-i}(\theta_{-i}) = (\sigma_j(\theta_j))_{j\in I\setminus\{i\}}$.[16] Note, for any fixed conjecture σ_{-i} about the decision functions of the co-players, the expected payoff being maximized in eq. (8.4.1) is well defined because we have specified an exogenous belief p^i. Without such specification (for each player), we could not ascertain whether a profile of decision functions $(\sigma_i)_{i\in I}$ is an equilibrium.

Example 33. (*First Price Auctions with independent private values*) A given object is put on sale by means of an auction. The set of participant is $I = \{1, ..., n\}$, the monetary value of the object for participant i is θ_i. Therefore the model has private values. The set of possible valuations is $\Theta_i = [0, 1]$ and the profile $(\theta_1, ..., \theta_n)$ is uniformly distributed on $[0, 1]^n$. (This means that each bidder i believes that his competitors' valuations are mutually independent, uniform random variables.) The object is assigned to the player that makes the highest offer. In case of a draw, it is randomly assigned among the highest bidders. Whoever is given the object has to pay his bid (First Price rule). The resulting payoff function is:

$$u_i(\theta,a) = \begin{cases} (\theta_i - a_i)\frac{1}{|\arg\max_j a_j|}, & \text{if } a_i = \max_j a_j, \\ 0, & \text{if } a_i < \max_j a_j. \end{cases}$$

It turns out that this game has a symmetric equilibrium in which each player bids $\sigma_i(\theta_i) = \frac{n-1}{n}\theta_i$.
How can we guess that these functions form an equilibrium? Consider the following heuristic derivation. If $\theta_i = 0$ it does not make sense (indeed, it is weakly dominated) to offer more than zero. Now assume $\theta_i > 0$. If bidder i conjectures that each competitor bids according to the linear rule

[16] Furthermore, even if the structure BG^s were not transparent (perhaps because the statistical distributions are not commonly known), the definition of equilibrium given in the text could perhaps be motivated as a steady state of an adaptive process. This is certainly the case under private values.

$\sigma_j(\theta_j) = k\theta_j$, where $k \in (0,1)$ *is a coefficient that we have to determine*, then i believes that, if he bids a_i, the probability of winning the object is

$$\mathbb{P}\left([\forall j \neq i, \sigma_j(\widetilde{\theta}_j) < a_i]\right) = \mathbb{P}\left(\left[\forall j \neq i, \widetilde{\theta}_j < \frac{a_i}{k}\right]\right)$$

($\widetilde{\theta}_j$ denotes the random valuation of j from i's viewpoint; we can neglect ties because, for each competitor j, the probability of event $[\widetilde{a}_j = a_i]$ is zero). Since we assumed independent and uniform distributions of values, we have

$$\mathbb{P}\left(\left[\forall j \neq i, \widetilde{\theta}_j < \frac{a_i}{k}\right]\right) = \begin{cases} \left(\frac{a_i}{k}\right)^{n-1}, & \text{if } \frac{a_i}{k} < 1, \\ 1, & \text{if } \frac{a_i}{k} \geq 1. \end{cases}$$

Thus, if bidder i is rational he offers the smallest of the following numbers: $\arg\max_{0 \leq a_i < k} \left(\frac{a_i}{k}\right)^{n-1} (\theta_i - a_i) = \frac{n-1}{n}\theta_i$ and k,[17] that is

$$a_i = \min\left\{k, \frac{n-1}{n}\theta_i\right\}.$$

In a symmetric equilibrium each player has a correct conjecture about the bidding functions of the competitors and each player has the same (best reply) bidding function, therefore $k = \frac{n-1}{n}$. This implies

$$\min\left\{k, \frac{n-1}{n}\theta_i\right\} = \min\left\{\frac{n-1}{n}, \frac{n-1}{n}\theta_i\right\} = \frac{n-1}{n}\theta_i.$$

Therefore, the optimal bid when the symmetric conjecture of i about each j is $\sigma_j(\theta_j) = \frac{n-1}{n}\theta_j$ turns out to be precisely $a_i = \sigma_i(\theta_i) = \frac{n-1}{n}\theta_i$. ▲

The best method to compute the equilibria of a simple Bayesian game depends on the specific game. In Example 33 we followed a "conjecture and verify" method to compute the symmetric equilibrium. For finite games, one can adopt the method of constructing the so called "strategic form" of the game and compute its Nash equilibria. We describe this method formally in our analysis of general Bayesian games.

[17] To compute the maximizer in $[0,1)$ use the first order condition:

$$\frac{(n-1)}{k}\left(\frac{a_i}{k}\right)^{n-2}(\theta_i - a_i) - \left(\frac{a_i}{k}\right)^{n-1} = 0.$$

8.5 The General Case: Bayesian Games

In order to provide a more general definition of equilibrium in games with incomplete information it is necessary to consider the case in which the beliefs of a generic player i on $\Theta_0 \times \Theta_{-i}$ are not known to his co-players $-i$. This means that i is not characterized by his private information θ_i alone, and the belief p^i has to be separately specified. From j's point of view the pair $(\theta_i, p^i) \in \Theta_i \times \Delta(\Theta_0 \times \Theta_{-i})$ is unknown. Why should j care about p^i? The problem, from j's point of view, is that j's payoff depends on i's choice a_i which in turn depends on θ_i *and* p^i. Thus, in order to extend the equilibrium concept to this more general case it is necessary to specify j's beliefs about (θ_i, p^i).

According to the subjective formulation, known as the **Bayesian approach**, a *decision maker forms probabilistic beliefs over all the relevant variables that are unknown to him.* In our example j has probabilistic beliefs about (θ_i, p^i).

To simplify the exposition, we first focus on the two-person case with distributed knowledge of parameter θ (Θ_0 is a singleton): the only opponent of i is j and *viceversa.* The beliefs of j about θ_i and p^i are therefore a joint probability measure $q^j \in \Delta(\Theta_i \times \Delta(\Theta_j))$ from which the beliefs $p^j \in \Delta(\Theta_i)$ can be recovered by computing the marginal distribution on Θ_i. The beliefs q^j are called **second-order beliefs**, whereas the beliefs p^j are called **first-order beliefs**: second-order beliefs are beliefs about the primitive uncertainty and the opponent's first-order beliefs. Hence, first-order beliefs can be derived from second-order beliefs.

For some applications there is no need to go beyond second-order beliefs. Suppose, for example that $\Theta_i = \{\theta_i^0, \theta_i^1\}$, $i = 1, 2$. Then we may think of p^i as i's subjective probability of θ_j^1: $\Delta(\Theta_j)$ is isomorphic to $[0, 1]$. Suppose that it is transparent that the second-order beliefs of each player j have the form $q^j = p^j \times \overline{q}^j$ where $\overline{q}^j$ is given by the uniform distribution on $[0, 1]$, which is the set of possible values of the subjective probability $p^i(\theta_j^1)$ (recall: by "transparent" we mean that second-order beliefs do indeed have this form and furthermore it is common belief that they have this form). Then each first-order belief p^j also pins down the second-order belief $q^j = p^j \times \overline{q}^j$ and there is no need to explicitly consider beliefs of higher order.

However, since beliefs are subjective, *we have to allow for more* *general situations where second-order beliefs are not pinned down by first-order beliefs.* Since behavior may depend on second-order beliefs, it is necessary to introduce **third-order beliefs** $r^j \in \Delta(\Theta_i \times \Delta(\Theta_j) \times \Delta(\Theta_j \times \Delta(\Theta_i)))$,[18] from which the first and second-order beliefs can be recovered by marginalization. At this point, the formalism is already quite complex. Furthermore, there is no compelling reason to stop at third-order beliefs. How should we proceed? Is it possible to use a formal and compact representation of strategic interactions with incomplete information which is not too complex, but at the same time allows a representation of beliefs of arbitrarily high order and a meaningful definition of equilibrium? The answer is Yes.

8.5.1 Bayesian Games

The solution relies on the adoption of a more abstract and self-referential approach proposed by John Harsanyi [40].[19] Starting from the game with payoff uncertainty $\langle I, \Theta_0, (\Theta_i, A_i, u_i)_{i\in I}\rangle$, let us consider a richer structure with a set of "states of the world" Ω (assumed *finite* for the sake of simplicity). A **state of the world** ω characterizes each player's knowledge and "interactive beliefs." This can be formalized mathematically introducing functions $\tau_i : \Omega \to T_i$ and $\vartheta_i : T_i \to \Theta_i$ for each player i, a function $\vartheta_0 : \Omega \to \Theta_0$ and a probability measure $p_i \in \Delta(\Omega)$. It is also implicitly assumed that the situation of strategic interaction represented by these elements is transparent.

To grasp the meaning of the above functions we follow Harsanyi and present a *metaphor*.[20] Imagine an *ex ante* state in which all players are equally ignorant. Each player i is endowed with a *prior* (subjective) probability measure $p_i \in \Delta(\Omega)$. Before choosing an action, each player i receives a "signal" t_i about the state of the world. From signal t_i player i infers that $\theta_i = \vartheta_i(t_i)$ and $\omega \in \tau_i^{-1}(t_i) = \{\omega' : \tau_i(\omega') = t_i\} \subseteq \Omega$. Assume

[18] Clearly, the symbol r^j used here is not to be confused with the symbol denoting the best reply correspondence.

[19] This fundamental contribution led to the award to John Harsanyi of the 1994 Nobel Prize for Economics, jointly with John Nash and Reinhard Selten.

[20] This is what Harsanyi calls the "random vector model" of the Bayesian game. The only difference is that such model (not his general analysis of incomplete information) posits a common prior ($p_i = p \in \Delta(\Omega)$ for every $i \in I$), which represents an objective probability distribution.

for simplicity that p_i assigns positive probability to all signals that i could conceivably receive,[21] i.e.,

$$p_i\left(\tau_i^{-1}(t_i)\right) = \sum_{\omega:\tau_i(\omega)=t_i} p_i(\omega) > 0$$

for all $t_i \in T_i$.[22] Then, for each $t_i \in T_i$, the corresponding conditional probability measure $p_i\left(\cdot|t_i\right)$ is well defined:

$$\forall E \subseteq \Omega,\, p_i\left(E|t_i\right) = \frac{p_i\left(E \cap \tau_i^{-1}(t_i)\right)}{p_i\left(\tau_i^{-1}(t_i)\right)}.$$

With this, the beliefs of player i at state of the world ω are given by probability measure $p_i\left(\cdot|\tau_i(\omega)\right) \in \Delta(\Omega)$. Since τ_i and p_i are common knowledge, the function $\omega \mapsto p_i\left(\cdot|\tau_i(\omega)\right)$ is also common knowledge. It follows that a state of the world determines a player's *beliefs about the parameter* θ *and about other players' beliefs about it.* To verify this claim, let us focus on the case of *two players and distributed knowledge of* Θ (this is done only to simplify the notation) and let us derive the beliefs of i about t_j given signal t_i:

$$\forall t_j \in T_j,\, p_i\left(t_j|t_i\right) = p_i\left(\tau_j^{-1}(t_j)|t_i\right),$$

where $\tau_j^{-1}(t_j) = \{\omega : \tau_j\left(\omega\right) = t_j\}$. Next we derive the **first-order beliefs** about θ (since i knows θ_i, his beliefs about θ are determined by his beliefs about Θ_j):

$$\forall \theta_j \in \Theta_j,\, p_i^1\left(\theta_j|t_i\right) = p_i\left(\vartheta_j^{-1}(\theta_j)|t_i\right),$$

where $\vartheta_j^{-1}(\theta_j) = \{t_j : \vartheta_j(t_j) = \theta_j\}$. The functions $t_j \mapsto p_j^1\left(\cdot|t_j\right) \in \Delta(\Theta_i)$ ($j \in \{1,2\}$) are also common knowledge. Hence, we can derive the **second-order beliefs** about θ, that is, the joint belief of a player about θ and about the opponent's first-order beliefs about θ:

$$\forall(\overline{\theta}_j, \overline{p}_j^1) \in \Theta_j \times \Delta(\Theta_i),\, p_i^2\left(\overline{\theta}_j, \overline{p}_j^1|t_i\right) = \sum_{t_j:\vartheta_j(t_j)=\overline{\theta}_j, p_j^1(\cdot|t_j)=\overline{p}_j^1} p_i\left(t_j|t_i\right).$$

[21] It can be shown that this comes without substantial loss of generality provided that we allow the subjective priors of different players to be different.

[22] We denote by $p_i\left(\cdot\right)$ and $p_i\left(\cdot|\cdot\right)$, respectively, the prior and conditional probabilities of events. For instance, $p_i\left(t_{-i}|t_i\right)$ is the probability that event $\{\omega : \tau_{-i}(\omega) = t_{-i}\}$ occurs conditional on the event $\{\omega : \tau_i(\omega) = t_i\}$.

[It can be verified that

$$\forall \overline{\theta}_j \in \Theta_j,\, p_i^1\left(\overline{\theta}_j|t_i\right) = \sum_{\overline{p}_j^1 : \exists t_j \in T_j, p_j^1(\cdot|t_j) = \overline{p}_j^1} p_i^2\left(\overline{\theta}_j, \overline{p}_j^1|t_i\right),$$

in words, $p_i^1(\cdot|t_i)$ is the marginal on Θ_j of the joint distribution $p_i^2(\cdot|t_i) \in \Delta(\Theta_j \times \Delta(\Theta_i))$.]

It should be clear by now that it is possible to iterate the argument and compute the functions that assign to each type the corresponding third-order beliefs about θ, fourth-order beliefs about θ, and so on. To sum up, we can conclude that the information and beliefs of all orders of player i about θ are determined by t_i according to the function $t_i \mapsto (\vartheta_i(t_i), p_i^1(\cdot|t_i), p_i^2(\cdot|t_i), ...)$. Signal t_i is called the **type**[23] of player i. Sequence $(p_i^1(\cdot|t_i), p_i^2(\cdot|t_i), ...)$ is called the **hierarchy of beliefs** of type t_i. The information and beliefs of each player i in a given state of the world ω are those corresponding to the type $t_i = \tau_i(\omega)$. The **information-type** of player i, $\theta_i = \vartheta_i(t_i)$, is just one component of his overall type t_i, which also specifies i's beliefs about all the relevant exogenous variables/parameters, θ_{-i}, p_{-i}^1, p_{-i}^2, etc.

Definition 38. *A (finite)* Bayesian Game *is a structure*

$$BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i\in I}\rangle$$

where each set in BG is finite; for every player $i \in I$, $p_i \in \Delta(\Omega)$, $\vartheta_0 : \Omega \to \Theta_0$, $\tau_i : \Omega \to T_i$, $\vartheta_i : T_i \to \Theta_i$, $p_i(t_i) = p_i\left(\tau_i^{-1}(t_i)\right) > 0$ for all $t_i \in T_i$,[24] *and $u_i : \Theta \times A \to \mathbb{R}$.*

The analysis of solution concepts for Bayesian games will clarify that *only the beliefs $p_i(\cdot|t_i)$ ($i \in I$, $t_i \in T_i$) really matter.* The priors p_i are just a convenient mathematical tool to represent the beliefs of each type.

In what follows, we will often use the phrase "type t_i chooses a_i" to mean that if player i *were* of type t_i *he would choose* action a_i.

Transparent Restrictions on First-Order Exogenous Beliefs

It is worth noting that Harsanyi's self-referential and implicit representation of hierarchical exogenous beliefs allows—in principle—to model the transparent restrictions on exogenous (first-order) beliefs

[23] The expression "type *à la* Harsanyi" is also used.
[24] This condition implies that $\tau_i : \Omega \to T_i$ is *onto*.

considered in Section 8.3. Fix a game with payoff uncertainty $\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i)_{i\in I}\rangle$. To obtain a Bayesian game, we append to $\hat{G}$ a belief structure $\langle \Omega, \vartheta_0, (T_i, \tau_i, \vartheta_i, p_i)_{i\in I}\rangle$ as per Definition 38. Fix a profile $\bar{\Delta} = \left(\bar{\Delta}_{i,\theta_i}\right)_{i\in I,\theta_i\in\Theta_i}$, where $\bar{\Delta}_{i,\theta_i} \subseteq \Delta\left(\Theta_0 \times \Theta_{-i}\right)$ for each i and θ_i. With a slight abuse of language, we say that $\bar{\Delta}$ is transparent to mean that it is true and commonly believed that, if the information type of player i is θ_i, then his first-order exogenous beliefs belong to set $\bar{\Delta}_{i,\theta_i}$.[25] Belief structure $\langle \Omega, \vartheta_0, (T_i, \tau_i, \vartheta_i, p_i)_{i\in I}\rangle$ yields the transparency of $\bar{\Delta}$ if the implied first-order beliefs satisfy these restrictions, that is, for every $i \in I$, $\theta_i \in \Theta_i$ and $t_i \in \vartheta_i^{-1}(\theta_i)$, we have $p^1(\cdot|t_i) \in \bar{\Delta}_{\theta_i}$. If, additionally, there is some $t_i \in \vartheta_i^{-1}(\theta_i)$ such that $p^1(\cdot|t_i) \in \bar{\Delta}_{\theta_i}$ for every $i \in I$ and $\theta_i \in \Theta_i$, then the given belief structure *exactly* captures the transparency of $\bar{\Delta}$.[26] Our analysis of the connection between Harsanyi's approach and the Directed Rationalizability approach in Section 8.5.5 relies on this observation.

8.5.2 Bayesian Equilibria

In general, players' choices depend not only on their private information on θ, but also more generally on their *types*. In equilibrium, each player has a correct conjecture about the dependence of each opponent's choice on his type, and maximizes his expected payoff given his type:

Definition 39. *A* ***Bayesian equilibrium*** *is a profile of decision functions* $(\sigma_i : T_i \to A_i)_{i\in I}$ *such that*

$$\forall i \in I, \forall t_i \in T_i,\ \sigma_i(t_i) \in \arg\max_{a_i\in A_i} \mathbb{E}_{p_i,\sigma_{-i}}(u_{i,\vartheta_i(t_i),a_i}|t_i),$$

where

$$\begin{aligned}&\mathbb{E}_{p_i,\sigma_{-i}}(u_{i,\vartheta_i(t_i),a_i}|t_i)\\ =\ &\sum_{\omega\in\Omega} p(\omega|t_i)u_i(\vartheta_0(\omega), \vartheta_i(t_i), \vartheta_{-i}(\tau_{-i}(\omega)), a_i, \sigma_{-i}(\tau_{-i}(\omega)))\end{aligned}$$

(for every ω and t_{-i}, $\tau_{-i}(\omega) = (\tau_j(\omega))_{j\neq i}$, $\vartheta_{-i}(t_{-i}) = (\vartheta_j(t_j))_{j\neq i}$, $\sigma_{-i}(t_{-i}) = (\sigma_j(t_j))_{j\neq i}$).

---

[25] We call this a "slight abuse of language" because we defined transparency as a property of *events*, while profile $\bar{\Delta}$ is not in itself an event. However, $\bar{\Delta}$ corresponds to an event concerning players' exogenous first-order beliefs, as explained in the text.

[26] If at least one set $\bar{\Delta}_{\theta_i}$ is infinite, then the beliefs structure has to be infinite as well, which requires some technical conditions.

Note that the expected payoff of type t_i of player i choosing action a_i can be re-written as

$$\mathbb{E}_{p_i,\sigma_{-i}}(u_{i,\vartheta_i(t_i),a_i}|t_i) = \sum_{(\theta_0,t_{-i})\in\Theta_0\times T_{-i}} p(\theta_0,t_{-i}|t_i)u_i(\theta_0,\vartheta_i(t_i),\vartheta_{-i}(t_{-i}),a_i,\sigma_{-i}(t_{-i})),$$

where $p\left(\theta_0,t_{-i}|t_i\right)$ is the probability that event

$$\{\omega : \vartheta_0\left(\omega\right) = \theta_0,\ \tau_{-i}(\omega) = t_{-i}\}$$

occurs conditional on the event $\{\omega : \tau_i(\omega) = t_i\}$. Again, as in we did in Section 8.4, we stress that without the specification of the belief structure (that is, of a specific exogenous belief for each Harsanyi type t_i of each player i), we cannot determine the expected payoff function maximized for each type t_i for some given conjecture σ_{-i} about co-players' decision functions. Thus, such specification is necessary to ascertain whether a profile of decision functions $(\sigma_i : T_i \to A_i)_{i\in I}$ is an equilibrium.

Example 34. We illustrate the concepts of Bayesian game and equilibrium elaborating on game $\hat{G}^2$ of Example 32. Suppose that Rowena does not know Colin's first-order beliefs. From her point of view such beliefs can assign either probability $\frac{1}{3}$ or $\frac{3}{4}$ to θ^a. The two possibilities are regarded as equally likely by Rowena and all this is common knowledge. This situation can be represented by the following Bayesian game:

$$\Omega = \{\alpha,\beta,\gamma,\delta\},$$

$$\Theta_1 \cong \Theta = \left\{\theta^a,\theta^b\right\},\ T_1 = \left\{t_1^a,t_1^b\right\},$$

$$\tau_1(\alpha) = \tau_1(\beta) = t_1^a,\ \tau_1(\gamma) = \tau_1(\delta) = t_1^b,\ \vartheta_1(t_1^a) = \theta^a,\ \vartheta_1(t_1^b) = \theta^b,$$

$$\forall\omega\in\Omega,\ p_1(\omega) = \frac{1}{4},$$

$$\Theta_2 = \left\{\hat{\theta}_2\right\},\ T_2 = \left\{t_2',t_2''\right\},$$

$$\tau_2(\alpha) = \tau_2(\gamma) = t_2',\ \tau_2(\beta) = \tau_2(\delta) = t_2'',$$

$$p_2(\alpha) = \frac{3}{8},\ p_2(\beta) = \frac{1}{6},\ p_2(\gamma) = \frac{1}{8},\ p_2(\delta) = \frac{1}{3},$$

and where the functions u_i are as in $\hat{G}^2$. The probabilistic structure is described in the table below:

	t_2'	t_2''
θ^a, t_1^a	$\alpha, \frac{1}{4}, \frac{3}{8}$	$\beta, \frac{1}{4}, \frac{1}{6}$
θ^b, t_1^b	$\gamma, \frac{1}{4}, \frac{1}{8}$	$\delta, \frac{1}{4}, \frac{1}{3}$

To verify that this Bayesian game represents the situation outlined above we compute the following:

$$\begin{aligned} p_2^1\left(\theta^a|t_2'\right) &= p_2\left(t_1^a|t_2'\right) = \frac{p_2(\alpha)}{p_2(\alpha)+p_2(\gamma)} = \frac{3/8}{3/8+1/8} = \frac{3}{4}, \\ p_2^1\left(\theta^a|t_2''\right) &= p_2\left(t_1^a|t_2''\right) = \frac{p_2(\beta)}{p_2(\beta)+p_2(\delta)} = \frac{1/6}{1/6+1/3} = \frac{1}{3}, \\ p_1\left(t_2'|t_1^a\right) &= p_1\left(t_2'|t_1^b\right) = \frac{1}{2}. \end{aligned}$$

This means that in every state of the world Rowena believes that the two events *[Colin assigns probability* $\frac{3}{4}$ *to* θ^a*]* and *[Colin assigns probability* $\frac{1}{3}$ *to* θ^a*]* are equally likely. We now derive the Bayesian equilibria. Let $\sigma = (\sigma_1, \sigma_2)$ be an equilibrium. Since a is dominant when Rowena knows that $\theta = \theta^a$, $\sigma_1(t_1^a) = a$. Hence, the equilibrium expected payoff accruing to type t_2' if he chooses c is

$$p_2\left(t_1^a|t_2'\right) u_2(\theta^a, a, c) + p_2\left(t_1^b|t_2'\right) u_2(\theta^b, \sigma_1(t_1^b), c) = \frac{3}{4} \times 0 + \frac{1}{4} \times 1 = \frac{1}{4};$$

the equilibrium expected payoff for t_2' if he chooses d is

$$p_2\left(t_1^a|t_2'\right) u_2(\theta^a, a, d) + p_2\left(t_1^b|t_2'\right) u_2(\theta^b, \sigma_1(t_1^b), d) = \frac{3}{4} \times 1 + \frac{1}{4} \times 0 = \frac{3}{4}.$$

It follows that in equilibrium $\sigma_2(t_2') = d$. The expected payoffs for the two actions of type t_2'' are

$$\begin{aligned} p_2\left(t_1^a|t_2''\right) u_2(\theta^a, a, c) + p_2\left(t_1^b|t_2''\right) u_2(\theta^b, \sigma_1(t_1^b), c) &= \frac{1}{3} \times 0 + \frac{2}{3} \times 1 = \frac{2}{3}, \\ p_2\left(t_1^a|t_2''\right) u_2(\theta^a, a, d) + p_2\left(t_1^b|t_2''\right) u_2(\theta^b, \sigma_1(t_1^b), d) &= \frac{1}{3} \times 1 + \frac{2}{3} \times 0 = \frac{1}{3}. \end{aligned}$$

Since the maximizing choice for type t_2'' is c, $\sigma_2(t_2'') = c$. We can now determine the equilibrium choice for type t_1^b:

$$\begin{aligned}
p_1\left(t_2'|t_1^b\right) u_1(\theta^b, a, d) + p_1\left(t_2''|t_1^b\right) u_1(\theta^b, a, c) &= \frac{1}{2}\times 0 + \frac{1}{2}\times 1 = \frac{1}{2}, \\
p_1\left(t_2'|t_1^b\right) u_1(\theta^b, b, d) + p_1\left(t_2''|t_1^b\right) u_1(\theta^b, b, c) &= \frac{1}{2}\times 2 + \frac{1}{2}\times 0 = 1.
\end{aligned}$$

Therefore, $\sigma_1(t_1^b) = b$. ▲

8.5.3 Bayesian Equilibrium and Nash Equilibrium

The concept of Bayesian equilibrium for a game of incomplete information BG can be restated in an equivalent way as a Nash equilibrium of two games with complete information that are associated with the original game: the *ex ante* strategic form and the *interim* strategic form.

The *ex ante* strategic form refers to the metaphor that was previously introduced to explain the elements of the Bayesian game: if there is an *ex ante* stage in which all players are "ignorant," then at such stage each player i can make a contingent plan of action, or strategy, $\sigma_i : T_i \to A_i$, that specifies the action to take for every possible signal t_i that player i might receive. If i believes that the other players follow the strategy profile σ_{-i}, the expected payoff of strategy σ_i is

$$U_i(\sigma_i, \sigma_{-i}) = \mathbb{E}_{p_i,\sigma_i,\sigma_{-i}}(u_i), \tag{8.5.1}$$

where (in a finite game)

$$\begin{aligned}
&\mathbb{E}_{p_i,\sigma_i,\sigma_{-i}}(u_i) \\
= &\sum_{\omega\in\Omega} p_i(\omega) u_i(\vartheta_0(\omega), \vartheta_i(\tau_i(\omega)), \vartheta_{-i}(\tau_{-i}(\omega)), \sigma_i(\tau_i(\omega)), \sigma_{-i}(\tau_{-i}(\omega))) \\
= &\sum_{t_i\in T_i} p_i(t_i) \sum_{(\theta_0,t_{-i})\in\Theta_0\times T_{-i}} p_i(\theta_0, t_{-i}|t_i)\, u_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), \sigma_i(t_i), \sigma_{-i}(t_{-i})).
\end{aligned}$$

Let $\Sigma_i = A_i^{T_i}$ (recall, this is the set of functions with domain T_i and codomain A_i). The **ex ante strategic form** of BG is the static game

$$\mathcal{AS}(BG) = \langle I, (\Sigma_i, U_i)_{i\in I}\rangle,$$

where U_i is defined by (8.5.1). The reader should be able to prove the following result by inspection of the definitions and using the assumption that each type t_i is assigned positive probability by the prior p_i.

Remark 30. *A profile* $(\sigma_i)_{i\in I}$ *is a Bayesian equilibrium of BG if and only if it is a Nash equilibrium of the game* $\mathcal{AS}(BG)$.

The *interim* strategic form is based on a different metaphor. Assume that for each role i in the game there is a set of potential players T_i. Assume for notational simplicity that $T_i \cap T_j = \emptyset$ for each $i, j \in I$, $i \neq j$ (this is just a matter of labelling). A potential player t_i is characterized by the payoff function $u_i(\vartheta_i(t_i), \cdot, \cdot, \cdot) : \Theta_0 \times \Theta_{-i} \times A \to \mathbb{R}$ and the beliefs $p_i(\cdot, \cdot|t_i) \in \Delta(\Theta_0 \times T_{-i})$. In the event that t_i is selected to play the game in the role of agent i, he will assign probability $p_i(\theta_0, t_{-i}|t_i)$ to the event that the residual uncertainty is θ_0 and that he is facing exactly the profile of "opponents" $t_{-i} = (t_j)_{j\neq i}$. The set of actions available to t_i is A_i, that is, $A_{t_i} = A_i$ ($i \in I$, $t_i \in T_i$). If each potential player $t_j \in T_j$ chooses the action $a_{t_j} \in A_{t_j}$, t_i's expected payoff is computed as follows:

$$\overline{u}_{t_i}((a_{t_j})_{j\in I, t_j\in T_j}) = \mathbb{E}_{p_i, \overrightarrow{a}_{-i}}\left(u_{i,a_{t_i}}|t_i\right), \tag{8.5.2}$$

where $\overrightarrow{a}_{-i} = \left(a_{t_j}\right)_{t_j\in T_j, j\neq i}$ and (in a finite game)

$$\mathbb{E}_{p_i, \overrightarrow{a}_{-i}}\left(u_{i,a_{t_i}}|t_i\right) = \sum_{(\theta_0, t_{-i})\in\Theta_0\times T_{-i}} p_i(\theta_0, t_{-i}|t_i)\, u_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_{t_i}, (a_{t_j})_{j\neq i}).$$

The **interim strategic form** of BG is the following static game with $\sum_{i\in I}|T_i|$ players:

$$\mathcal{IS}(BG) = \left\langle \bigcup_{i\in I} T_i, (A_{t_i}, \overline{u}_{t_i})_{i\in I, t_i\in T_i} \right\rangle,$$

where $\overline{u}_{t_i}$ is defined by (8.5.2). The reader should be able to prove the following result by inspection of the definitions:

Remark 31. *A profile* $(\sigma_i^*)_{i\in I}$ *is a Bayesian equilibrium of BG if and only if the corresponding profile* $(a_{t_i}^*)_{i\in I, t_i\in T_i}$ *such that* $a_{t_i}^* = \sigma_i^*(t_i)$ *(*$i \in I$, $t_i \in T_i$*) is a Nash equilibrium of the game* $\mathcal{IS}(BG)$.

Example 35. Consider the simple Bayesian game of Example 32, where $T_1 = \Theta_1 \cong \Theta = \{\theta^a, \theta^b\} = \Omega$, $\tau_1(\theta) = \theta$ for each θ (player 1 knows θ), $\tau_2(\theta^a) = \tau_2(\theta^b)$ (player 2 does not know θ). To ease notation, assume $p_1 = p_2$ (this does not affect the strategic analysis of the game), and let

$p = p_i(\theta^a)$. The ex ante strategic form of the game is a 4×2 matrix game ($a.a$ means "a for each type," $a.b$ means "a if θ^a and b otherwise," $b.a$ means the opposite, etc.).

$\sigma_1 \backslash \sigma_2$	c	d
$a.a$	$3p+1, 1-p$	$2p, p$
$a.b$	$4p, 1-p$	$2, p$
$b.a$	$2p+1, 1$	$p, 0$
$b.b$	$3p, 1$	$2-p, 0$

If $p < 1/2$, the unique equilibrium of the matrix game is $(\sigma_1, \sigma_2) = (a.a, c)$, which can be obtained by iterated dominance: first note that $a.a$ dominates $b.a$, and $a.b$ dominates $b.b$, then the computation unravels. The interim strategic form is formally a three-person game, where the prior probability $p = p_2$ matters only to compute the payoffs of (the unique type of) player 2. By Remark 31, if $p < 1/2$, this three-person game has a unique equilibrium where both players/types θ^a and θ^b choose a, and player 2 chooses c. Also in this case, the equilibrium can be obtained by iterated dominance: one starts noticing that b is dominated by a for type θ^a, then the computation unravels (see the analysis in Example 32). ▲

In the previous example, iterated dominance in the ex ante strategic form gives the same result as iterated dominance in the interim strategic form. This is not always true. In the Appendix of this chapter, we analyze in detail the relationships between iterated dominance in the ex ante and interim strategic form and different notions of rationalizability for Bayesian games.

8.5.4 Bayesian Equilibrium and Correlated Equilibrium

Even though it may seem counter-intuitive, the definition of Bayesian game admits as a special case that there is common knowledge of the payoff functions and yet there are multiple types for some players. This is the case when there are functions $\bar{u}_i \in \mathbb{R}^A$ ($i \in I$) such that $u_i(\theta, \cdot) = \bar{u}_i$ for each θ and i.[27] In other words, *we can define a Bayesian game with many*

[27] When the sets Θ_0 and Θ_i are all singletons we obtain a special case of this special case!

types even starting from a game with complete information! An even more special case can be analyzed: the one in which, besides having complete information, all players are characterized by the same prior belief ($p_i = p$ for each $i \in I$), which we refer to as **"common prior."** What can we say about Bayesian equilibria of a game with complete information?

To make the exposition more precise, we have to introduce some terminology and notation: given a complete-information game $G = \langle I, (A_i, \bar{u}_i)_{i\in I}\rangle$ we call **Bayesian elaboration** of G any Bayesian game

$$BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i\in I}\rangle$$

such that, Θ_0 and every Θ_i ($i \in I$) contain a single element, say $\bar{\theta}_0$ and $\bar{\theta}_i$ ($i \in I$). Thus, for every profile $a \in A$, we can write $u_i(\bar{\theta}, a) = \bar{u}_i(a)$. The following remark follows directly from the definitions:

Remark 32. *Let BG be a Bayesian elaboration of a finite game with complete information G and suppose that there is a common prior. Then, every Bayesian equilibrium of BG corresponds to a correlated equilibrium of G.*[28]

8.5.5 Bayesian Equilibrium and Rationalizability

The equivalence between the correlated equilibria of a game with complete information G and the Bayesian equilibria of its Bayesian elaboration BG holds under the assumption that players share a common prior. Removing this common prior assumption, one obtains the notion of subjective correlated equilibrium: a **subjective correlated equilibrium** of a complete-information game G is a Bayesian equilibrium of a Bayesian elaboration of G.

Theorem 27. *An action profile $(a_i)_{i\in I}$ of the finite complete-information game G is rationalizable if and only if it is selected in a subjective correlated equilibrium, that is, if and only if there are a Bayesian elaboration BG of G, an equilibrium $(\sigma_i)_{i\in I}$ of BG, and a state of the world ω in BG such that $a_i = \sigma_i(\tau_i(\omega))$ for every $i \in I$.*

Proof. (**If**) Let σ be an equilibrium of a Bayesian elaboration of $G = \left\langle I, (A_i, u_i)_{i\in I}\right\rangle$. To show that every image of σ is a profile of

[28] In a way, one could say that Harsanyi [40] had implicitly defined the correlated equilibrium concept before Aumann [5], but without being aware of it!

rationalizable actions, define the sets $C_i = \sigma_i(T_i) = \sigma_i(\tau_i(\Omega))$ $(i \in I)$. We are going to verify that the cross-product $C = \times_{i \in I} C_i$ has the best reply property $C \subseteq \rho(C)$, which implies that each action profile of $\sigma(\tau(\Omega)) \subseteq C$ is rationalizable in G (Theorem 3).[29] For any i, $a_i^* \in C_i = \sigma_i(T_i)$ and $t_i \in \sigma_i^{-1}(a_i^*)$, let $\mu^{t_i} = p_i(\cdot|t_i) \circ (\tau_{-i} \circ \sigma_{-i})^{-1} \in \Delta(A_{-i})$ denote the conjecture induced by the beliefs on Ω of type t_i given σ_{-i}. More explicitly,

$$\forall a_{-i} \in A_{-i},\ \mu^{t_i}(a_{-i}) = p_i\left((\sigma_{-i} \circ \tau_{-i})^{-1}(a_{-i})\,|t_i\right) = \sum_{\omega:\sigma_{-i}(\tau_{-i}(\omega))=a_{-i}} p_i(\omega|t_i).$$

By construction, $u_i(a_i, \mu^{t_i}) = \mathbb{E}_{p_i,\sigma_{-i}}(u_{i,a_i}|t_i)$ for all $a_i \in A_i$ (recall that u_i is independent of θ). Since σ is an equilibrium,

$$a_i^* = \sigma_i(t_i) \in \arg\max_{a_i \in A_i} \mathbb{E}_{p_i,\sigma_{-i}}(u_{i,a_i}|t_i) = \arg\max_{a_i \in A_i} u_i\left(a_i, \mu^{t_i}\right).$$

Furthermore, again by construction, $\mathrm{supp}\mu^{t_i} \subseteq \sigma_{-i}(T_{-i}) = C_{-i}$. Thus, $C_i \subseteq r_i(\Delta(C_{-i}))$ for each $i \in I$, that is, $C \subseteq \rho(C)$.

(**Only if**) It is sufficient to show that there exists a subjective canonical correlated equilibrium in which every rationalizable profile is played in some state of the world. Let $C = \rho^\infty(A)$ be the set of rationalizable profiles. Then $C = \rho(C)$ (Theorem 2), and for every $i \in I$ and $a_i \in C_i$ there exists a conjecture $\mu^{a_i} \in \Delta(C_{-i})$ such that $a_i \in r_i(\mu^{a_i})$. We construct the Bayesian elaboration of G as follows. Types are "copies" of rationalizable actions (they can be interpreted as "suggested actions"); if the type of i is a copy of action a_i, then the belief of i is derived from the justifying conjecture μ^{a_i}. Formally, $\Omega = C = T$; for every $i \in I$, $T_i = C_i$ and τ_i is the identity function, $\tau_i = \mathrm{Id}_{C_i}$; for every rationalizable action/type $t_i \in C_i = T_i$, let the interim belief be $p_i(\cdot|t_i) = \mu^{t_i}$;[30] define the prior belief of i as follows: for every $t = (t_j)_{j \in I} \in C = T = \Omega$, $p_i(t) = \frac{1}{|C_i|} p_i(t_{-i}|t_i)$. Then, the profile of identity functions (for every $i \in I$ and for every $a_i \in T_i = A_i$, $\sigma_i(a_i) = a_i$) is an equilibrium of this Bayesian elaboration.[31]

■

[29] Note that, for each i, $C_i = \mathrm{proj}_{A_i} \sigma(\tau(\Omega))$ because τ_i is onto; but $\sigma(\tau(\Omega))$ may be not Cartesian. Hence, in general, $\sigma(\tau(\Omega)) \subseteq C$.

[30] Recall that $\mu^{t_i} \in \Delta(C_{-i}) = \Delta(T_{-i})$.

[31] The "type" $t_i = a_i$ is not to be interpreted as the action that i necessarily has to play, but rather as a specification of i's belief about the opponents and of the best reply to such belief, a best reply that i *freely chooses* in equilibrium.

This result can be generalized. Fix a game with payoff uncertainty

$$\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i : \Theta \times A \to \mathbb{R})_{i\in I}\rangle ;$$

a Bayesian game based on $\hat{G}$ is a Bayesian game obtained by appending a state space, signal functions etc. to $\hat{G}$, that is,

$$BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i\in I}\rangle .$$

Using arguments that generalize the proof of Theorem 27, one can prove the following result.

Theorem 28. *A profile of information-types and actions $(\theta_i, a_i)_{i\in I}$ is rationalizable in the game with payoff uncertainty $\hat{G}$ if and only if there is a Bayesian game BG based on $\hat{G}$, an equilibrium σ of BG and a state of the world ω in BG such that $(\theta_i, a_i)_{i\in I} = (\vartheta_i(\tau_i(\omega)), \sigma_i(\tau_i(\omega)))_{i\in I}$.*

These results show that, without restrictions on players' exogenous beliefs, the Bayesian equilibrium assumption that players best respond to correct conjectures about their opponents' decision functions has no behavioral implications beyond those that can be derived from rationality and common belief in rationality.[32] In a sense, *rationalizability gives the "robust" implications of Bayesian equilibrium analysis.*

Suppose that the modeler is confident that contextual considerations make some restrictions on exogenous first-order beliefs transparent. Then the robust implications of Bayesian equilibrium are captured by Directed Rationalizability. More formally, consider a Bayesian game BG obtained by appending to game with payoff uncertainty $\hat{G}$ a belief structure $\langle \Omega, \vartheta_0, (T_i, \tau_i, \vartheta_i, p_i)_{i\in I}\rangle$. Say that BG yields the restrictions on exogenous beliefs Δ if, for all $i \in I$, $\theta_i \in \Theta_i$, and $t_i \in \vartheta_i^{-1}(\theta_i)$, there is some $\mu^i \in \Delta_{i,\theta_i}$ such that $p_i^1(\cdot|t_i) = \mathrm{marg}_{\Theta_0\times\Theta_{-i}}\mu^i$ (see Section 8.5.1). With this, we obtain a variation of the previous result.

Theorem 29. *Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of restrictions on* exogenous *beliefs. A profile of information-types and actions $(\theta_i, a_i)_{i\in I}$ is Δ-rationalizable in the game with payoff uncertainty $\hat{G}$ if and only if there is a Bayesian game BG based on $\hat{G}$ that yields the restrictions on exogenous beliefs Δ, an equilibrium σ of BG, and a state of the world ω in BG such that $(\theta_i, a_i)_{i\in I} = (\vartheta_i(\tau_i(\omega)), \sigma_i(\tau_i(\omega)))_{i\in I}$.*

[32] See Brandenburger and Dekel [24], and Battigalli and Siniscalchi [14].

8.6 Incomplete Information and Asymmetric Information

Recall from Chapter 1 that is important to distinguish between the mathematical structure used to represent a real world interactive decision situation[33] and the situation itself. With this in mind, let us go back to Harsanyi's metaphor, presented in Section 8.5.1: the information players have, the residual uncertainty and players' payoffs depend on a random variable, or chance move.[34] Let $\omega \in \Omega$ denote a typical realization of this random variable and assume that this realization occurs before players make their choices. For example, ω may be the order of cards in a deck before some of these cards are distributed to the players. The payoff of player i is determined by a function $\hat{u}_i : \Omega \times A \to \mathbb{R}$, that is, how payoffs depend on actions is determined by the initial chance move. Let $p_i \in \Delta(\Omega)$ denote the subjective probability measure of i over the possible realizations. Each player first receives a signal regarding the initial chance move, then chooses an action $a_i \in A_i$ simultaneously to the other players. Let $\tau_i : \Omega \to T_i$ denote player i's signal function. The prior belief and signal function of each player i are such that $p_i\left(\tau_i^{-1}(t_i)\right) > 0$ for all $t_i \in T_i$. We call such interactive situation a **game with asymmetric information** about an initial move by chance.

In this context, it may make sense to suppose the subjective priors p_i coincide with an objective probability measure $p \in \Delta(\Omega)$ and that this is commonly believed. More generally, even if the priors do not coincide, it is assumed that the profile of subjective measures $(p_i)_{i\in I}$ is transparent.[35]

An **equilibrium** of such game with asymmetric information is a strategy profile $(\sigma_i : T_i \to A_i)_{i\in I}$ such that the strategy of each player is a best reply to the strategy profile of the other players. More explicitly, define the payoff function in strategic form

$$U_i((\sigma_i)_{i\in I}) = \sum_{\omega\in\Omega} p_i(\omega)\hat{u}_i(\omega, (\sigma_j(\tau_j(\omega)))_{j\in I}).$$

An equilibrium of a given game of asymmetric information is a Nash

[33]Or a class of situations of the same kind.

[34]As in Section 8.5.4, it is also interesting to consider the possibility that payoffs do not depend on the random variable.

[35]In line with our definition of "transparency" in Section 8.3, this means that it is true and commonly believed that, for each $i \in I$, the subjective prior belief of i is p_i.

equilibrium of the corresponding game in strategic form $\langle I, (\Sigma_i, U_i)_{i\in I}\rangle$, where $\Sigma_i = A_i^{T_i}$ for each $i \in I$.

At this point, one may ask: How is this different from a game of incomplete information, modeled as a Bayesian game? If we just look at the mathematical formulation without considering how it is supposed to map to the real world, the two situations may seem indistinguishable. At the interim stage of the asymmetric information game in which all players have received their private signals about the chance move, the interactive situation is indeed very similar to one with incomplete information: the way payoffs depend on actions is not commonly known, and each player has private information about such dependence, along with probabilistic beliefs regarding the private information and (implicitly) the beliefs of other players.

We can establish an explicit formal relationship between the two models, deriving a Bayesian game as per Definition 38 from the mathematical structure representing a game with asymmetric information about the realization of an initial chance move. For the sake of simplicity, we restrict our attention to the case in which the joint signal function $\tau = (\tau_i)_{i\in I} : \Omega \to \bigtimes_{i\in I} T_i$ is one-to-one (injective), that is, $\omega' \neq \omega''$ implies that there is at least one player i for whom $\tau_i(\omega') \neq \tau_i(\omega'')$. Such simplification rules out any residual uncertainty about payoffs. With this, we formally obtain a Bayesian game by defining $\Theta_i = T_i$, letting ϑ_i be the identity function on T_i, so that $\tau(\Omega) \subseteq \Theta = \bigtimes_{i\in I} \Theta_i$, and defining state-dependent payoffs as follows:

$$\forall \theta \in \tau(\Omega),\ u_i(\theta, a) = \hat{u}_i(\tau^{-1}(\theta), a). \tag{8.6.1}$$

Note that this is enough, because it is common belief that realizations θ outside $\tau(\Omega)$ are impossible, therefore, the definition of u_i on $\Theta\backslash\tau(\Omega)$ is irrelevant for expected payoff computations. Keeping in mind Remark 30, it is easy to verify that *a profile of functions* $(\sigma_i)_{i\in I}$ *is a Bayesian equilibrium of this Bayesian game if and only if it is an equilibrium of the asymmetric information game.*

Recall that in order to introduce the constituent elements of the Bayesian game structure we used the metaphor of the *ex ante* stage in describing interactive situations with incomplete information. The use of such metaphor and the formal similarity between Bayesian games and the mathematical description of games with asymmetric information (about an initial chance move) has induced many scholars to neglect the differences between games with incomplete information and games with asymmetric information. We should stress, however, that they are indeed different situations. In games with incomplete information the *ex ante* stage *does not exist*, it is just a useful theoretical fiction: Ω only represents a set of possible states of the world, where by "state of the world" we mean a configuration of payoff state, information, and subjective exogenous beliefs.[36] The profile $(p_i)_{i\in I} \in [\Delta(\Omega)]^I$ of so called prior beliefs is simply a useful *mathematical tool* to determine (along with the functions ϑ_i and τ_i) players' interactive beliefs in a given state of the world. Instead, in the interactive decision problems that we just called "games with asymmetric information" the *ex ante* stage is real and the players' priors represent the expectations they hold at that stage.

These differences in the interpretation of the formal structure are not innocuous. Indeed, the interpretation (asymmetric information *vs.* incomplete information) determines to what extent some given assumptions are meaningful and plausible, and what solution concepts are appropriate. For instance, for the case of asymmetric information about the realization of an initial chance move, it is meaningful and often plausible to assume that there exists a common prior given by an objective distribution over Ω. On the other hand, in the case of incomplete information the meaning of the common prior assumption is not obvious and its plausibility is even less obvious. Furthermore, we will see that different notions of self-confirming equilibrium are appropriate for situations with incomplete information and for situations with asymmetric information about an initial chance move, even though the two situations can be represented by the same mathematical structure. In the appendix of this chapter, we show that similar considerations apply to the appropriate notion rationalizability for Bayesian games, as long as one is willing to assume some independence restrictions on players' conjectures.

[36] Recall, we called "state of nature" the configuration $\theta = \boldsymbol{\theta}\,\theta_0, (\theta_i)_{i\in I}\big)$ of payoff state and private information interpreted as a fixed imperfectly known parameter. A "state of the world" also include a specification of subjective beliefs.

8.7 Self-Confirming Equilibrium and Incomplete Information

As previously argued, the self-confirming (or conjectural) equilibrium concept characterizes patterns of behavior that are stable with respect to learning processes in situations of recurrent interaction, taking into account players' information feedback. The appropriate definition of self-confirming equilibrium for games with incomplete information depends on the relevant scenario, and in particular on the interpretation of the mathematical structure we use to represent incomplete information.

The first question to address is whether the set of interacting players and their characteristics are (a) fixed once and for all (long-run interaction), or (b) randomly determined in each period via draws from large populations (anonymous recurrent interaction). In both cases it is necessary to specify what information players are able to obtain ex post about the actions and characteristics of co-players. As in Section 6.3, we describe this with feedback functions. As we did earlier, we assume for the sake of simplicity that the players' goal is to maximize the expected payoff of the current period, without worrying about future payoffs.

8.7.1 Long-Run Interaction

This is the easiest case to analyze. It is sufficient to consider the game with payoff uncertainty

$$\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i : \Theta \times A \to \mathbb{R})_{i\in I} \rangle \, .$$

The state of nature (profile of parameters) θ is *fixed* once and for all; obviously, we continue to assume that every player i knows only θ_i, that the sets of possible values for θ is $\Theta = \Theta_0 \times \left(\bigtimes_{i\in I} \Theta_i\right)$, and that *ex post* i gets further information about θ_0, θ_{-i} and a_{-i} according to a feedback function f_i. In general, the signal received may depend also on θ, that is $f_i : \Theta \times A \to M_i$; for instance, the signal could be the payoff obtained by the player.[37] Every player i holds a belief $\mu^i \in \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i})$ about the residual uncertainty θ_0 and the opponents' information-types and actions.

[37]However, we need to be aware that the payoff u_i does not necessarily represent a material gain which can be cashed. Therefore in the theoretical analysis we are not forced to assume that the realization of u_i is observed by i (observable payoffs).

The question is whether, for any *given* θ, a profile of actions and beliefs $(a_i, \mu^i)_{i\in I} \in \bigtimes_{i\in I} (A_i \times \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i}))$ satisfies the conditions of rationality and confirmed conjectures. These conditions can be obtained as a simple generalization of Definition 25. As in Section 6.3, we give the definition of self-confirming equilibrium for games with payoff uncertainty and feedback $(\hat{G}, f)$, where $f = (f_i)_{i\in I}$ is the profile of feedback functions.

Definition 40. *Fix a game with payoff uncertainty and feedback* $(\hat{G}, f)$ *and a state of nature* $\theta^* \in \Theta$. *A profile of actions and beliefs* $(a_i^*, \mu^i)_{i\in I} \in \bigtimes_{i\in I} (A_i \times \Delta(\Theta_0 \times \Theta_{-i} \times A_{-i}))$ *is a* ***self-confirming equilibrium (SCE) at*** θ^* *if for every* $i \in I$:[38]
*(1) (*rationality*)* $a_i^* \in r_i(\mu^i, \theta_i^*)$,
*(2) (*confirmed conjectures*):*

$$\mu^i\left(\{(\theta_0, \theta_{-i}, a_{-i}) : f_i(\theta_0, \theta_i^*, \theta_{-i}, a_i^*, a_{-i}) = f_i(\theta^*, a^*)\}\right) = 1.$$

$(a_i^*)_{i\in I}$ *is an SCE action profile at* θ^* *if there is some profile of conjectures* $(\mu^i)_{i\in I}$ *such that* $(a_i^*, \mu^i)_{i\in I}$ *is an SCE at* θ^*.

In what follows, for each game with payoff uncertainty

$$\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i : \Theta \times A \to \mathbb{R})_{i\in I}\rangle$$

and state of nature $\theta \in \Theta$, we let $\hat{G}_\theta$ denote the game

$$\hat{G}_\theta = \langle I, (A_i, u_{i,\theta} : A \to \mathbb{R})_{i\in I}\rangle .$$

Thus, for a game with payoff uncertainty and feedback $(\hat{G}, f)$ (where $f_i : \Theta \times A \to M_i$ for each $i \in I$) and fixed θ, $(\hat{G}_\theta, f_\theta)$ denotes the "true" game with feedback at state of nature θ.

Remark 33. *Fix a game with payoff uncertainty* $\hat{G}$. *For each state of nature* $\theta^* \in \Theta$ *and each profile of feedback functions* f, *every Nash equilibrium of* $\hat{G}_{\theta^*}$ *is an SCE action profile at* θ^* *of the game with payoff uncertainty and feedback* $(\hat{G}, f)$.

[38] Recall that $r_i(\mu^i, \theta_i)$ is the set of actions of i that maximize her expected payoff given θ_i and μ^i.

The following remark highlights the fact that in the private-values case (when u_i does not depend on θ_0 and θ_{-i}) we get the notion of self-confirming equilibrium already introduced in Section 6.3, coherently with the claim made there according to which the self-confirming equilibrium concept only presumes that a player knows her own payoff function, not those of the co-players.

Remark 34. *Fix a game with payoff uncertainty and feedback* $(\hat{G}, f)$*, a state of nature* θ^* *and an action profile* a^**, and suppose that the game with payoff uncertainty* $\hat{G}$ *has* private values*. Then* a^* *is an SCE action profile at* θ^* *if and only if* a^* *is an SCE action profile of the game with feedback* $(\hat{G}_{\theta^*}, f_{\theta^*})$*.*

8.7.2 Anonymous Recurrent Interaction

Assume for simplicity that the given game with payoff uncertainty is *finite* and consider the following scenario: There are n heterogeneous populations, $i \in I = \{1, ..., n\}$; the fraction of agents with characteristic θ_i is $q_i(\theta_i) > 0$; n agents are drawn at random, one from each population i, and play a static game with payoffs determined by the parameterized functions $u_i : \Theta \times A \to \mathbb{R}$ $(i \in I)$, where Θ_0 represents a set of possible values of a random shock that affect agents' utility: each time agents play the game, $\theta_0 \in \Theta_0$ is drawn with probability $q_0(\theta_0) > 0$ independently of previous plays. This is precisely the scenario that motivated the definition of equilibrium for a "simple" Bayesian game (with type-independent beliefs)

$$BG = \langle I, \Theta_0, q_0, (\Theta_i, q_i, A_i, u_i)_{i \in I} \rangle \tag{8.7.1}$$

where $q_0 \in \Delta(\Theta_0)$, $q_i \in \Delta(\Theta_i)$, $u_i : \Theta \times A \to \mathbb{R}$ $(i \in I)$.

In this case, however, it is only assumed that each agent in each population i knows θ_i and $u_{i,\theta_i} : \Theta_0 \times \Theta_{-i} \times A \to \mathbb{R}$, while the interactive situation represented by BG is *not* assumed to be common knowledge. Agents play the game recurrently, each time with a different θ_0 drawn from Θ_0 and with different co-players randomly drawn from the respective populations. If the play stabilizes, at least in a statistical sense, for every j, θ_j, a_j, the fraction of agents in sub-population j with characteristic θ_j that choose action a_j remains constant. Let $\alpha_j(a_j|\theta_j)$ denote this fraction and write

$$\alpha_{-i} = (\alpha_j(\cdot|\theta_j))_{j\neq i,\theta_j\in\Theta_j} \in \bigtimes_{j\neq i} \Delta\left(A_j\right)^{\Theta_j}$$

to denote the profile of fractions for the populations different from i. By random matching (and the law of large numbers), the long-run frequency of each profile $(\theta_0, \theta_{-i}, a_{-i})$ is $q_0\left(\theta_0\right)\prod_{j\neq i}\alpha_j(a_j|\theta_j)q_j(\theta_j)$.

The ex post information feedback of an agent playing in role i is described by a feedback function $f_i : \Theta \times A \to M_i$. Hence, in a steady state, an agent of population i with characteristic θ_i that (always) chooses a_i, observes that the long-run frequency of each message m_i is

$$\mathbb{P}^{f_i}_{a_i,\alpha_{-i},q_{-i}}\left(m_i|\theta_i\right) := \sum_{(\theta_0,\theta_{-i},a_{-i}):f_i(\theta_0,\theta_i,\theta_{-i},a_i,a_{-i})=m_i} q_0\left(\theta_0\right)\prod_{j\neq i}\alpha_j(a_j|\theta_j)q_j(\theta_j).$$

Given θ_i and a_i, conjecture $\mu^i \in \Delta(\Theta_0\times\Theta_{-i}\times A_{-i})$ assigns to each message m_i the probability

$$\mathbb{P}^{f_i}_{a_i,\mu^i}\left(m_i|\theta_i\right) := \sum_{(\theta_0,\theta_{-i},a_{-i}):f_i(\theta_0,\theta_i,\theta_{-i},a_i,a_{-i})=m_i} \mu^i(\theta_0,\theta_{-i},a_{-i}).$$

A conjecture is confirmed in the long run for a player of type θ_i who keeps playing a_i if the subjective probability of each message is equal to the observed frequency. If $a_i \in r_i(\mu^i, \theta_i)$ and μ^i is confirmed, then the agents of type θ_i choosing a_i have no reason to switch to another action. In the following definition we let $\mu^i_{\theta_i,a_i}$ denote a conjecture that justifies action a_i for an agent with characteristic (information-type) θ_i.

Definition 41. *Fix a simple Bayesian game with feedback* (BG, f)*. A profile of mixed actions and conjectures:*

$$\left(\alpha_i(\cdot|\theta_i), (\mu^i_{(a_i,\theta_i)})_{a_i\in\mathrm{supp}\alpha_i(\cdot|\theta_i)}\right)_{i\in I,\theta_i\in\Theta_i}$$

(a mixed action $\alpha_i(\cdot|\theta_i)$ *for each* i *and* θ_i*, and a conjecture for each* i*,* θ_i *and* $a_i \in \mathrm{supp}\alpha_i(\cdot|\theta_i)$*) is an* ***anonymous self-confirming equilibrium*** *of* (BG, f) *if for every* $i \in I$*,* $\theta_i \in \Theta_i$*,* $a_i \in A_i$*, the following conditions hold:*

*(1) (*rationality*) if* $\alpha_i(a_i|\theta_i) > 0$*, then* $a_i \in r_i(\mu^i_{(a_i,\theta_i)}, \theta_i)$*,*

*(2) (*confirmed conjectures*)* $\mathbb{P}^{f_i}_{a_i,\mu^i}\left(\cdot|\theta_i\right) = \mathbb{P}^{f_i}_{a_i,\alpha_{-i},q_{-i}}\left(\cdot|\theta_i\right).$

Remark 35. *The definition needs to be modified and made more stringent if the distributions q_j are known. In this case equilibrium conjectures must be consistent with $q = (q_j)_{j \in I_0}$. For instance, in the two-person case,* $\mathrm{marg}_{\Theta_j}\mu^i = q_j$.[39]

Remark 36. *The definition of self-confirming equilibrium* at θ^* *(Definition 40) is obtained as a special case when $q(\theta^*) = 1$.*

Remark 37. *Every mixed equilibrium of the interim strategic form of BG is a self-confirming equilibrium profile of mixed actions.*

Adapting in the obvious way the definitions of observable payoffs and own-action independence of feedback of Section 6.3, we obtain the analog of Theorem 16: if (BG, f) satisfies observable payoffs and own-action independence of feedback, then every (anonymous) self-confirming equilibrium profile of mixed actions $(\alpha_i(\cdot|\theta_i))_{i \in I, \theta_i \in \Theta_i}$ is also a mixed Bayes-Nash equilibrium of BG, therefore the two equilibrium concepts coincide under these assumptions about feedback.

Analogous considerations apply to defining SCE in games with asymmetric information about the realization of an initial chance move. Also in this case, the confirmed-conjectures condition must require that, for each player, the observed distribution of messages coincides with the subjectively expected distribution of messages.

8.8 Appendix

The cleanest and easiest extension of the rationalizability idea to incomplete-information environments is the notion of (directed) rationalizability for games with payoff uncertainty studied in Section 8.2 (and 8.3). However, most theorists have analyzed incomplete information restricting their attention to Bayesian games. Therefore, the

[39] A special case of equilibrium of this kind obtains when actions are observed ($f_i(\theta, a) = a$) and conjectures are "naive" as they do not take into account that opponents' actions depend on their type. For example in the two-agents case without residual uncertainty, $\mu^i(\theta_j, a_j) = q_j(\theta_j)\nu^i(a_j)$ for some $\nu^i \in \Delta(A_j)$. This is called *cursed equilibrium*. Such behavior may explain, for example, the so called "winners' curse" in common value auctions. This refinement of self-confirming equilibrium was proposed by Eyster and Rabin [31], although the link to the self-confirming equilibrium concept was not made explicit.

rationalizability idea has been first extended to such games. Thus, rather than defining what is rationalizable for an information-type θ_i, theorists first tried to define what is rationalizable for a Harsanyi type t_i; but the meaning of "action a_i is rationalizable for type t_i" is not entirely transparent, because Harsanyi's notion of "type" is self-referential, hence elusive. Let us make a preliminary observation. Often the Bayesian games considered in applications *conflate Harsanyi types t_i with information-types θ_i*. Formally, this means that, for each player i, $\Theta_i = T_i$ and the information-type map ϑ_i is the identity on Θ_i (equivalently, Θ_i and T_i are isomorphic, so that ϑ_i is a bijection). For such simple Bayesian games, we just have to consider a special case of Directed Rationalizability: for each i and each θ_i, let

$$\Delta_{i,\theta_i} = \left\{\mu^i \in \Delta\left(\Theta_0 \times \Theta_{-i} \times A_{-i}\right) : \mathrm{marg}_{\Theta_0 \times \Theta_{-i}}\mu^i = p_i^1\left(\cdot|\theta_i\right)\right\},$$

where $p_i^1\left(\cdot|\theta_i\right)$ is the exogenous first-order belief of type θ_i. Let $\Delta = \left(\Delta_{i,\theta_i}\right)_{i\in I,\theta_i\in\Theta_i}$ denote the corresponding profile of belief restrictions. With this, we compute what actions are Δ-rationalizable for each type θ_i of each player i.

What we explain below applies instead to *general* Bayesian games. We first delve into the conceptual subtleties of defining rationalizability for Bayesian games and then we illustrate the concepts with the so called "electronic mail game." To simplify the analysis, we restrict our attention to games with *finite* action sets and finite or *countable* type sets.

8.8.1 Rationalizability in Bayesian Games

A game with payoff uncertainty, i.e., the mathematical structure

$$\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i)_{i\in I}\rangle,$$

does not specify the exogenous interactive beliefs of players, that is, their beliefs about private information and beliefs of the opponents (recall that we call such beliefs "exogenous" because they are not explained, nor partially determined, by strategic reasoning). A specification of exogenous beliefs is necessary to extend the traditional definition of equilibrium due to Nash.[40] The richer structure

$$BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i \in I} \rangle ,$$

known as "Bayesian game," was introduced with this purpose, allowing the definition of Bayesian equilibrium.

It was also pointed out that, for traditional equilibrium analysis, it makes no difference whether the given mathematical structure BG represents a game situation where there is no common knowledge of the payoff functions (incomplete information), or a situation where there is asymmetric information about an initial chance move. The interpretation of BG may make some specific assumptions about interactive beliefs more or less plausible (think about the common prior assumption), but the computation of equilibria is not affected.[41]

However, the interpretation seems to matter when it comes to define the appropriate concept of rationalizability, assuming that the interactive situation represented by BG is transparent to the players. The crucial point is the following: in an incomplete information game, the rationality and strategic reasoning of a player need to be evaluated according to his type. Specifically, we consider the so called *interim* stage and, for each action a_i and type t_i, we ask: Is a_i a justifiable choice for t_i?[42] On the other hand, in a game with asymmetric information about an initial chance move, rationality and strategic reasoning need to be assessed at the *ex ante* stage. For each decision function $\sigma_i : T_i \to A_i$, we ask: Is σ_i a justifiable decision function for player i? By answering these different questions, game theorists obtained two different concept of rationalizability for Bayesian games.

Definition 42. *Let BG be a Bayesian game. An action $a_i \in A_i$ is* interim rationalizable *for type $t_i \in T_i$ of player $i \in I$, if a_i is rationalizable for t_i in the* interim *strategic form $\mathcal{IS}(BG)$. A decision function $\sigma_i : T_i \to A_i$ is* ex ante rationalizable *if σ_i is a rationalizable decision function of the* ex ante *strategic form $\mathcal{AS}(BG)$.*

[40] As we argue in Section 8.7, this does not hold for other definitions of equilibrium that characterize the rest points of adaptive processes.

[41] As we pointed out, this is not true for self-confirming equilibrium.

[42] Recall that we say that a choice is "*justifiable*" if it is a best reply to some belief. A choice is "*rationalizable*" if it survives the iterated elimination of non justifiable choices.

Lemma 22. *If a decision function σ_i^* is justifiable in the* ex ante *strategic form, then for every type t_i the action $a_{t_i}^* = \sigma_i^*(t_i)$ is justifiable in the* interim *strategic form.*

Proof. First observe that, for each type t_i, according to eq. (8.5.2), the actions of the other types t_i' of player i are completely irrelevant for the determination of the *interim* payoff. Hence, one can re-define the conjecture of type t_i in the *interim* strategic form as a probability distribution over the action profiles of the types of players $j \neq i$. Notice that such action profiles coincide with the profiles of decision functions of the opponents of player i in the *ex ante* strategic form:

$$(a_{t_j})_{j\neq i, t_j \in T_j} \in \bigtimes_{j\neq i} \left(A^{T_j}\right) = \bigtimes_{j\neq i} \Sigma_j = \Sigma_{-i}.$$

Hence, the set of conjectures of any type/player t_i in the *interim* strategic form coincides with the sets of conjectures of player i in the *ex ante* strategic form. To be consistent with the notation used for games of complete information, we denote by $U_i(\sigma_i, \mu^i)$ and $\overline{u}_{t_i}(a_i, \mu^i)$ the expected payoffs of player i and of type t_i in the corresponding strategic forms, given conjecture $\mu^i \in \Delta(\Sigma_{-i})$.

Let us fix arbitrarily a decision function σ_i^* and a conjecture μ^i. We now prove that σ_i^* is a best reply to μ^i if and only if, for every type $t_i \in T_i$, action $\sigma_i^*(t_i)$ is a best reply to μ^i. This implies the thesis. Fix $\mu^i \in \Delta(\Sigma_{-i})$. For every decision function σ_i, the expected payoff of σ_i given μ^i is

$$U_i(\sigma_i, \mu^i) = \sum_{\sigma_{-i}} \mu^i(\sigma_{-i}) U_i(\sigma_i, \sigma_{-i})$$

$$\begin{aligned}
&= \sum_{\sigma_{-i}} \mu^i(\sigma_{-i}) \sum_{t_i} p_i\left(t_i\right) \sum_{\theta_0, t_{-i}} p_i\left(\theta_0, t_{-i}|t_i\right) u_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), \sigma_i(t_i), \sigma_{-i}(t_{-i})) \\
&= \sum_{t_i} p_i\left(t_i\right) \sum_{\sigma_{-i}} \mu^i(\sigma_{-i}) \sum_{\theta_0, t_{-i}} p_i\left(\theta_0, t_{-i}|t_i\right) u_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), \sigma_i(t_i), \sigma_{-i}(t_{-i})) \\
&= \sum_{t_i} p_i(t_i) \overline{u}_{t_i}(\sigma_i(t_i), \mu^i).
\end{aligned}$$

Recall: $p_i\left(t_i\right) > 0$ for each $t_i \in T_i$. It follows that

$$\sigma_i^* \in \arg\max_{\sigma_i} U_i(\sigma_i, \mu^i) = \arg\max_{\sigma_i} \sum_{t_i} p_i\left(t_i\right) \overline{u}_{t_i}(\sigma_i(t_i), \mu^i)$$

if and only if

$$\forall t_i \in T_i,\ \sigma_i^*(t_i) \in \arg\max_{a_i} \overline{u}_{t_i}(a_i, \mu^i).$$

The average of the expected payoffs for the different types of i is maximized if and only if[43] the expected payoff of every type of i is maximized. ■

Corollary 5. *If a decision function $\sigma_i \in \Sigma_i$ is rationalizable in the* ex ante *strategic form, then for every type $t_i \in T_i$ the action $\sigma_i(t_i)$ is rationalizable for t_i in the* interim *strategic form.*

From the previous considerations it looks like interim rationalizability is the appropriate solution concept if we interpret the mathematical structure BG as an interactive situation with incomplete information, where the *ex ante* stage is only a useful notational device and therefore the states ω represent only possible configurations of private information and beliefs. *Ex ante* rationalizability is instead the appropriate solution concept if we interpret BG as an interactive situation with asymmetric information over some initial chance move (which directly affects the payoffs only through the payoff-relevant components of θ).

Corollary 5 states that *ex ante* rationalizability is at least as strong a solution concept as *interim* rationalizability. It can be shown by example that it is actually stronger. The formal reason may be understood already from the proof of Lemma 22. Assume that, for every t_i, action $\sigma_i^*(t_i)$ is justifiable. Then there exists a profile of conjectures $(\mu^{t_i})_{t_i \in T_i}$ (with $\mu^{t_i} \in \Delta(\Sigma_{-i})$) such that $\sigma_i^*(t_i)$ is a best reply to μ^{t_i} for each t_i. But this does not imply that there exists a *unique* conjecture $\mu^i \in \Delta(\Sigma_{-i})$ such that $\sigma_i^*(t_i)$ is a best reply to μ^i for every t_i. The difference between the two solutions concept is clearly illustrated by the following numerical example.

Example 36. Let $\Omega = \{\omega', \omega''\}$; player 1 (Rowena) knows the true state, $\tau_1(\omega') \neq \tau_1(\omega'')$, whereas player 2 (Colin) does not, $\tau_2(\omega') = \tau_2(\omega'')$, and considers the two states equally likely: $p_2(\omega') = p_2(\omega'') = \frac{1}{2}$. The payoff functions are as in the matrixes below (note that the payoff of Rowena is independent of the state, this simplifies the example):

[43] Of course, the "only if" part holds because we assume that $p_i(t_i) > 0$ for each $t_i \in T_i$.

ω'	c	d
a	$3,3$	$0,2$
m	$2,0$	$2,2$
b	$0,0$	$3,2$

ω''	c	d
a	$3,0$	$0,2$
m	$2,0$	$2,2$
b	$0,3$	$3,2$

First, observe that every action by Colin is justifiable. In particular, c is justifiable by the conjecture that Rowena chooses a if ω' and b if ω''. It is easy to verify that for each type of Rowena all the actions are justifiable by some conjecture about Colin. Hence, if choices are evaluated at the *interim* stage it is not possible to exclude any action, everything is *interim* rationalizable. Instead if choices are evaluated at the *ex ante* stage, then the decision function $\sigma_1(\omega') = a$, $\sigma_1(\omega'') = b$, denoted by ab, can be excluded. Indeed, since Rowena's payoff does not depend on the state, ab could be a best reply to some conjecture μ^1 only if both a and b were best replies to μ^1. But if Rowena believes $\mu^1(c) \geq \frac{1}{2}$, then b is **not** a best reply; if Rowena believes $\mu^1(c) \leq \frac{1}{2}$, then a is **not** a best reply. Therefore ab is not justifiable by any μ^1. The same argument shows that ba is not justifiable either. The *ex ante* justifiable decision functions of Rowena are: *aa,am,ma,mm,bb,bm,mb*.[44] Given any belief μ^2 such that $\mu^2(ab) = 0$, action c yields an expected payoff less or equal than $\frac{3}{2}$ (check this); action d instead yields $2 > \frac{3}{2}$. Then, the only rationalizable action for Colin in the *ex ante* strategic form is d. It follows that the only rationalizable decision function of Rowena in the *ex ante* strategic form is bb. ▲

Although the gap between *ex-ante* and *interim* rationalizability was at first just accepted as a fact, it should be disturbing. Conceptually, rationalizability is meant to capture the assumptions of rationality (i.e., expected utility maximization) and common belief in rationality given the background transparency of the Bayesian Game BG. Since both strategic forms are based on the same Bayesian game and since *ex-ante* maximization is equivalent to *interim* maximization,[45] why does not *ex-*

[44] Given that Row's payoff does not depend on ω, those decision functions that select different actions for the two states are justifiable only by beliefs that make Row indifferent between these two actions; the decision functions am and ma are among the best replies to the conjecture $\mu^1(c) = \frac{2}{3}$, the decision functions bm and mb are among the best replies to the conjecture $\mu^1(c) = \frac{1}{3}$.

[45] This equivalence holds under the assumptions that for every player i, all types t_i

ante rationalizability coincide (in terms of behavioral predictions) with *interim* rationalizability? Which features of the strategic form create the gap highlighted in Example 36? Before providing an answer to these questions, we introduce another puzzling feature of rationalizability in Bayesian environments: the dependence of *interim* rationalizability on apparently irrelevant details of the state space. To understand this problem, consider the following example.

Example 37. [Dekel et al. [28]] Rowena and Colin are involved in a betting game. Each player can decide to bet (action B) or not to bet (action N). There is an unknown parameter $\theta_0 \in \Theta_0 = \{\theta_0', \theta_0''\}$ and players have no private information ($\Theta_i = \left\{\bar{\theta}_i\right\}$ for every agent i). Rowena (Colin) wins if both players bet and $\theta_0 = \theta_0'$ ($\theta_0 = \theta_0''$). The decision to bet entails a deadweight loss of \$4 independently of the opponent's action; if both agents bet, the loser gives \$12 to the winner. The corresponding game with payoff uncertainty $\hat{G}$ can be represented as follows:

θ_0'	B	N
B	$8, -16$	$-4, 0$
N	$0, -4$	$0, 0$

θ_0''	B	N
B	$-16, 8$	$-4, 0$
N	$0, -4$	$0, 0$

Let us assume that there is common belief that each agent assigns probability $\frac{1}{2}$ to each payoff state. Thus, each player i has only one hierarchy of beliefs about θ_0.
The simplest state space representing this situation is the following: $\Omega = \{\omega', \omega''\}$,,$\vartheta_0(\omega') = \theta_0'$, $\vartheta_0(\omega'') = \theta_0''$ and for every agent i, $T_i = \{\bar{t}_i\}$, functions ϑ_i and τ_i are trivially defined and $p_i(\omega') = \frac{1}{2}$. In this case, the *ex-ante* and the *interim* strategic forms coincide and are represented by the following game:

	B	N
B	$-4, -4$	$-4, 0$
N	$0, -4$	$0, 0$

have positive probability.

It is immediate to see that betting is not rationalizable. Thus, N is the only *interim* rationalizable action.

Now, consider an alternative state space: $\Omega = \left\{\omega^1, \omega^2, \dots, \omega^8\right\}$,

$$\vartheta_0(\omega) = \begin{cases} \theta_0', & \text{if } \omega \in \left\{\omega^1, \omega^2, \omega^3, \omega^4\right\}, \\ \theta_0'', & \text{if } \omega \in \left\{\omega^5, \omega^6, \omega^7, \omega^8\right\}, \end{cases}$$

for every agent i, $T_i = \{t_i', t_i''\}$, ϑ_i is trivially defined,

$$\tau_1(\omega) = \begin{cases} t_1', & \text{if } \omega \in \left\{\omega^1, \omega^2, \omega^5, \omega^6\right\}, \\ t_1'', & \text{if } \omega \in \left\{\omega^3, \omega^4, \omega^7, \omega^8\right\}, \end{cases}$$

$$\tau_2(\omega) = \begin{cases} t_2', & \text{if } \omega \in \left\{\omega^1, \omega^3, \omega^5, \omega^7\right\}, \\ t_2'', & \text{if } \omega \in \left\{\omega^2, \omega^4, \omega^6, \omega^8\right\}, \end{cases}$$

and $p_1 = p_2 = p$:

State of the world, ω	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8
$p[\omega]$	$\frac{1}{4}$	0	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{4}$	0

One can easily check that this common prior induces a distribution over residual uncertainty and types which can be represented as follows:

θ_0'	t_2'	t_2''
t_1'	$\frac{1}{4}$	0
t_1''	0	$\frac{1}{4}$

θ_0''	t_2'	t_2''
t_1'	0	$\frac{1}{4}$
t_1''	$\frac{1}{4}$	0

and that for both types of both players the following holds: (i) the player has no private information, (ii) he assigns probability $\frac{1}{2}$ to each state θ_0, and (iii) there is common belief of this. Thus, this second state space represents exactly the same belief hierarchies than the former one: each i assigns probability $\frac{1}{2}$ to each θ_0, each i is certain that $-i$ assigns probability $\frac{1}{2}$ to each θ_0, and so on. If we write down the *ex-ante* strategic form derived from this state space, we get the following game:

	BB	BN	NB	NN
BB	$-4,-4$	$-4,-2$	$-4,-2$	$-4,0$
BN	$-2,-4$	$1,-5$	$-5,1$	$-2,0$
NB	$-2,-4$	$-5,1$	$1,-5$	$-2,0$
NN	$0,-4$	$0,-2$	$0,-2$	$0,0$

Notice that, in this game, strategy *NB* (*BN*) is a best response for Rowena to the conjecture that Colin is playing *NB* (*BN*), while *NB* (*BN*) is Colin's best response to the conjecture that Rowena is playing *BN* (*NB*). Thus, the set $\{BN,NB\}\times\{BN,NB\}$ has the best response property and, consequently, betting is *ex-ante* and, thanks to Corollary 5, *interim* rationalizable. ▲

In Example 37, the two state spaces generate two different Bayesian Games representing the same information and belief hierarchies over Θ; thus, they seem to represent the same background common knowledge and beliefs concerning the game with payoff uncertainty $\hat{G}$. Then, should not we expect them to lead to the same *interim* rationalizable prediction? Why is this intuition false? Why does the choice of one type space over the other matter?

It turns out that these two *puzzles* are related: they both *depend on independence assumptions that are implicitly made in the construction of the strategic forms.* Once these assumptions are removed and the proper (*ex ante* and *interim*) notions of *correlated rationalizability* are defined, the puzzles disappear: (i) the set of *interim correlated rationalizable* actions is not affected by the choice among different state spaces representing the same information and belief hierarchies over Θ, and (ii) there is *no gap* between *interim correlated* and *ex-ante correlated rationalizability.*

First, let us focus on the puzzle highlighted by Example 37. Consider the second state space and recall that in such state space each player has two types representing the same belief hierarchy over Θ.[46] In this case, if type t_1' conjectures that t_2' plays B and t_2'' plays N, this conjecture, together with the common belief prior p, will lead t_1' to believe that Colin will bet only when the residual uncertainty is θ_0'; this belief

[46] In Example 37, players have no private information.

will justify the decision of type t_1' to bet.[47] To put it differently, in this state space, Rowena's types hold beliefs in which Colin's type is correlated with residual uncertainty and this can, in turn, introduce correlation between Colin's behavior, $\sigma_2(\cdot)$, and residual uncertainty, θ_0. On the contrary, this correlation is not possible with the first state space in which each belief hierarchy is represented only by one type; this happens because the construction of the *interim* strategic form imposes an implicit independence constraint on players' beliefs: each player has to regard his opponents' behavior, $\sigma_2(\cdot)$, as independent of the residual uncertainty, θ_0, conditional on the opponents' types. If there are relatively few types representing the same information and belief hierarchy, this independence constraint may bind, preventing players from holding correlated beliefs and restricting the set of rationalizable actions.[48] To address this problem, Dekel et al. [28] define a solution concept, **Interim Correlated Rationalizability**, in which beliefs are explicitly allowed to exhibit correlation between opponents' actions and residual uncertainty. Since the independence restriction imposed by the construction of the *interim* strategic form is not related to the assumption of rationality, we can immediately conclude that *interim correlated rationalizability* represents a step forward in the search for a solution concept *characterized by Rationality and Common Belief in Rationality (RCBR) given the background transparency of the given Bayesian game BG*. Actually, it can be shown that *interim* correlated rationalizability fully characterizes the behavioral implications of these epistemic assumptions. To provide a formal definition of this solution concept, fix a Bayesian Game $BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i\in I}\rangle$. Then, for every i, t_i and $\mu^i \in \Delta(\Omega \times A_{-i})$ let

$$r_i\left(t_i, \mu^i\right) = \arg\max_{a_i\in A_i} \sum_{\omega, a_{-i}} \mu^i(\omega, a_{-i})\, u_i\left(\vartheta_0(\omega), \vartheta_i(t_i), \vartheta_{-i}(\tau_{-i}(\omega)), a_i, a_{-i}\right).$$

The set of *interim* correlated rationalizable actions is iteratively defined

[47] It is easy to see that a similar reasoning holds for any other type.

[48] This problem would not arise if for every player i and every information and belief hierarchy he may hold, there were at least $|A_i|$ types representing this information and belief hierarchy. Stricter requirement can be provided, but this would go beyond the scope of this textbook.

as follows: for every i and t_i, let $ICR_i^{0,BG}(t_i) = A_i$ and for each $k \geq 1$, let

$$ICR_i^{k,BG}(t_i) = \left\{ \begin{array}{c} a_i \in A_i : \exists \mu^i \in \Delta(\Omega \times A_{-i}),\ \exists \varphi_{-i} : \Theta_0 \times T_{-i} \to \Delta(A_{-i}), \\ 1)\ a_i \in r_i\left(t_i, \mu^i\right) \\ 2)\ \forall (\theta_0, t_{-i}),\ \varphi_{-i}(\theta_0, t_{-i})(a_{-i}) > 0 \Rightarrow a_{-i} \in ICR_{-i}^{k-1,BG}(t_{-i}) \\ 3)\ \forall (\omega, a_{-i}),\ \mu^i(\omega, a_{-i}) = p(\omega|t_i) \cdot \varphi_{-i}(\vartheta_0(\omega), \tau_{-i}(\omega))(a_{-i}) \end{array} \right\},$$

where $ICR_{-i}^{k-1,BG}(t_{-i}) = \times_{j \neq i} ICR_j^{k-1,BG}(t_j)$.

In the previous definition, function $\varphi_{-i}(\theta_0, t_{-i})$ represents the conjecture of player i concerning the behavior of his opponents given their types are t_{-i} and residual uncertainty is θ_0. Since such functions depend both on θ_0 and on t_{-i}, representing conjectures with such functions allows to introduce correlation between a_{-i} and θ_0 even after conditioning on t_{-i}; this correlation is not allowed by *interim* rationalizability.[49]

Definition 43. *Fix a Bayesian game BG. For player $i \in I$, action $a_i \in A_i$ and type $t_i \in T_i$, a_i is* interim *correlated rationalizable for t_i if $a_i \in ICR_i^{BG}(t_i) = \bigcap_{k \geq 0} ICR_i^{k,BG}(t_i)$.*

Since *interim* correlated rationalizability allows for correlation between opponents' behavior and residual uncertainty, it does not impose any implicit independence restriction and it captures the assumptions of players' rationality and common belief in rationality given the transparency of the Bayesian game BG (as usual, the k-th step in the iterative definition of interim correlated rationalizability corresponds to the $k-1$-th level of mutual beliefs in rationality). As a consequence, given a game with payoff uncertainty $\hat{G}$, the set of interim correlated rationalizable actions of type t_i depends only on the information and *belief hierarchy* entailed by t_i. From now on, the latter will be denoted by $\hat{p}_i(t_i)$. The following Theorem formalizes this result.

Theorem 30. *Fix $\hat{G} = \left\langle I, \Theta_0, (\Theta_i, A_i, u_i)_{i \in I} \right\rangle$ and consider two Bayesian Games based on it: $BG' = \langle I, \Omega', \Theta_0, \vartheta'_0, (\Theta_i, T'_i, A_i, \tau'_i, \vartheta'_i, p'_i, u_i)_{i \in I} \rangle$ and $BG'' = \langle I, \Omega'', \Theta_0, \vartheta''_0, (\Theta_i, T''_i, A_i, \tau''_i, \vartheta''_i, p''_i, u_i)_{i \in I} \rangle$. Let $t'_i \in T'_i$ and $t''_i \in T''_i$. Then, if $\hat{p}_i(t'_i) = \hat{p}_i(t''_i)$, $ICR_i^{BG'}(t'_i) = ICR_i^{BG''}(t''_i)$.*

[49] Indeed, it is possible to show that the set of *interim* rationalizable actions for type t_i is equivalent to the set of actions obtained through an iterative construction similar to the one we just described, but in which $\varphi_{-i} : T_{-i} \to \Delta(A_{-i})$.

Proof. Take $t'_i \in T'_i$ and $t''_i \in T''_i$ and suppose $\hat{p}_i(t'_i) = \hat{p}_i(t''_i)$. We will prove by induction that, for every i and for every $k \geq 0$, $ICR_i^{k,BG'}(t'_i) = ICR_i^{k,BG''}(t''_i)$. To simplify notation, we will not specify the dependence of the set of *interim* correlated rationalizable actions on Bayesian Game (notice that this is also justified by the statement of the Theorem). The result is trivially true for $k = 0$. Now, suppose that $ICR_i^s(t'_i) = ICR_i^s(t''_i)$ for every $s \leq k-1$. We need to show that $ICR_i^k(t'_i) = ICR_i^k(t''_i)$.

Take any $a_i \in ICR_i^k(t'_i)$. By definition, we can find $\mu^i_{t'_i} \in \Delta(\Omega' \times A_{-i})$ and $\varphi_{t'_i} : \Theta_0 \times T'_{-i} \to \Delta(A_{-i})$ such that (i) $a_i \in r_i\left(t'_i, \mu^i_{t'_i}\right)$, (ii) $\varphi_{t'_i}(\theta_0, t_{-i})(a_{-i}) > 0$ implies $a_{-i} \in ICR_{-i}^{k-1}(t_{-i})$ and (iii) for every pair (ω, a_{-i}),

$$\mu^i_{t'_i}(\omega, a_{-i}) = p\left(\omega|t'_i\right) \cdot \varphi_{t'_i}\left(\vartheta'_0(\omega), \tau'_{-i}(\omega)\right)(a_{-i}).$$

Let

$$P_{-i}^{k-1} = \left\{\left(\vartheta'_j(t_j), p_j^{k-1}(t_j)\right)_{j \neq i} : t_j \in T'_j\right\}$$

be the set of possible profiles of private information and $(k-1)$-th order beliefs that agents other than i can have. For every $\theta_0 \in \Theta_0$, $\pi_{-i}^{k-1} \in P_{-i}^{k-1}$ and $a_{-i} \in A_{-i}$, define[50]

$$[\theta_0]_{BG'} = \left\{(\omega, a_{-i}) \in \Omega' \times A_{-i} : \vartheta_0(\omega) = \theta_0\right\},$$

$$[a_{-i}]_{BG'} = \left\{\left(\omega, a'_{-i}\right) \in \Omega' \times A_{-i} : a'_{-i} = a_{-i}\right\},$$

and

$$\left[\pi_{-i}^{k-1}\right]_{BG'} = \left\{(\omega, a_{-i}) \in \Omega' \times A_{-i} : \left(\vartheta'_j(\tau_j(\omega)), p_j^{k-1}(\tau_j(\omega))\right)_{j \neq i} = \pi_{-i}^{k-1}\right\}.$$

To ease notation, we also write

$$\left[\theta_0, \pi_{-i}^{k-1}, a_{-i}\right]_{BG'} = [\theta_0]_{BG'} \cap \left[\pi_{-i}^{k-1}\right]_{BG'} \cap \left[a'_{-i}\right]_{BG'},$$

$$\left[\theta_0, \pi_{-i}^{k-1}\right]_{BG'} = [\theta_0]_{BG'} \cap \left[\pi_{-i}^{k-1}\right]_{BG'},$$

[50] As in the previous sections, we let symbols in square brackets denote corresponding events in the relevant state space.

and

$$\left[\theta_0, a'_{-i}\right]_{BG'} = [\theta_0]_{BG'} \cap \left[a'_{-i}\right]_{BG'}.$$

Finally, for every $\omega \in \Omega'$, let

$$\hat{\pi}_{-i}^{k-1,BG'}(\omega) = \left(\vartheta'_j\left(\tau'_j(\omega)\right), p_j^{k-1}\left(\tau'_j(\omega)\right)\right)$$

and, similarly, for every $\omega \in \Omega''$, let

$$\hat{\pi}_{-i}^{k-1,BG''}(\omega) = \left(\vartheta''_j\left(\tau''_j(\omega)\right), p_j^{k-1}\left(\tau''_j(\omega)\right)\right).$$

Then for every $\theta_0 \in \Theta_0$ and $\pi_{-i}^{k-1} \in P_{-i}^{k-1}$, define

$$\mu_{t'_i}\left(\left[\theta_0, \pi_{-i}^{k-1}\right]\right) = \sum_{(\omega,a_{-i})\in\left[\theta_0,\pi_{-i}^{k-1}\right]_{BG'}} \mu_{t'_i}(\omega, a_{-i}).$$

$\mu_{t'_i}\left(\left[\theta_0, \pi_{-i}^{k-1}\right]\right)$ is the probability that type t'_i assigns to the event "θ_0 is the case and other agents have private information and $(k-1)$-th order beliefs given by π_{-i}^{k-1}." If $\mu_{t'_i}\left(\left[\theta_0, \pi_{-i}^{k-1}\right]\right) > 0$, let

$$\varphi_{t_i}\left(\theta_0, \pi_{-i}^{k-1}\right)\left[a'_{-i}\right] = \frac{\sum_{(\omega,a_{-i})\in\left[\theta_0,\pi_{-i}^{k-1},a'_{-i}\right]_{BG'}} \mu_{t'_i}(\omega, a_{-i})}{\mu_{t'_i}\left(\left[\theta_0, \pi_{-i}^{k-1}\right]\right)};$$

if $\mu_{t_i}\left(\left[\theta_0, \pi_{-i}^{k-1}\right]\right) = 0$, take any $t'_{-i} = \left(t'_j\right)_{j\neq i} \in T'_{-i}$ such that $\left(\vartheta'_j\left(t'_j\right), p_j^{k-1}\left(t'_j\right)\right)_{j\neq i} = \pi_{-i}^{k-1}$ and let

$$\varphi_{t'_i}\left(\theta_0, \pi_{-i}^{k-1}\right)(a_{-i}) = \begin{cases} \frac{1}{\left|ICR_{-i}^{k-1}\left(t'_{-i}\right)\right|}, & \text{if } a_{-i} \in ICR_{-i}^{k-1}\left(t'_{-i}\right), \\ 0, & \text{otherwise} \end{cases}$$

(by the inductive hypothesis, the definition of $\varphi_{t'_i}$ does not depend on the actual choice of t'_{-i}).
For every $(\omega, a_{-i}) \in \Omega'' \times A_{-i}$ define:

$$\mu_{t''_i}(\omega, a_{-i}) = p\left(\omega|t''_i\right)\varphi_{t'_i}\left(\vartheta''_0(\omega), \hat{\pi}_{-i}^{k-1,BG''}(\omega)\right)(a_{-i}),$$

where the previous expression is well defined since BG' and BG'' are based on the same $\hat{G}$ and $\hat{p}_i(t'_i) = \hat{p}_i(t''_i)$. Moreover, since $\hat{p}_i(t'_i) = \hat{p}_i(t''_i)$:

$$\begin{aligned}
\mu_{t'_i}\left(\left[\theta_0, \pi^{k-1}_{-i}\right]\right) &= \sum_{(\omega,a_{-i})\in\left[\theta_0,\pi^{k-1}_{-i}\right]_{BG'}} \mu_{t'_i}(\omega, a_{-i}) \\
&= p\left(\left\{\omega \in \Omega' : \vartheta'(\omega) = \theta_0,\ \hat{\pi}^{k-1,BG'}_{-i}(\omega) = \pi^{k-1}_{-i}\right\} | t'_i\right) \\
&= p\left(\left\{\omega \in \Omega'' : \vartheta''(\omega) = \theta_0,\ \hat{\pi}^{k-1,BG''}_{-i}(\omega) = \pi^{k-1}_{-i}\right\} | t''_i\right).
\end{aligned}$$

Then, for every $\left(\theta_0, a'_{-i}\right)$, we have

$$\sum_{(\omega,a_{-i})\in\left[\theta_0,a'_{-i}\right]_{BG''}} \mu_{t''_i}(\omega, a_{-i}) = \sum_{\omega:\vartheta''_0(\omega)=\theta_0} p\left(\omega|t''_i\right)\varphi_{t'_i}\left(\vartheta''_0(\omega), \hat{\pi}^{k-1}_{-i}\left(\tau''_{-i}(\omega)\right)\right)\left(a'_{-i}\right)$$

$$= \sum_{\pi^{k-1}_{-i}\in P^{k-1}_{-i}} p\left(\left\{\omega : \vartheta''_0(\omega) = \theta_0,\ \hat{\pi}^{k-1,BG''}_{-i}(\omega) = \pi^{k-1}_{-i}\right\} \mid t''_i\right)\varphi_{t'_i}\left(\theta_0, \pi^{k-1}_{-i}\right)\left(a'_{-i}\right)$$

$$= \sum_{\pi^{k-1}_{-i}\in P^{k-1}_{-i}} \mu_{t'_i}\left(\left[\theta_0, \pi^{k-1}_{-i}\right]\right) \frac{\sum_{(\omega,a_{-i})\in\left[\theta_0,\pi^{k-1}_{-i},a'_{-i}\right]_{BG'}} \mu_{t_i}(\omega, a_{-i})}{\mu_{t'_i}\left(\left[\theta_0, \pi^{k-1}_{-i}\right]\right)}$$

$$= \sum_{(\omega,a_{-i})\in\left[\theta_0,a'_{-i}\right]_{BG'}} \mu_{t'_i}(\omega, a_{-i}),$$

where the definition of $[\theta_0, a_{-i}]_{BG''}$ is analogous to the one of $[\theta_0, a_{-i}]_{BG'}$. Thus, $\mu_{t'}$ and $\mu_{t''}$ have the same marginal distribution over $\Theta_0 \times A_{-i}$ implying that $a_i \in r_i\left(t''_i, \mu^i_{t''_i}\right)$. Furthermore we already constructed a function $\varphi_{-i} : \Theta_0 \times T''_{-i} \to \Delta(A_{-i})$ (namely, $\varphi_{t'_i}$) such that for every (ω, a_{-i}), $\mu^i_{t''_i}(\omega, a_{-i}) = p(\omega|t''_i)\varphi_{-i}\left(\vartheta''_0(\omega), \tau''_{-i}(\omega)\right)(a_{-i})$. Finally, by construction and by the inductive hypothesis, $\mu_{t''_i}(\omega, a_{-i}) > 0$ implies $a_{-i} \in ICR^{k-1}_{-i}\left(\tau''_{-i}(\omega)\right)$. We conclude that $a_i \in ICR^k_i(t''_i)$. The statement of the Theorem follows by induction. ■

Notice that, by definition, the conditional independence restriction implicit in the construction of the *interim* strategic form has no bite when there is distributed knowledge of the state (Θ_0 is a singleton); in this case it is easy to verify that *interim* correlated rationalizability is equivalent to *interim* rationalizability. On the contrary, suppose that there is some residual uncertainty ($|\Theta_0| > 1$) and that we insist on using *interim* rationalizability. A natural question arises: can we at least provide an *expressible* characterization of this solution concept and, in particular, of the independence restriction implied by it? A characterization is deemed **expressible** *if it can be stated in a language based on primitives (that is, elements contained in the description of the game with payoff uncertainty, $\hat{G}$) and terms derived from them (such as, hierarchies of beliefs over these primitives).*

The answer to the previous question is affirmative only in some particular cases and the reason goes to the very heart of Harsanyi's approach. To understand why, recall that *interim* rationalizability requires players to regard opponents' actions as independent of the residual uncertainty *conditional on their types.* But what is a type? A *type* is a self-referential object: it *is a private information and a belief over residual uncertainty and other players' types.* Thus, unless we can establish a one-to-one mapping between types and agents' information and belief hierarchies over primitives, the conditional independence assumption is not expressible. Obviously, in order to assess the existence of such a mapping, we can use both payoff-relevant and payoff-irrelevant primitives. Thus, even though two types represent the same private information and belief hierarchy over payoff-relevant parameters, they can still be distinguished by the information and beliefs hierarchies over payoff-irrelevant elements that they capture. Whenever this is the case, the conditional independence assumption is expressible and we can show that, given the transparency of BG, *interim* rationalizability characterizes the behavioral implications of the following epistemic assumptions: (R) players are rational, (CI) their beliefs satisfy *independence* between of the opponents' behavior and the residual uncertainty *conditional* on the information and belief hierarchy over primitives of the opponents, and (CB($R \cap CI$)) there is common belief of R and CI.

Notice that, insofar players may believe that (i) their opponents' behavior may depend on payoff-irrelevant information, and (ii) this information is correlated with θ_0, the information and belief hierarchies over payoff-irrelevant parameters may still be *strategically relevant.* For instance, in Example 37, Rowena may be "superstitious" and believe that the payoff-relevant state, θ_0, is correlated with a particular dream

Colin may have had (payoff-irrelevant information), which could also affect Colin's decision to bet. Similarly, in Example 26, a firm may believe that the quantity produced by its competitor depends on the analysis carried out by its marketing department and that this (payoff-irrelevant) information may be correlated with the actual position of the demand function (θ_0).

As a special case, in which *interim* rationalizability admits an expressible characterization, we can consider **simple Bayesian Games**.[51]

Definition 44. *Fix a game with payoff uncertainty*

$$\hat{G} = \langle I, \Theta_0, (\Theta_i, A_i, u_i)_{i\in I}\rangle \, .$$

A Bayesian Game

$$BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i\in I}\rangle$$

based on $\hat{G}$ is ***simple*** *if $T_i = \Theta_i$ for each $i \in I$.*[52] *In this case, the functions $(\vartheta_i(\cdot) : T_i \to \Theta_i)_{i\in I}$ are trivially defined: $\vartheta_i = \mathrm{Id}_{\Theta_i}$ for each $i \in I$.*

Focusing on simple Bayesian games, one can show that *interim* rationalizability characterizes the behavioral implications of the following epistemic assumptions: (R) rationality, $(CIPI)$ independence between players' actions and residual uncertainty conditional on players' private information, and $(\mathrm{CB}(R\cap CIPI))$ common belief of R and $CIPI$. Besides being interesting *per se*, simple Bayesian games are also widely used in the study of many relevant problems in information economics, such as adverse selection, moral hazard, mechanism design.

Now, let us turn to the other puzzling result: the gap between *ex-ante* and *interim* rationalizability. As already pointed out, the gap highlighted by Example 36 arises because the *interim* strategic form regards different types as different agents and, as a consequence, different types playing in Rowena's role may hold different beliefs concerning Colin's behavior (one of Rowena's type can believe Colin will play c with probability greater than $\frac{1}{2}$, while the other type can believe that this will happen only with probability

[51] They are slight generalization of "simple Bayesian games with type-independent beliefs" of Section 8.4 because, in the latter, beliefs about θ_{-i} are independent of θ_i, whereas in a simple Bayesian game $p_i \in \Delta(\Theta)$ and $p_i(\cdot|\theta_i)$ may depend on θ_i.

[52] More formally, the requirement is that T_i and Θ_i are isomorphic.

lower than $\frac{1}{2}$); since Rowena's conjecture concerning Colin's behavior can vary with her own type, her beliefs can exhibit correlation between the state of the world, ω (which determines residual uncertainty and players' type) and Colin's action. Instead, the *ex-ante* strategic form implicitly requires Rowena to regard Colin's decision function as independent from the state of the world.[53] To understand this, let us analyze the strategic-form expected payoff of player i when he plays "strategy" σ_i and holds conjecture $\mu^i \in \Delta(\Sigma_{-i})$ about the strategies of other players:

$$U_i(\sigma_i, \mu^i) = \sum_{\sigma_{-i} \in \Sigma_{-i}} \mu^i(\sigma_{-i}) U_i(\sigma_i, \sigma_{-i})$$

$$= \sum_{\sigma_{-i} \in \Sigma_{-i}} \mu^i(\sigma_{-i}) \sum_{\omega \in \Omega} p_i(\omega) u_i(\vartheta_0(\omega), \vartheta_i(\tau_i(\omega)), \vartheta_{-i}(\tau_{-i}(\omega)), \sigma_i(\tau_i(\omega)), \sigma_{-i}(\tau_{-i}(\omega)))$$

$$= \sum_{\omega \in \Omega} \sum_{\sigma_{-i} \in \Sigma_{-i}} p_i(\omega) \mu^i(\sigma_{-i}) \bar{U}_i(\omega, \sigma_i, \sigma_{-i}),$$

where we let $\bar{U}_i(\omega, \sigma_i, \sigma_{-i})$ denote the payoff of i when the state of the world is ω and profile (σ_i, σ_{-i}) is played. As the formula shows, the expected payoff of i is computed under the assumption that the probability of each pair (ω, σ_{-i}) is the product of the marginal probability of ω, $p_i(\omega)$ and the marginal probability of σ_{-i}, $\mu^i(\sigma_{-i})$. This means that i regards ω, the choice of nature, as independent of σ_{-i}, the choice of $-i$.

This implicit independence restriction reduces the set of Rowena's admissible beliefs and makes *ex-ante* rationalizability more demanding (hence, stronger) than *interim* rationalizability. Battigalli et al. [19] address this problem by defining **Ex-Ante Correlated Rationalizability**, a solution concept that allows correlation, according to players' subjective beliefs, between the initial chance move and players' decision functions. To introduce *ex-ante* correlated rationalizability, consider a simple Bayesian game; although such restriction is not necessary from the mathematical point of view, the *ex-ante* strategic form interprets types as actual information received by players concerning the initial chance move and, consequently, the assumption of simple Bayesian games

[53] If we interpret ω as the choice made by an external and neutral player called Nature, the independence restriction on beliefs is equivalent to requiring independence between the strategy chosen by Colin and the one chosen by Nature.

is the most, if not the only, reasonable one. For every i and $\mu^i \in \Delta(\Omega \times \Sigma_{-i})$, let[54]

$$r_i\left(\mu^i\right) = \arg\max_{\sigma_i \in \Sigma_i} \sum_{\omega, \sigma_{-i}} \mu^i(\omega, \sigma_{-i})\, \hat{u}_i(\omega, \sigma_i(\tau_i(\omega)), \sigma_{-i}(\tau_{-i}(\omega))).$$

Then we can inductively define the set of *ex-ante* correlated rationalizable decision functions as follows: for every i, $ACR_i^{0,BG} = \Sigma_i$ and for every $k \geq 1$,

$$ACR_i^{k,BG} = \left\{ \begin{array}{c} \sigma_i \in \Sigma_i : \exists \mu^i \in \Delta(\Omega \times \Sigma_{-i}), \\ 1)\ \sigma_i \in r_i\left(\mu^i\right),\ 2)\ \mathrm{marg}_{\Omega}\mu^i = p, \\ 3)\ \mu^i(\omega, \sigma_{-i}) > 0 \Longrightarrow \sigma_{-i} \in ACR_{-i}^{k-1} \end{array} \right\},$$

where $ACR_{-i}^{k-1} = \times_{j \neq i} ACR_j^{k-1}$.

Definition 45. *Fix a simple Bayesian game BG. A decision function for player i is* ex-ante *correlated rationalizable if $\sigma_i \in ACR_i^{BG} = \bigcap_{k \geq 0} ACR_i^{k,BG}(\Sigma_{-i})$.*

From the iterative definition of *ex-ante* correlated rationalizability, it is easy to see that the justifying belief μ^i introduces correlation between the state of the world and opponents' behavior. Since different states ω can be associated with different types $t_i = \tau_i(\omega)$, which can hold different beliefs about the behavior of the other players, this correlation can also be introduced by *interim* correlated rationalizability. On the contrary, such correlation is not allowed by *ex-ante* rationalizability that forces each player to hold the same conjecture over opponents' decision functions independently of her type.

Thus, *interim* correlated rationalizability and *ex-ante* correlated rationalizability avoid any independence restriction implicitly entailed by the *interim* and *ex-ante* strategic form; as a consequence, both these solution concepts are characterized only by the epistemic assumptions of rationality and common belief in rationality given the background common knowledge of BG. Then, it should not come as a surprise that, if we use these solution concepts, the gap highlighted in Example 36 disappears and the following theorem holds.

[54] Recall that function $\hat{u}_i : \Omega \times A \to \mathbb{R}$ was introduced in Section 8.6 to analyze games with asymmetric information about an initial move by chance and that $\hat{u}_i$ can be related to u_i according to 8.6.1.

Theorem 31. *Let* $BG = \langle I, \Omega, \Theta_0, \vartheta_0, (\Theta_i, T_i, A_i, \tau_i, \vartheta_i, p_i, u_i)_{i\in I}\rangle$ *be a simple Bayesian game. Then, for every* i*,*

$$ACR_i^{BG} = \left\{\sigma_i \in \Sigma_i : \forall t_i \in T_i,\ \sigma_i\left(t_i\right) \in ICR_i^{BG}\left(t_i\right)\right\}.$$

Proof. We will prove that a stronger result holds, namely that for every i and every $k \geq 0$

$$ACR_i^{k,BG} = \left\{\sigma_i \in \Sigma_i : \forall t_i \in \Theta_i,\ \sigma_i\left(t_i\right) \in ICR_i^{k,BG}\left(t_i\right)\right\}.$$

The result is trivial for $k = 0$. Now suppose that the result holds for every $s \leq k-1$; we will show that it holds also for k.

Take an agent i and an arbitrary $\sigma_i \in ACR_i^{k,BG}$. By definition we can find a rationalizing belief $\mu^i \in \Delta\left(\Omega \times \Sigma_{-i}\right)$. To ease notation, we keep denoting with symbols in square brackets corresponding events in the relevant state space. Thus, let

$$[\theta_0] = \left\{(\omega, \sigma_{-i}) : \vartheta_0\left(\omega\right) = \theta_0\right\},$$

$$[t] = \left\{(\omega, \sigma_{-i}) : \tau\left(\omega\right) = t\right\},$$

$$[a_{-i}] = \left\{(\omega, \sigma_{-i}) : \sigma_{-i}\left(\tau_{-i}\left(\omega\right)\right) = a_{-i}\right\},$$

$$[\theta_0, t, a_{-i}] = [\theta_0] \cap [t] \cap [a_{-i}],$$

and

$$p\left([\theta_0, t]\right) = p\left(\left\{\omega : \vartheta_0\left(\omega\right) = \theta_0, \tau\left(\omega\right) = t\right\}\right).$$

For every $t_i \in T_i$, construct a function $\varphi_{t_i} : \Theta_0 \times T_{-i} \to \Delta\left(A_{-i}\right)$ in the following way: if $p\left[\theta_0, t\right] > 0$,

$$\varphi_{t_i}\left(\theta_0, t_{-i}\right)\left(a_{-i}\right) = \frac{\sum_{(\omega,\sigma_{-i})\in[\theta_0,t,a_{-i}]} \mu\left(\omega, \sigma_{-i}\right)}{p\left([\theta_0, t]\right)};$$

if $p\left([\theta_0, t]\right) = 0$,

$$\varphi_{t_i}\left(\theta_0, t_{-i}\right)\left(a_{-i}\right) = \begin{cases} \frac{1}{\left|ICR_{-i}^{k-1,BG}(t_{-i})\right|}, & \text{if } a_{-i} \in ICR_{-i}^{k-1,BG}\left(t_{-i}\right), \\ 0, & \text{if } a_{-i} \notin ICR_{-i}^{k-1,BG}\left(t_{-i}\right). \end{cases}$$

For every t_i and pair (ω, a_{-i}), let

$$\hat{\mu}_{t_i}(\omega, a_{-i}) = p(\omega|t_i) \cdot \varphi_{t_i}(\vartheta_0(\omega), \tau_{-i}(\omega))(a_{-i}).$$

Observe that:

$$\begin{aligned}\sum_{(\omega,\sigma_{-i})\in[\theta_0,t,a_{-i}]} \mu(\omega, \sigma_{-i}) &= p([\theta_0, t]) \cdot \varphi_{t_i}(\theta_0, t_{-i})(a_{-i}) \\ &= p(t_i)\, p(\theta_0, t_{-i}|t_i) \cdot \varphi_{t_i}(\theta_0, t_{-i})(a_{-i}).\end{aligned}$$

Thus:

$$\begin{aligned}&\sum_{\omega,\sigma_{-i}} \mu^i(\omega, \sigma_{-i})\, \hat{u}_i(\omega, \sigma_i(\tau_i(\omega)), \sigma_{-i}(\tau_{-i}(\omega))) \\ &= \sum_{\theta_0,t,a_{-i}} \sum_{(\omega,\sigma_{-i})\in[\theta_0,t,a_{-i}]} \mu(\omega, \sigma_{-i})\, \hat{u}_i(\omega, \sigma_i(\tau_i(\omega)), \sigma_{-i}(\tau_{-i}(\omega))) \\ &= \sum_{t_i} p(t_i) \sum_{\theta_0,t_{-i}} p(\theta_0, t_{-i}|t_i) \sum_{a_{-i}} \varphi_{t_i}(\theta_0, t_{-i})(a_{-i}) \cdot u_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), \sigma_i(t_i), a_{-i}).\end{aligned}$$

Since $\sigma_i \in r_i(\mu^i)$ and $p(t_i) > 0$, we conclude that $\sigma_i(t_i) \in r_i(t_i, \hat{\mu}_{t_i})$. Finally, if $p([\theta_0, t]) = 0$, by construction $\varphi_{t_i}(\vartheta_0(\omega), t_{-i})(a_{-i}) > 0$ only if $a_{-i} \in ICR_{-i}^{k-1,BG}(t_{-i})$. If instead, $p([\theta_0, t]) > 0$, the same result follows from the definition of *ex ante* correlated rationalizability and the inductive hypothesis. Thus, for every t_i, we can conclude that $\sigma_i(t_i) \in ICR_i^{k,BG}(t_i)$ and, consequently, that:

$$ACR_i^{k,BG} \subseteq \left\{\sigma_i \in \Sigma_i : \forall t_i \in \Theta_i,\ \sigma_i(t_i) \in ICR_i^{k,BG}(t_i)\right\}.$$

Now we will prove the other inclusion. For every $t_i \in T_i$, let (a_{t_i}, t_i) be a pair such that $a_{t_i} \in ICR_i^{k,BG}(t_i)$. Thus for every a_{t_i} we can find a rationalizing belief $\mu_{a_{t_i}} \in \Delta(\Omega \times A_{-i})$ and a function $\varphi_{a_{t_i}} : \Theta_0 \times T_{-i} \to \Delta(A_{-i})$ satisfying the properties stated in the iterative definition of *interim* correlated rationalizability. Now construct belief $\hat{\mu}_i \in \Delta(\Omega \times \Sigma_{-i})$ as follows: for every pair (ω, σ_{-i}),

$$\hat{\mu}_i(\omega, \sigma_{-i}) = \frac{p(\omega) \cdot \varphi_{a_{\tau_i(\omega)}}(\vartheta_0(\omega), \tau_{-i}(\omega))(\sigma_{-i}(\tau_{-i}(\omega)))}{\left|\left\{\sigma'_{-i} \in ACR_{-i}^{k-1,BG} : \sigma'_{-i}(\tau_{-i}(\omega)) = \sigma_{-i}(\tau_{-i}(\omega))\right\}\right|}$$

whenever the denominator is positive,[55] and $\hat{\mu}_i(\omega, \sigma_{-i}) = 0$ otherwise. It is immediate to see that $\hat{\mu}_i(\omega, \sigma_{-i}) > 0$ only if $\sigma_{-i} \in ACR_{-i}^{k-1,BG}$. Now, let

$$[\omega, a_{-i}] = \left\{ \left(\omega', \sigma_{-i}\right) : \omega' = \omega,\ \sigma_{-i}\left(\tau_{-i}(\omega)\right) = a_{-i} \right\}.$$

With this, if $\left\{\sigma'_{-i} \in ACR_{-i}^{k-1,BG} : \sigma'_{-i}\left(\tau_{-i}(\omega)\right) = a_{-i}\right\} \neq \emptyset$, then

$$\sum_{(\omega',\sigma_{-i})\in[\omega,a_{-i}]} \hat{\mu}_i\left(\omega', \sigma_{-i}\right) =$$

$$\frac{\sum_{\sigma_{-i}\in ACR_{-i}^{k-1,BG}:\sigma_{-i}(\tau_{-i}(\omega))=a_{-i}} p(\omega)\cdot\varphi_{a_{\tau_i(\omega)}}\left(\vartheta_0(\omega), \tau_{-i}(\omega)\right)(a_{-i})}{\left|\left\{\sigma'_{-i} \in ACR_{-i}^{k-1,BG} : \sigma'_{-i}\left(\tau_{-i}(\omega)\right) = a_{-i}\right\}\right|} =$$

$$p(\omega)\cdot\varphi_{a_{\tau_i(\omega)}}\left(\vartheta_0(\omega), \tau_{-i}(\omega)\right)(a_{-i}).$$

Also notice that for every (ω, a_{-i}), if $\left\{\sigma'_{-i} \in ACR_{-i}^{k-1,BG} : \sigma'_{-i}\left(\tau_{-i}(\omega)\right) = a_{-i}\right\} = \emptyset$, the inductive hypothesis implies that $a_{-i} \notin ICR_{-i}^{k,BG}\left(\tau_{-i}(\omega)\right)$ and consequently $\varphi_{a_{\tau_i(\omega)}}\left(\vartheta_0(\omega), \tau_{-i}(\omega)\right)(a_{-i}) = 0$. Therefore, for every (ω, a_{-i}),

$$\sum_{(\omega',\sigma_{-i})\in[\omega,a_{-i}]} \hat{\mu}_i(\omega, \sigma_{-i}) = p(\omega)\cdot\varphi_{a_{\tau_i(\omega)}}\left(\vartheta_0(\omega), \tau_{-i}(\omega)\right)(a_{-i})$$

and

$$\begin{aligned}
\sum_{\sigma_{-i}} \hat{\mu}_i(\omega, \sigma_{-i}) &= \sum_{a_{-i}} \sum_{(\omega',\sigma_{-i})\in[\omega,a_{-i}]} \hat{\mu}_i(\omega, \sigma_{-i}) \\
&= \sum_{a_{-i}} p(\omega)\cdot\varphi_{a_{\tau_i(\omega)}}\left(\vartheta_0(\omega), \tau_{-i}(\omega)\right)(a_{-i}) = p(\omega).
\end{aligned}$$

Notice that we have already proved two of the properties in the definition of *ex-ante* correlated rationalizability. Finally, for every $\sigma_i \in \Sigma_i$

$$\begin{aligned}
&\sum_{\omega,\sigma_{-i}} \hat{\mu}^i(\omega, \sigma_{-i})\, \hat{u}_i\left(\omega, \sigma_i\left(\tau_i(\omega)\right), \sigma_{-i}\left(\tau_{-i}(\omega)\right)\right) \\
&= \sum_{\omega,a_{-i}} p(\omega)\cdot\varphi_{a_{\tau_i(\omega)}}\left(\vartheta_0(\omega), \tau_{-i}(\omega)\right)(a_{-i})\cdot u_i\left(\vartheta_0(\omega), \vartheta_i\left(\tau_i(\omega)\right), \vartheta_{-i}\left(\tau_{-i}(\omega)\right), \sigma_i\left(\tau_i(\omega)\right), a_{-i}\right).
\end{aligned}$$

[55] That is, if $\left\{\sigma'_{-i} \in ACR_{-i}^{k-1,BG} : \sigma'_{-i}\left(\tau_{-i}(\omega)\right) = \sigma_{-i}\left(\tau_{-i}(\omega)\right)\right\} \neq \emptyset$.

Since $p(t_i) > 0$ for every t_i, the decision function σ_i defined by $\sigma_i(t_i) = a_{t_i}$ for every t_i is such that $\sigma_i \in r_i(\hat{\mu}^i)$. We conclude that $\sigma_i \in ACR_i^{k,BG}$. Since pairs (a_{t_t}, t_i) were chosen arbitrarily, we can write:

$$ACR_i^{k,BG} \supseteq \left\{\sigma_i \in \Sigma_i : \forall t_i \in \Theta_i,\ \sigma_i(t_i) \in ICR_i^{k,BG}(t_i)\right\}.$$

The statement of the theorem follows by induction. ■

8.8.2 The Electronic Mail Game

Consider the following incomplete information game where *player 1* (Rowena) *knows the true payoff matrix*, whereas 2 (Colin) does not know it. The probabilities and the payoffs are represented in the table below:

$(1-\lambda)$:

θ^α	a	b
a	M, M	$1, -L$
b	$-L, 1$	$0, 0$

λ:

θ^β	a	b
a	$0, 0$	$1, -L$
b	$-L, 1$	M, M

$$L > M > 1, \qquad \lambda < 1/2$$

If we had complete information, (a, a) would be the dominant strategy equilibrium for matrix θ^α, and (b, b) would be the Pareto efficient equilibrium for the matrix θ^β. The second matrix has also another equilibrium, (a, a).

Notice that the game has distributed knowledge; thus, our previous analysis implies that the set of *interim* rationalizable actions is equivalent to the set of *interim* correlated rationalizable actions. In particular, it can be easily verified that any Bayesian game representing this situation would have only one *interim* rationalizable profile: (aa, a). As a matter of fact, for type θ^α of Rowena a is dominant. If Colin believes that Rowena is rational, so that type θ^α chooses a, then $a_2 = a$ yields a larger expected utility than $a_2 = b$. More precisely let $\mu^2 \in \Delta(aa, ab, ba, bb)$ be Colin's conjecture about Rowena's strategy and let $u_2(\mu^2, a_2)$ be the expected utility of action $a_2 \in \{a, b\}$ given conjecture μ^2. Since $\mu^2(ba) = \mu^2(bb) = 0$

(as θ^α certainly chooses a), we obtain

$$\begin{aligned}
u_2(\mu^2, a) &= (1-\lambda)u_2(\theta^\alpha, a, a) + \lambda\mu^2(aa)u_2(\theta^\beta, a, a) + \lambda\mu^2(ab)u_2(\theta^\beta, b, a) \\
&\geq (1-\lambda)M \\
&> -(1-\lambda)L + \lambda M \\
&\geq (1-\lambda)u_2(\theta^\alpha, a, b) + \lambda\mu^2(aa)u_2(\theta^\beta, a, b) + \lambda\mu^2(ab)u_2(\theta^\beta, b, b) \\
&= u_2(\mu^2, b).
\end{aligned}$$

Hence, the only rationalizable action for Colin is a. This implies that also for type θ^β the only *interim* rationalizable choice is a.

Consider now the following more complex variation of the game.[56] Obviously the players would find it convenient to communicate and coordinate on (a, a) if $\theta = \theta^\alpha$ and on (b, b) if $\theta = \theta^\beta$. Let us then consider the following form of communication via electronic mail. Rowen and Colin sit in front of their respective computer screens. If the payoff-type of Rowena is θ^β, her computer, C_1, *automatically* sends a message to the computer of Colin, C_2. Furthermore, if computer C_i receives a message, it *automatically* sends a confirmation message to computer C_{-i}. However, any time a message is sent,it gets *lost* with probability ε and thus it does not reach the receiver.

This information structure yields a corresponding Bayesian game. A generic state of the world is given by a pair of numbers $\omega = (q, r)$ where $q = t_1$ is the number of messages *sent* by C_1 and $r = t_2$ is the number of messages received (and therefore also sent) by C_2. Hence, the set of states of the world is given by

$$\begin{aligned}
\Omega &= \{(0,0), (1,0), (1,1), (2,1), (2,2), (3,2), (3,3), ...\} \\
&= \{(q, r) \in \mathbb{N} \times \mathbb{N} : q = r \text{ or } q = r + 1\}.
\end{aligned}$$

If the state is $(q, q-1)$ it means that the last message (among those sent either by C_1 or by C_2) has been sent by C_1. If the state is (r, r) (with $r > 0$) it means that the last message sent by C_1 has reached C_2, but the confirmation by C_2 has not reached C_1. The signal functions are $\tau_1(q, r) = q$, $\tau_2(q, r) = r$. The function that determines θ is $\vartheta_1(0) = \theta^\alpha$, $\vartheta_1(q) = \theta^\beta$ if $q > 0$.[57] There is a common prior given by $p(0, 0) = (1-\lambda)$,

[56] See [58] (or the textbook by Osborne and Rubinstein [53], pp. 81-84). The analysis in terms of *interim* rationalizability is not contained in the original work.

[57] Θ_2 is a singleton, hence ϑ_2 is a constant.

$p(r+1,r) = \lambda(1-\varepsilon)^{2r}\varepsilon$, $p(r+1,r+1) = \lambda(1-\varepsilon)^{2r+1}\varepsilon$ for any $r \geq 0$ (that is, $p(q,r) = \lambda(1-\varepsilon)^{q+r-1}\varepsilon$ for any $(q,r) \in \Omega \setminus \{(0,0)\}$). This information is summed up in the following table:

$t_1{=}q \backslash t_2{=}r$	0	1	2	3	4	5	…
$0,\ \theta^\alpha$	$1-\lambda$	–	–	–	–	–	…
$1,\ \theta^\beta$	$\lambda\varepsilon$	$\lambda(1-\varepsilon)\varepsilon$	–	–	–	–	…
$2,\ \theta^\beta$	–	$\lambda(1-\varepsilon)^2\varepsilon$	$\lambda(1-\varepsilon)^3\varepsilon$	–	–	–	…
$3,\ \theta^\beta$	–	–	$\lambda(1-\varepsilon)^4\varepsilon$	$\lambda(1-\varepsilon)^5\varepsilon$	–	–	…
$4,\ \theta^\beta$	–	–	–	$\lambda(1-\varepsilon)^6\varepsilon$	$\lambda(1-\varepsilon)^7\varepsilon$	–	…
$5,\ \theta^\beta$	–	–	–	–	$\lambda(1-\varepsilon)^8\varepsilon$	$\lambda(1-\varepsilon)^9\varepsilon$	…
…	…	…	…	…	…	…	…

The resulting beliefs, or conditional probabilities, are then determined as follows:

$$\begin{aligned} p_1[(0|0) &= 1, \\ p_1[(r|r+1) &= \frac{\lambda(1-\varepsilon)^{2r}\varepsilon}{\lambda(1-\varepsilon)^{2r}\varepsilon + \lambda(1-\varepsilon)^{2r+1}\varepsilon} = \frac{1}{2-\varepsilon}, \\ p_2[(0|0) &= \frac{1-\lambda}{1-\lambda+\lambda\varepsilon}, \\ p_2[(r+1|r+1) &= \frac{\lambda(1-\varepsilon)^{2r+1}\varepsilon}{\lambda(1-\varepsilon)^{2r+1}\varepsilon + \lambda(1-\varepsilon)^{2r+2}\varepsilon} = \frac{1}{2-\varepsilon}\ (r>0). \end{aligned}$$

In other words, any player—having received a certain number of messages—computes the probability that his confirmation message does not reach the other computer, which is $\frac{1}{2-\varepsilon} > \frac{1}{2}$.

If ε is very small, the Bayesian game is in some sense "close" to the game with common knowledge of θ. First, notice that in all states (q,r) with $r > 0$ both players know that the true state is θ^β. Moreover, for any n and any $\delta > 0$, there always exists an ε sufficiently small such that, given $\theta = \theta^\beta$, the probability that there exists mutual knowledge of degree n of the true value of θ is bigger than $1-\delta$.[58] Given that (b,b) is an equilibrium of the complete information game in which $\theta = \theta^\beta$, we could be induced

[58] For instance, consider $n = 2$. For any $q \geq 2$, in state (q,r) ($r = q$, or $r = q-1$) every player knows that $\theta = \theta^\beta$ and every player knows that the other knows it, therefore

to think that if ε is very small, then action b is *interim* rationalizable for some type of Rowena and Colin. But the following result shows that this intuition is incorrect:

Proposition 1. *In the Electronic Mail game there exists a unique* interim *rationalizable profile: every type of every player chooses action a.*

The crucial point in the proof is to realize that when a player i knows that $\theta = \theta^\beta$, but s/he is sure that $-i$ plays a whenever s/he has not received her/his last message, then i prefers to play a. On the other hand, it is easy to show that when i *does not know* that $\theta = \theta^\beta$ then a is the only rationalizable action. It follows by induction that a is the only rationalizable action in all states. Here is the formal argument:

Proof of Proposition 1. Clearly, the only rationalizable action for $t_1 = 0$ is a, as this type of Rowena knows that in the true payoff matrix a is dominant. An argument similar to the one used for the simple case discussed above shows that the only rationalizable action for $t_2 = 0$ is also a.

Consider now the rationalizable choices for the types $t_i = r > 0$, that is for those types for which player i *knows that* $\theta = \theta^\beta$. We first prove two intuitive intermediate steps:

(i) *If the only rationalizable action for type $t_2 = r - 1$ is a, then the only rationalizable action for type $t_1 = r$ is a.* Any rationalizable action for $t_1 = r$ needs to be justified by a rationalizable conjecture about Colin (see Theorem 2). By assumption, such conjecture must assign probability 1 to the set of profiles $\sigma_2 \in \{a, b\}^{T_2}$ such that $\sigma_2(r-1) = a$. Hence, the expected utility for type $t_1 = r$ if he chooses a is at least 0 (this value is realized if type $t_2 = r$ also chooses a), whereas the expected utility from choosing b is at most $-p_1(r-1|r)\,L + p_1(r|r)\,M$ (this value is realized if type $t_2 = r$ chooses b). Since $p_1(r-1|r) = \frac{1}{2-\varepsilon} > \frac{1}{2}$ and $L > M$, it follows that $0 > -p_1(r-1|r)\,L + p_1(r|r)\,M$.

An analogous argument[59] shows that:

there is mutual knowledge of degree two of the true payoff matrix. The probability of being in one of these states conditional on the event $\theta = \theta^\beta$ is

$$1 - \frac{p(\{(1,0),(1,1)\})}{\lambda} = 1 - 2\varepsilon + \varepsilon^2.$$

To guarantee that such probability is larger than $1 - \delta$ we need $\varepsilon < 1 - \sqrt{1-\delta}$.

[59] Just reverse the roles and modify indexes accordingly.

(ii) *If the only rationalizable action for type* $t_1 = r$ *is* a*, the only rationalizable action for type* $t_2 = r$ *is also* a*.*

From steps (i) and (ii) it follows that, for every $r > 0$, *if the only rationalizable profile in state* $(r-1, r-1)$ *is* (a, a) *then the only rationalizable profile in state* $(r, r-1)$ *is* (a, a); *if the only rationalizable profile in state* $(r, r-1)$ *is* (a, a) *then the only rationalizable profile in state* (r, r) *is* (a, a). Since we have shown that the only rationalizable profile in state $(0, 0)$ is (a, a), it follows by induction that (a, a) is the only rationalizable profile in every state. ∎

Part II

Sequential Games

In this part we extend the analysis of strategic thinking to interactive decision situations where players move sequentially. As the play unfolds, players obtain new information. We consider a world inhabited by Bayesian players. When a player receives a piece of information that was possible (had positive probability) according to his previous beliefs, then he simply updates his beliefs according to the standard rules of conditional probabilities. When the new piece of information is completely unexpected, the player forms new subjective beliefs.

The observation of some previous moves by the co-players may provide information about the strategies they are implementing and/or their type. How such information is interpreted depends, for example, on beliefs about the co-players rationality, or whether they are carrying out their plan. This introduces a new fascinating dimension to strategic thinking.

As in Part I, we first analyze games with complete information and then move on to games with incomplete information. Chapter 9 provides a mathematical description of **multistage games with observable actions**, i.e., games where at the beginning of each stage all the actions taken in previous stages are publicly observed and hence are common knowledge. We then define the central concepts of **strategy** and **strategic form**. Chapter 10 analyzes rational planning for given beliefs about the behavior of the co-players. For games with finite horizon (i.e., games that end in bounded, finite time), we represent rational planning with the **folding back** procedure. It turns out that a strategy is consistent with folding-back if and only if it is **sequentially optimal,** and if and only if it satisfies the **one-deviation** property. Sequential optimality and the one-deviation property, which are equivalent in finite-horizon games, can be defined also for games with infinite horizon, and the equivalence between sequential optimality and the one-deviation property—called **one-deviation principle**—extends to a very large class of "sufficiently regular" infinite-horizon games. Chapter 11 puts forward notions of rationalizability for multistage games justified by different assumptions about strategic reasoning. In many games of interest, these solution concepts admit useful strategic-form characterizations. Chapter 12 analyzes equilibrium concepts. A first definition of equilibrium is given by the Nash equilibrium of the strategic form. This, however, is unsatisfactory. The more demanding concept of **subgame perfect equilibrium** is based on the details of the sequential representation of the multistage game: it requires that the strategy of each player be sequentially optimal given correct conjectures about the strategies of the co-players. The one-deviation principle implies that, in finite games where only one player is active at each stage and payoffs are "generic," there is only one subgame perfect equilibrium that can be computed with the so called **backward induction** algorithm, a kind of inter-personal folding-back procedure. In games with finite horizon, one can use a "case-by-case backward induction" method to find the subgame perfect equilibria. Chapters 13 and 14 apply the one-deviation principle to the analysis of repeated games and bargaining games. Chapters 15 and 16 extend the analysis to sequential games with incomplete information and with imperfectly observable actions.

9

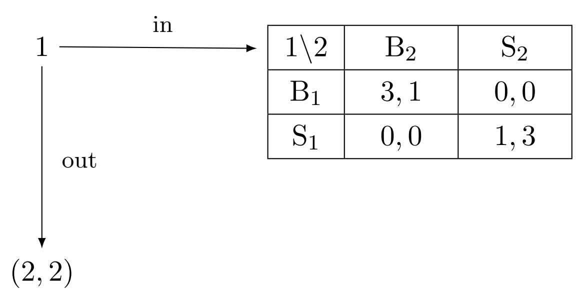
Multistage Games with Complete Information

We start by providing a mathematical definition of multistage games whereby, at the beginning of each stage, all the actions taken in previous stages are publicly observed, hence are common knowledge (Section 9.1). In a single stage, actions are chosen simultaneously, but some players may be inactive. Inactive players will be formally represented as players with only one feasible action. It is possible that in a given stage all the players but one are inactive. The set of feasible actions in a given stage may be endogenous, that is, it may depend on the actions taken in previous stages. The length of the game may also be endogenous. Armed with such formal notation, we move on to define **strategies** and the (reduced) **strategic form** of a game (Section 9.3). We can interpret a strategy as (i) the complete description of the contingent behavior of a player, and (ii) the contingent plan of a player. The strategic form of a game is based on the first interpretation. The analysis of rational planning (formally developed in Chapter 10) considers the second interpretation. For a player who would always carry out his plan, (i) and (ii) coincide. Any solution concept for static games can be applied to the strategic form of a multistage game and thus yields a "candidate solution." In Section 9.3.1, we will discuss whether some solutions concepts for static games—such as rationalizability or Nash equilibrium—make sense as solutions of a multistage game, when applied to its strategic form. Finally, we extend the analysis by introducing two different notions of randomized strategic behavior, **mixed strategies** and **behavior strategies**; the latter are better suited to define subgame perfect equilibrium in randomized strategies, as we will see in Chapter 12.

9.1 Preliminary Definitions

It is useful to start with a preliminary example before we plunge into the mathematical representation of multistage games with observable actions.

Example 38. Ann (player 1) and Bob (player 2) play the **Battle of the Sexes (BoS) with an Outside Option**. Ann first chooses between playing the BoS with Bob (action in) or not (action out). If she chooses in, this choice becomes common knowledge and then the simultaneous-move "subgame" BoS is played. If she chooses out the game ends. Payoff are displayed in Figure 9.1 below.



$1\backslash 2$	B_2	S_2
B_1	$3, 1$	$0, 0$
S_1	$0, 0$	$1, 3$

Figure 9.1: Battle of the Sexes with an Outside Option.

The set of feasible actions of a player depends on the position reached: Ann's feasible set is $\{\mathrm{in}, \mathrm{out}\}$ at the beginning of the game, and it is $\{\mathrm{B}_1, \mathrm{S}_1\}$ if she moves in. At the beginning of the game, Bob can only "wait and see" what Ann does; therefore his feasible set is the singleton $\{\mathrm{wait}\}$. If Ann moves in, Bob's feasible set is $\{\mathrm{B}_2, \mathrm{S}_2\}$. There are 5 possible plays of the game: either the pair of actions $(\mathrm{out}, \mathrm{wait})$ is played and the game ends, or one of the following 4 sequences of actions pairs is played: $((\mathrm{in}, \mathrm{wait}), (\mathrm{B}_1, \mathrm{B}_2))$, $((\mathrm{in}, \mathrm{wait}), (\mathrm{B}_1, \mathrm{S}_2))$, $((\mathrm{in}, \mathrm{wait}), (\mathrm{S}_1, \mathrm{B}_2))$ and $((\mathrm{in}, \mathrm{wait}), (\mathrm{S}_1, \mathrm{S}_2))$. A pair of payoffs $(u_1(z), u_2(z))$ is associated with each possible play z. ▲

A possible play of the game is called **terminal history**. Possible partial plays like (in, wait) are called **non-terminal** (or **partial**) histories. The rules of the game specify what sequences of action profiles are terminal or non-terminal histories. For each terminal history, viz., z,[1] the rules of the game specify an outcome (consequence) $y = g(z)$, e.g., a distribution of money among the players. Each player i assigns to each outcome y utility $v_i(y)$. As in the first part, v_i represents i's preferences over lotteries of outcomes (consequences) $\lambda \in \Delta(Y)$ by way of expected utility calculations. Given the rules and the utility functions, each terminal history z is mapped to a profile of "payoffs" (induced utilities) $(u_i(z))_{i\in I} = (v_i(g(z)))_{i\in I}$. For example, the payoff pair attached to terminal history $z = ((\text{in}, \text{wait}), (\text{S}_1, \text{S}_2))$ is $(u_1(z), u_2(z)) = (3, 1)$. Under the complete information assumption, the rules of the game (including the outcome function g) and players' preferences over lotteries of consequences (represented by the utility functions v_i, $i \in I$) are common knowledge. Thus, also the payoff functions $z \mapsto u_i(z)$ are common knowledge. In this chapter, we assume complete information and we directly specify the payoff functions. But, as in the analysis of static games, it is important to keep in mind that such payoff functions are derived from preferences and the outcome function g.

When discussing a *specific example* like the BoS with an Outside Option it would be simpler to represent the possible plays as sequences like (out), $((\text{in}), (\text{B}_1, \text{B}_2))$, $((\text{in}), (\text{B}_1, \text{S}_2))$, etc. But we use this awkward notation for a reason. Having each inactive player automatically choose "wait" simplifies the *abstract* notation for *general* games: possible plays of the game are just sequences of action profiles $(a^1, a^2, ...)$ where each element a^t of the sequence is a profile $(a_i^t)_{i\in I}$ and there is no need to keep track in the formal notation of who is active. Thus, if A denotes the set of all action profiles, histories are just sequences of elements from set A. To allow for the theoretical possibility of games that can go on forever, we also consider infinite sequences of elements from A. The rules of the game specify which sequences are possible, i.e., they specify the set of "histories."

Since sequences of elements from a given domain are a crucial ingredient of the formal representation of a game, it is useful to introduce some preliminary concepts and notation about sequences.

[1] Since z is the last (i.e., terminal) letter of the alphabet, it seems like a good choice as a symbol to denote terminal histories.

9.1.1 Sequences and Trees

Fix a set X. The set of finite sequences of length ℓ of elements from set X is the ℓ-fold Cartesian product

$$X^{\ell} = \underbrace{X \times ... \times X}_{\ell \text{ times}}.$$

Such sequences can be thought as functions from the set of the first ℓ positive integers, $\{1, ..., \ell\}$ to X; thus, X^{ℓ} is isomorphic to the set of functions $X^{\{1,...,\ell\}}$. Now consider the union of all such sets, i.e., the set of all finite sequences of elements from X, plus the singleton $X^0 = \{\varnothing\}$ containing only the **empty sequence** $\varnothing$ (a convenient mathematical object analogous to the empty set $\emptyset$).[2] With a descriptive notation, such union is represented by the symbol $X^{<\mathbb{N}_0}$, i.e.,

$$X^{<\mathbb{N}_0} = \bigcup_{\ell \in \mathbb{N}_0} X^{\ell}$$

where $\mathbb{N}_0 = \{0, 1, 2, ...\}$ is the set of natural numbers including 0. It may help to think of X as an *alphabet* for some language Λ. Then, the set of *words* in language Λ is a subset of $X^{<\mathbb{N}_0}$. The *rules* of language Λ determine which elements of $X^{<\mathbb{N}_0}$ are words of Λ.

Similarly, the rules of a given game Γ determine which sequences of actions profiles are feasible **histories.** But, unlike words in natural languages, some histories of games may have an infinite length. This means that the game does not necessarily end in a finite number of stages. This is a useful abstraction. Consider, for example, a bargaining game where the parties go on making offers and counteroffers until they reach a binding agreement: it is possible that they never agree and hence bargain forever. Countably infinite sequences of elements from X are like functions from $\mathbb{N} = \{1, 2, ...\}$ to X. Therefore the set of all such sequences is denoted $X^{\mathbb{N}}$:

$$X^{\mathbb{N}} = \{(x^k)_{k=1}^{\infty} : \forall k \in \mathbb{N}, x^k \in X\}.$$

The set of all finite and infinite sequences from X (empty sequence included) is

$$X^{\leq \mathbb{N}_0} = X^{<\mathbb{N}_0} \cup X^{\mathbb{N}}.$$

[2] Similarly, the power set of X contains the empty set, that is, $\emptyset \in 2^X$, or equivalently $\{\emptyset\} \subseteq 2^X$.

The histories of a game form a **rooted tree**, i.e., a partially ordered subset[3] $\bar{H} \subseteq X^{\leq \mathbb{N}_0}$ where (i) the order relation $x \prec y$ is "x is a **prefix** (initial subsequence) of y" (by convention, $\varnothing$ is a prefix of every sequence $\left(x^k\right)_{k=1}^{\ell}$, $\ell = 1, 2, ..., \infty$), (ii) every prefix of a sequence in $\bar{H}$ (including the empty sequence) is also a sequence in $\bar{H}$, so that the empty sequence is the root of the tree, and (iii) (in games that may never end) if $z \in X^{\mathbb{N}}$ and every (finite) prefix of z is in $\bar{H}$ then also $z \in \bar{H}$.[4]

9.2 Multistage Games with Observable Actions

Consider a game that proceeds through stages. At each stage there is a subset of active players, and this set may depend on previous moves; each active player chooses an action in a feasible set with two or more alternatives, and the choices of the active players are simultaneous. The rules of the game are such that at the beginning of each stage the action profiles chosen in previous stages become public information. The rules determine which actions are feasible for each player according to previous choices of every player. The only feasible action of an inactive player is to wait, hence his feasible set is a singleton. When the game ends, players have empty feasible sets: not only they cannot choose between alternatives, they also stop waiting. The consequences for the players implied by the rules of the game, in general, may depend on the actions taken since the game started, and on the order in which they were taken, that is, the consequences depend on the sequence of action profiles that led to the end of the game. Game forms with simultaneous moves, or static game forms, are a special case whereby the first actions profile ends the game.

As we did in the case of simultaneous moves games, we do not represent the rules directly, but rather we describe the players' set, their feasible sets of actions (including how they depend on previous moves) and the consequences of their actions. We do this by providing

[3] A set P is partially ordered by the binary relation $\prec \subseteq P \times P$ if $\prec$ is transitive and asymmetric, that is, for all $p, q, r \in P$, $(p \prec q \wedge q \prec r)$ implies $p \prec r$, and $p \prec q$ implies $\neg(q \prec p)$. The reflexive closure of $\prec$ is the relation $\preceq$ given by $(\prec \cup =)$, that is, $p \preceq q$ if either $p \prec q$ or $p = q$. If $\prec$ is transitive and asymmetric, then $\preceq$ is transitive, reflexive, and antisymmetric (which means that $p \preceq q \wedge q \preceq p$ implies $p = q$). A set P can be equivalently said to be partially ordered by $\preceq$ if $\preceq$ is transitive, reflexive, and antisymmetric. See [52, A.1.4].

[4] Note that (ii) and (iii) do not hold for the words of natural languages.

an abstract representation of the possible sequences of actions profiles, called "histories;" from this we obtain a game tree with a set of terminal histories; then we introduce a consequence map from terminal histories to collective consequences, thus obtaining a multistage game form, which is a mathematical description of the rules of the game. Finally, we introduce players' preferences and obtain a multistage game. If the rules of the game and players' preferences are common knowledge, this provides a complete mathematical description of those aspects if the game that we deem relevant for the analysis of strategic interaction.

A **multistage game tree** with observable actions is a structure

$$\langle I, (A_i, \mathcal{A}_i(\cdot))_{i\in I}\rangle$$

given by the following components:

- For each $i \in I$, A_i is a nonempty set of **potentially feasible actions**.
- Let $A = \times_{i\in I} A_i$ and consider the set $A^{<\mathbb{N}_0}$ of finite sequences of action profiles; then, for each $i \in I$, $\mathcal{A}_i(\cdot) : A^{<\mathbb{N}_0} \rightrightarrows A_i$ is a **feasibility correspondence** that assigns to each finite sequence of action profiles $h^\ell = (a^t)_{t=1}^{\ell}$ the set $\mathcal{A}_i(h^\ell)$ of actions of i that are feasible immediately after h^ℓ. It is assumed that $\mathcal{A}_i(\varnothing) \neq \emptyset$ and, for all $h \in A^{<\mathbb{N}_0}$, $\mathcal{A}_i(h) = \emptyset$ if and only if $\mathcal{A}_j(h) = \emptyset$ for every $j \in I$ (the reason for this assumption will be explained below).

The reason why we say that this structure has observable actions is that this is the intended interpretation. In particular, if the feasible set $\mathcal{A}_i(h)$ depends on h, the assumption that h is observed as soon as it occurs implies that player i always knows what his feasible set of action is. Indeed, we assume more, i.e., that when a history h occurs it becomes common knowledge that h has occurred. If this were not the case then we would have to describe what players know about what other players may have observed.

Let $\mathcal{A}(h) = \times_{i\in I} \mathcal{A}_i(h)$ denote the set of feasible action profiles given $h \in A^{<\mathbb{N}_0}$. A sequence $(a^t)_{t=1}^{\ell}$ ($\ell = 1, 2, ..., \infty$) is a (feasible) **history** if $a^1 \in \mathcal{A}(\varnothing)$ and $a^{t+1} \in \mathcal{A}(a^1, a^2, ..., a^t)$ for each positive integer $t < \ell$. Thus, a history is a sequence of action profiles whereby each action profile is feasible given the previous ones. By convention, the empty sequence—denoted $\varnothing$—is a history. Let $\bar{H} \subseteq A^{\leq \mathbb{N}_0}$ denote the set of histories. A history $z = (a^t)_{t=1}^{\ell} \in \bar{H}$ is **terminal** if either $z \in A^{\mathbb{N}}$ (i.e., z is an infinite history) or $\mathcal{A}(z) = \emptyset$. Let

$$Z = \left\{ z \in \bar{H} : z \in A^{\mathbb{N}} \text{ or } \mathcal{A}(z) = \emptyset \right\}$$

denote the set of terminal histories, and let

$$H = \bar{H} \backslash Z$$

denote the set of **non-terminal** (or **partial**) histories.[5]

Consider the restriction to $\bar{H}$ of the prefix-of relation $\prec$ on $A^{\leq \mathbb{N}_0}$: let $h = \left(a^1, ..., a^k\right)$ and $\bar{h} = \left(\bar{a}^1, ..., \bar{a}^{\ell}\right)$, then $h \prec \bar{h}$ if $k < \ell$ and $\left(a^1, ..., a^k\right) = \left(\bar{a}^1, ..., \bar{a}^k\right)$, that is, $\bar{h} = \left(a^1, ..., a^k, \bar{a}^{k+1}, ...\right)$; in this case we say that h **precedes** $\bar{h}$, or—equivalently—that $\bar{h}$ **follows** h. We write $h \preceq \bar{h}$ if either $h \prec \bar{h}$ or $h = \bar{h}$, and we say that history h **weakly precedes** history $\bar{h}$ (or that $\bar{h}$ **weakly follows** h). We also write $\bar{h} \succ h$ ($\bar{h} \succeq h$) to mean that $\bar{h}$ (weakly) follows h. The following result implies that $\bar{H}$ ordered by the precedence relation $\prec$ (restricted to $\bar{H}$) is a tree with distinguished root $\varnothing$:

Remark 38. *The set of histories $\bar{H}$ has the following properties:*
(1) $\varnothing \in \bar{H}$ and, for each $h \in \bar{H} \backslash \{\varnothing\}$, $\varnothing \prec h$;
(2) for each sequence $h \in A^{<\mathbb{N}_0}$ and each history $h' \in \bar{H}$, if $h \prec h'$ then $h \in \bar{H}$;
(3) for each infinite sequence $z \in A^{\mathbb{N}}$, if every predecessor (prefix) of z is in $\bar{H}$, then $z \in \bar{H}$.

Proof. (1) holds by convention. To verify (2), fix arbitrarily a history $h' = \left(a^t\right)_{t=1}^{\ell} \in \bar{H}$. Then, by definition, $a^1 \in \mathcal{A}(\varnothing)$ and $a^{t+1} \in \mathcal{A}(a^1, a^2, ..., a^t)$ for each positive integer $t < \ell$. If $h \prec h'$ then either $h = \varnothing \in \bar{H}$, or $h = \left(a^k\right)_{k=1}^{t}$ for some positive integer $t < \ell$. In the latter case, we must have $a^1 \in \mathcal{A}(\varnothing)$ and $a^{k+1} \in \mathcal{A}(a^1, a^2, ..., a^k)$ for each $k \in \{1, ..., t-1\}$. Therefore h must be a history as well, i.e., $h \in \bar{H}$. To

[5] It makes sense to assume that the effective range of the feasibility correspondence of a player is his action set, that is $\cup_{h \in H} \mathcal{A}_i(h) = A_i$. But this is not necessary for our analysis.

verify (3), let $z = \left(a^k\right)_{k=1}^{\infty}$ be such that $h = \left(a^k\right)_{k=1}^{\ell} \in \bar{H}$ for each $\ell \in \mathbb{N}$, i.e., for each prefix of z. Then $a^1 \in \mathcal{A}(\varnothing)$ and $a^{t+1} \in \mathcal{A}(a^1, a^2, ..., a^t)$ for each $t \in \mathbb{N}$, which implies that $z \in \bar{H}$. ■

Next we add to our description a set Y of outcomes, or consequences (e.g., monetary payoffs for all the players), and an **outcome** (or consequence) **function** $g : Z \to Y$, thus obtaining a **multistage game form** with observable actions

$$\langle I, Y, g, (A_i, \mathcal{A}_i(\cdot))_{i \in I} \rangle .$$

This structure describes the rules of the game, but not how each player would like this game to end. Even when $Y \subseteq \mathbb{R}^I$ and $g : Z \to Y$ describes monetary payoffs, it may be the case that players do not just care about their own money, and even in the latter case, we still lack a specification of risk attitudes. Thus, in order to analyze interaction between a specific group of individuals for given rules of the game, we add their personal preferences over lotteries, represented by a profile of von Neumann-Morgenstern utility functions $(v_i : Y \to \mathbb{R})_{i \in I}$ according to expected utility calculations.

With this, we obtain a **multistage game** with observable actions

$$\Gamma = \langle I, Y, g, (A_i, \mathcal{A}_i(\cdot), v_i)_{i \in I} \rangle .$$

For each player $i \in I$, the composition

$$u_i = v_i \circ g : Z \to \mathbb{R}$$

is called the **payoff function** of i. As we did for static games, we will often neglect the specification of the outcome function and utility functions, and look at a simplified, reduced representation $\langle I, (A_i, \mathcal{A}_i(\cdot), u_i)_{i \in I} \rangle$ showing only the payoff functions. In this chapter, we maintain the informal assumption of *complete information*: the rules of the game and players' preferences over lotteries are common knowledge. As we have seen for static games, if there is incomplete information we have to enrich our description of the game.

We consider both games that end within a given finite time and games that can last for an arbitrarily long time:

Definition 46. *Game Γ has **finite horizon** if $\bar{H} \subseteq \bigcup_{t=1}^{T} A^t$ for some $T \in \mathbb{N}$, otherwise Γ has **infinite horizon**. Game Γ is **finite** if $\bar{H}$ is finite, otherwise Γ is **infinite**. Game Γ is a **static**, or simultaneous-move game if $Z = \mathcal{A}(\varnothing)$. Game Γ is **compact-continuous** if A is a compact subset of a Euclidean space, $\bar{H}$ is compact (in the direct-sum product topology)*[6] *and u_i is continuous for each $i \in I$.*

Remark 39. *If Γ is finite, then Γ has finite horizon.*

Proof. We prove the result by contraposition, showing that if Γ has infinite horizon, then Γ is infinite. Suppose that, for each $T \in \mathbb{N}$, there is some history $\left(a^t\right)_{t=1}^{\ell} \in \bar{H}$ with length $\ell > T$, where $\ell \in \mathbb{N} \cup \{\infty\}$. Then, there are two (not mutually exclusive) cases: (i) $\bar{H}$ is not finite, which means that Γ is not finite, or (ii) $\bar{H}$ contains an infinite (necessarily terminal) history, $z = \left(a^t\right)_{t=1}^{\infty}$. In the latter case, by Remark 38, the countably infinite set of predecessors $\left(a^t\right)_{t=1}^{T}$ ($T \in \mathbb{N}$) of z is included in $\bar{H}$ as well; hence Γ is not finite. ■

For each $h \in \bar{H}$, we let $H(h)$ and $Z(h)$ respectively denote the set of non-terminal and terminal histories that weakly follow h:[7]

$$H(h) = \{h' \in H : h \preceq h'\},$$

$$Z(h) = \{z \in Z : h \preceq z\}.$$

Note that the set of histories $H(h) \cup Z(h)$ and the restriction of the precedence relation $\preceq$ on $H(h) \cup Z(h)$ form a sub-tree with root h.

For any non-terminal history $h \in H$, the sub-tree given by $H(h) \cup Z(h)$ and the restriction of g (hence, of each u_i) to sub-domain $Z(h)$ determine the **subgame with root** h, denoted by $\Gamma(h)$. The maximal length of histories in this subgame, called **height** of $\Gamma(h)$, is

$$L(\Gamma(h)) := \max_{z \in Z(h)} \ell(z) - \ell(h).$$

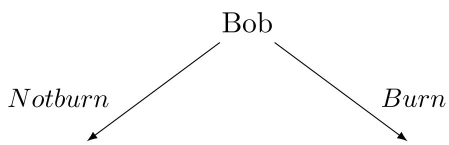
[6]For any $t \in \mathbb{N} \cup \{\infty\}$, we say that $h_n = \boldsymbol{\theta}\, a^{k,n}\big)_{k=1}^{t}$ converges to $\boldsymbol{\theta}\, a^{k}\big)_{k=1}^{t}$ (in the product topology), written $h_n \to h$, if $a^{k,n} \to a^k$ for each k. In the finite-dimentional case, this is the standard notion of convergence. The set $\bar{H}$ is closed (in the direct-sum product topology) if, for every $t \in \mathbb{N} \cup \{\infty\}$ and $(h_n)_{n=1}^{\infty} \in \boldsymbol{\theta}\, A^t\big)^{\mathbb{N}}$ with $\{h_n\}_{n=1}^{\infty} \subseteq \bar{H}$, $\lim_{n\to\infty} h_n = h$ implies $h \in \bar{H}$.

[7]If $h \in Z$, then $H(h) = \emptyset$ and $Z(h) = \{h\}$.

For notational convenience, we extend $L(\Gamma(\cdot))$ to terminal histories as well: $L(\Gamma(z)) = 0$ for each $z \in Z$.

We illustrate our notation and the definitions with a leading example: the "**Battle of the Sexes (BoS) with Dissipative Action.**"

Example 39. The game depicted in the picture below is the Battle of the Sexes preceded by an observed "money-burning" (dissipative) move by Bob:[8] if Bob chooses the dissipative action *Burn*, he burns 2 dollars.[9]



Ann\Bob	L	R
U	$4, 1$	$0, 0$
D	$0, 0$	$1, 4$

Ann\Bob	l	r
u	$4, -1$	$0, -2$
d	$0, -2$	$1, 2$

The game tree is described by the following elements:

– $\mathcal{A}(\varnothing) = \{w\} \times \{N, B\}$, $\mathcal{A}((w, N)) = \{U, D\} \times \{L, R\}$, $\mathcal{A}((w, B)) = \{u, d\} \times \{l, r\}$, where $\mathcal{A}(h) = \mathcal{A}_a(h) \times \mathcal{A}_b(h)$ for each history/node h, w is the "wait" action of an inactive player, and we omit one order of parentheses when this causes no confusion (e.g., we write $\mathcal{A}((w, B))$ instead of $\mathcal{A}(((w, B)))$ for the feasible action pairs at $h = ((w, B))$, a sequence of action pairs of length one);

– $H = \{\varnothing, ((w, N)), ((w, B))\}$;

– $Z = \{(w, N)\} \times \mathcal{A}((w, N)) \cup \{(w, B)\} \times \mathcal{A}((w, B))$, $\bar{H} = H \cup Z$;

– there are eight maximal chains[10] of histories/nodes starting with the root and ending with a terminal history, the following is an example:

$$\varnothing \prec ((w, B)) \prec ((w, B), (u, l)) \in Z;$$

on the other hand

$$((w, N)) \not\prec ((w, B), (u, l)).$$

[8]See, e.g., Ben-Porath and Dekel [20].

[9]Utility is measured in dollars. In the figure, payoff pairs follow the alphabetical order: the first number is Ann's payoff.

[10]A chain in a partially ordered set $(X, \preceq)$ is a subset $C \subseteq X$ that is totally ordered by $\preceq$.

▲

Sometimes the definition of games with observable actions takes the set $\bar{H}$ as primitive and postulates properties (1)-(3) of Remark 38 (cf. Osborne and Rubinstein [53, pp 89-90, 102.]). The most common definition of game with sequential moves takes the tree-structure as primitive, but deals with simultaneous actions in a somewhat arbitrary and un-intuitive way (see, e.g., Fudenberg and Tirole [34, pp 80-81]). We will take advantage of the tree-structure to provide graphical representations of some games, in particular the games where at each stage t at most one player has at least two feasible actions; such games are said to have "perfect information":

Definition 47. *Player i is **active** at history $h \in H$ (that is, immediately after h) if i has at least two feasible actions at h: $|\mathcal{A}_i(h)| \geq 2$. A multistage game with observable actions has **perfect information** if, for each $h \in H$, at most one player is active at h.*

Note, perfect information is a formal property of multistage games, and it should not be confused with complete information, an assumption with a different meaning that here we do not express formally.[11]

9.2.1 Comments

Active and Inactive Players

Our definition of game tree allows for the possibility that all players are inactive at a given history h. We do not even exclude that along some path z all players become inactive after some stage k; such path z may even be infinite. In this case, the path can be truncated at the stage where all players become inactive, and this may transform a formally infinite game into a finite one. Of course, such transformations are innocuous.

According to the present notation, when $\mathcal{A}_i(h) = \emptyset$ for some player i, it means that the game is over. Therefore it is assumed that $\mathcal{A}_i(h) = \emptyset$ for *some* i if and only if $\mathcal{A}_i(h) = \emptyset$ for *all* i. An **inactive** player in an ongoing game is a player with only *one* feasible action, say action "*wait.*"

One could give a more elegant definition in which there is a set $I(h)$ of active players at each partial history h and feasible actions are specified

[11] Recall that a game features complete information is there is common knowledge of the rules of the game and of players' preferences over lotteries of outcomes.

for active players only. This would have the advantage of simplifying the notation for perfect information games and also for specific examples of games. But such a definition would be more complex[12] without essential gains in generality.

Streams of Outcomes and Intertemporal Preferences

As a matter of interpretation, $y = g(z)$ need not be an outcome that realizes at the end of the game. Indeed, this cannot be literally true when z is an infinite history (i.e., $z \in Z \cap A^{\mathbb{N}}$). In many applications, $y = g(z)$ is a sequence, or *stream*, of outcomes that realize at the end of each time period with periods being composed by one or more stages. Suppose, for example, that each stage coincides with a time period; furthermore, to ease notation suppose that the game has a fixed—finite or infinite—duration $T \in \mathbb{N} \cup \{\infty\}$. In this case, g may be derived from a sequence of functions $\left(g_t : A^t \to Y_t\right)_{t=1}^{T}$, where and $g_t\left(a^1, ..., a^t\right)$ is the period-t outcome, which in general may depend on the whole history up to period t. With this, we let $Y = \bigtimes_{t=1}^{T} Y_t$ and

$$\begin{array}{rccc} g: & Z & \to & Y, \\ & \left(a^t\right)_{t=1}^{T} & \mapsto & \left(g_t\left(\left(a^\tau\right)_{\tau=1}^{t}\right)\right)_{t=1}^{T}, \end{array}$$

that is, $y = g(z)$ is the sequence of end-of-period outcomes realized through terminal history z. For example, $g_t\left(\left(a^\tau\right)_{\tau=1}^{t}\right) \in \mathbb{R}^I$ could be the profile of period-t profits in an oligopoly with a set of firms I. If $y = (y_t)_{t=1}^{T}$ is a stream of outcomes, $v_i : Y \to \mathbb{R}$ aggregates the outcomes of different periods and represents the intertemporal preferences of player i. The most common intertemporal aggregator in economic applications satisfies time separability and exponential discounting, that is, there are a discount factor $\delta_i \in (0,1)$ and a sequence of utility functions $(v_{i,t} : Y_t \to \mathbb{R})_{t=1}^{T}$ such

[12]With this alternative definition, action profiles have the form $a_J = (a_j)_{j\in J}$ with $\emptyset \neq J \subseteq I$; and the set of such profiles is

$$\bar{A} := \bigcup_{\emptyset \neq J \subseteq I} (\bigtimes_{i\in J} A_i);$$

an active players correspondence $I(\cdot) : \bar{A}^{<\mathbb{N}} \rightrightarrows 2^I$ has to be introduced among the primitive elements defining the game; the feasibility correspondence $\mathcal{A}_i(\cdot)$ is defined on the subset $H_i = \{h \in \bar{A}^{<\mathbb{N}} : i \in I(h)\}$.

that

$$\forall y \in Y,\, v_i\left(y\right) = \sum_{t=1}^{T} \delta_i^{t-1} v_{i,t}\left(y_t\right).$$

When periods comprise several stages the explicit description of histories and outcome functions requires a more complex notation. We do not pursue this issue here.

9.2.2 Graphical Representation

It is traditional to represent games of *perfect information* (see Definition 47) as **game trees** in the sense of graph theory: each $h \in \bar{H}$ is a **node**; each $z \in Z$ is a **terminal node** (or leaf); each $h \in H$ is a **decision node**; each terminal history/node z is associated with payoff profile $(u_i(z))_{i \in I}$; each partial history/decision node h is associated with the only active player, denoted by $i = \iota(h)$; if $h' = (h, a)$ (the concatenation of h and a), then there is a **directed arc** (arrow) from h to h', this directed arc is associated with action $a_{\iota(h)}$, the action of the active player in profile $a = (a_{\iota(h)}, \text{wait}_{-\iota(h)})$. Sometimes it is notationally useful to distinguish between "physically identical" actions if they are taken after different histories. In this case, there is a 1-1 correspondence between arcs and actions in perfect information games. The following example illustrates such graphical representation.

Example 40. Consider the following game **Take-it-Or-Leave-it** game of length T between player 1 and player 2. A referee (or mechanical device) puts a dollar on the table. Player 1 can take it or leave it, player 2 just observes and waits. If player 1 takes the dollar the game is over, otherwise the referee puts another dollar on the table, player 2 can take the two dollars or leave them on the table, player 1 observes and waits. If player 2 takes the dollars the game is over, otherwise the referee puts another dollar on the table. The game goes on like this until the referee has exhausted his T dollars or someone has taken the dollars on the table. If there are T dollars on the table and the active player (player 1 if T is odd) leave them, they go to the other player. It is common knowledge that players only care about money and are risk-neutral. The action set of each player i is $A_i = \{\text{Take}, \text{Leave}, \text{Wait}\}$. If $T = 4$, the game is represented as in Figure 9.2.

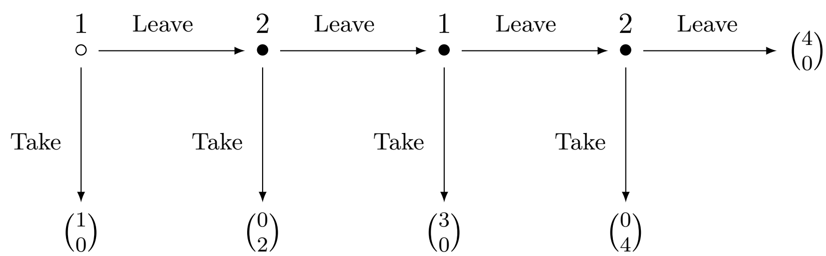


Figure 9.2: Game for $T = 4$.

We have

$$H = \{\varnothing, ((\text{Leave}, \text{Wait})), ((\text{Leave}, \text{Wait}), (\text{Wait}, \text{Leave})), \\ ((\text{Leave}, \text{Wait}), (\text{Wait}, \text{Leave}), (\text{Leave}, \text{Wait}))\},$$

$\mathcal{A}_i(h) = \{\text{Take}, \text{Leave}\}$ if i is active at $h \in H$ and $\mathcal{A}_i(h) = \{\text{Wait}\}$ otherwise. The vertical (respectively, horizontal) directed arcs are labeled by the action Take (respectively Leave) of the active player. ▲

In Example 40, we explicitly included the pseudo-action Wait only to clarify the abstract, general notation introduced above; but, from now on, we will identify histories with sequences of actions by *active players only*. For example, (Leave, Wait) will be simply written as (Leave), ((Leave, Wait),(Wait, Leave)) will be written as (Leave, Leave), etc.

More generally, *all games with observable actions can be graphically represented as game trees*. Consider the BoS with an Outside Option of Example 38: if history (in) occurs, then the four action pairs—that is, (B_1, B_2), (B_1, S_2), (S_1, B_2), and (S_1, S_2)—can be represented by directed arcs (arrows) from (in) to the terminal nodes, where each directed arc is associated with the corresponding action pair $(a_1, a_2) \in \mathcal{A}((\text{in}))$. See Figure 9.3 below.

Pictures with trees such as in Figure 9.3 are perfectly legitimate and they faithfully represent the simultaneity of moves at some stages of games *without* perfect information, but are not very common in the game-theoretic literature. The reason is twofold. First, they are not easily readable. Second, they are not traditional. Indeed, the traditional approach—which we do *not* pursue in this chapter—to represent multistage games without perfect information is to pretend that

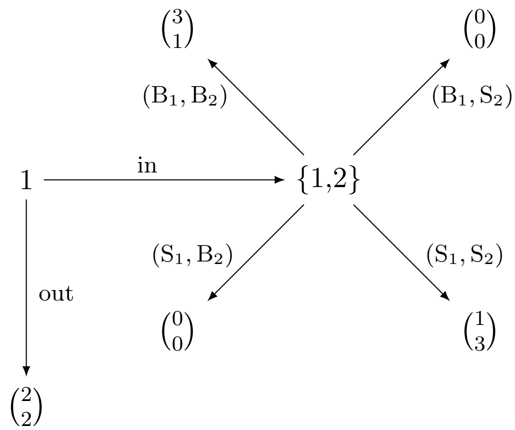


Figure 9.3: BoS with an Outside Option as game tree.

simultaneous moves are, instead, sequential, and use so called "information sets" to make such misrepresentation innocuous.[13] Consider again the BoS with an Outside Option of Example 38. If Ann (player 1) goes in, then the simultaneous-move "subgame" BoS is played. The approach via "information sets" is to pretend that, after history (in), one of the players moves first, and the co-player moves second, but does not observe the action chosen by the first mover. In Figure 9.4, the representation of the BoS "subgame" starting after history (in) is that Ann is the first mover. To represent the assumption that Bob (player 2) cannot observe Ann's choice, i.e., that he does not know whether he is at history/node (in, B_1) or (in, S_1), a dashed line joins the two decision nodes of Bob. The set of nodes that a player cannot distinguish is called **information set.** That said, we do not follow such approach for the representation of games with observable actions. Information sets will be useful for the analysis of multistage games with incomplete information (Chapter 15) or imperfectly observable actions (Chapter 16).[14]

[13] See, in particular, von Neumann and Morgenstern [68] and Kuhn [43], who set the stage for the following literature on dynamic games.

[14] Recall that—unfortunately—"(im)perfect information" and "(in)complete information" have very different meanings in the language of game theory (see Chapters 1.4.3, 8, and Definition 47), despite the fact that they are essentially synonymous in the natural language.

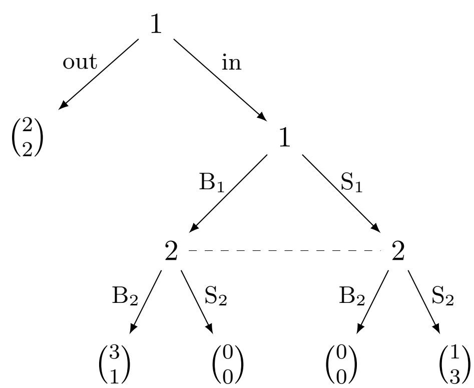


Figure 9.4: Alternative representation of BoS with an Outside Option.

9.3 Strategies, Plans, Strategic Form, Reduced Strategic Form

A strategy is a complete contingent plan which includes “instructions” contingent on every non-terminal history, including histories where the opponents make unexpected moves as well as histories inconsistent with the plan itself. More formally:

Definition 48. *A **strategy** for player i is an element of the product set $\times_{h\in H}\mathcal{A}_i(h)$, that is, a function $s_i : H \to A_i$ such that $s_i(h) \in \mathcal{A}_i(h)$ for each $h \in H$. The set of strategies for player i is denoted S_i, that is, $S_i = \times_{h\in H}\mathcal{A}_i(h)$. We let $S = \times_{i\in I} S_i$ and $S_{-i} = \times_{j\neq i} S_j$.*

Remark 40. *In a finite game the number of strategies of each player $i \in I$ is $|S_i| = \prod_{h\in H}|\mathcal{A}_i(h)|$.*

A typical element of S_i is denoted by $s_i = (s_i(h))_{h\in H}$. Note that we can equivalently define the set of strategies neglecting the histories where player i is not active: let

$$H_i = \{h \in H : |\mathcal{A}_i(h)| \geq 2\}$$

denote the set of histories (nodes of the game tree) where player i is active. Then $\times_{h\in H}\mathcal{A}_i(h)$ is isomorphic to $\times_{h\in H_i}\mathcal{A}_i(h)$ and it makes sense to write $S_i = \times_{h\in H_i}\mathcal{A}_i(h)$.

The set of **sub-strategies in the sub-tree with root** $h \in H$ is denoted by $S_i^{\succeq h}$, that is,

$$S_i^{\succeq h} = \bigtimes_{h' \in H(h)} \mathcal{A}_i(h').$$

A generic element of $S_i^{\succeq h}$ is denoted by $s_i^{\succeq h}$. The sub-strategy induced by $s_i \in S_i$ in the sub-tree with root h is denoted by

$$(s_i|h) = (s_i(h'))_{h' \in H(h)} \in S_i^{\succeq h}.$$

Given S, it is possible to derive a fictitious static game $G = \mathcal{N}(\Gamma)$, called the "strategic form," or "normal form" of Γ, that captures some important aspects of the game Γ.[15] The idea is the following: each player instructs a trustworthy agent, or writes a computer program for a machine, so that the agent/machine implements a particular strategy. Thus, each player commits to (*chooses*) a strategy, rather than just planning it in his mind. Each profile of strategies $s = (s_i)_{i \in I}$ induces a terminal history and the associated payoff vector. Of course, the players' choices of strategies are simultaneous (or, at least, no player can observe anything about the strategies chosen by the opponents before he chooses his own strategy).

The strategic form of the Take-it-Or-Leave-it game of length $T = 4$ of Figure 9.2 (henceforth, TOL4) is represented in Figure 9.5. Strategies are labelled by sequences of letters separated by dots, with the following convention: the first letter denotes the action to be taken at the first history where the player is active ($\varnothing$ for player 1 and $h = (\text{Leave})$ for player 2), the second letter denotes the action to be taken at the second history where the player is active ($h = (\text{Leave}, \text{Leave})$ for player 1, $h = (\text{Leave}, \text{Leave}, \text{Leave})$ for player 2). For example, $s_1 = T.T$ is the strategy [Take; Take if $(\text{Leave}, \text{Leave})$] of player 1.

Definition 49. *The **path function** $\zeta : S \to Z$ associates each strategy profile $s \in S$ with the induced terminal history $\zeta(s) = (a^t)_{t=1}^T$ ($T \in \mathbb{N} \cup \{\infty\}$) defined recursively as follows:*

$$\begin{aligned} a^1 &= (s_i(\varnothing))_{i \in I}, \\ a^{t+1} &= \left(s_i\left((a^k)_{k=1}^t\right)\right)_{i \in I} \end{aligned}$$

[15] According to some theorists—but not ourselves—all the relevant aspects of a game are captured by this derived representation.

$1\backslash 2$	$T.T$	$T.L$	$L.T$	$L.L$
$T.T$	$1,0$	$1,0$	$1,0$	$1,0$
$T.L$	$1,0$	$1,0$	$1,0$	$1,0$
$L.T$	$0,2$	$0,2$	$3,0$	$3,0$
$L.L$	$0,2$	$0,2$	$0,4$	$4,0$

Figure 9.5: Strategic form of ToL4.

for every $t \in \mathbb{N}$ such that $(a^k)_{k=1}^t$ is not terminal. The ***strategic*** *or* ***normal form*** *of Γ is the static game*

$$\mathcal{N}(\Gamma) = \langle I, (S_i, U_i)_{i \in I} \rangle$$

where $U_i = u_i \circ \zeta : S \to \mathbb{R}$ for every $i \in I$.

As stated in Definition 49, "normal form" and "strategic form" are (for us) synonyms. We use mostly the latter because it is more self-explanatory, yet symbol $\mathcal{N}(\Gamma)$ refers to the former.

Consider the strategies $T.T$ and $T.L$ of player 1 in TOL4. Both of them prescribe to take one dollar immediately, an action that terminates the game. They differ only for the instruction corresponding to the history $h = (\text{Leave}, \text{Leave})$, that is, a history prevented by both strategies. These two strategies yield the same terminal history, (Take), independently of the strategy adopted by player 2. A similar observation holds for player 2: the two strategies $T.T$ and $T.L$ differ only for the instruction corresponding to the history $h = (\text{Leave}, \text{Leave}, \text{Leave})$**,** that both of them prevent. Which terminal history is reached depends on the strategy of player 1, but does not depend on which of these two strategies is implemented by player 2. This is the reason why these strategies correspond to identical rows (player 1) or columns (player 2).

These considerations motivate the following definitions. For every pair of histories $h, \bar{h} \in \bar{H}$ with $h \prec \bar{h}$, let $\alpha_i(h, \bar{h})$ denote the (unique) **action of i at h that allows $\bar{h}$**, that is, for each $a_i \in \mathcal{A}_i(h)$,

$$a_i = \alpha_i(h, \bar{h}) \Leftrightarrow (\exists a_{-i} \in \mathcal{A}_{-i}(h), (h, (a_i, a_{-i})) \preceq \bar{h}).$$

With this, the set of non-terminal **histories not prevented by** s_i can be defined as follows:

$$H_i(s_i) = \{\bar{h} \in H : \forall h \in H, h \prec \bar{h} \Rightarrow s_i(h) = \alpha_i(h, \bar{h})\}.$$

In words, $H_i(s_i)$ is the set of non-terminal histories $\bar{h}$ such that s_i selects the action leading to $\bar{h}$ at each predecessor $h \prec \bar{h}$.

For each strategy s_i, we can characterize the set $H_i(s_i)$ as follows. Formally, history h is not prevented, or is allowed by s_i if there is a profile of strategies of the opponents s_{-i} such that (s_i, s_{-i}) induces h. Since (s_i, s_{-i}) induces h if and only if it induces a terminal history preceded by h, we obtain the following characterization of the set of non-terminal histories allowed by s_i.

Lemma 23. $H_i(s_i) = \{\bar{h} \in H : \exists s_{-i} \in S_{-i}, \bar{h} \prec \zeta(s_i, s_{-i})\}$.

Proof. Fix $\bar{h}$ and suppose that $\bar{h} \prec \zeta(s_i, s_{-i})$ for some s_{-i}. Then, by definition of ζ, $\alpha_i(h, \bar{h}) = s_i(h)$ for each $h \prec \bar{h}$, which implies $\bar{h} \in H_i(s_i)$. This shows

$$\{\bar{h} \in H : \exists s_{-i} \in S_{-i}, \bar{h} \prec \zeta(s_i, s_{-i})\} \subseteq H_i(s_i).$$

To show that

$$H_i(s_i) \subseteq \{\bar{h} \in H : \exists s_{-i} \in S_{-i}, \bar{h} \prec \zeta(s_i, s_{-i})\},$$

pick $\bar{h} \in H_i(s_i)$ arbitrarily and construct a strategy profile $s_{-i} \in S_{-i}$ as follows: $s_j(h) = \alpha_j(h, \bar{h})$ for each $h \prec \bar{h}$ and $j \neq i$, otherwise $s_j(h)$ is arbitrarily chosen in $\mathcal{A}_j(h)$. Since $\bar{h} \in H_i(s_i)$, we have $\alpha_i(h, \bar{h}) = s_i(h)$ for each $h \prec \bar{h}$. Hence, by construction, $(s_j(h))_{j \in I} = (\alpha_j(h, \bar{h}))_{j \in I}$ for each $h \prec \bar{h}$. Therefore $\bar{h} \prec \zeta(s_i, s_{-i})$. ■

Definition 50. *Two strategies s_i and s_i' are **realization-equivalent** if*

$$\forall s_{-i} \in S_{-i},\ \zeta(s_i, s_{-i}) = \zeta(s_i', s_{-i});$$

*they are **behaviorally equivalent** if*

$$H_i(s_i) = H_i(s_i') \text{ and } \forall h \in H_i(s_i), s_i(h) = s_i'(h).$$

In words, two strategies s_i and s_i' are realization-equivalent if, for every given strategy profile of the co-players, they induce the same terminal history. Strategies s_i and s_i' are behaviorally equivalent if they allow the same non-terminal histories and prescribe the same actions at such histories.

Lemma 24. *Two strategies are realization-equivalent if and only if they are behaviorally equivalent.*

Proof. We prove both directions of the statement by contraposition. Suppose first that s_i and s_i' are *not* realization equivalent, then there is some s_{-i} such that $\zeta(s_i, s_{-i}) \neq \zeta(s_i', s_{-i})$. We must show that s_i and s_i' are not behaviorally equivalent. Let $\bar{h}$ denote the maximal element of the chain of common predecessors of $\zeta(s_i, s_{-i})$ and $\zeta(s_i', s_{-i})$ (the chain is not empty because it contains the root $\varnothing$, and it is finite, hence it has a maximal element); this is the node at which the two paths "bifurcate." By Lemma 23, $\bar{h} \in H_i(s_i) \cap H_i(s_i')$. The two strategy profiles (s_i, s_{-i}) and (s_i', s_{-i}) specify the same action $s_j(\bar{h})$ at $\bar{h}$ for each co-player $j \neq i$; hence, it must be the case that $s_i(\bar{h}) \neq s_i'(\bar{h})$ (otherwise, there would be no "bifurcation" at $\bar{h}$). Therefore, strategies s_i and s_i' differ at a history not precluded by either one of them, which implies that they are not behaviorally equivalent.

Now suppose that s_i and s_i' are not behaviorally equivalent. To show that they are not realization-equivalent, we consider the following exhaustive cases, and we show that in both cases Lemma 23 implies that $\zeta\left(s_i, s_{-i}\right) \neq \zeta\left(s_i', s_{-i}\right)$ for some s_{-i}:

1. $H_i(s_i) \neq H_i(s_i')$,
2. $H_i(s_i) = H_i(s_i')$ and $s_i(h) \neq s_i'(h)$ for some $h \in H_i(s_i)$.

In **case 1**, there is some $h \in (H_i(s_i) \backslash H_i(s_i')) \cup (H_i(s_i') \backslash H_i(s_i))$. By Lemma 23, if $h \in H_i(s_i) \backslash H_i(s_i')$, there is some s_{-i} such that $h \prec \zeta(s_i, s_{-i})$ and $h \not\prec \zeta(s_i', s_{-i})$, that is, $\zeta(s_i, s_{-i}) \in Z\left(h\right)$ and $\zeta(s_i', s_{-i}) \in Z \backslash Z\left(h\right)$. Case $h \in H_i(s_i') \backslash H_i(s_i)$ is analogous. Whatever the case, $\zeta(s_i, s_{-i}) \neq \zeta(s_i', s_{-i})$ for some s_{-i}.

In **case 2**, there is $h \in H_i(s_i) = H_i(s_i')$ such that $s_i(h) \neq s_i'(h)$. By Lemma 23 there some s_{-i} such that $h \prec \zeta(s_i, s_{-i})$ and $h \prec \zeta(s_i', s_{-i})$. Since $s_i(h) \neq s_i'(h)$, we have $\zeta(s_i, s_{-i}) \neq \zeta(s_i', s_{-i})$. ■

In particular, the proof of Lemma 24 **(case 1)** shows that if s_i and s_i' are realization-equivalent they allow for the same histories, that is, $H_i\left(s_i\right) = H_i\left(s_i'\right) = H_i(s_i) \cap H_i(s_i')$. Furthermore, from the first part of the proof we obtain another characterization of behavioral equivalence:

Remark 41. *Two strategies s_i and s_i' are behaviorally equivalent if and only if $s_i\left(h\right) = s_i'\left(h\right)$ for every $h \in H_i(s_i) \cap H_i(s_i')$.*

The behavioral equivalence relation for player i corresponds to a partition of S_i (the quotient set of S_i with respect to the equivalence relation): two strategies belong to the same cell of the partition if and only if they are behaviorally equivalent.[16]

Definition 51. *A **reduced strategy** for player i is an element of the partition of S_i induced by the behavioral equivalence relation.*

We let $\mathbf{S}_i^r$ denote the set of reduced strategies of player i, and we let $\mathbf{S}^r = \times_{i\in I}\mathbf{S}_i^r$ denote the set of reduced strategy profiles. Note, each $\mathbf{s}_i^r \in \mathbf{S}_i^r$ is a *subset* $\mathbf{s}_i^r \subseteq S_i$ of behaviorally equivalent strategies. By Lemma 24, we can define a reduced strategic-form path function $\zeta^r : \mathbf{S}^r \to Z$ as follows:

$$\zeta^r((\mathbf{s}_i^r)_{i\in I}) = \zeta((s_i)_{i\in I}) \text{ if } s_i \in \mathbf{s}_i^r \text{ for each } i \in I.$$

Given the payoff functions $(u_i : Z \to \mathbb{R})_{i\in I}$, we obtain the reduced-form payoff functions $(U_i^r = u_i \circ \zeta^r : \mathbf{S}^r \to \mathbb{R})_{i\in I}$, where $U_i^r(\mathbf{s}^r) = U_i(s)$ for each $\mathbf{s}^r \in \times_{j\in I}\mathbf{S}_j^r$ and $s \in \mathbf{s}^r$.

Definition 52. *The **reduced strategic** (or normal) **form** of a game Γ is the static game $\mathcal{N}^r(\Gamma) = \langle I, (\mathbf{S}_i^r, U_i^r)_{i\in I}\rangle$.*

The reduced strategic form of the TOL4 game is represented in Figure 9.6.

$1\backslash 2$	T	$L.T$	$L.L$
T	$1,0$	$1,0$	$1,0$
$L.T$	$0,2$	$3,0$	$3,0$
$L.L$	$0,2$	$0,4$	$4,0$

Figure 9.6: Reduced strategic form of ToL4.

We illustrate the previous concepts and results with the BoS with a Dissipative Action of Example 39.

[16] The quotient of a set S with respect to an equivalence relation $\approx$ on S is the collection of equivalence classes:

$$S/\approx := \{\mathbf{s} \subseteq S : \forall (s,t) \in S \times S, \{s,t\} \subseteq \mathbf{s} \Leftrightarrow s \approx t\}.$$

See Ok [52, A.1.3].

Example 41. We denote strategies as lists of actions separated by dots, we denote strategy sets (profiles) as lists (ordered lists) of strategies separated by commas:

$$\begin{aligned} S_a &= \{U, D\} \times \{u, d\} = \{U.u, U.d, D.u, D.d\}, \\ S_b &= \{N, B\} \times \{L, R\} \times \{l, r\} \\ &= \{N.L.l, N.R.l, N.L.r, N.R.r, B.L.l, B.R.l, B.L.r, B.R.r\}. \end{aligned}$$

The following are examples of actions leading from a history h to a successor $\bar{h}$:

$$\alpha_b(\varnothing, (B, (u, l))) = B, \ \alpha_a((B), (B, (u, l))) = u.$$

The set of non-terminal histories is

$$H = \{\varnothing, (N), (B)\},$$

and the set of non-terminal histories allowed by $s_b = N.L.r$ is

$$H_b(N.L.r) = \{\varnothing, (N)\}.$$

The sub-strategy induced by $s_b = N.L.r$ in the subgame with root $h = (B)$ is $s_b^{\succ h} = r$;
The terminal history induced by strategy pair $(U.d, B.L.r)$ is

$$\zeta(U.d, B.L.r) = (B, (d, r)).$$

Finally, the following pairs of strategies of Bob are realization-equivalent:

$N.L.l$ and $N.L.r$ (reduced strategy $N.L$),
$N.R.l$ and $N.R.r$ (reduced strategy $N.R$),
$B.L.l$ and $B.R.l$ (reduced strategy $B.l$),
$B.L.r$ and $B.R.r$ (reduced strategy $B.r$);

therefore the reduced strategic form of the game is

$a\backslash b$	$N.L$	$N.R$	$B.l$	$B.r$
$U.u$	4,1	0,0	4,-1	0,-2
$U.d$	4,1	0,0	0,-2	1,2
$D.u$	0,0	1,4	4,-1	0,-2
$D.d$	0,0	1,4	0,-2	1,2

▲

A reduced strategy is sometimes called "plan of action" (see Rubinstein [59]). The reason is that "plan of action" corresponds to the intuitive idea that a player has to plan for all *external* contingencies, i.e., contingencies that not depend on his behavior, but he does not have to plan for contingences that cannot occur if he follows his plan. For example, in the TOL4 game, $T.T$ and $T.L$ are realization-equivalent and they correspond to the reduced strategy, or "plan" T, that is, "take at the first opportunity."

We avoid this terminology, and stick to the more neutral "reduced strategy," because there is a perfectly meaningful notion of "planning" that yields the specification of a whole strategy, not just a reduced strategy. To see this, consider the BoS with an Outside Option of Example 38. Suppose Ann, player 1, believes that, conditional on choosing in, the probability that Bob chooses B_2 is $\mu^1(\mathrm{B}_2|\mathrm{in}) = q$. Then Ann can compute a dynamically optimal plan with the following **folding back** procedure. First, she computes the values of choosing, respectively, B_1 and S_1 given in:

$$\begin{aligned} V_1^q(\mathrm{B}_1|\mathrm{in}) &= 3q, \\ V_1^q(\mathrm{S}_1|\mathrm{in}) &= 1-q. \end{aligned}$$

Next, assuming that she would maximize her expected payoff in the subgame if she chose in, Ann computes the value of in:

$$V_1^q(\mathrm{in}) = \max\{3q, 1-q\} = \begin{cases} 3q, & \text{if } q \geq \frac{1}{4}, \\ 1-q, & \text{if } q < \frac{1}{4}. \end{cases}$$

This means that Ann "plans" to choose B_1 (resp. S_1) in the subgame if $q > 1/4$ (resp. $q < 1/4$).[17] Finally, Ann compares $V_1^q(\mathrm{in})$ with the value

[17] If $q = 1/4$, then Ann is indifferent, and the optimal plan for the subgame is arbitrary.

of the outside option $V_1(\text{out}) = 2$ and "plans" to choose in (resp. out) if $V_1^q(\text{in}) > 2$ (resp. $V_1^q(\text{in}) < 2$), that is, if $q > 2/3$ (resp. $q < 2/3$). The upshot of all this is that, given her subjective beliefs q, Ann can use the folding back procedure to form a dynamically optimal plan, which corresponds to a *full* strategy, not a reduced strategy:[18]

$$s_1 = \begin{cases} \text{out.S}_1, & \text{if } q < \frac{1}{4}, \\ \text{out.B}_1, & \text{if } \frac{1}{4} < q < \frac{2}{3}, \\ \text{in.B}_1, & \text{if } q > \frac{2}{3}. \end{cases}$$

On the other hand, one may object that specifying a complete strategy is not necessary for rational planning: if q is not high enough, i.e., if $\max\{3q, 1-q\} < 2$, Ann has no need to plan ahead for the subgame, as she can see that it is not worth her while to reach it. Specifically, she can look forward for the payoff consequences of choosing out, in and B_1, or in and S_1, and her best "forward plan" is just out, a reduced strategy. Since both notions of rational planning—forward and backward—are meaningful and intuitive, we are not going to endorse only the second one of them by calling "plan of action" the reduced strategies. The solution concepts presented in the next section further clarify why the notion of full strategy (as opposed to reduced strategy) is important in game theory. We are going to elaborate on the folding back procedure of dynamic programming in Chapter 10. While the folding back procedure refers to a *single* player and his arbitrarily given *subjective* beliefs, the standard equilibrium theory of multistage games looks for *"interpersonal"* elaborations of the folding back procedure, see Chapter 12. Yet, as we did in the part on static games, before moving on to standard equilibrium theory we analyze multistage versions of the rationalizability solution concept in Chapter 11. The distinction between full and reduced strategies will be dicussed again in that context.

We considered above two notions of equivalence between (full) strategies—behavioral and realization equivalence—that depend only on the game tree, and we observed in Lemma 24 that they are congruent, that is, for each $i \in I$, they correspond to the same subset of $S_i \times S_i$. Now we consider a weaker equivalence relation that depends on the whole game Γ as it does not rely on induced behavior, but rather on induced payoffs.

[18]Again, we ignore ties.

Definition 53. *Two strategies s_i and s_i' are* ***payoff-equivalent****, written $s_i \sim_i s_i'$ if, for each strategy profile of the co-players, they yield the same profile of payoffs:*

$$\forall s_{-i} \in S_{-i},\ (s_i \sim_i s_i') \Longleftrightarrow (\forall j \in I, U_j(s_i, s_{-i}) = U_j(s_i', s_{-i})).$$

Since $U_j = u_j \circ \zeta$ $(j \in I)$, it is clear that realization equivalence implies payoff equivalence. Thus, Lemma 24 implies the following:

Remark 42. *If two strategies are behaviorally equivalent, then they are payoff-equivalent.*

In the games of Examples 38 and 40, as well as in many other games, there is no difference between payoff-equivalence and behavioral equivalence. A difference may arise only if the game features some ties between payoffs at distinct terminal histories, that is, if there are $z, z' \in Z$ such that $z \neq z'$ and yet $(u_i(z))_{i\in I} = (u_i(z'))_{i\in I}$. Such ties are "structural" when z and z' yield the same outcome, $g(z) = g(z')$, e.g., the same allocation of resources; otherwise, they are due to "non generic" ties between utility profiles. This happens if, for some $z, z' \in Z$, $g(z) \neq g(z')$ and $(v_i(g(z)))_{i\in I} = (v_i(g(z')))_{i\in I}$. Unfortunately, the outcome function g is mostly overlooked in game theory and all ties between payoffs at distinct terminal histories are called "non generic." Therefore, this terminology has to be taken with a grain of salt.

Comment: The most common definition of "reduced strategic form" refers to classes of *payoff*-equivalent strategies, rather than realization-equivalent strategies. This is fine if one is only interested in computing equilibria (or other solutions) of the strategic form. Instead, we are interested in the strategic form $\mathcal{N}(\Gamma)$ only as an auxiliary tool that (sometimes) helps analyzing Γ itself (as we will see in Section 9.3.1). Therefore we chose to emphasize the concept of realization-equivalence and reduced strategy, which is based on a meaningful notion of plan, and the corresponding concept of reduced strategic form.

9.3.1 Old Wine in New Bottles: Applying Solution Concepts to the Strategic Form

Any solution concept for static games can be applied to the strategic (or normal) form $\mathcal{N}(\Gamma)$ of a multistage game Γ and thus yields a "candidate solution" for Γ. For example, we can find the rationalizable strategies or the Nash equilibria of $\mathcal{N}(\Gamma)$ and ask ourselves if they make sense as solutions of Γ.

Consider first the following very stylized **Entry Game**: a firm, player 1, has the opportunity to enter a (so far) monopolistic market. The incumbent, player 2, may fight the entry with a price war that damages both, or it may "acquiesce." The game is (summarily) represented in Figure 9.7.

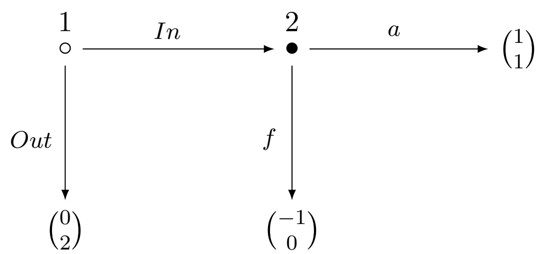


Figure 9.7: An "entry game."

The strategic form of this game is represented in the following table:

$1\backslash 2$	f	a
Out	0, 2	0, 2
In	-1, 0	1, 1

It is easily checked that every strategy profile of the strategic form of this game is rationalizable, and there are two Nash equilibria, (Out, f) and (In, a) (there is also a continuum of mixed equilibria where player 2 plays f with probability larger than $\frac{1}{2}$). Yet it should be noted that, under complete information, if the potential entrant believes that the incumbent is rational (in an obvious sense), then she should expect acquiescence and therefore she should enter. So, the only "reasonable" solution should be (In, a).

Consider now the strategic form of the TOL4 game (Figure 9.5). Again, it is easily checked that every strategy profile is rationalizable, and there are multiple Nash equilibria, namely, all the pairs in the set $\{T.T, T.L\} \times \{T.T, T.L\}$. Note that all Nash equilibria "correspond" to the unique Nash equilibrium (T,T) of the reduced strategic form (Figure 9.6), and they induce the terminal history (Take). This should not be surprising, as strategies $T.T$ and $T.L$ are realization-equivalent. Indeed:

Remark 43. Payoff-equivalences (hence, also realization-equivalences) induce a trivial multiplicity of Nash equilibria*: if a strategy profile s^* is a Nash equilibrium, every profile $\bar{s}^*$ obtained from s^* by replacing some strategy s_i^* with an equivalent strategy $\bar{s}_i^*$ is also a Nash equilibrium. In other words, to find Nash equilibria we may just look at the equilibria $\mathbf{s}^{*,r}$ of reduced strategic form and then take into account that any profile s^* of the non-reduced form that corresponds to $\mathbf{s}^{*,r}$ is a Nash equilibrium of the multistage game. Similar considerations apply to strategic-form rationalizability: if a strategy s_i^* is rationalizable in $\mathcal{N}(\Gamma)$ every payoff-equivalent (hence, every realization-equivalent) strategy $\bar{s}_i^*$ is also rationalizable in $\mathcal{N}(\Gamma)$.*

The problem with rationalizability and Nash equilibrium of the strategic form is that such solution concepts require that players maximize their expected payoff "on the path," i.e., along the play induced by the strategy profile, but they also allow players to "plan" non maximizing actions at histories that are "off the path." For instance, in the Nash equilibrium (Out, f) of the Entry Game player 2 plans to choose a non maximizing action at the "off the path" history (In). Does this make sense?

In a multistage game, a player has initial beliefs which he updates as the play unfolds. Sometimes what he observes may be unexpected (i.e., have zero probability) according to his previous beliefs. In this case, he will not be able to pin down his new beliefs by updating his previous ones according to the rules of conditional probability, but he will still form new beliefs about his opponents' behavior. For example, suppose that in the TOL4 game of Figure 9.2 player 2 initially expects that player 1 will immediately take one dollar. If instead player 1 leaves the dollar on the table, player 2 will revise his beliefs about player 1, because he now knows that he is implementing one of the following strategies: $L.T$ (Leave, then Take if also player 2 leaves) or $L.L$ (Leave, then Leave again if also player 2 leaves). The revised beliefs must assign zero probability to every other strategy of player 1.[19]

[19] In this informal discussion we are taking for granted that players actually implement

We say that a player in a multistage game is *rational* if he would make expected utility maximizing choices given his (updated) beliefs *for every possible history of observed choices of his opponents*. Consider ToL4. Can one say that player 1 is irrational if he leaves three dollars on the table? No. Player 1 may hope that then player 2 will be "generous" and leave him four dollars. We might argue that such a belief is not very reasonable, indeed it is inconsistent with the rationality of player 2; yet, if player 1 had this belief, leaving three dollars on the table would be rational, i.e., expected utility maximizing.

In the analysis of static games, one can characterize with a solution concept the behavioral implications of the following assumptions: all players are rational and there is common belief of rationality. The solution concept is rationalizability. Furthermore, an action is rationalizable if and only if it is iteratively undominated.

What assumptions about rationality and strategic reasoning are worth considering in the case of multistage games? How should we extend the notion of rationalizability from static to multistage games so as to characterize the behavioral implications of such assumptions? How does such extension relate to rationalizability in the strategic form?

In Chapter 10 we analyze rationality, i.e., the implementation of strategies obtained by rational planning given conjectures about the behavior of other players. In Chapter 11 we introduce different forms of rationalizability for multistage games justified by different assumptions about strategic reasoning. Of course, these solution concepts have to coincide with rationalizability in the strategic form $\mathcal{N}(\Gamma)$ when the given game Γ has only one stage, and therefore is static. But in games with two or more stages they yield refinements of rationalizability in the strategic form, that is, some strategies that are rationalizable in the strategic form $\mathcal{N}(\Gamma)$ are deleted by the appropriate version of the rationalizability concept for multistage game Γ. Does this mean that the strategic form is useless? Not exactly: in many games of interest, notions of rationalizability for multistage games admit a useful characterization that relies only on the strategic form.

Consider first an elementary example, the Entry Game. It is pretty clear that if the players are rational in the sense specified above and if the first

---

the strategies they have in mind, and that this is "transparent." We will come back to this point in our analysis of solution concepts for multistage games.

mover believes that also the opponent is rational, then the "reasonable" solution (In, a) obtains. Furthermore, (In, a) can be obtained in the strategic form of the game by first eliminating the *weakly* dominated strategies (just f in this case) and then eliminating the (strictly) dominated strategies (just Out in this case) of the residual strategic form. This procedure works in all two-stage games with perfect information with "no relevant ties."

Now consider more complex games—such as the BoS with an Outside Option—and the associated strategic forms. We will argue in Chapters 10 and 11 that, in many games of interest, **iterated admissibility**, i.e., the iterated deletion of all weakly dominated strategies,[20] coincides with **strong rationalizability**. The latter is a solution concept for multistage games which captures the following **best rationalization principle**: every player always ascribes to his opponents the highest degree of "strategic sophistication" consistent with their observed behavior. To illustrate this point, consider the BoS with an Outside Option (Figure 9.1). The highest degree of strategic sophistication that Bob can ascribe to Ann if he observes action "in" is that Ann is rational, because strategy in.B_1 is justifiable. On the other hand, strategy in.S_1 is not justifiable (indeed, it is dominated by "out"). Thus, at history (in) Bob should believe that Ann is rational and that she is going to choose B_1 in the subgame. The best reply to such belief is B_2. Anticipating this, Ann implements strategy in.B_1. Hence, the unique solution consistent with the best rationalization principle is $(\text{in}.\mathrm{B}_1, \mathrm{B}_2)$.

It can be easily checked that $(\text{in}.\mathrm{B}_1, \mathrm{B}_2)$ can be obtained *in the strategic form* of the game by iterated admissibility. Chapter 11 will provide a formal definition of strong rationalizability for multistage games with observable actions. Here we just note that iterated admissibility on the strategic form is a useful solution concept that yields in many games "reasonable" solutions capturing the best rationalization principle.

In Chapter 12 we show how to extend Nash's classical idea that players best respond to correct conjectures. This will be a rather direct application of the analysis of rational planning of Chapter 10. The resulting solution concept is called **subgame perfect** (Nash) **equilibrium (SPE)**, because it is characterized by the property of inducing a Nash equilibrium in every

[20] Iterated admissibility is a **maximal** iterated elimination procedure, that is, in each step *all* the weakly dominated strategies of *all* players are eliminated.

"subgame." For example, the Entry Game has only one SPE, (In, a), because this is the only strategy pair that is both a Nash equilibrium of the game and also induces an equilibrium of the (trivial) one-person game following history (In). The BoS with an Outside Option instead has two SPEs, $(\text{in.B}_1, \text{B}_2)$ and $(\text{out.S}_1, \text{S}_2)$, each one of them induces a Nash equilibrium of the BoS subgame following history (in). The standard way to compute SPEs relies on all the details of the description of the multistage game Γ.

What about self-confirming equilibrium (SCE)? Intuitively, a strategy profile s^* is a (pure) SCE if each s_i^* can be justified as a best reply to a conjecture about the other players that is confirmed by the evidence obtained ex post by i if everybody plays according to s^*. Of course, to give a definition of SCE we have to specify players' feedback, which can be described, for each player i, by a function $f_i : Z \to M_i$. Given the profile of feedback functions $f = (f_i)_{i \in I}$, we obtain a multistage game with feedback (Γ, f). Assume that players recall whatever they observe. In the context of multistage games with observable actions, it makes sense to assume that either f_i is one-to-one, because players observe *ex post* each other actions also at the end of the last stage, or—at least—player i remembers all the actions played in all stages but the last one, and remembers what he did in the last stage. Be as it may, the key observation for SCE is that, unlike SPE, *we can analyze the essential aspects of SCE for multistage games with feedback* (Γ, f) *by looking at the strategic (or normal) form*

$$\mathcal{N}(\Gamma, f) = \left\langle I, (U_i, F_i)_{i \in I} \right\rangle = \left\langle I, (u_i \circ \zeta, f_i \circ \zeta)_{i \in I} \right\rangle,$$

where, for each player $i \in I$, $F_i : S \to M_i$ is the strategic-form feedback function of i. Specifically, any fixed strategy profile s^* is an SCE of (Γ, f) if and only if there is an SCE $\bar{s}^*$ of $\mathcal{N}(\Gamma, f)$ inducing the same terminal history, i.e., such that $\zeta(s^*) = \zeta(\bar{s}^*)$.

To get some intuition for this result, let Γ be the Entry Game and assume (reasonably in this case) that players observe ex post the terminal history; in particular, if player 1 goes In, he observes ex post whether 2 acquiesces or fights. We argued that, whatever the conjecture of player 2 about 1, the only reasonable "best reply" should be a. Thus, if player 1—knowing 2's payoff function—anticipates this, we obtain (In, a). Yet, *SCE does not rely on the idea that players reason strategically given common knowledge of the game*, they only have to best respond to confirmed justifying conjectures, given that each i knows his own payoff function u_i and feedback function f_i. Therefore, player 1 may be afraid of a fight and go Out, which makes his fears unverifiable. Hence, there are two (pure) SCE strategy pairs, (In, a) and (Out, a). The latter is supported by every conjecture of player 1 assigning at least 50% probability to a fight, because staying Out is a best reply and, given this, the conjecture—even if inaccurate—is trivially confirmed. The strategic form feedback is such that $F_i\,(Out, a) = Out = F_i\,(Out, f)$ for each i. Hence, there is a third (pure) SCE pair in the strategic form, the (imperfect) Nash equilibrium (Out, f) (player 2 is *ex ante* indifferent if he is certain of Out), but this additional strategic-form SCE yields the same terminal history as (Out, a). Note that the equality between Nash and SCE outcomes in this game is not a coincidence: indeed, one can show that, *in every two-person game with feedback where players observe ex post the terminal history, pure SCE and Nash outcomes coincide.*[21]

9.4 Randomized Strategies

In this section we introduce randomization in multistage games with observable actions. To simplify the probabilistic analysis we focus on *finite* games. In the context of multistage games (and more generally of all games with a sequential structure) one can think of two types of randomization:

(1) player i implements a pure strategy at random according to a probability measure $\sigma_i \in \Delta(S_i)$, which is called a **mixed strategy**;

(2) for each non terminal history $h \in H$, player i would choose an action at random according to a probability measure $\beta_i(\cdot|h) \in \Delta(\mathcal{A}_i(h))$; the array of probability measures $\beta_i = (\beta_i(\cdot|h))_{h\in H}$ is called **behavior strategy**.[22]

Note that *we can always regard a pure strategy as a degenerate randomized strategy*: s_i can be identified with the mixed strategy σ_i such

[21] The result extends to the mixed SCEs such that all pure strategies in the support can be justified by the same confirmed conjecture. See, e.g., Battigalli et al. [15].

[22] We use Kuhn's [43] original terminology. Some authors, e.g., Osborne and Rubinstein [53], say "behavioral strategy." Kuhn introduced the concept and the terminology in his analysis of general games with a sequential structure. Obviously, $\beta_i(\cdot|h)$ is non-trivial only at histories where i is active.

that $\sigma_i(s_i) = 1$ and with the behavior strategy β_i such that $\beta_i(s_i(h)|h) = 1$ for every $h \in H$.

9.4.1 Realization-Equivalence between Mixed and Behavior Strategies

The literal interpretation of randomized strategies is that players spin roulette wheels or toss coins and let their actions be decided by the outcomes of such randomization devices. But this seems a bit farfetched, especially in the case of mixed strategies. Furthermore, agents who maximize expected utility have no use for randomization, as pure choices always do at least as well as randomized choices.

As in static games, a randomized strategy of i can be interpreted as a representation of the opponents' beliefs about the pure strategy of i. For example, suppose that game Γ is played by agents drawn at random from large populations; let $\sigma_i(s_i)$ be the fraction of agents in population i (the population of agents that may play in role i) that would implement strategy s_i. If j knows the statistical distribution σ_i, this will also represent the belief of j about the strategy of i. On the other hand, $\beta_i(a_i|h)$ may be interpreted as the conditional probability assigned by j to action a_i given history h.

This belief interpretation suggests a relationship between mixed and behavior strategies. Suppose that h has just realized; what does anyone learn about the pure strategy implemented by i? She learns that i is implementing a strategy that allows (does not prevent) h. Let $S_i(h)$ be the set of such strategies. Formally,[23]

$$S_i(h) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, h \prec \zeta(s_i, s_{-i})\}.$$

Next we define the set of strategies of i that allow h *and* select a_i at h:

$$S_i(h, a_i) = \{s_i \in S_i(h) : s_i(h) = a_i\}.$$

Now suppose that β_i is derived from the statistical distribution σ_i under the "*independence-across-players*" assumption that what is observed about the behavior of other players does not affect the beliefs about the strategy

[23] If the game has chance moves the definition in the text must be adapted by including in s_{-i} also s_0, the "strategy" of the chance player.

of i. Then the probability of a_i given h is just the fraction of agents in population i implementing a strategy that allows h and selects a_i at h, divided by the fraction of agents in population i implementing a strategy that allows h (if positive), that is,

$$\forall h \in H,\, \sigma_i(S_i(h)) > 0 \Rightarrow \beta_i(a_i|h) = \frac{\sigma_i(S_i(h, a_i))}{\sigma_i(S_i(h))} \tag{9.4.1}$$

(for every subset $X \subseteq S_i$, we write $\sigma_i(X) = \sum_{s_i \in X} \sigma_i(s_i)$). If $\sigma_i(S_i(h)) = 0$, $\beta_i(\cdot|h)$ can be specified *arbitrarily.* Note that (9.4.1) can be written in a more compact, but slightly less transparent form:

$$\forall h \in H,\, \beta_i(a_i|h)\sigma_i(S_i(h)) = \sigma_i(S_i(h, a_i)).$$

The formula above, or (9.4.1), says that σ_i and β_i are mutually consistent in the sense that they jointly satisfy a kind of chain rule for conditional probabilities.

Definition 54. *A mixed strategy $\sigma_i \in \Delta(S_i)$ and a behavior strategy $\beta_i = (\beta_i(\cdot|h))_{h\in H} \in \times_{h\in H} \Delta(\mathcal{A}_i(h))$ are mutually consistent if they satisfy (9.4.1).*

Remark 44. *If mixed strategy σ_i is such that $\sigma_i(S_i(h)) = 0$ for some h where player i is active, then there is a continuum of behavior strategies β_i consistent with σ_i. If $\sigma_i(S_i(h)) > 0$ for every h where i is active, then there is a unique β_i consistent with σ_i. If i is active at more than one history, then there is a continuum of mixed strategies σ_i consistent with any given behavior strategy β_i.*

The last statement in the remark can be understood with a counting-dimensionality argument. Let $H_i = \{h \in H : |A_i(h)| \geq 2\}$ denote the set of histories where i is active. In a finite game, the number of elements of S_i is $|S_i| = \prod_{h\in H_i} |\mathcal{A}_i(h)|$, thus the dimensionality of $\Delta(S_i)$ is $|S_i| - 1 = \prod_{h\in H_i} |\mathcal{A}_i(h)| - 1$. On the other hand, the dimensionality of the set of behavior strategies is $\sum_{h\in H_i}(|\mathcal{A}_i(h)| - 1) = \sum_{h\in H_i} |\mathcal{A}_i(h)| - |H_i|$. It can be shown (by induction on the cardinality of H_i) that $\prod_{h\in H_i} |\mathcal{A}_i(h)| \geq \sum_{h\in H_i} |\mathcal{A}_i(h)|$ because $|\mathcal{A}_i(h)| \geq 2$ for each $h \in H_i$.[24] Thus the dimensionality of $\Delta(S_i)$ is higher than the dimensionality of $\times_{h\in H} \Delta(\mathcal{A}_i(h))$ if $|H_i| \geq 2$.

[24] First note that for every integer $L \geq 1$, we have $2^{L-1} \geq L$. (This can be easily proved by induction: it is trivially true for $L = 1$; suppose it is true for some $L \geq 1$,

Example 42. In the game "tree" in Figure 9.8, Rowena (player 1) has 4 strategies, $S_1 = \{\text{out}.u, \text{out}.d, \text{in}.u, \text{in}.d\}$; two of them, out.$u$ and out.d, are realization-equivalent and correspond to the reduced strategy out. The figure labels terminal histories: $v = (\text{out})$, $w = (\text{in}, (u, l))$ etc. The set of non-terminal histories is $H = \{\varnothing, (\text{in})\}$ (recall that $\varnothing$ denotes the initial history) and $S_1(\varnothing) = S_1$, $S_1(\text{in}) = \{\text{in}.u, \text{in}.d\}$.

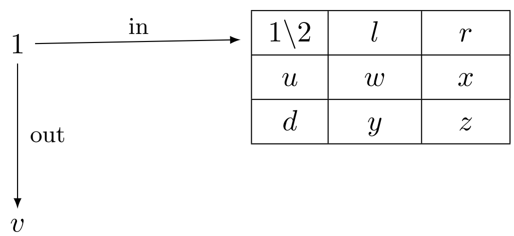


Figure 9.8: A game "tree."

Consider the following mixed strategy, parameterized by $p \in [0, 1]$:

$$\begin{aligned} \sigma_1^p(\text{out}.u) &= \frac{p}{2},\ \sigma_1^p(\text{out}.d) = \frac{1-p}{2}, \\ \sigma_1^p(\text{in}.u) &= \frac{1}{6},\ \sigma_1^p(\text{in}.d) = \frac{2}{6}. \end{aligned}$$

Since no history in H is ruled out by σ_1, there is only one behavior strategy consistent with σ_1:

$$\begin{aligned} \beta_1(\text{in}|\varnothing) &= \frac{\sigma_1(S_1(\text{in}))}{\sigma_1(S_1)} = \frac{1}{6} + \frac{2}{6} = \frac{1}{2}, \\ \beta_1(u|\text{in}) &= \frac{\sigma_1(S_1(\text{in}.u))}{\sigma_i\,(S_1(\text{in}))} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{2}{6}} = \frac{1}{3}. \end{aligned}$$

---

then $2^L = 2 \times 2^{L-1} \geq 2 \times L = L + L \geq L + 1$.)

Let $n_k \geq 2$ for each $k = 1, 2, ...$ We show that $\prod_{k=1}^{L} n_k \geq \sum_{k=1}^{L} n_k$ for each $L = 1, 2, ...$ Let $n^* = \max\{n_1, ..., n_\ell\}$. Then

$$\prod_{k=1}^{L} n_k \geq n^* \times 2^{L-1} \geq n^* \times L \geq \sum_{k=1}^{L} n_k.$$

Note also that β_1 does not depend on the distribution $p : 1-p$ of the probability mass $\sigma_1^p(S_1(\text{out})) = \frac{1}{2}$ between the realization-equivalent strategies out.u and out.d. The reason is that the only difference between these two strategies is the choice between u and d contingent on in, a counterfactual contingency if Rowena plans to go out. Now consider the mixed strategy $\overline{\sigma}_1^p$ with $\overline{\sigma}_1^p(\text{out}.u) = p$, $\overline{\sigma}_1^p(\text{out}.d) = 1-p$, which amounts to the deterministic plan of going out. Since $\overline{\sigma}_1^p(S_1(\text{in})) = 0$, there is a continuum of behavior strategies consistent with $\overline{\sigma}_1^p$, all those in the set

$$\{\overline{\beta}_1 \in \Delta(\{\text{out}, \text{in}\}) \times \Delta(\{u,d\}) : \overline{\beta}_1(\text{out}|\varnothing) = 1\}.$$

▲

Next suppose that player i does indeed randomize according to behavior strategy β_i. Is there a method to obtain a mixed strategy σ_i consistent with β_i? Here is an answer: It is natural, if not entirely obvious, to consider the case where the randomization devices that i would use at different histories are stochastically independent. We call this assumption "*independence across agents*," as one can regard the contingent choice of i at each $h \in H$ as carried out by an agent (i,h) of i who only operates under contingency h. A pure strategy s_i is just a collection $(s_i(h))_{h\in H}$ of contingent choices of the agents (i,h), $h \in H$. If the random choices at different histories are mutually independent, then the probability of pure strategy s_i is the product of the probabilities of the contingent choices $s_i(h)$, $h \in H$. Therefore one can derive from β_i the mixed strategy σ_i that satisfies

$$\forall s_i \in S_i,\ \sigma_i(s_i) = \prod_{h\in H} \beta_i(s_i(h)|h). \tag{9.4.2}$$

It is boring, but rather straightforward to show that such σ_i is consistent with β_i:[25]

Remark 45. *For all $i \in I$, $\sigma_i \in \Delta(S_i)$, $\beta_i \in \bigtimes_{h\in H} \Delta(\mathcal{A}_i(h))$, if (9.4.2) holds, then also (9.4.1) holds, that is, σ_i and β_i are mutually consistent.*

Example 43. In the game "tree" of Figure 9.8, the mixed strategy associated with $\beta_1(\text{in}|\varnothing) = \beta_1(\text{out}|\varnothing) = \frac{1}{2}$, $\beta_1(u|\text{in}) = \frac{1}{3}$, $\beta_1(d|\text{in}) = \frac{2}{3}$

[25] This is one part of what is known as Kuhn's theorem for mixed and behavior strategies. For a proof, see the appendix.

under the "independence-across-agents" assumption is σ_1:

$$\begin{aligned}\sigma_1(\text{out}.u) &= \frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6},\ \sigma_1(\text{out}.d)=\frac{1}{2}\cdot\frac{2}{3}=\frac{2}{6},\\ \sigma_1(\text{in}.u) &= \frac{1}{2}\cdot\frac{1}{3}=\frac{1}{6},\ \sigma_1(\text{in}.d)=\frac{1}{2}\cdot\frac{2}{3}=\frac{2}{6}.\end{aligned}$$

Note that β_1 was derived in Example 42 from an arbitrary mixed strategy with the parametric representation σ_1^p, whereas here we derived exactly $\sigma_1 = \sigma_1^{1/3}$. The reason is that only σ_1^p with $p = \frac{1}{3}$ is consistent with "independence across agents." ▲

Remark 44 says that many behavior strategies may be consistent with a given mixed strategy and that many mixed strategies may be consistent with a given behavior strategy, although only one satisfies the assumption of "independence across agents." Then why is consistency between mixed and behavior strategies important? Intuitively, the reason is that mutually consistent mixed and behavior strategies yield the same probabilities of histories, which is all that matter to expected payoff maximizing players. To make the intuition precise, note that a profile of randomized strategies ξ ($\xi = \sigma = (\sigma_i)_{i\in I}$ or $\xi = \beta = (\beta_i)_{i\in I}$) induces a probability measure $\hat{\zeta}(\cdot|\xi) \in \Delta(Z)$ on terminal histories, which is determined by the following formulas (which implicitly assume "independence across players"):

$$\forall z \in Z,\ \hat{\zeta}(z|\sigma) = \sum_{s:\zeta(s)=z}\left(\prod_{i\in I}\sigma_i(s_i)\right),$$

$$\forall z \in Z,\ \hat{\zeta}(z|\beta) = \prod_{t=1}^{\ell(z)}\prod_{i\in I}\beta_i(\mathbf{a}_i^t(z)|\mathbf{h}^{t-1}(z)).$$

(In the second formula, $\mathbf{a}_i^t(z)$ is the action played by i at stage t in history z, and $\mathbf{h}^{t-1}(z)$ is the prefix of length $t-1$ of z.)

Theorem 32. *(Kuhn)*[26] *For any two profiles of mixed and behavior strategies σ and β the following statements hold:*

[26] In his seminal article ([43]), Harold Kuhn proved two important results about general games with a sequential structure (so called "extensive form games"), one concerns the realization-equivalence of mixed and behavior strategies under the assumption of perfect recall (a generalization of the observable actions assumption), the other concerns the existence of equilibria.

(a) Fix any $i \in I$; if σ_i and β_i satisfy either (9.4.1) or (9.4.2), then

$$\forall s_{-i} \in S_{-i},\ \hat{\zeta}(\cdot|\sigma_i, s_{-i}) = \hat{\zeta}(\cdot|\beta_i, s_{-i}).$$

(b) If, for all $i \in I$, σ_i and β_i satisfy either (9.4.1) or (9.4.2), then

$$\hat{\zeta}(\cdot|\sigma) = \hat{\zeta}(\cdot|\beta).$$

Example 44. Again, the example of Figure 9.8 illustrates. For Rowena (player 1) consider the randomized strategies σ_1^p and β_1 of Example 42. Colin (player 2) is active at only one history, hence there is an obvious isomorphism between his mixed and behavior strategies. Let $\sigma_2(l) = \beta_2(l|\text{in}) = q$. Then $\hat{\zeta}(\cdot|\beta_1, l) = \hat{\zeta}(\cdot|\sigma_1, l)$, $\hat{\zeta}(\cdot|\beta_1, r) = \hat{\zeta}(\cdot|\sigma_1, r)$ and $\hat{\zeta}(\cdot|\beta) = \hat{\zeta}(\cdot|\sigma)$. In particular

$$\begin{aligned}
\hat{\zeta}(v|\beta) &= \frac{1}{2} = \hat{\zeta}(v|\sigma),\ \hat{\zeta}(w|\beta) = \frac{q}{6} = \hat{\zeta}(w|\sigma),\ \hat{\zeta}(x|\beta) = \frac{1-q}{6} = \hat{\zeta}(x|\sigma), \\
\hat{\zeta}(y|\beta) &= \frac{q}{3} = \hat{\zeta}(y|\sigma),\ \hat{\zeta}(z|\beta) = \frac{1-q}{3} = \hat{\zeta}(z|\sigma).
\end{aligned}$$

▲

9.5 Appendix

Proof of Remark 45. Fix i, σ_i, β_i, $\hat{h} \in H$ and $\hat{a}_i \in \mathcal{A}_i(\hat{h})$ (it is notationally convenient to "put a hat" on the fixed history and action). Suppose that (9.4.2) holds; it must be shown that $\sigma_i(S_i(\hat{h})) > 0$ implies $\beta_i(\hat{a}_i|\hat{h}) = \sigma_i(S_i(\hat{h}, \hat{a}_i))/\sigma_i(S_i(\hat{h}))$. For every $h \prec \hat{h}$, let $\hat{\mathbf{a}}_i(h)$ denote the action taken by i at h to reach $\hat{h}$ from h (in other words, $\hat{\mathbf{a}}(h) = (\hat{\mathbf{a}}_j(h))_{j \in I}$ is the unique action profile such that $(h, \hat{\mathbf{a}}(h)) \preceq \hat{h}$). Then $S_i(\hat{h}) = \{s_i : \forall h \prec \hat{h}, s_i(h) = \hat{\mathbf{a}}_i(h)\}$ and $S_i(\hat{h}, \hat{a}_i) = \{s_i : s_i(\hat{h}) = \hat{a}_i, \forall h \prec \hat{h}, s_i(h) = \hat{\mathbf{a}}_i(h)\}$, so that, for each $s_i \in S_i(\hat{h})$,

$$\prod_{h \in H} \beta_i(s_i(h)|h) = \left(\prod_{h \in H : h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h) \right) \cdot \left(\prod_{h \in H : h \nprec \hat{h}} \beta_i(s_i(h)|h) \right),$$

and for each $s_i \in S_i(\hat{h}, \hat{a}_i)$,

$$\prod_{h \in H} \beta_i(s_i(h)|h) = \left(\prod_{h \in H : h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h) \right) \cdot \beta_i(\hat{a}_i|\hat{h}) \cdot \left(\prod_{h \in H : h \npreceq \hat{h}} \beta_i(s_i(h)|h) \right).$$

Therefore,

$$\begin{aligned}\sigma_i(S_i(\hat{h})) &= \sum_{s_i \in S_i(\hat{h})} \sigma_i(s_i) = \sum_{s_i \in S_i(\hat{h})} \prod_{h \in H} \beta_i(s_i(h)|h) \\ &= \left(\prod_{h \in H: h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h) \right) \cdot \left(\sum_{s_i \in S_i(\hat{h})} \prod_{h \in H: h \not\prec \hat{h}} \beta_i(s_i(h)|h) \right),\end{aligned}$$

$$\begin{aligned}\sigma_i(S_i(\hat{h}, \hat{a}_i)) &= \sum_{s_i \in S_i(\hat{h}, a_i)} \sigma_i(s_i) = \sum_{s_i \in S_i(\hat{h}, a_i)} \prod_{h \in H} \beta_i(s_i(h)|h) \\ &= \left(\prod_{h \in H: h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h) \right) \cdot \beta_i(\hat{a}_i|\hat{h}) \cdot \left(\sum_{s_i \in S_i(\hat{h}, \hat{a}_i)} \prod_{h \in H: h \not\preceq \hat{h}} \beta_i(s_i(h)|h) \right).\end{aligned}$$

Count the partial histories that do not weakly precede $\hat{h}$ in any order, e.g., $\{h \in H : h \not\preceq \hat{h}\} = \{h_1, ..., h_L\}$. Now note that there is a canonical bijection between $S_i(\hat{h})$ and $\mathcal{A}_i(\hat{h}) \times \left(\bigtimes_{k=1}^{L} A_i(h_k) \right)$, and between $S_i(\hat{h}, \hat{a}_i)$ and $\bigtimes_{k=1}^{L} \mathcal{A}_i(h_k)$. Given this, we can re-order terms as follows:

$$\begin{aligned}\sum_{s_i \in S_i(\hat{h}, \hat{a}_i)} \prod_{h \in H: h \not\preceq \hat{h}} \beta_i(s_i(h)|h) &= \sum_{k=1}^{L} \sum_{a_{i,k} \in \mathcal{A}_i(h_k)} \beta_i(a_{i,k}|h_k) = 1, \\ \sum_{s_i \in S_i(\hat{h})} \prod_{h \in H: h \not\prec \hat{h}} \beta_i(s_i(h)|h) &= \sum_{a_i \in \mathcal{A}_i(\hat{h})} \beta_i(a_i|\hat{h}) \sum_{k=1}^{L} \sum_{a_{i,k} \in \mathcal{A}_i(h_k)} \beta_i(a_{i,k}|h_k) = 1.\end{aligned}$$

Then

$$\begin{aligned}\sigma_i(S_i(\hat{h})) &= \prod_{h \in H: h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h), \\ \sigma_i(S_i(\hat{h}, \hat{a}_i)) &= \left(\prod_{h \in H: h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h) \right) \cdot \beta_i(\hat{a}_i|\hat{h}),\end{aligned}$$

and

$$\beta_i(\hat{a}_i|\hat{h}) \sigma_i(S_i(\hat{h})) = \beta_i(\hat{a}_i|\hat{h}) \left(\prod_{h \in H: h \prec \hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h) \right) = \sigma_i(S_i(\hat{h}, \hat{a}_i)).$$

■

10

Rational Planning

In this chapter we analyze rational planning from the perspective of a single player with a subjective probabilistic conjecture about the behavior of co-players. We introduce and analyze several dynamic programming properties for strategies in finite multistage games with observable actions. In particular, we focus on **folding-back optimality** as a representation of rational planning: the decision maker computes his subjectively optimal strategy starting from the last stage of the game and factoring into the decision problems of earlier stages the expected payoffs computed for the later stages. One-Step Optimality is a property of a strategy given the conjecture: at every history, the decision maker has no incentive to change the prescribed action, keeping fixed the actions prescribed for the future stages. Sequential Optimality takes a forward-planning approach: at every history, the continuation strategy maximizes the player's expected payoff conditional on reaching the history. We show that folding-back optimality is equivalent to One-Step Optimality (*Folding-Back Principle*) and to Sequential Optimality (*Optimality Principle*). It follows that Sequential Optimality is equivalent to One-Step Optimality (*One-Deviation Principle*). In light of these equivalence results, we adopt Sequential Optimality for at least one conjecture as a notion of justifiability for multistage games. Extending the analogous result obtained for static games (Lemma 2 of Chapter 3), we characterize justifiability with the notion of **conditional dominance**, according to which a strategy s_i is conditionally dominated if there exists a history h (consistent with s_i) such that s_i is dominated conditional on observing h. To simplify the analysis, we assume that *the game tree is finite*, that is, there is finite horizon and the cardinality of each feasible set of actions is finite. This maintained assumption is not explicitly mentioned in the results below. The finiteness assumption is removed in Section 10.5, where we show that the One-Deviation Principle also holds for all games with *infinite horizon* that satisfy a regularity property called "continuity at infinity."

10.1 Conditional Beliefs and Decision Trees

In the analysis of best replies in static games (Chapter 3), we modeled players' conjectures as probability measures over the actions of the co-players. There are at least two conceivable ways to extend the definition of conjecture to multistage games: (i) Define conjectures as (possibly correlated) probability measures over the strategies of the co-players, where such strategies are interpreted as descriptions of how co-players (would) behave at each non-terminal history. Note that conjectures have to be updated, or revised, as the play unfolds. (ii) Define conjectures as arrays of (possibly correlated) probability measures over co-players' feasible actions sets, one such measure for each non-terminal history. These two approaches are essentially equivalent. To see this, consider the easier case of two-person games, where there is only one co-player. Return to Section 9.4: An array of probability measures $\beta^i \in \times_{h\in H} \Delta(\mathcal{A}_{-i}(h))$ can also be interpreted as a behavior strategy of co-player $-i$. By Theorem 32, β^i is realization equivalent to a probability measure $\mu^i \in \Delta(S_{-i})$. Now consider any $\mu^i \in \Delta(S_{-i})$ such that every non-terminal history can be reached with positive probability, that is, $\mu^i(S_{-i}(h)) > 0$ for every $h \in H$. Then we can recover a realization-equivalent β^i letting

$$\beta^i(a_{-i}|h) = \frac{\mu^i(S_{-i}(h, a_{-i}))}{\mu^i(S_{-i}(h))}$$

for all $h \in H$ and $a_{-i} \in \mathcal{A}_{-i}(h)$.[1] If instead $\mu^i(S_{-i}(h)) = 0$ for some h, we should assume that upon observing h player i forms a new conjecture, which can be used to derive $\beta^i(\cdot|h)$. Thus, we obtain $\beta^i(\cdot|h)$ for every

[1] Recall from Section 9.4 of Chapter 9 that $S_j(h, a_j)$ is the set of strategies of player j consistent with h that select action a_j at h. Here, player j is i's opponent, that is, $j = -i$.

$h \in H$. We omit the details.[2]

To analyze rational planning, it is easier to model conjectures as arrays of probability measures over co-players' feasible actions. Arrays of probability measures over strategies will be more convenient to analyze how conjectures are formed through strategic reasoning—we will introduce them in Section 10.4 before using them in Chapter 11. From now on, we reserve the term **conjecture** for a system of conditional beliefs about co-players' actions

$$\beta^i = (\beta^i(\cdot|h))_{h\in H} \in \bigtimes_{h\in H} \Delta(\mathcal{A}_{-i}(h)),$$

where $\beta^i(a_{-i}|h)$ denotes the subjective probability of a_{-i} conditional on h. In this section, *we focus on the subjectively optimal behavior of player i given his conjectures*, hence we fix β^i. Nonetheless, we make the dependence on β^i explicit, because in a game-theoretic analysis such conjectures are endogenous, i.e., they must be consistent with the strategic analysis of the game. With multiple co-players, β^i is a kind of "correlated behavior strategy."

Let $\beta_i^{s_i}$ denote the degenerate behavior strategy determined by pure strategy s_i according to the obvious rule:

$$\forall h \in H, \forall a_i \in \mathcal{A}_i(h),\ \beta_i^{s_i}(a_i|h) = 1 \Leftrightarrow a_i = s_i(h).$$

The probability of action profile $a = (a_i, a_{-i})$ conditional on h given strategy s_i and conjecture β^i is

$$\mathbb{P}^{s_i,\beta^i}(a|h) := \beta_i^{s_i}(a_i|h)\beta^i(a_{-i}|h) = \begin{cases} 0, & \text{if } a_i \neq s_i(h), \\ \beta^i(a_{-i}|h), & \text{if } a_i = s_i(h). \end{cases} \tag{10.1.1}$$

Equation 10.1.1 can be interpreted in two ways. The first is that player i will certainly implement his plan s_i. The second is that s_i is an objective description of how player i would behave conditional on reaching each history h. In this chapter we adopt the first interpretation, from the viewpoint of player i: he is certain that he is going to implement his plan s_i; therefore, for the continuation of the game, he assigns probability 1 to the

[2] The key assumption is that player i updates or revises his beliefs in compliance with the chain rule of conditional probabilities. See below.

actions precribed by s_i. Using the **chain rule** of conditional probabilities,[3] we can define the subjective probability of h' conditional on prefix h ($h \prec h'$), given that s_i is played from h onward: let $h = (a^1, ..., a^{\ell(h)})$, $h' = (a^1, ..., a^{\ell(h)}, ..., a^{\ell(h')})$; then[4]

$$\mathbb{P}^{s_i,\beta^i}(h'|h) = \prod_{t=\ell(h)+1}^{\ell(h')} \beta_i^{s_i}(a_i^t|h, ..., a^{t-1})\beta^i(a_{-i}^t|h, ..., a^{t-1}). \qquad (10.1.2)$$

Recall from Section 9.3 that $S_i^{\succeq h}$ is the set of continuation strategies in the subgame with root h, and that $(s_i|h) \in S_i^{\succeq h}$ is the continuation strategy implied by s_i (that is, the projection of s_i onto $S_i^{\succeq h}$). With this, note that the above conditional probability is well defined and meaningful even if s_i *precludes* h, because $\mathbb{P}^{s_i,\beta^i}(h'|h)$ only depends on $(s_i|h) \in S_i^{\succeq h}$, i.e., it depends only on how s_i behaves in the sub-tree with root h. The subjective unconditional probability of reaching h' given s_i is

$$\mathbb{P}^{s_i,\beta^i}(h') = \mathbb{P}^{s_i,\beta^i}(h'|\varnothing).$$

Comment: If $-i$ is just one player, then β^i—as a mathematical object—is a behavior strategy. For example, $-i = 0$ may be the Chance player,[5] in this case $\beta^i = \pi_0$ represents the objective probabilities of chance moves in a decision tree; or $-i = 0$ may be Nature[6] and then $\beta^i = \beta_0$ represents the subjective conditional probabilities assigned by i to moves by Nature. Such subjective probabilities may be obtained *via* updating from some prior $\mu_0 \in \Delta\left(\bigtimes_{h\in H} \Delta(\mathcal{A}_0(h))\right)$ on objective probability models:

---

[3]The chain rule of conditional probabilities states that, given a chain of three events $D \subseteq E \subseteq F$,

$$\mathbb{P}(D|F) = \mathbb{P}(D|E)\mathbb{P}(E|F).$$

Here we can look at chains of events in Z corresponding to chains of histories: if $h \prec h' \prec h''$, then $Z(h'') \subseteq Z(h') \subseteq Z(h)$.

[4]Equation 10.1.2 and other equations in the continuation of this chapter contain the following abuse of notation: When $t = \ell(h) + 1$, we have $a^{t-1} = a^{\ell(h)}$, which is already included in h, therefore $h, ..., a^{t-1}$ shall be read as just h.

[5]A pseudo-player choosing some moves with known objective probabilities.

[6]A pseudo-player selecting a state of nature with unknown objective probabilities.

for example, if $\mathrm{supp}\mu_0 = \{\pi_0^1, ..., \pi_0^K\}$ and no h is ruled out by μ_0,[7] then

$$\forall h \in H,\ \forall a_0 \in \mathcal{A}_0(h),\ \beta_0(a_0|h) = \frac{\sum_{k=1}^K \mathbb{P}^{s_i,\pi_0^k}(h,(s_i(h),a_0))\mu_0(\pi_0^k)}{\sum_{k=1}^K \mathbb{P}^{s_i,\pi_0^k}(h)\mu_0(\pi_0^k)}$$

for every s_i consistent with h.[8]

The structure $\left\langle(\bar{H},\preceq),u_i,\beta^i\right\rangle$ forms a **subjective decision tree for player** i: $(\bar{H},\preceq)$ is the tree, $u_i : Z \to \mathbb{R}$ is the payoff function of i, and β^i is a subjective assessment of the probabilities of actions not controlled by i.

Example 45. The BoS with Dissipative Action is a two-person game; therefore a conjecture of player i is—as a mathematical object—a behavior strategy of the co-player $-i$. We will focus on the following examples of conditional beliefs (they are not meant to form an equilibrium):

$$\begin{array}{lcl} \beta^b & : & \beta^b(U|N) = \frac{3}{4},\ \beta^b(u|B) = \frac{1}{4}; \\ \beta^a & : & \beta^a(N|\varnothing) = 1,\ \beta^a(L|N) = 0,\ \beta^a(l|B) = 1; \end{array}$$

hence β^a corresponds to the pure strategy $s_b = N.R.l$ of Bob. Next we give a few examples of realization probabilities, where s_b, β^b and β^a are specified above:

$$\begin{array}{rcl} \mathbb{P}^{s_b,\beta^b}(B,(u,l)) & = & 0 = \mathbb{P}^{s_b,\beta^b}(N,(U,L)), \\ \mathbb{P}^{s_b,\beta^b}(N,(U,R)) & = & \frac{3}{4}, \\ \mathbb{P}^{D.u,\beta^a}(B,(u,l)) & = & 0,\ \mathbb{P}^{D.u,\beta^a}(B,(u,l)|B) = 1. \end{array}$$

▲

[7] That is, $\sum_{k=1}^K \mathbb{P}^{s_i,\pi_0^k}(h)\mu_0(\pi_0^k) > 0$ for every $h \in H$ and s_i consistent with h. This is true, for example, if there is at least one model $\pi_0^k \in \mathrm{supp}\mu_0$ such that $\pi_0^k(a_0|\bar{h}) > 0$ for every $\bar{h} \in H$ and $a_0 \in \mathcal{A}_0(\bar{h})$.

[8] Recall that strategy s_i is consistent with history h if $h \preceq \zeta(s_i, s_{-i})$ for some $s_{-i} \in S_{-i}$. The set of such strategies is denoted by $S_i(h)$ (see Section 9.4).

10.2 Subjective Values

For all $s_i \in S_i$, $\beta^i \in \times_{h\in H} \Delta(\mathcal{A}_{-i}(h))$ and $h \in H$, we can determine the subjective **value of** h given that s_i is followed from h onward:

$$V_i^{s_i,\beta^i}(h) := \sum_{z\in Z(h)} u_i(z)\mathbb{P}^{s_i,\beta^i}(z|h).$$

Like $\mathbb{P}^{s_i,\beta^i}(z|h)$, also $V_i^{s_i,\beta^i}(h)$ is well defined and meaningful even if s_i precludes h, because it depends only on $(s_i|h)$, the sub-strategy induced by s_i in the sub-tree with root h. Therefore, it makes sense to write (with a slight abuse of notation) $V_i^{s_i,\beta^i}(h) = V_i^{(s_i|h),\beta^i}(h)$, and to write $V_i^{t_i,\beta^i}(h)$ for any $t_i \in S_i^{\succeq h}$; that is, $V_i^{t_i,\beta^i}(h)$ is the value of playing sub-strategy t_i in the decision sub-tree with root h.

Similarly, for every $a_i \in \mathcal{A}_i(h)$ we can define the **value of** taking **action** a_i at h given that s_i will be followed from the next stage:

$$V_i^{s_i,\beta^i}(h, a_i) := \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|h)V_i^{s_i,\beta^i}(h,(a_i,a_{-i})).$$

Finally, the **ex ante value** of playing s_i is

$$V_i^{s_i,\beta^i}(\varnothing) := \sum_{z\in Z} u_i(z)\mathbb{P}^{s_i,\beta^i}(z).$$

Note: since $\beta_i^{s_i}(a_i|h) = 1$ if $a_i = s_i(h)$ and $\beta_i^{s_i}(a_i|h) = 0$ otherwise, by inspection of the definitions we have

$$\begin{aligned}
& V_i^{s_i,\beta^i}(h, s_i(h)) \\
= & \sum_{(a_i,a_{-i})\in\mathcal{A}(h)} V_i^{s_i,\beta^i}(h,(a_i,a_{-i}))\beta_i^{s_i}(a_i|h)\beta^i(a_{-i}|h) \\
= & \sum_{(a_i,a_{-i})\in\mathcal{A}(h)} \beta_i^{s_i}(a_i|h)\beta^i(a_{-i}|h) \sum_{z\in Z(h,(a_i,a_{-i}))} u_i(z)\mathbb{P}^{s_i,\beta^i}(z|h,(a_i,a_{-i})) \\
= & \sum_{z\in Z(h)} u_i(z)\mathbb{P}^{s_i,\beta^i}(z|h) = V_i^{s_i,\beta^i}(h)
\end{aligned}$$

(the second to last equality follows from the fact that $Z(h) = \bigcup_{a\in\mathcal{A}(h)} Z(h,a)$ and from the chain rule). Thus we obtain:

Remark 46. *For every $s_i \in S_i$ and $h \in H$, $V_i^{s_i,\beta^i}(h, s_i(h)) = V_i^{s_i,\beta^i}(h)$.*

10.3 Rational Planning

The values defined above take a particular strategy as given. Next we define recursively the subjective value of reaching a history h and of taking an action a_i at h, *under the presumption that the behavior of player i will be subjectively rational in the following stages.* We use the symbol $\hat{V}_i^{\beta^i}$ to denote such values to emphasize that they are optimal given conjecture β^i.

The recursion is based on $L(\Gamma(h))$, the height of the subgame starting at $h \in \bar{H}$:[9]

- If $L(\Gamma(h)) = 0$ (that is, $h \in Z$) let
$$\hat{V}_i^{\beta^i}(h) = u_i(h).$$
- Suppose $\hat{V}_i^{\beta^i}(h')$ has been defined for every h' with $L(\Gamma(h')) \leq k$. Then if $L(\Gamma(h)) = k + 1$ let
$$\hat{V}_i^{\beta^i}(h, a_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \hat{V}_i^{\beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h),$$
$$\hat{V}_i^{\beta^i}(h) = \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h, a_i). \tag{10.3.1}$$

This determination of optimal values by backward recursion is often called "**folding back**" in the literature on dynamic programming. We can interpret folding back as a method (an algorithm) to compute an optimal plan, which "collects" actions that solve the maximization problems of equation (10.3.1) at every history. We call a strategy obtained in this way **folding-back optimal.**

Definition 55. *A strategy $\bar{s}_i \in S_i$ is **folding-back optimal given** β^i if*
$$\forall h \in H,\ \hat{V}_i^{\beta^i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h, a_i) = \hat{V}_i^{\beta^i}(h).$$

For every conjecture there is a folding-back optimal strategy: by finiteness of the game, the maximization problem of equation (10.3.1) has a solution at every history.

[9] Recall, $L(\Gamma(h)) := \max_{z \in Z(h)} \ell(z) - \ell(h)$.

Remark 47. *There exists at least one folding-back optimal strategy for any given* β^i.

Of course, there can be more than one folding-back optimal strategy given the same conjecture. This happens whenever the maximazion problem of equation (10.3.1) has more than one solution at a history. When there are multiple solutions at more than one history, a natural question arises: is every combination of optimal actions a folding-back optimal strategy? The answer is affirmative, because folding back only depends on the maximal expected payoff computed at the future histories, not on the specific maximizers.

Remark 48. *For every two folding-back optimal strategies* $\bar{s}_i, \bar{s}'_i$ *given* β^i, *and for every strategy* $\bar{s}''_i$ *such that* $\bar{s}''_i(h) \in \{\bar{s}_i(h), \bar{s}'_i(h)\}$ *for each* $h \in H$, $\bar{s}''_i$ *is folding-back optimal given* β^i.

Remark 48 contrasts with Example 49, where we will show that the set of strategies that are optimal given *some* conjecture does not feature every combination of actions that are prescribed by some strategy in the set.[10]

Next we move to alternative notions of optimal planning. **One step-optimality** takes the strategy $\bar{s}_i$ as given and makes sure that at every history h player i has no incentive to deviate to an action $a_i \neq \bar{s}_i(h)$ given β^i. **Sequential optimality** is a notion of forward planning: conditional on every history, the continuation strategy must maximize the expected payoff given β^i. Under this idea of forward planning, it makes sense to introduce the weaker notion of **weak sequential optimality**, which only considers the histories that can be reached given the plan. Finally, we introduce **ex-ante optimality**, which only requires optimality from the viewpoint of the initial history.

Definition 56. *A strategy* $\bar{s}_i \in S_i$ *is*

- ***one-step optimal given*** β^i *if*

$$\forall h \in H,\ V_i^{\bar{s}_i,\beta^i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} V_i^{\bar{s}_i,\beta^i}(h, a_i); \qquad (10.3.2)$$

[10] By Remark 48, the set of folding-back optimal strategies given β^i can be written as the intersection over all histories h of the sets of strategies that prescribe a folding-back optimal action at h given β^i, and an intersection of Cartesian sets is Cartesian. By contrast, the union over all conjectures β^i of the Cartesian sets of folding-back optimal strategies given β^i need not be Cartesian.

- ***sequentially optimal given*** β^i *if*

$$\forall h \in H,\ V_i^{\bar{s}_i,\beta^i}(h) = \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h); \tag{10.3.3}$$

- ***weakly sequentially optimal given*** β^i *if*[11]

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\beta^i}(h) = \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h); \tag{10.3.4}$$

- ***ex ante optimal given*** β^i *if*

$$\bar{s}_i \in \arg\max_{s_i \in S_i} V_i^{s_i,\beta^i}(\varnothing). \tag{10.3.5}$$

We will sometimes omit the phrase "given β^i" when the reference to a specific conjecture is clear from the context. We will sometimes refer to the maximization property at h for a strategy $\bar{s}_i$ that satisfies one-step optimality as the **local optimality of** $\bar{s}_i$ **at** h.

Ex-ante optimality implies that the strategy is optimal not just from the viewpoint of the initial history, but also at every history that player i initially expects to reach with positive probability.

Proposition 2. *A strategy* $\bar{s}_i$ *is ex ante optimal if and only if*

$$\forall h \in H,\ \mathbb{P}^{\bar{s}_i,\beta^i}(h) > 0 \Rightarrow (\bar{s}_i|h) \in \arg\max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h). \tag{10.3.6}$$

We will state a more general result for beliefs over strategies in the next section (Proposition 25), thus we refer the reader to the proof of that result for an intuition of how to prove Proposition 2.

In the rest of this section, we investigate the relations among the notions of optimal plan introduced in Definitions 55 and 56. By inspection of Definition 56 and Remark 46, we immediately obtain the following relation between sequential optimality and one-step optimality (and ex-ante optimality as well).

Remark 49. *If a strategy is sequentially optimal given* β^i*, then it is one-step optimal and ex ante optimal given* β^i*.*

[11] Recall from Chapter 9 that $H_i(\bar{s}_i)$ is the set of non-terminal histories not precluded by $\bar{s}_i$.

The following results show that folding-back optimality and sequential optimality are equivalent to one-step optimality. First we establish the equivalence between the two properties defined by local optimization conditions, that is, folding-back optimality and one-step optimality; this is called the "Folding Back Principle." Next we show that folding-back optimality is equivalent to sequential optimality; this is the so called "Optimality Principle." As a corollary we obtain the "One-Deviation Principle" which states the equivalence between one-step optimality and sequential optimality.

Proposition 3. (Folding Back Principle)
(I) A strategy $\bar{s}_i$ *is one-step optimal given* β^i *if and only if*

$$\begin{aligned} V_i^{\bar{s}_i,\beta^i}(h,a_i) &= \hat{V}_i^{\beta^i}(h,a_i), \\ V_i^{\bar{s}_i,\beta^i}(h) &= \hat{V}_i^{\beta^i}(h) \end{aligned} \tag{10.3.7}$$

for every $h \in H$ *and* $a_i \in \mathcal{A}_i(h)$.
(II) A strategy $\bar{s}_i$ *is folding-back optimal given* β^i *if and only if it is one-step optimal given* β^i.

Proof. We prove by induction on the height of subgames that if $\bar{s}_i$ is one-step optimal given β^i then (10.3.7) holds for every $h \in H$ and $a_i \in \mathcal{A}_i(h)$. It is easy to show that also the converse of this statement holds, thus that (I) holds, and that (I) is equivalent to (II). We leave the proof of these steps as an exercise.

Suppose that $\bar{s}_i$ is one-step optimal given β^i.

Basis step. Consider any $h \in H$ such that $L(\Gamma(h)) = 1$. Then $(h,a) \in Z$ for every $a \in \mathcal{A}(h)$; therefore

$$V_i^{\bar{s}_i,\beta^i}(h,a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} u_i(h,(a_i,a_{-i}))\beta^i(a_{-i}|h) = \hat{V}_i^{\beta^i}(h,a_i)$$

for every $a_i \in \mathcal{A}_i(h)$; hence

$$\begin{aligned} V_i^{\bar{s}_i,\beta^i}(h) &= V_i^{\bar{s}_i,\beta^i}(h,\bar{s}_i(h)) \overset{(loc.opt.)}{=} \max_{a_i\in\mathcal{A}_i(h)} V_i^{\bar{s}_i,\beta^i}(h,a_i) \\ &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h,a_i) = \hat{V}_i^{\beta^i}(h), \end{aligned}$$

where the first equality holds by Remark 46 and the second equality holds because $\bar{s}_i$ is locally optimal (loc.opt.) at h.

Inductive step. Suppose that (10.3.7) holds for each $h \in H$ with $L(\Gamma(h)) \leq k$. Now, fix h with $L(\Gamma(h)) = k+1$. Then $L(\Gamma(h,a)) \leq k$ for each $a \in \mathcal{A}(h)$. Therefore the inductive hypothesis (I.H.) implies

$$V_i^{\bar{s}_i,\beta^i}(h,a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} V_i^{\bar{s}_i,\beta^i}(h,(a_i,a_{-i}))\beta^i(a_{-i}|h)$$

$$\overset{(I.H.)}{=} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \hat{V}_i^{\beta^i}(h,(a_i,a_{-i}))\beta^i(a_{-i}|h) =: \hat{V}_i^{\beta^i}(h,a_i)$$

for every $a_i \in \mathcal{A}_i(h)$; hence

$$\begin{aligned} V_i^{\bar{s}_i,\beta^i}(h) &= V_i^{\bar{s}_i,\beta^i}(h,\bar{s}_i(h)) \overset{(loc.opt.)}{=} \max_{a_i\in\mathcal{A}_i(h)} V_i^{\bar{s}_i,\beta^i}(h,a_i) \\ &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h,a_i) = \hat{V}_i^{\beta^i}(h), \end{aligned}$$

where the first equality holds by Remark 46 and the second equality holds because $\bar{s}_i$ is locally optimal at h. ■

With this, we can prove the Optimality Principle.

Theorem 33. (Optimality Principle) *A strategy of player i is sequentially optimal given β^i if and only if it is folding-back optimal given β^i.*

Proof. To ease notation, in the proof we use the symbols t_i, t_i' to denote generic *sub*-strategies.

(If) Let $\bar{s}_i$ be folding-back optimal given β^i. We will prove by induction on the height of subgames that

$$\forall h \in H,\, \hat{V}_i^{\beta^i}(h) \geq \max_{t_i\in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h). \tag{10.3.8}$$

Since $\bar{s}_i$ is folding-back optimal, it satisfies one-step optimality (Proposition 3 II), therefore $V_i^{\bar{s}_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$ (Proposition 3 I); this means that $\bar{s}_i$ is sequentially optimal given β^i.

Basis step. Let $L(\Gamma(h)) = 1$. Then $(h,a) \in Z$ for every $a \in \mathcal{A}(h)$; therefore

$$\max_{t_i\in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h) = \max_{a_i\in\mathcal{A}_i(h)} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} u_i(h,(a_i,a_{-i}))\beta^i(a_{-i}|h);$$

furthermore

$$\hat{V}_i^{\beta^i}(h) \geq \hat{V}_i^{\beta^i}(h, a_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} u_i(h, (a_i, a_{-i}))\beta^i(a_{-i}|h)$$

for every $a_i \in \mathcal{A}_i(h)$, where both the inequality and the equality hold by definition. Hence (10.3.8) holds at h.

Inductive step. Suppose that $\hat{V}_i^{\beta^i}(h) \geq \max_{t_i \in S_i^{\succeq h}} V_i^{t_i, \beta^i}(h)$ for every $h \in H$ with $L(\Gamma(h)) \leq k$. Now fix h with $L(\Gamma(h)) = k + 1$. Then $L(\Gamma(h, a)) \leq k$ for each $a \in \mathcal{A}(h)$, and the inductive assumption (I.H.) yields

$$\begin{aligned} \hat{V}_i^{\beta^i}(h) &\geq \hat{V}_i^{\beta^i}(h, a_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \hat{V}_i^{\beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h) \\ &\overset{(I.H.)}{\geq} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{t_i \in S_i^{\succeq (h,(a_i,a_{-i}))}} V_i^{t_i, \beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h) \end{aligned}$$

for every $a_i \in \mathcal{A}_i(h)$. Therefore, for every sub-strategy $t_i' \in S_i^{\succeq h}$,

$$\begin{aligned} \hat{V}_i^{\beta^i}(h) &\geq \max_{a_i \in \mathcal{A}_i(h)} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{t_i \in S_i^{\succeq (h,(a_i,a_{-i}))}} V_i^{t_i, \beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h) \\ &\geq \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{t_i \in S_i^{\succeq (h,(t_i'(h),a_{-i}))}} V_i^{t_i, \beta^i}(h, (t_i'(h), a_{-i}))\beta^i(a_{-i}|h) \\ &\geq \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} V_i^{t_i', \beta^i}(h, (t_i'(h), a_{-i}))\beta^i(a_{-i}|h) = V_i^{t_i', \beta^i}(h). \end{aligned}$$

The latter set of inequalities yield

$$\hat{V}_i^{\beta^i}(h) \geq \max_{t_i' \in S_i^{\succeq h}} V_i^{t_i', \beta^i}(h)$$

as desired.

(Only if) If $\bar{s}_i$ is sequentially optimal, then it satisfies one-step optimality (Remark 49), which in turn implies that it is folding-back optimal (Proposition 3 II). ■

Proposition 3 and Theorem 33 yield the following result, known as the **One-Deviation (OD) Principle**:

Corollary 6. (One-Deviation Principle) *A strategy of player i is sequentially optimal given β^i if and only if it is one-step optimal given β^i.*

Since folding-back optimality, sequential optimality, and one-step optimality are equivalent, we can generically refer to any of them as "**rational planning.**" The following result relates rational planning to weak sequential optimality.

Proposition 4. *A strategy $\bar{s}_i$ is weakly sequentially optimal given β^i if and only if there is a strategy s_i that is behaviorally equivalent to $\bar{s}_i$ and is sequentially optimal given β^i.*

Proof. Suppose that $\bar{s}_i$ is weakly sequentially optimal. Let s_i^* be any strategy that is folding-back optimal given β^i (by finiteness, there is at least one such strategy). Define s_i as follows:

$$s_i(h) = \begin{cases} s_i^*(h), & \text{if } h \in H \backslash H_i(\bar{s}_i), \\ \bar{s}_i(h), & \text{if } h \in H_i(\bar{s}_i). \end{cases}$$

By construction, s_i is behaviorally equivalent to $\bar{s}_i$ and satisfies $V_i^{s_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$ for every $h \in H \backslash H_i(\bar{s}_i)$. Since $\bar{s}_i$ is weakly sequentially optimal and behaviorally equivalent to s_i,

$$V_i^{s_i,\beta^i}(h) = V_i^{\bar{s}_i,\beta^i}(h) = \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$$

for every $h \in H_i(\bar{s}_i)$, where the first equality follows from behavioral equivalence, which implies realization-equivalence, hence, value-equivalence, and the third equality follows from the proof of the optimality principle (Theorem 33). Thus, s_i is folding-back optimal, because $V_i^{s_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$ for every $h \in H$. By the optimality principle, s_i is also sequentially optimal.

The converse follows by inspection of the definitions. ∎

The following example illustrates all the notions of optimal plan introduced in Definitions 55 and 56.

Example 46. Using the conditional beliefs of Example 45 we obtain the following values and optimal strategies. First consider Ann:

$$\begin{aligned}\hat{V}_a^{\beta^a}(N,U) &= 0,\ \hat{V}_a^{\beta^a}(N,D)=1,\\ \hat{V}_a^{\beta^a}(B,u) &= 4,\ \hat{V}_a^{\beta^a}(B,d)=0,\\ \hat{V}_a^{\beta^a}(N) &= 1=\hat{V}_a^{\beta^a}(\varnothing),\ \hat{V}_a^{\beta^a}(B)=4.\end{aligned}$$

Strategy $D.u$ of Ann is folding-back, sequentially, and ex ante optimal, and it satisfies one-step optimality; strategy $D.d$ is ex ante optimal, but it does not satisfy one-step optimality and is not folding-back, nor sequentially optimal. The difference between ex ante optimality and the three kinds of dynamic optimality depends on the fact that Ann initially does not expect that the subgame with root B can be reached. Since Ann cannot move more than once, there is no difference between sequential and weakly sequential optimality in her case. Next consider Bob:

$$\begin{aligned}\hat{V}_b^{\beta^b}(N,L) &= \frac{3}{4},\ \hat{V}_b^{\beta^b}(N,R)=1,\\ \hat{V}_b^{\beta^b}(B,l) &= -\frac{7}{4},\ \hat{V}_b^{\beta^b}(B,r)=1,\\ \hat{V}_b^{\beta^b}(B) &= \hat{V}_b^{\beta^b}(N)=\hat{V}_b^{\beta^b}(\varnothing)=1;\end{aligned}$$

strategies $N.R.r$ and $B.R.r$ of Bob are folding-back, sequentially, and ex ante optimal, and they satisfy one-step optimality; strategies $N.R.l$ (realization-equivalent to $N.R.r$) and $B.L.r$ (realization-equivalent to $B.R.r$) are ex ante optimal and weakly sequentially optimal, but do not satisfy one-step optimality and are not folding-back, nor sequentially optimal. ▲

10.4 Conditional Probability and Dominance

We have seen how, in order to plan rationally, player i must form conjectures about the behavior of the co-players at each non-terminal history. A deterministic conjecture specifies an action profile $a_{-i} \in \mathcal{A}_{-i}(h)$ for each $h \in H$. Formally, such specification is an element of $S_{-i} = \bigtimes_{h\in H} \mathcal{A}_{-i}(h)$, the set of functions $s_{-i} : H \to A_{-i}$ such that $s_{-i}(h) \in \mathcal{A}_{-i}(h)$ for each $h \in H$. Elements of S_{-i} may also be interpreted as profiles of plans in the minds of the co-players, but this is *not* the interpretation used here, because the payoff of i is affected by co-players' behavior, not by their plans. (Of course, if we assume that plans are necessarily carried out, the two interpretations are equivalent.) Thus, a probabilistic conjecture can be expressed by assigning probabilities to elements of S_{-i}. Such assignments must take into account the information about s_{-i} revealed by the play as it unfolds, and condition on such information. This representation in terms of conditional beliefs allows us to characterize rational planning with a notion of dominance.

At any given history $h \in H$, a probability measure over the elements of S_{-i} that are consistent with h suffices for our player to solve his decision problem at h. Then, rational planning is possible if the beliefs at different histories satisfy natural consistency rules. Recall from Chapter 9 that $S_{-i}(h)$ is the set of s_{-i} that do not prevent h from occurring, that is, the s_{-i} such that $h \prec \zeta(s_i, s_{-i})$ for some s_i. Thus, let

$$\mathcal{H}_{-i} = \{C_{-i} \subseteq S_{-i} : \exists h \in H, C_{-i} = S_{-i}(h)\}$$

denote the collection of "relevant hypotheses" or "conditioning events" about co-players' behavior.

Definition 57. *A* ***Conditional Probability System*** *is an array of probability measures* $\mu^i = (\mu^i(\cdot|C_{-i}))_{C_{-i}\in\mathcal{H}_{-i}} \in (\Delta(S_{-i}))^{\mathcal{H}_{-i}}$ *such that:*

1. *for every* $C_{-i} \in \mathcal{H}_{-i}$*,* $\mu^i(C_{-i}|C_{-i}) = 1$*;*
2. *for all* $C_{-i}, D_{-i} \in \mathcal{H}_{-i}$ *with* $D_{-i} \subseteq C_{-i}$*, and for every* $E_{-i} \subseteq D_{-i}$*,*

$$\mu^i(E_{-i}|C_{-i}) = \mu^i(E_{-i}|D_{-i}) \cdot \mu^i(D_{-i}|C_{-i}).$$

A Conditional Probability System (henceforth, CPS) is an array of probability measures over S_{-i} that conditions on information (property 1) and satisfies the *chain rule of conditional probabilities* (property 2). The chain rule imposes coherency between the belief of player i at a history h and the belief at a future history h' that player i deems possible at h: player i should not "change his mind" from h to h', but just update his belief. To see this clearly, note that the equality in property 2 is equivalent to the following condition:

$$\mu^i(D_{-i}|C_{-i}) > 0 \Rightarrow \mu^i(E_{-i}|D_{-i}) = \frac{\mu^i(E_{-i}|C_{-i})}{\mu^i(D_{-i}|C_{-i})}.$$

Considering that C_{-i} and D_{-i} are elements of $\mathcal{H}_{-i}$, this can also be written as follows. Let $h \prec h'$ and $s_{-i} \in S_{-i}(h')$, then

$$\mu^i(S_{-i}(h')\,|S_{-i}(h)) > 0 \Rightarrow \mu^i(s_{-i}|S_{-i}(h')) = \frac{\mu^i(s_{-i}|S_{-i}(h))}{\mu^i(S_{-i}(h')\,|S_{-i}(h))},$$

because $S_{-i}(h') \subseteq S_{-i}(h)$. Given a CPS μ^i, at every history h that player i initially deems possible (i.e., $\mu^i(S_{-i}(h)|S_{-i}) > 0$, where $S_{-i} = S_{-i}(\varnothing)$), player i simply updates the initial belief $\mu^i(\cdot|S_{-i})$ using the standard rule of conditional probabilities. Instead, at a history h that was deemed impossible (i.e., $\mu^i(S_{-i}(h)|S_{-i}) = 0$), player i has to come up with a new belief $\mu^i(\cdot|S_{-i}(h))$. Then, player i starts updating this new belief exactly as he was doing with the initial one, until also this new belief is "falsified" by the observed behavior of the co-players, and so on. To ease notation, from now on we write

$$\mu^i(\cdot|S_{-i}(h)) = \mu^i(\cdot|h),$$

and we let $\Delta^H(S_{-i})$ denote the set of all CPSs of player i on S_{-i}.[12]

We now redefine rational planning with respect to a CPS. Of the different representations of rational planning provided for conjectures, we seek to adapt to CPSs the definition of weak sequential optimality, for two reasons. First, sequential optimality is easier to define than folding-back optimality with respect to a CPS, because a sequentially optimal strategy is defined via a comparison of continuation plans at each history, and the belief specified by the CPS for the history suffices for this comparison. Second, we look for just *weakly* sequentially optimal strategies because, *under the assumption that players always execute their plans*, they are indistinguishable from the realization-equivalent sequentially optimal strategies, based on the observation of co-players' behavior. (We will elaborate on this issue in the next chapter.)

Given the belief $\mu^i(\cdot|h)$ that player i holds at history h, and given a strategy s_i that is consistent with h, we can compute the probability of reaching a terminal history $z \in Z(h)$ as follows:

$$\mathbb{P}^{s_i,\mu^i}(z|h) = \begin{cases} 0, & \text{if } s_i \notin S_i(z), \\ \mu^i(S_{-i}(z)|h), & \text{if } s_i \in S_i(z). \end{cases}$$

[12] Note that, for all $\mu^i \in \Delta^H(S_{-i})$ and $h', h'' \in H$, $S_{-i}(h') = S_{-i}(h'')$ implies $\mu^i(\cdot|h') = \mu^i(\cdot|h'')$. In other words, beliefs about co-players depend only on what the history reveals about their behavior, and do not depend on what it reveals about own behavior.

With this, we can recompute the subjective value of reaching h when s_i is followed from h onwards as follows:

$$V_i^{s_i,\mu^i}(h) = \sum_{z\in Z(h)} u_i(z)\mathbb{P}^{s_i,\mu^i}(z|h).$$

Then, we can adapt to CPSs the definition of weak sequential optimality provided for conjectures. Recall that $H_i(\bar{s}_i)$ is the set of non-terminal histories consistent with (not precluded by) playing strategy $\bar{s}_i$.

Definition 58. *A strategy $\bar{s}_i$ is* ***weakly sequentially optimal given*** μ^i *if*[13]

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i\in S_i(h)} V_i^{s_i,\mu^i}(h). \tag{10.4.1}$$

Condition (10.4.1) must hold at every history not precluded by $\bar{s}_i$, but given the coherency between beliefs at different histories imposed by the chain rule, it is sufficient to check that it holds at the histories that are deemed impossible by player i until they are actually reached.

Lemma 25. *Fix any strategy $\bar{s}_i$. For all $h, h' \in H_i(\bar{s}_i)$ such that $h \prec h'$ and $\mu^i(S_{-i}(h')|S_{-i}(h)) > 0$, if*

$$V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i\in S_i(h)} V_i^{s_i,\mu^i}(h) \tag{10.4.2}$$

then

$$V_i^{\bar{s}_i,\mu^i}(h') = \max_{s_i\in S_i(h')} V_i^{s_i,\mu^i}(h').$$

Proof. Fix $s_i \in S_i(h')$. Define $s_i' \in S_i(h')$ as $s_i'(\widetilde{h}) = \bar{s}_i(\widetilde{h})$ for each $\widetilde{h} \not\succeq h'$, and $s_i'(\widetilde{h}) = s_i(\widetilde{h})$ for each $\widetilde{h} \succeq h'$. Then, we have $\zeta(s_i', s_{-i}) = \zeta(s_i, s_{-i})$ for each $s_{-i} \in S_{-i}(h')$, and $\zeta(s_i', s_{-i}) = \zeta(\bar{s}_i, s_{-i})$

[13]We have previously defined the continuation value at h given a conjecture β^i considering the *continuation* strategies in $S_i^{\succeq h}$. To ease notation, here we define the continuation value at h given a CPS μ^i for the strategies that are consistent with h, that is, $S_i(h)$. The two representations are equivalent because $\mu^i(S_{-i}(h)|h) = 1$, therefore only the actions prescribed by each $s_i \in S_i(h)$ from h onwards matter for the expected payoff.

for each $s_{-i} \notin S_{-i}(h')$. Hence,

$$\sum_{s_{-i}\in S_{-i}(h)\setminus S_{-i}(h')} u_i(\zeta(s_i', s_{-i}))\mu^i(s_{-i}|h) \qquad (10.4.3)$$
$$= \sum_{s_{-i}\in S_{-i}(h)\setminus S_{-i}(h')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h),$$
$$\sum_{s_{-i}\in S_{-i}(h')} u_i(\zeta(s_i', s_{-i}))\mu^i(s_{-i}|h) \qquad (10.4.4)$$
$$= \sum_{s_{-i}\in S_{-i}(h')} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h).$$

By (10.4.2) and (10.4.3), we get

$$\sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(s_i', s_{-i}))\mu^i(s_{-i}|h) \le \sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h),$$

so by (10.4.4),

$$\sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h) \le \sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h).$$

Finally, dividing both sides by $\mu^i(S_{-i}(h')|h)$, we get

$$\sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h') \le \sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h'),$$

as desired. ■

Now we state and prove formally the equivalence between weak sequential optimality given a CPS and weak sequential optimality given the conjecture derived from the CPS. Recall that, for all $h \in H$ and $a_{-i} \in \mathcal{A}_{-i}(h)$, we let $S_{-i}(h, a_{-i})$ denote the set of all $s_{-i} \in S_{-i}(h)$ such that $s_{-i}(h) = a_{-i}$.

Proposition 5. *For each CPS $\mu^i \in \Delta^H(S_{-i})$, define the corresponding conjecture $\beta^i \in \times_{h\in H} \Delta(\mathcal{A}_{-i}(h))$ as follows:*

$$\forall h \in H,\ \forall a_{-i} \in \mathcal{A}_{-i}(h),\ \beta^i(a_{-i}|h) = \mu^i\left(S_{-i}\left(h, a_{-i}\right)|h\right).$$

Then, a strategy is weakly sequentially optimal given μ^i if and only if it is weakly sequentially optimal given β^i.

Proof. A strategy $\bar{s}_i \in S_i$ is weakly sequentially optimal given β^i if

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\beta^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i,\beta^i}(h),$$

where

$$V_i^{s_i,\beta^i}(h) = \sum_{z \in Z(h)} u_i(z)\mathbb{P}^{s_i,\beta^i}(z|h).$$

Thus, we have only to show that, for all $s_i \in S_i$, $h \in H_i(s_i)$, and $z \in Z(h)$,

$$\mathbb{P}^{s_i,\mu^i}(z|h) = \mathbb{P}^{s_i,\beta^i}(z|h).$$

Since h is a prefix of every $z \in Z(h)$, there is some $n \geq 1$ and some $(a^1, ..., a^n)$ such that $(h, a^1, ..., a^n)$. Case $n = 1$ is trivial. This, suppose that $n \geq 2$. Previously, we defined

$$\mathbb{P}^{s_i,\beta^i}(z|h) = \mathbb{P}^{s_i,\beta^i}\left(a^1|h\right)\prod_{t=2}^{n}\mathbb{P}^{s_i,\beta^i}\left(a^t|(h, a^1, ..., a^{t-1})\right),$$

where, for each $\widetilde{h} \in H$ and $a = (a_i, a_{-i}) \in \mathcal{A}(\widetilde{h})$,

$$\mathbb{P}^{s_i,\beta^i}\left(a|\widetilde{h}\right) = \begin{cases} 0, & \text{if } a_i \neq s_i(\widetilde{h}), \\ \beta^i(a_{-i}|\widetilde{h}), & \text{if } a_i = s_i(\widetilde{h}). \end{cases}$$

Suppose first that $s_i \notin S_i(z)$. Since $s_i \in S_i(h)$, there exists t such that $a_i^t \neq s_i(h)$. Then we have

$$\mathbb{P}^{s_i,\beta^i}(z|h) = 0 = \mathbb{P}^{s_i,\mu^i}(z|h).$$

If $s_i \in S_i(z)$, then $a_i^t = s_i(h)$ for every $t = 1, ..., n$. So we have

$$\mathbb{P}^{s_i,\beta^i}(z|h) = \beta^i(a_{-i}^1|h)\prod_{t=2}^{n}\beta^i\left(a_{-i}^t|(h, a^1, ..., a^{t-1})\right). \tag{10.4.5}$$

By the chain rule, we can write

$$\begin{aligned}
\mathbb{P}^{s_i,\mu^i}(z|h) &= \mu^i(S_{-i}(z)|h) \\
&= \mu^i(S_{-i}((h, a^1))|h) \cdot \mu^i(S_{-i}(z)|(h, a^1)) \\
&= \mu^i(S_{-i}((h, a^1))|h) \cdot \mu^i(S_{-i}((h, a^1, a^2))|(h, a^1)) \cdot \mu^i(S_{-i}(z)|(h, a^1, a^2)) \\
&= ... \\
&= \mu^i(S_{-i}((h, a^1))|h) \cdot \prod_{t=2}^{n}\mu^i(S_{-i}((h, a^1, ..., a^t))|(h, a^1, ..., a^{t-1})). \qquad (10.4.6)
\end{aligned}$$

For each $\widetilde{h} \in H$ and $a = (a_i, a_{-i}) \in \mathcal{A}(\widetilde{h})$, by definition we have $\beta^i(a_{-i}|\widetilde{h}) = \mu^i(S_{-i}(\widetilde{h}, a_{-i})|\widetilde{h})$, and obviously $S_{-i}(\widetilde{h}, a_{-i}) = S_{-i}(\widetilde{h}, a)$. Hence, (10.4.5) and (10.4.6) coincide, as desired. ∎

Weak sequential optimality given a conjecture is weaker than sequential optimality, which in turn is equivalent to folding-back optimality by the optimality principle (Theorem 33). Hence, the existence of a folding-back optimal strategy guarantees the existence of a weakly sequentially optimal strategy given the conjecture derived from the CPS; thus, by Proposition 5, we obtain the existence of a weakly sequentially optimal strategy given the CPS. Therefore, we can adopt weak sequential optimality for at least one CPS as our extension of the notion of justifiability of Chapter 3 to multistage games.

Definition 59. *A strategy $\bar{s}_i$ is* ***justified by*** *a CPS μ^i if it is weakly sequentially optimal given μ^i. We will say that a strategy is* ***justifiable*** *if it is justified by some CPS.*

We let
$$r_i : \Delta^H(S_{-i}) \rightrightarrows S_i$$
denote the correspondence that assigns to each CPS μ^i the set of strategies justified by μ^i. With this, the set of justifiable strategies is

$$r_i\left(\Delta^H(S_{-i})\right) = \bigcup_{\mu^i \in \Delta^H(S_{-i})} r_i\left(\mu^i\right).$$

The set of strategies justified by a CPS satisfies the following property, which we will use in the next chapter. Fix a CPS μ^i. If a history h is consistent with some weakly sequentially optimal strategy and with a strategy $\bar{s}_i$ that satisfies Condition (10.4.1) at h, then there is a weakly sequentially optimal strategy that is consistent with h and prescribes $\bar{s}_i(h)$. (We omit the proof.)

Lemma 26. *For each $h \in \bigcup_{s_i \in r_i(\mu^i)} H_i(s_i)$ and $\bar{s}_i \in S_i(h)$ such that*

$$V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i,\mu^i}(h),$$

there exists $s'_i \in r_i(\mu^i) \cap S_i(h)$ such that $s'_i(h) = \bar{s}_i(h)$.

In Chapter 3, we stated and proved an equivalence result relating (un)justifiability and strict dominance (Lemma 2). Here, we extended the notion of justifiability to multistage games. Can we extend the notion of strict dominance to multistage games in such a way to preserve the equivalence between justifiability and not being dominated? The answer is yes.

Definition 60. *A strategy $\bar{s}_i$ is **conditionally dominated** if there exist a history $h \in H_i(\bar{s}_i)$ and a mixed strategy $\sigma_i \in \Delta(S_i)$ with $\sigma_i(S_i(h)) = 1$ such that*

$$\forall s_{-i} \in S_{-i}(h), \sum_{s_i \in S_i(h)} \sigma_i(s_i) u_i(\zeta(s_i, s_{-i})) > u_i(\zeta(\bar{s}_i, s_{-i})). \qquad (10.4.7)$$

The set of strategies of agent i that are not conditionally dominated *is denoted by NCD_i.*

A strategy $\bar{s}_i$ is conditionally dominated if there exists a mixed strategy that yields a higher expected payoff conditional on reaching some history consistent with $\bar{s}_i$, no matter what the co-players do. We restrict attention to histories that are consistent with the strategy itself, because we want to relate conditional dominance to *weak* sequential optimality.[14] We can state the equivalence result.

Lemma 27. *A strategy is justifiable if and only if it is not conditionally dominated, that is, $r_i\left(\Delta^H(S_{-i})\right) = NCD_i$.*

Proof. Fix a strategy $\bar{s}_i$. Suppose first that $\bar{s}_i$ is justifiable. Then there exists $\mu^i \in \Delta^H(S_{-i})$ such that

$$\forall h \in H_i(\bar{s}_i),\, V_i^{\bar{s}_i, \mu^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i, \mu^i}(h),$$

which means that

$$\forall h \in H_i(\bar{s}_i),\, \bar{s}_i \in \arg\max_{s_i \in S_i(h)} \sum_{s_{-i} \in S_{-i}(h)} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h). \qquad (10.4.8)$$

[14] One can establish an analogous equivalence between sequential optimality and a notion of conditional dominance that also considers the histories that are precluded by the dominated strategy (focusing on the continuation strategy).

Next, fix any $h \in H_i(\bar{s}_i)$. Consider an auxiliary simultaneous-move game with, for each $j \in I$, action set $A_j^h = S_j(h)$ and payoff function

$$\begin{array}{rcl} v_j : \bigtimes_{k\in I} A_k^h & \to & \mathbb{R} \\ s & \mapsto & u_j(\zeta(s)). \end{array}$$

By (10.4.8), $\bar{s}_i$ is a best reply to the conjecture $\nu^i \in \Delta(S_{-i}(h))$ such that $\nu^i(s_{-i}) = \mu^i(s_{-i}|h)$ for each $s_{-i} \in S_{-i}(h)$. Then, by Lemma 2, no mixed action $\sigma_i \in \Delta(S_i(h))$ dominates $\bar{s}_i$. Since this is true for every $h \in H_i(\bar{s}_i)$, it follows that $\bar{s}_i$ is not conditionally dominated.

Conversely, suppose that $\bar{s}_i$ is not conditionally dominated. Fix any $h \in H_i(\bar{s}_i)$ and construct an auxiliary simultaneous-move game as above. Thus, $\bar{s}_i$ is not dominated by any mixed action in this game. Then, by Lemma 2, $\bar{s}_i$ is a best reply among the strategies in $S_i(h)$ to some conjecture $\nu^{i,h} \in \Delta(S_{-i}(h))$. Order the conjectures $(\nu^{i,h})_{h\in H_i(\bar{s}_i)}$ in such a way that if history h precedes history h' in the game, $\nu^{i,h}$ comes before $\nu^{i,h'}$ in the ordering. Next, fix some residual measure $\nu \in \Delta(S_{-i})$ such that $\nu(s_{-i}) > 0$ for every $s_{-i} \in S_{-i}$. With this, define $\mu^i = (\mu^i(\cdot|h))_{h\in H} \in (\Delta(S_{-i}))^H$ as follows. For each $h \in H$, derive $\mu^i(\cdot|h)$ by considering prefixes $\bar{h}$ of h and conditioning the first $\nu^{i,\bar{h}}$ in the ordering such that $\nu^{i,\bar{h}}(S_{-i}(h)) > 0$, *if any*; otherwise, derive $\mu^i(\cdot|h)$ by conditioning the residual measure ν.

First we check that μ^i is a CPS. Fix $C_{-i}, D_{-i} \in \mathcal{H}_{-i}$ and $h, h' \in H$ such that $D_{-i} \subseteq C_{-i}$, $C_{-i} = S_{-i}(h)$, $D_{-i} = S_{-i}(h')$, and $h \neq h'$. Suppose that $\mu^i(D_{-i}|h) > 0$. If $\mu^i(\cdot|h)$ was derived by conditioning ν, so was $\mu^i(\cdot|h')$, therefore condition 2 of a CPS is satisfied. If $\mu^i(\cdot|h)$ was derived by conditioning some $\nu^{i,\bar{h}}$, $\mu^i(D_{-i}|h) > 0$ implies $\nu^{i,\bar{h}}(D_{-i}) > 0$. Moreover, for each $\nu^{i,\widetilde{h}}$ that precedes $\nu^{i,\bar{h}}$ in the ordering, we must have $\nu^{i,\widetilde{h}}(S_{-i}(h)) = 0$, otherwise $\mu^i(\cdot|h)$ could not be derived from $\nu^{i,\bar{h}}$. Hence, $\mu^i(\cdot|h')$ was derived by conditioning $\nu^{i,\bar{h}}$ as well, and condition 2 of a CPS is satisfied.

Now we want to show that $\bar{s}_i$ is weakly sequentially optimal under μ_i. We want to show that, for each $h \in H_i(\bar{s}_i)$, $\bar{s}_i$ satisfies (10.4.1). For every $\widetilde{h} \in H_i(\bar{s}_i)$ with $\widetilde{h} \not\preceq h$ and $h \not\preceq \widetilde{h}$, there is no $s_{-i} \in S_{-i}$ such that $\widetilde{h} \prec \zeta(\bar{s}_i, s_{-i})$ and $h \prec \zeta(\bar{s}_i, s_{-i})$. Hence, $S_{-i}(\widetilde{h}) \cap S_{-i}(h) = \emptyset$ and $\mu^i(\cdot|h)$ cannot be derived by conditioning $\nu^{i,\widetilde{h}}$. So, since $\nu^{i,h}(S_{-i}(h)) = 1$, $\mu^i(\cdot|h)$ is derived by conditioning some $\nu^{i,\bar{h}}$ with $\bar{h} \preceq h$ and $\nu^{i,\bar{h}}(S_{-i}(h)) > 0$. Note that $\mu^i(\cdot|\bar{h}) = \nu^{i,\bar{h}}$, therefore $\bar{s}_i$ satisfies (10.4.1) at $\bar{h}$. Then, by Lemma 25, $\bar{s}_i$ satisfies (10.4.1) at h. ■

To better understand conditional dominance, it is useful to compare it to strict dominance and weak dominance applied to the *strategic form* of the game. Strict dominance in the strategic form clearly implies conditional dominance. Indeed, if a strategy is strictly dominated, then it is conditionally dominated given the empty (initial) history $h = \varnothing$, because $S_{-i}(\varnothing) = S_{-i}$.

Remark 50. *If a strategy is strictly dominated in the strategic form of the game, then it is conditionally dominated.*

Furthermore, conditional dominance implies weak dominance: if $\bar{s}_i$ is worse than some mixed strategy σ_i from some history h onwards, then we can construct a mixed strategy that imitates σ_i from h onwards and $\bar{s}_i$ everywhere else, so that it is equivalent to $\bar{s}_i$ if h is not reached.

Proposition 6. *If a strategy is conditionally dominated, then it is weakly dominated in the strategic form of the game.*

Proof. We use Lemma 5: a strategy is weakly dominated if (and only if) it is not a best response to any conjecture assigning *strictly positive* probability to every profile s_{-i}. Suppose that $\bar{s}_i$ is conditionally dominated at some history $\bar{h}$. Pick any conjecture $\mu^i \in \Delta(S_{-i})$ such that $\mu^i(s_{-i}) > 0$ for each $s_{-i} \in S_{-i}$, and let ν^i be the conjecture over $S_{-i}(\bar{h})$ obtained from μ^i by conditioning. Now, focus on the auxiliary simultaneous-move game where the action set of each player j coincides with the subset of strategies of the multistage game $S_j(\bar{h})$, and for each profile of actions s the payoff of j is given by $u_j(\zeta(s))$. We can view ν^i as a conjecture of player i in this game. Since $\bar{s}_i$ is conditionally dominated at $\bar{h}$ in the multistage game, it is dominated in this simultaneous-move game, therefore by Lemma 2 it is not a best reply to ν^i. It follows that there exists $s_i \in S_i(\bar{h})$ such that

$$\sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\bar{s}_i, s_{-i}))\nu^i(s_{-i}) < \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(s_i, s_{-i}))\nu^i(s_{-i}).$$

Now go back to the multistage game. Define $\widetilde{s}_i$ as $\widetilde{s}_i(h) = s_i(h)$ for all $h \succeq \bar{h}$ and $\widetilde{s}_i(h) = \bar{s}_i(h)$ for all $h \not\succeq \bar{h}$. Hence,

$$\sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\widetilde{s}_i, s_{-i}))\nu^i(s_{-i}) = \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(s_i, s_{-i}))\nu^i(s_{-i}),$$

$$\forall s_{-i} \in S_{-i} \backslash S_{-i}(\bar{h}), \quad \zeta(\widetilde{s}_i, s_{-i}) = \zeta(\bar{s}_i, s_{-i}).$$

Thus, we obtain

$$\begin{aligned}
& \sum_{s_{-i} \in S_{-i}} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}) \\
= \;& \mu^i(S_{-i}(\bar{h})) \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\bar{s}_i, s_{-i}))\nu^i(s_{-i}) + \sum_{s_{-i} \in S_{-i} \backslash S_{-i}(\bar{h})} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}) \\
< \;& \mu^i(S_{-i}(\bar{h})) \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\widetilde{s}_i, s_{-i}))\nu^i(s_{-i}) + \sum_{s_{-i} \in S_{-i} \backslash S_{-i}(\bar{h})} u_i(\zeta(\widetilde{s}_i, s_{-i}))\mu^i(s_{-i}) \\
= \;& \sum_{s_{-i} \in S_{-i}} u_i(\zeta(\widetilde{s}_i, s_{-i}))\mu^i(s_{-i}).
\end{aligned}$$

This argument proves that $\bar{s}_i$ is not a best reply to any strictly positive conjecture in the strategic form of the multistage game. Therefore, by Lemma 5, s_i is weakly dominated. ■

The reason why weak dominance does not imply conditional dominance can be seen in a simultaneous-move game, where conditional dominance coincides with strict dominance. However, it is often the case that all the weak dominance relations in the strategic form of a multistage game imply a strict dominance relation from some history h onwards, as in the definition of conditional dominance. For example, in all leader-follower games[15] a strategy of the follower is weakly dominated in the strategic form of the game if and only if it does not select a best reply for at least one action of the leader, and it is therefore conditionally dominated given this action. Indeed, in a large class of games, weak dominance in the strategic form coincides with conditional dominance. Next we quantify how large this class is.

Suppose that a mixed strategy $\bar{\sigma}_i$ weakly dominates strategy $\bar{s}_i$ in the strategic form of the game. Fix a history $h \in H_i(\bar{s}_i)$ such that (i) $\bar{\sigma}_i(S_i(h)) = 1$ and (ii) $\bar{\sigma}_i(s_i) > 0$ for some s_i with $s_i(h) \neq \bar{s}_i(h)$. By weak dominance, $\bar{\sigma}_i$ does weakly better than $\bar{s}_i$ against every $s_{-i} \in S_{-i}(h)$. Now decrease by $\varepsilon > 0$ the (only) payoffs of player i associated with terminal histories that are consistent with $\bar{s}_i$. Then, by (ii), $\bar{\sigma}_i$ now

[15] Formally, a leader-follower game is a two-person, two-stage game with perfect information such that the first-mover cannot move again in the second stage.

does *strictly* better than $\bar{s}_i$ against every $s_{-i} \in S_{-i}(h)$, and thus $\bar{s}_i$ is conditionally dominated. So, if $\bar{s}_i$ is not conditionally dominated in the original game, it is conditionally dominated in "neighboring games" with the same tree and slightly different payoffs. The bottom line is that games where weak dominance in the strategic form does not imply conditional dominance represent "an exception" from a mathematical perspective, as long as it makes sense to slightly modify the payoffs of some terminal histories and not others, which is certainly the case when the outcome function $g : Z \to Y$ is one-to-one. We formalize this concept using the notion of *Lebesgue measure*. In $\mathbb{R}^n$, the Lebesgue measure is the generalization of the notions of length, area, and volume for $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$, respectively. Recall that the set $\mathbb{R}^Z$ of all payoff functions $u_i : Z \to \mathbb{R}$ is essentially the same as the Euclidean space $\mathbb{R}^{|Z|}$: fix arbitrarily a bijection $n : Z \to \{1, ..., |Z|\}$, then $u_i \in \mathbb{R}^Z$ corresponds to the vector $\left(u_i\left(n^{-1}(1)\right), ..., u_i\left(n^{-1}(|Z|)\right)\right) \in \mathbb{R}^{|Z|}$.

Lemma 28. *Fix a finite multistage game tree. The closure*[16] *of the set of payoff vectors $(u_i(z))_{z \in Z}$ such that some strategy of player i is weakly dominated in the strategic form but not conditionally dominated has Lebesgue measure zero in $\mathbb{R}^Z$.*

The proof of Lemma 28 is in the appendix of Chapter 11, where we provide a more general version of this result.

In words, we will say that a statement (such as the equivalence between conditional dominance and weak dominance) holds **generically** when, given every multistage game tree, it is true for all payoff vectors except possibly for a subset with zero-measure closure. Now we provide the intuition for Lemma 28 with an example.

Example 47. Consider a modified version of the Battle of the Sexes of Example 38, where Bob's payoff at terminal history $((\text{in}), (\text{B}_1, \text{B}_2))$ is substituted by a parameter $v \in \mathbb{R}$. Strategy B_2 of Bob is conditionally dominated if and only if $v < 0$. If $v < 0$, then B_2 is conditionally dominated by S_2 at history (in). If $v \geq 0$, then B_2 is not conditionally dominated: for every mixed strategy $\sigma_2 \in \Delta(\{\text{S}_2, \text{B}_2\})$, we have

$$\begin{aligned}\sigma_2(\text{S}_2)u_2(\zeta(\text{in.B}_1, \text{S}_2)) + \sigma_2(\text{B}_2)u_2(\zeta(\text{in.B}_1, \text{B}_2)) &= \sigma_2(\text{B}_2) \cdot v \qquad (10.4.9)\\ &\leq v = u_2(\zeta(\text{in.B}_1, \text{B}_2)),\end{aligned}$$

[16] The closure of a set is the smallest closed set that contains it.

therefore condition (10.4.7) is not satisfied for s_1 = in.B$_1$ (at either non-terminal history).

Strategy B$_2$ is weakly dominated in the strategic form of the game if and only if $v \leq 0$. If $v \leq 0$, then B$_2$ is weakly dominated by S$_2$—note that B$_2$ is not strictly dominated because the two strategies yield the same payoff if Ann plays out. If $v > 0$, then B$_2$ is not weakly dominated, because then inequality (10.4.9) is strict.

So, for $v \neq 0$, B$_2$ is weakly dominated if and only if it is conditionally dominated. There is only one value of v in the entire real line, $v = 0$, such that B$_2$ is weakly dominated but not conditionally dominated. ▲

10.5 Infinite Horizon and Continuity at Infinity

Although the previous analysis is focused on finite games, it applies more generally to games with finite horizon.[17] In games with infinite horizon, folding-back planning does not apply, because there is no last stage from which the computation can start. However, the OD Principle still holds under a mild regularity property, called "continuity at infinity," implied by compactness-continuity of the game.

For any history $h = (a^1, a^2, ...) \in \bar{H}$ and number of stages $t \in \mathbb{N}$, let h^t denote the prefix of h of length t if $\ell(h) > t$ (recall that $\ell(h)$ is the length of h, and $\ell(h) = \infty$ if h is an infinite history), otherwise let $h^t = h$; that is, if $h = \left(a^k\right)_{k=1}^{\ell(h)}$, then $h^t = \left(a^k\right)_{k=1}^{\min\{\ell(h),t\}}$.

Definition 61. *Game Γ is* ***continuous at infinity*** *if, for every player $i \in I$,*

$$\lim_{t\to\infty}\left[\sup\left\{\left|u_i(h) - u_i(\widetilde{h})\right| : h, \widetilde{h} \in Z, h^t = \widetilde{h}^t\right\}\right] = 0$$

where we let $\sup\emptyset = 0$ by convention.

Note that every game with finite horizon is trivially continuous at infinity. Indeed, let $T < \infty$ denote the horizon, then for every $t \geq T$, $h^t = \widetilde{h}^t$ implies $h = \widetilde{h}$, hence the limsup must be 0. Thus, Γ is continuous

[17] In infinite games with finite horizon, the optimization problems defining the dynamic programming properties studied above may not have a solution, but the main equivalence results (Folding-Back Principle, Optimality Principle, OD Principle) still hold. Other results, such as Proposition 4, do not hold for all infinite games, even if they have finite horizon.

at infinity if and only if either Γ has finite horizon, or Γ has infinite horizon and for all $\epsilon > 0$ there is some sufficiently large positive integer T_ϵ such that, for all $t \geq T_\epsilon$ and all pairs of histories $h, \widetilde{h} \in Z$ satisfying $h^t = \widetilde{h}^t$, we have $\left| u_i(h) - u_i(\widetilde{h}) \right| < \epsilon$. This means that what happens in the distant future has little impact on overall payoffs.

Remark 51. *Every compact-continuous game is continuous at infinity.*

Proof. Let Γ be compact-continuous and fix a terminal history $h \in Z$ arbitrarily; we must show that

$$\lim_{t\to\infty} \left[\sup \left\{ \left| u_i(h) - u_i(\widetilde{h}) \right| : \widetilde{h} \in Z, h^t = \widetilde{h}^t \right\} \right] = 0.$$

The equality holds trivially if h is finite. Suppose that h is infinite. For every $t \in \mathbb{N}$, the set of infinite continuations of the prefix h^t, $\left\{ \widetilde{h} \in Z : h^t = \widetilde{h}^t \right\}$, is compact; therefore, by continuity of u_i, there is some terminal history $\widetilde{h}_t \in \left\{ \widetilde{h} \in Z : h^t = \widetilde{h}^t \right\}$ such that

$$\sup \left\{ \left| u_i(h) - u_i(\widetilde{h}) \right| : \widetilde{h} \in Z, h^t = \widetilde{h}^t \right\} = \left| u_i(h) - u_i(\widetilde{h}_t) \right| < \infty.$$

By construction, the sequence of terminal histories $\left(\widetilde{h}_t \right)_{t=1}^{\infty}$ converges to h, because $\widetilde{h}_t^\ell = h^\ell$ for $\ell \leq t$, therefore $\lim_{t\to\infty} \widetilde{h}_t^\ell = h^\ell$ for each $\ell \in \mathbb{N}$, which implies $\lim_{t\to\infty} \widetilde{h}_t = h$. By continuity of u_i, $\lim_{t\to\infty} u_i \left(\widetilde{h}_t \right) = u_i(h)$, which implies the claim. ■

Example 48. Suppose that the game has infinite horizon and, at every stage, all the action profiles in the Cartesian set $A = \times_{i\in I} A_i$ are feasible, that is, $H = A^{<\mathbb{N}_0}$ and $Z = A^{\mathbb{N}}$. For every $i \in I$ and $t = 1, 2, ...$, there is a period-t (istantaneous, or flow) payoff function $u_{i,t} : A^t \to \mathbb{R}$.[18] Furthermore, there is some $\overline{v} \in \mathbb{R}$ such that for all $i \in I$,

$$\sup_{h^t \in A^t} |u_{i,t}(h^t)| \leq \overline{v}.$$

[18] The period-t payoff functions $u_{i,t}$ may be the composition of period-t outcome functions $g_t : A^t \to Y_t$ and period-t utility functions $v_{i,t} : Y_t \to \mathbb{R}$, that is, $u_{i,t} = v_{i,t} \circ g_t : A^t \to \mathbb{R}$. See the discussion in Section 9.2.1.

Payoffs are assigned to terminal histories as follows: for each player i there is a discount factor $\delta_i \in (0,1)$ such that

$$u_i(h^\infty) = (1-\delta_i)\sum_{t=1}^{\infty}(\delta_i)^{t-1}u_{i,t}(h^t)$$

for all $h^\infty \in Z$, where h^t denotes the length-t prefix of h^∞. One can verify that such games are continuous at infinity. Infinitely repeated games with discounting are a special case where $u_{i,t}(a^1,...,a^{t-1},a^t) = u_i(a^t)$ is independent of $\left(a^1,...,a^{t-1}\right)$. ▲

Theorem 34. (OD Principle with continuity at infinity). *Suppose that* Γ *is continuous at infinity. Then, for every* $s_i \in S_i$ *and* $\beta^i \in \times_{h\in H}\Delta(\mathcal{A}_{-i}(h))$*, the following conditions are equivalent:*
(1) s_i *is sequentially optimal given* β^i*;*
(2) s_i *is one-step optimal given* β^i*.*

Proof. (1) implies (2) by Remark 49. To show that (2) implies (1), we prove the contrapositive: if s_i is *not* sequentially optimal given β^i, then there must be incentives for one-step deviations. If β^i is non-deterministic, the proof requires some measure-theoretic technicalities. Therefore we provide the proof for the deterministic case: for some $s_{-i} \in S_{-i}$ and every $h \in H$, $\beta^i\left(s_{-i}\left(h\right)|h\right) = 1$.[19] Thus, we replace β^i with s_{-i} in the formulas, and we have $V_i^{s_i,s_{-i}}(h) = u_i\left(\zeta\left(s_i,s_{-i}|h\right)\right)$, where $\zeta\left(s_i,s_{-i}|h\right)$ denotes the terminal history induced by (s_i,s_{-i}) starting from h, which depends only on $(s_i|h) \in S_i^{\succeq h}$ and $(s_{-i}|h) \in S_{-i}^{\succeq h}$.

Suppose that s_i is *not* sequentially optimal given some $s_{-i} \in S_{-i}$, that is, there are $h \in H$, $\widehat{s}_i \in S_i$ and $\epsilon > 0$ such that

$$u_i\left(\zeta\left(s_i,s_{-i}|h\right)\right) = u_i\left(\zeta\left(\widehat{s}_i,s_{-i}|h\right)\right) - \epsilon. \tag{10.5.1}$$

For every positive integer T and for every strategy profile $s \in S$, let us define the auxiliary "truncated" game $\Gamma^{T,s} = \left\langle I,(\mathcal{A}_i^T(\cdot),u_i^{T,s})_{i\in I}\right\rangle$ with horizon T derived from Γ as follows:

- $\mathcal{A}_i^T(h') = \mathcal{A}_i(h')$ if $\ell(h') < T$ and $\mathcal{A}_i^T(h') = \emptyset$ otherwise, so that $\bar{H}^T = \{h' \in \bar{H} : \ell(h') \leq T\}$ and $Z^T = \{h' \in \bar{H}^T$: either $\ell(h') = T$, or $\ell(h') < T$ and $h' \in Z\}$,

[19] The proof of this special case is sufficient for the analysis of pure equilibria in games without chance moves.

- for all $\bar{h} \in Z \cap Z^T$, $u_i^{T,s}(\bar{h}) = u_i(\bar{h})$,
- for all $\bar{h} \in Z^T \backslash Z$, $u_i^{T,s}(\bar{h}) = u_i\left(\zeta\left(s_i, s_{-i}|\bar{h}\right)\right)$.

Intuitively, the game is truncated so that it has at most T stages. If a terminal history h' of a truncated game $\Gamma^{T,s}$ is not terminal for the original game Γ, then the imputed payoffs are those that would obtain in Γ if the players followed the given strategy profile s after history h'.

By continuity at infinity there is a "large enough" T (which depends on ϵ) such that, for all $\widetilde{h}, \widehat{h} \in Z$, if $\widetilde{h}^T = \widehat{h}^T$ (i.e., if these two histories coincide in the first T stages) then

$$|u_i(\widetilde{h}) - u_i(\widehat{h})| < \epsilon.$$

Let $\widetilde{s}_i$ be the strategy that behaves as s_i after T and as $\widehat{s}_i$ before T. More precisely:

$$\widetilde{s}_i(h') = \begin{cases} s_i(h'), & \text{if } \ell(h') > T, \\ \widehat{s}_i(h'), & \text{if } \ell(h') \leq T. \end{cases}$$

Let $\widehat{h} = \zeta\left(\widehat{s}_i, s_{-i}|h\right)$ and $\widetilde{h} = \zeta\left(\widetilde{s}_i, s_{-i}|h\right)$. Then

$$u_i\left(\zeta\left(\widetilde{s}_i, s_{-i}|h\right)\right) = u_i(\widetilde{h}) > u_i(\widehat{h}) - \epsilon = u_i\left(\zeta\left(\widehat{s}_i, s_{-i}|h\right)\right) - \epsilon$$

(the equalities hold by definition and the inequality by choice of T and by construction). Taking (10.5.1) into account, we obtain

$$V_i^{\widetilde{s}_i, s_{-i}}(h) = u_i\left(\zeta\left(\widetilde{s}_i, s_{-i}|h\right)\right) > u_i\left(\zeta\left(s_i, s_{-i}|h\right)\right) = V_i^{s_i, s_{-i}}(h).$$

This means that the restriction of s_i to $\bar{H}^{T+1}$ is not sequentially optimal given s_{-i} in the finite horizon game $\Gamma^{T+1,s}$. Thus, by Corollary 6,[20] strategy s_i is not one-step optimal given s_{-i}. By definition of $\Gamma^{T+1,s}$, this implies that also in game Γ player i has an incentive to make one-step deviations after history h. ■

[20] Corollary 6 pertains to finite games, but the result can be extended to games with finite horizon.

11

Rationalizability in Multistage Games

In this chapter we analyze the behavior of rational players who form their subjective beliefs about the behavior of co-players by means of strategic thinking. As we did in Chapter 4 for simultaneous-move games, the cornerstone of players' reasoning is the belief of a rational player that the co-players are rational as well. Therefore, we seek to extend the idea of rationalizability for simultaneous-move games to multistage games. However, there is no obvious way to do so. The reason is that, as the game unfolds, the initial belief about the co-players may be falsified by the observed behavior. Depending on how players revise their beliefs after unexpected moves, we will define two notions of belief in the rationality of co-players, and, correspondingly, two notions of rationalizability for multistage games. **Initial rationalizability** captures the idea that players believe in the rationality of co-players at the beginning of the game and form a belief about their (present and future) behavior accordingly; yet, if this belief is falsified, they are free to entertain the possibility that their co-players are not rational, even if their past behavior does not contradict their rationality. **Strong rationalizability**, instead, captures a principle of **best rationalization**: players always ascribe to their co-players the highest level of strategic sophistication that is consistent with the observed behavior. In particular, they believe in the rationality of the co-players as long as the evidence is consistent with this hypothesis. A similar assumption holds for higher levels of strategic sophistication. This idea is used to rationalize unexpected moves of co-players and predict their future moves accordingly—an instance of "forward-induction reasoning." We assume throughout this chapter that the given multistage game Γ is *finite*.[1]

11.1 Rationality in Multistage Games

Before dealing with interactive beliefs about rationality, we have to be clear about the notion of rationality we adopt. We want "rationality" to mean two things: (i) players make "rational plans;" and (ii) they actually implement what they planned. In Chapter 10 we introduced different, but equivalent notions of rational planning for multistage games: sequential optimality, folding-back optimality, and one-step optimality. We also introduced the behaviorally equivalent notion of weak sequential optimality, and we adapted its definition to the case in which players' beliefs about co-players' behavior are represented by conditional probability systems (CPSs). As anticipated, in this chapter we adopt CPSs and weak sequential optimality.

The reason for choosing *weak* sequential optimality is the following. We assume to be transparent (i.e., always commonly believed) that each player i always executes his plan. Thus, when a co-player j observes a deviation from the expected behavior of i, he does *not* believe that player i committed a mistake in execution and will revert to his original plan, but rather that player i has a different plan in his mind and is executing this different plan. Accordingly, there is no need to specify the moves that a player's strategy prescribes after deviations from the strategy itself. In fact, the whole analysis of this chapter goes through verbatim with reduced strategies in place of strategies, as the elimination procedures we will consider do not make any distinction between any two behaviorally-equivalent (hence, realization-equivalent) strategies. It is worth stressing that other ways of reasoning in a multistage game (which we do not illustrate in this chapter)—such as so called "backwards rationalizability"—are based on the opposite idea: upon observing a

[1] The analysis can be extended to all **simple** multistage games, whereby the set of feasible actions of each player is finite for all histories except, possibly, those of height 1, for which the set of feasible actions must be, if not finite, at least compact. This implies that the set H of non-terminal histories is finite or at least countable. See Battigalli [7].

deviation from the expected plan of i, the co-players believe that player i committed a mistake and will revert to his original plan.

The reason for working with CPSs is that, when a player reasons strategically about the game, the "external state" he is interested in is "how the co-players would play conditional on the moves I can make," and this is conveniently described by strategies (see the discussion in Section 10.4 of Chapter 10). Just keeping track of the actions the co-players could play at different histories and directly taking conjectures over such actions would ultimately expand the set of conjectures our player may have, introducing additional optimal strategies. Here is an example.

Example 49. Consider the common interest game with perfect information in Figure 11.1 (the numbers at terminal nodes represent the common payoff).

We want to compute the set of justifiable strategies of Ann. (As usual, we omit action "wait" from the description of players' strategies.)

▶ Consider first any $s_a \in S_a$ such that $s_a(\varnothing) = out$; then s_a is justified by any CPS μ^a such that $\mu^a(\{m\} \times S_c|\varnothing) = 1$.

▶ Strategy $in.b.b'$ is justified by a CPS μ^a such that $\mu^a((\ell, d.d')|\varnothing) = 1$ and $\mu^a((r, d.d')|(in, r)) = 1$.

▶ Strategy $in.b.a'$ is justified by a CPS μ^a such that $\mu^a((\ell, d.d')|\varnothing) = 1$ and $\mu^a((r, d.c')|(in, r)) = 1$.

▶ Strategy $in.a.b'$ is justified by a CPS μ^a such that $\mu^a((r, d.d')|\varnothing) = 1$ and $\mu^a((\ell, c.c')|(in, \ell)) = 1$.

▶ Strategy $in.a.a'$ is not justifiable. To see this, observe that $in.a.a'$ is conditionally dominated by *out* at the root of the game: with $in.a.a'$ Ann obtains payoff 1 for sure, with *out* she obtains 2. Therefore, by Lemma 27, $in.a.a'$ is not justifiable.

Thus, every action of Ann is prescribed by at least one justifiable strategy. However, not every combination of these actions is a justifiable strategy.

Take now the viewpoint of Bob at history (in) with a belief over the justifiable strategies of Ann and Colin that are consistent with (in). Hence, Bob assigns probability one to $d.d'$ and to strategies of Ann that prescribe either b, or b', or both. With this, we obtain a conjecture of Bob that assigns probability at least $1/2$ to either b or b'. Therefore, his optimal action can be ℓ or r, but never m. Consider now instead a conjecture of Bob that, at each history of the game, assigns probability 1 to the

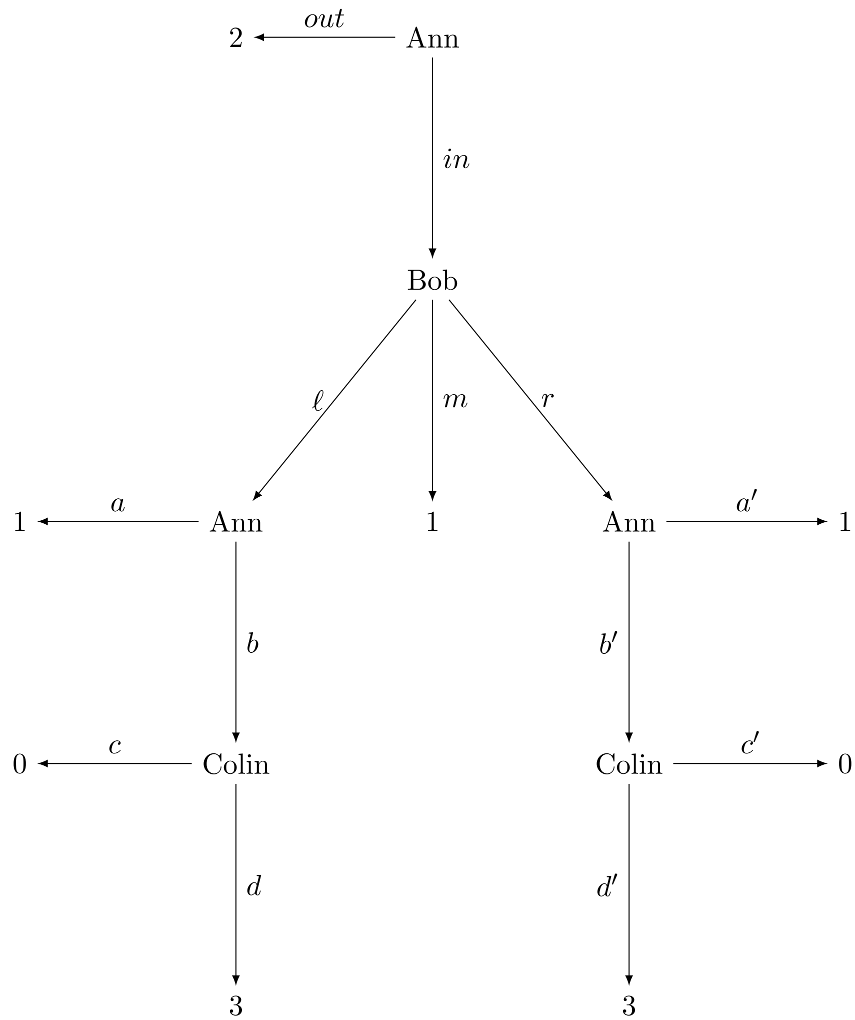


Figure 11.1: A common interest game.

actions of Ann and Colin that are prescribed by some justifiable strategy consistent with the history. Such a conjecture can assign probability one to a at (in, ℓ) and to a' at (in, r). So, Bob could optimally choose m. The problem just described arises because, if we pick for each history h an action $a_{i,h}$ consistent with i's rationality (the action "wait" if i is not active) and we form the strategy $s_i = (a_{i,h})_{h \in H}$, it is possible that s_i is *in*consistent with i's rationality. This is the case for strategy $in.a.a'$ of Ann.[2] ▲

Finally, we provide a characterization of weak sequential optimality that will come in handy later in this chapter.

Lemma 29. *Fix a player $i \in I$ and a CPS $\mu^i \in \Delta^H(S_{-i})$. Consider a nonempty subset of strategies $C_i \subseteq S_i$ such that*

$$r_i(\mu^i) \subseteq C_i. \tag{11.1.1}$$

A strategy $\bar{s}_i \in C_i$ is justified by μ^i if and only if

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in C_i \cap S_i(h)} V_i^{s_i,\mu^i}(h). \tag{11.1.2}$$

Proof. The "only if" part is obvious. The proof of the "if" part is by induction on the length of histories. Fix any $\bar{s}_i \in C_i$ such that (11.1.2) holds. We want to show that, for every $h \in H_i(\bar{s}_i)$,

$$V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i,\mu^i}(h). \tag{11.1.3}$$

At the initial history we have

$$V_i^{\bar{s}_i,\mu^i}(\varnothing) \overset{(11.1.2)}{=} \max_{s_i \in C_i} V_i^{s_i,\mu^i}(\varnothing) \overset{(11.1.1)}{=} \max_{s_i \in S_i} V_i^{s_i,\mu^i}(\varnothing).$$

Now, fix $n > 0$ and $h' \in H_i(\bar{s}_i)$ of length n. Suppose by way of induction that for every $h \in H_i(\bar{s}_i)$ of length smaller than n, $r_i(\mu^i) \cap S_i(h) \neq \emptyset$ and condition (11.1.3) holds. Now let h be the immediate predecessor of h'. Then, by Lemma 26, there exists $\widetilde{s}_i \in r_i(\mu^i) \cap S_i(h)$ such that $\widetilde{s}_i(h) = \bar{s}_i(h)$, thus $\bar{s}_i \in S_i(h')$. Hence,

$$V_i^{\widetilde{s}_i,\mu^i}(h') = \max_{s_i \in S_i(h')} V_i^{s_i,\mu^i}(h').$$

[2] In other words, the set of justifiable strategies

$$r_i\left(\Delta^H(S_{-i})\right) \subseteq S_i = \bigtimes_{h \in H} \mathcal{A}_i(h)$$

is not the Cartesian product of its projections onto set sets $\mathcal{A}_i(h)$ for $h \in H$. Strategy $in.a.a'$ belongs to the Cartesian product of the projections of $r_a\left(\Delta^H(S_{-a})\right)$, but $in.a.a' \notin r_a\left(\Delta^H(S_{-a})\right)$.

By (11.1.1), we have $\widetilde{s}_i \in C_i$, therefore

$$\max_{s_i \in S_i(h')} V_i^{s_i,\mu^i}(h') = \max_{s_i \in C_i \cap S_i(h')} V_i^{s_i,\mu^i}(h'). \tag{11.1.4}$$

So

$$V_i^{\bar{s}_i,\mu^i}(h') \overset{(11.1.2)}{=} \max_{s_i \in C_i \cap S_i(h')} V_i^{s_i,\mu^i}(h') \overset{(11.1.4)}{=} \max_{s_i \in S_i(h')} V_i^{s_i,\mu^i}(h').$$

■

11.2 Initial Rationalizability

The first natural extension of common belief in rationality from simultaneous-move games to multistage games is common *initial* belief in rationality. At the beginning of the game, players believe that co-players are rational; believe that co-players believe that everybody else is rational; and so on. To derive the behavioral implications of rationality and common initial belief in rationality, as we did in Chapter 4, we introduce a rationalization operator that characterizes the behavioral implications for rational players of each step of this reasoning procedure. Consider the collection $\mathcal{C}$ of all Cartesian subsets of S. For each $C = \times_{i\in I} C_i \in \mathcal{C}$, and for each $i \in I$, we define

$$\Delta^H_\varnothing(C_{-i}) = \left\{\mu^i \in \Delta^H(S_{-i}) : \mu^i(C_{-i}|\varnothing) = 1\right\},$$

$$\rho_i(C_{-i}) = \left\{s_i \in S_i : \exists \mu^i \in \Delta^H_\varnothing(C_{-i}), s_i \in r_i(\mu^i)\right\} = r_i(\Delta^H_\varnothing(C_{-i})),$$

$$\rho(C) = \bigtimes_{i\in I} \rho_i(C_{-i}).$$

The set $\Delta^H_\varnothing(C_{-i})$ is the subset of CPSs such that player i initially believes C_{-i}, i.e., C_{-i} is assigned probability one at the initial history. Given this, $\rho_i(C_{-i})$ is the set of strategies of i that are justified by a CPS in $\Delta^H_\varnothing(C_{-i})$. Finally, $\rho(C)$ is the set of strategy profiles that can be "rationalized," i.e., justified, by a CPS with initial belief in C_{-i} for each player $i \in I$. We use the same symbols r_i and ρ_i we used for best replies and the rationalization operator in simultaneous-move games, because we are considering an extension of these notions to multistage games: when the game has simultaneous moves, $\Delta^H_\varnothing(C_{-i})$ coincides with $\Delta(C_{-i})$, and weakly sequentially optimal strategies coincide with best replies.

Equipped with the operator $\rho : \mathcal{C} \to \mathcal{C}$, we define *initial rationalizability* with the standard iteration of a self-map as in Chapter 4. Let $\rho^0(S) = S$, and for each $k > 0$, let $\rho^k(S) = \rho(\rho^{k-1}(S))$. Note that whenever $C' \subseteq C$, we also have $\Delta^H_\varnothing(C'_{-i}) \subseteq \Delta^H_\varnothing(C_{-i})$ for each $i \in I$. Then, ρ is *monotone*. Therefore, since $\rho^0(S) = S$, $\left(\rho^k(S)\right)_{k=0}^{\infty}$ must be a weakly decreasing sequence, and it makes sense to define:[3]

$$\rho^\infty(S) = \bigcap_{k\geq 1} \rho^k(S).$$

Definition 62. *A strategy profile $s \in S$ is* ***initially rationalizable*** *if $s \in \rho^\infty(S)$.*

At every step of reasoning, initial rationalizability preserves the strategies of each player that are justified by a CPS assigning probability 1 at the initial history to the co-players' strategies that survived the previous steps. So, for instance, the strategies that survive step 2 are justified by a CPS that, at the initial history, assigns probability 1 to the justifiable strategies of the co-players. Then, the belief in the rationality of the co-players is maintained at all histories that a player deems possible given the initial belief. However, at histories that are deemed subjectively impossible, initial rationalizability does not impose any restriction on the beliefs over co-players' strategies, even if such histories are consistent with some or all steps of the algorithm—Example 50 below is case in point.

Leveraging on the monotonicity of operator ρ, one can prove that extensions of Theorems 2 and 3 of Chapter 4 hold for initial rationalizability in multistage games. Also, in Chapter 4 we reformulated rationalizability as a *reduction* procedure, by iterating a reduction operator $\overline{\rho}$ that only considers the strategies in the reduced game as candidate best replies. Can we do the same with initial rationalizability? The answer is again yes, and a possible way to construct the reduction procedure is identical to that in Chapter 4, once the set of conjectures over the reduced set of co-players' action is replaced by the set of CPSs that initially believe in the reduced set of co-players' strategies. We leave these proofs as exercises.

[3] Compare with Chapter 4.

In Chapter 10, Lemma 27, we have shown that weak sequential optimality for some CPS is characterized by the notion of *conditional dominance*. With this, it is relatively easy to see that initial rationalizability can be given the following dominance characterization. The first step coincides with deletion of all conditionally dominated strategies. The following steps coincide with the iterated deletion of strategies that are strictly dominated at the initial history, or equivalently, with iterated strict dominance in the residual strategic form obtained after the deletion of conditionally dominated strategies. To formalize, let $NCD = \times_{i\in I} NCD_i$ denote the set of profiles of conditionally undominated (not conditionally dominated) strategies. Also, let $\mathrm{ND}\,(C)$ denote the set of profile of strategies that are not dominated (by mixed strategies) in the restricted Cartesian set $C \subseteq S$, i.e., the undominated profiles in the restricted strategic form $\left\langle I, (U_i|_C, C_i)_{i\in I}\right\rangle$. As in Chapter 4, $\mathrm{ND} : \mathcal{C} \to \mathcal{C}$ is a restriction operator: $\mathrm{ND}\,(C) \subseteq C$ for every $C \in \mathcal{C}$. As such, it yields a decresing sequence of subsets $\left(\mathrm{ND}^k\left(\bar{S}\right)\right)_{k\in\mathbb{N}}$ starting from any subset $\bar{S}$. With this, we can state our result.

Proposition 7. $\rho^\infty(S) = \mathrm{ND}^\infty(NCD)$.

We omit the proof. Note the order of elimination: *first* one has to delete all the conditionally dominated strategies of the multistage game Γ, obtaining the subset $NCD \subseteq S$, *then* one continues with the iterated deletion of dominated strategies starting from NCD. Inverting the two operations is not possible: the elimination of the conditionally dominated strategies does not yield the strategic form of a multistage game. This is because, as shown by Example 49, the set of justifiable strategies is not the Cartesian product of subsets of actions at the different histories, therefore it cannot be represented as a game tree.

Finally, we illustrate initial rationalizability in the BoS with an outside option (Example 38).

Example 50. Consider the BoS with an Outside Option. We omit action "wait" from the description of player 2's strategies, and we illustrate the steps of initial rationalizability. For player 1, we have $r_1(\mu^1) = \{\text{out.B}_1, \text{out.S}_1\}$ for each CPS μ^1 such that $\mu^1(\text{S}_2|\varnothing) = 1$, and $r_1(\mu^1) = \{\text{in.B}_1\}$ for each $\mu^1 \in \Delta^H(S_2)$ such that $\mu^1(\text{B}_2|\varnothing) = 1$. For player 2, we have $r_2(\mu^2) = \{\text{B}_2\}$ for each CPS μ^2 such that $\mu^2(\text{in.B}_1|\varnothing) = 1$, and

$r_2(\mu^2) = \{S_2\}$ for each CPS μ^2 such that $\mu^2(\text{in.S}_1|\varnothing) = 1$. Thus,

$$\rho^1(S) = \{\text{out.B}_1, \text{out.S}_1, \text{in.B}_1\} \times \{B_2, S_2\}.$$

At second step, no strategy for player 1 is deleted. For player 2, we have $r_2(\mu^2) = \{B_2\}$ for each CPS μ^2 such that $\mu^2(\text{in.B}_1|\varnothing) = 1$, and $r_2(\mu^2) = \{S_2\}$ for each CPS μ^2 such that $\mu^2(\{\text{out.B}_1, \text{out.S}_1\}|\varnothing) = 1$ and $\mu^2(\text{in.S}_1|(\text{in})) = 1$. Hence,

$$\rho^2(S) = \{\text{out.B}_1, \text{out.S}_1, \text{in.B}_1\} \times \{B_2, S_2\}.$$

No strategy has been eliminated with the second step of the procedure. Therefore

$$\rho^\infty(S) = \{\text{out.B}_1, \text{out.S}_1, \text{in.B}_1\} \times \{B_2, S_2\}.$$

▲

Note that, in Example 50, the set of initially rationalizable strategies of player 1 includes in.B_1 but not in.S_1. Despite this, the set of initially rationalizable strategies of player 2 includes not only B_2 but also S_2. This is so because player 2 can assign probability 1 to the initially rationalizable strategies of player 1 that prescribe action out; if this belief is contradicted by the observation of in, then player 2 is free to think that player 1 is not carrying out a rational plan, and that player 1 will then play S_1. The consequence is that both strategies of player 2 remain justifiable under initial belief in rationality of player 1. Thus, in this game, initial rationalizability yields the same strategies as rationalizability in the strategic form of the game (cf. Chapter 9, Section 9.3.1),[4] where both B_2 and S_2 survive simply because the strategy of player 2 is irrelevant for the outcome when player 1 plays out. In the next section, we will see what happens if instead player 2 believes that in was part of a rational plan (and player 1 believes that player 2 reasons in this way).

11.3 Strong Rationalizability

When the initial belief about co-players' behavior is falsified by observation, a player may still try to interpret unexpected moves as part

[4] Rationalizability in the strategic form of the game coincides with the iteration elimination of strateies that are note ex-ante optimal for some conjecture consisent with the previous steps—cf. Chapter 10.

of a rational plan. Obviously, this is not always possible as there may be observable behavior that is not part of any rational plan of the co-player. But when possible, this rationalization may narrow down what a player can expect from the co-players in the continuation of the game. This activity of looking at past moves to predict future moves is called *forward-induction reasoning*. Forward-induction reasoning is enabled precisely by the hypothesis that past and future moves of a co-player must be part of the same rational plan, which the co-player does not fail to carry out. An alternative way of reasoning is to think that the unexpected moves were mistakes that disrupted the implementation of co-players' plans; in this case, the future moves can at best be predicted from their optimality in the continuation of the game, a form of *backward-induction reasoning* that will be analyzed in the next chapter.

Thus, forward-induction reasoning rests on maintaining the belief that the co-players are rational (i.e., they have a rational plan and implement it) at all histories that are consistent with their rationality. This persistency of the belief in an event as long as not contradicted by observation is called *strong belief*.

Definition 63. *Fix a player $i \in I$, a CPS μ^i and a subset of co-players' strategy profiles $\overline{S}_{-i} \subseteq S_{-i}$.*

*We say that μ^i **strongly believes** $\overline{S}_{-i}$ if $\mu^i(\overline{S}_{-i}|h) = 1$ for every $h \in H$ such that $\overline{S}_{-i} \cap S_{-i}(h) \neq \emptyset$.*

We let $\Delta^H_{\mathrm{sb}}(\overline{S}_{-i})$ denote the set of CPSs of player i that strongly believe $\overline{S}_{-i}$.

Strong belief in rationality is the cornerstone of strategic reasoning. Here we represent this as strong belief in the set of co-players' profiles of justifiable strategies.[5] More generally, we impose a **best rationalization principle**: players always ascribe to co-players the highest level of "strategic sophistication" that is consistent with the observed behavior. For example, on top of strong belief in rationality, we also require that each player strongly believes that *his co-players are rational and* that *they strongly believe in rationality.* The latter is the second order of

[5]Recall that ratonality is a relationship between beliefs and behavior. Yet, here we represent formally only beliefs about behavior, not beliefs about others' beliefs. Thus, we represent belief in rationality as belief in behavior consistent with rationality, i.e., belief in justifiable behavior.

strategic sophistication of co-players. As we consider further iterations of this form of strategic thinking, we want to study the behavioral implications of *rationality and common strong belief in rationality*. Such behavioral implications are characterized by a solution concept called **strong rationalizability**.

Definition 64. *Consider the following elimination procedure.*

(Step $n = 0$**)** *For each $i \in I$, let $S_i^0 = S_i$. Also, let $S_{-i}^0 = \times_{j \neq i} S_j$ and $S^0 = S$.*

(Step $n > 0$**)** *For each $i \in I$, let*

$$\begin{aligned} \Delta_i^n &= \bigcap_{m=0}^{n-1} \Delta_{\text{sb}}^H(S_{-i}^m); \\ S_i^n &= \left\{ s_i \in S_i : \exists \mu^i \in \Delta_i^n, s_i \in r_i(\mu^i) \right\}. \end{aligned}$$

Also, let $S_{-i}^n = \times_{j \neq i} S_j^n$ and $S^n = \times_{i \in I} S_i^n$.

*Finally, for each $i \in I$, let $S_i^\infty = \bigcap_{n>0} S_i^n$, and $S^\infty = \times_{i \in I} S_i^\infty$. For each $i \in I$, the strategies in S_i^∞ are called **strongly rationalizable**.*

Let us compare strong and initial rationalizability. To ease the comparison, write the sequence of subsets obtained with the initial rationalizability procedure as $\left(S_\varnothing^n\right)_{n \in \mathbb{N}} = \left(\rho^n\left(S\right)\right)_{n \in \mathbb{N}}$ with $S_\varnothing^n = \times_{i \in I} S_{i,\varnothing}^n$. Since $\Delta_i^1 = \Delta^H\left(S_{-i}\right)$, the first step ($n = 1$) of strong rationalizability eliminates only the strategies of player i that are not justifiable, exactly like initial rationalizability: $S_i^1 = S_{i,\varnothing}^1$. At the second step of reasoning ($n = 2$), Δ_i^2 captures strong belief that the co-players are rational, i.e., strong belief in S_{-i}^1: $\mu^i(S_{-i}^1|h) = 1$ for every $h \in H$ such that $S_{-i}^1 \cap S_{-i}\left(h\right) \neq \emptyset$. Strong belief in an event E_{-i} clearly implies initial belief in E_{-i}, as the initial history is always consistent with the believed event. Hence,

$$\Delta_i^2 = \Delta_{\text{sb}}^H\left(S_{-i}^1\right) \subseteq \Delta_\varnothing^H\left(S_{-i}^1\right) = \Delta_\varnothing^H\left(S_{-i,\varnothing}^1\right)$$

($i \in I$), so that $S^2 \subseteq S_\varnothing^2 = \rho^2\left(S\right)$. Iterating this argument, we obtain the following result.

Remark 52. *Initial rationalizability is weaker than strong rationalizability, that is,*

$$S^n \subseteq S_\varnothing^n = \rho^n\left(S\right)$$

for all $n \in \mathbb{N} \cup \{\infty\}$.

Proof. We already know that the inclusion holds for the first two steps. Suppose, by way of induction, that $S^m \subseteq S^m_\varnothing$ for each $m < n$ (IH). Then, for each $i \in I$,

$$\Delta_i^n \overset{\text{(def)}}{=} \bigcap_{m=0}^{n-1} \Delta_{\text{sb}}^H \left(S_{-i}^m\right) \overset{\text{(sb}-\text{ib)}}{\subseteq} \bigcap_{m=0}^{n-1} \Delta_\varnothing^H \left(S_{-i}^m\right) \overset{\text{(IH,mon)}}{\subseteq} \Delta_\varnothing^H \left(S_{-i,\varnothing}^{n-1}\right),$$

where the equality holds by definition, the first inclusion holds because strong belief implies initial belief, and the second holds by the inductive hypothesis and the monotonicity of initial belief. Thus, for each $i \in I$,

$$S_i^n = r_i\left(\Delta_i^n\right) \subseteq r_i\left(\Delta_\varnothing^H\left(S_{-i,\varnothing}^{n-1}\right)\right) = S_{i,\varnothing}^n,$$

which implies $S^n \subseteq \rho^n(S)$ for all $n \in \mathbb{N}$. The latter implies $S^\infty \subseteq \rho^\infty(S)$. ■

Example 51. In Example 50, we have seen that in the BoS with an outside option the initially rationalizable strategies coincide with the justifiable strategies. In particular, Bob could play S_2 even though in.S_1 is not justifiable for Ann because Bob can initially believe that Ann will rationally play out, and then think that in has not been played as part of a rational plan. Strong rationalizability, instead, requires Bob to believe that in has been played as part of the rational plan in.B_1. Therefore, as anticipated in Chapter 9, the observation of in leads Bob to believe, by forward induction, that Ann will play B_1 next. Thus, we have $S_2^2 = \{\text{B}_2\} = S_2^\infty$. Moreover, Ann anticipates this, because she (strongly) believes that (i) Bob is rational, and (ii) Bob strongly believes that she is rational. Thus, we have $S_1^3 = \{\text{in.B}_1\} = S_1^\infty$. ▲

A more involved example of forward-induction reasoning captured by strong rationalizability is provided by the BoS with Dissipative Action.

Example 52. Consider the game in Example 39, Chapter 9, where Bob (b) is the first-mover. All strategies of Ann (a) are justifiable. Note that for Bob, given any CPS μ^b, the weakly sequentially optimal strategies are entirely determined by $\mu^b(\cdot|\varnothing)$; this is because Bob never plays after observing an action of Ann, hence Bob never has to revise $\mu^b(\cdot|\varnothing)$ before the game is over. Now, for any initial belief, it is not optimal for Bob to burn and then aim for the least preferred coordination outcome. So, we have

$$S_b^1 = \{B.L.r, B.R.r\} \cup \{s_b \in S_b : s_b(\varnothing) = N\}.$$

No strategy of Ann can be eliminated at the first step. Hence, no strategy of Bob can be eliminated at the second step, and so on; therefore we can focus on just one player at each step of reasoning. For Ann, strong belief in S_b^1 entails that playing u at history (B) is suboptimal. So, we have

$$\begin{aligned} \Delta_a^2 &= \left\{\mu^a \in \Delta^H(S_b) : \mu^a(\{B.L.r, B.R.r\} \,|(B)) = 1\right\}, \\ S_a^2 &= \{U.d, D.d\}. \end{aligned}$$

Then, for Bob, (initial) belief in S_a^2 guarantees that burning will effectively signal the intention to coordinate on his preferred outcome. Then, not burning makes sense for him only if he assigns sufficiently high probability to Ann playing d also after not burning. Hence, we have

$$\begin{aligned} \Delta_b^3 &= \left\{\mu^b \in \Delta^H(S_a) : \mu^b(\{U.d, D.d\} \,|\varnothing) = 1\right\}, \\ S_b^3 &= \{N.R.l, N.R.r, B.L.r, B.R.r\}. \end{aligned}$$

So, for Ann, strong belief in S_b^3 translates into certainty of R at history (N), on top of certainty of r at history (B). Thus, we have

$$\begin{aligned} \Delta_a^4 &= \left\{ \begin{array}{ll} \mu^a \in \Delta^H(S_b) : & \mu^a(\{B.L.r, B.R.r\} \,|(B)) = \\ & \mu^a(\{N.R.l, N.R.r\} \,|(N)) = 1 \end{array} \right\}, \\ S_a^4 &= \{D.d\} = S_a^\infty. \end{aligned}$$

For Bob, (initial) belief in S_a^4 guarantees that his preferred coordination outcome can be obtained regardless of his initial move, and then, obviously, it is optimal for him not to burn:

$$\begin{aligned} \Delta_b^5 &= \left\{\mu^b \in \Delta^H(S_a) : \mu^b(\{D.d\} \,|\varnothing) = 1\right\}, \\ S_b^5 &= \{N.R.l, N.R.r\}. \end{aligned}$$

Forward-induction reasoning, as captured by strong rationalizability, pins down a unique path: $(N, (R, D))$. The mere possibility of burning allows Bob to induce coordination on his preferred outcome even without actually using the burning option. ▲

11.3.1 Strong Rationalizability as a Reduction Procedure

An attentive reader probably noticed that we have not defined strong rationalizability with the iterated application of a rationalization operator, as we did for rationalizability in simultaneous-move games, and for initial rationalizability. Why is it so? The reason is that the best rationalization principle requires each player i, at any step of reasoning $n > 2$, to strongly believe in all the sets/events $S_{-i}^{n-1}, ..., S_{-i}^{1}$, not just S_{-i}^{n-1}. If instead we were to impose only strong belief in S_{-i}^{n-1}, then Δ_i^n would not necessarily be a subset of Δ_i^{n-1}, and thus we would not have a well-defined elimination procedure. The point is that strong belief in S_{-i}^{n-1} is not a stronger condition than strong belief in S_{-i}^{n-2}: although $S_{-i}^{n-1} \subseteq S_{-i}^{n-2}$, the set $\Delta_{\mathrm{sb}}^{H}(S_{-i}^{n-1})$ is not necessarily a subset of $\Delta_{\mathrm{sb}}^{H}(S_{-i}^{n-2})$, because the set of histories consistent with S_{-i}^{n-1} can be a strict subset of those consistent with S_{-i}^{n-2}, and thus strong belief in S_{-i}^{n-1} may restrict i's conditional beliefs at fewer histories than strong belief in S_{-i}^{n-2}. (In general, the smaller an event, the fewer the contingencies where one can believe the event is true.)

An alternative route is to reformulate strong rationalizability as a *reduction* procedure. By focusing on the strategies that survived the previous steps of elimination, we can impose strong belief in S_{-i}^{n-1} only, and rely on the fact that the strategies in S_i^{n-1} already capture the best rationalization principle for the lower orders of belief, i.e., they already capture strong belief in $S_{-i}^{n-2}, ..., S_{-i}^{1}$. In this way, we overcome the non-mononicity of strong belief and we can write the reduction procedure with a *constrained* rationalization operator. To this end, fix any $C = \times_{i \in I} C_i \in \mathcal{C}$ and let

$$H\,(C) = \{h \in H : \exists s \in C, h \prec \zeta\,(s)\} = \{h \in H : S\,(h) \cap C \neq \emptyset\}$$

denote the set of non-terminal histories consistent with some strategy profile in C. We define a constrained optimality correspondence as follows: for each $i \in I$ and $\mu^i \in \Delta_{\mathrm{sb}}^{H}(C_{-i})$, let

$$r_i(\mu^i|C) = \left\{ \overline{s}_i \in C_i : \forall h \in H_i(\overline{s}_i) \cap H\,(C)\,, V_i^{\overline{s}_i,\mu^i}(h) = \max_{s_i \in C_i \cap S_i(h)} V_i^{s_i,\mu^i}(h) \right\}.$$

Thus, $r_i(\mu^i|C)$ is the set of strategies that, among the strategies in C_i, maximize the continuation value given $\mu^i(\cdot|h)$ at every history h consistent with C (and with the strategy itself). The correspondence $r_i(\cdot|C)$ would be well defined also if the domain contained all CPSs, but conceptually (and for our purpose) the CPSs that strongly believe C_{-i} are the interesting ones: it is under these CPSs that player i can focus exclusively on the reduced set C, as long as no deviation from the paths in $\zeta(C)$ is observed.

With this, we introduce the constrained rationalization operators:

$$\begin{aligned}\overline{\rho}_{i,\mathrm{sb}}(C) &= r_i(\Delta^H_{\mathrm{sb}}(C_{-i})|C),\\ \overline{\rho}_{\mathrm{sb}}(C) &= \bigtimes_{i\in I}\overline{\rho}_{i,\mathrm{sb}}(C).\end{aligned}$$

Finally, we iterate $\overline{\rho}_{\mathrm{sb}}(S)$ in the usual way:

$$\forall k>0,\qquad \begin{aligned}&\overline{\rho}^{0}_{\mathrm{sb}}(S)=S,\\ &\overline{\rho}^{k}_{\mathrm{sb}}(S)=\overline{\rho}_{\mathrm{sb}}(\overline{\rho}^{k-1}_{\mathrm{sb}}(S)),\\ &\overline{\rho}^{\infty}_{\mathrm{sb}}(S)=\textstyle\bigcap_{k\ge 1}\overline{\rho}^{k}_{\mathrm{sb}}(S).\end{aligned}$$

We want to prove that the reduction procedure $(\overline{\rho}^{k}_{\mathrm{sb}}(S))^{\infty}_{k=0}$ coincides with strong rationalizability. Compared to the analogous result for rationalizability in Chapter 4, there is here a difficulty: even if $\overline{\rho}^{k-1}_{\mathrm{sb}}(S)=S^{k-1}$, the set of CPSs $\Delta^H_{\mathrm{sb}}(\overline{\rho}^{k-1}_{-i,\mathrm{sb}}(S))$ is larger than the set Δ^k_i we use in the definition of strong rationalizability, which is, the set of CPSs that strongly believe $S^{k-1}_{-i},...,S^0_{-i}$. Nonetheless, the expected equivalence holds.

Proposition 8. *For each* $k\in\mathbb{N}\cup\{\infty\}$, $\overline{\rho}^{k}_{\mathrm{sb}}(S)=S^k$.

Proof. Obviously, $\overline{\rho}^{0}_{\mathrm{sb}}(S)=S^0$. So, fix $k\ge 0$ and suppose by way of induction that $\overline{\rho}^{k}_{\mathrm{sb}}(S)=S^k$. We need to show that, for each $i\in I$,

$$r_i(\Delta^H_{\mathrm{sb}}(\overline{\rho}^{k}_{-i,\mathrm{sb}}(S))|\overline{\rho}^{k}_{\mathrm{sb}}(S))=r_i(\Delta^{k+1}_i);$$

then, $\overline{\rho}^{k+1}_{i,\mathrm{sb}}(S)=S^{k+1}_i$ follows by definition.

We first show that $r_i(\Delta^{k+1}_i)\subseteq r_i(\Delta^H_{\mathrm{sb}}(\overline{\rho}^{k}_{-i,\mathrm{sb}}(S))|\overline{\rho}^{k}_{\mathrm{sb}}(S))$. Fix any $s_i\in r_i(\Delta^{k+1}_i)$. Then, by definition, there exists μ^i that strongly believes S^k_{-i} such that $s_i\in r_i(\mu^i)$. By the induction hypothesis, $S^k_{-i}=\overline{\rho}^{k}_{-i,\mathrm{sb}}(S)$. Hence $\mu^i\in\Delta^H_{\mathrm{sb}}(\overline{\rho}^{k}_{-i,\mathrm{sb}}(S))$. Moreover, $s_i\in S^k_i=\overline{\rho}^{k}_{i,\mathrm{sb}}(S)$. With this, we obtain

$$s_i\in r_i(\Delta^H_{\mathrm{sb}}(\overline{\rho}^{k}_{-i,\mathrm{sb}}(S))|\overline{\rho}^{k}_{\mathrm{sb}}(S)).$$

We now show the opposite inclusion. Fix any $\overline{s}_i \in r_i(\Delta^H_{\text{sb}}(\overline{\rho}^k_{-i,\text{sb}}(S))|\overline{\rho}^k_{\text{sb}}(S))$. Then, by definition, there exists μ^i that strongly believes $\overline{\rho}^k_{-i,\text{sb}}(S)$ such that $\overline{s}_i \in r_i(\mu^i|\overline{\rho}^k_{\text{sb}}(S))$. By the induction hypothesis, $S^k_{-i} = \overline{\rho}^k_{-i,\text{sb}}(S)$, hence μ^i strongly believes S^k_{-i}. By the induction hypothesis, $\overline{s}_i \in S^k_i$. Therefore, there is $\hat{\mu}^i$ that strongly believes $S^{k-1}_{-i}, ..., S^0_{-i}$ such that $\overline{s}_i \in r_i(\hat{\mu}^i)$. Then, we can construct a CPS $\bar{\mu}^i \in \Delta^{k+1}_i$ as follows: for every $h \in H$,

$$\bar{\mu}^i(\cdot|h) = \begin{cases} \mu^i(\cdot|h), & \text{if } S^k_{-i} \cap S_{-i}(h) \neq \emptyset, \\ \hat{\mu}^i(\cdot|h), & \text{otherwise.} \end{cases}$$

As $\bar{\mu}^i$ is a well-defined CPS,[6] it remains to show that $\overline{s}_i \in r_i(\bar{\mu}^i)$. For each $h \in H_i(\overline{s}_i)$, we have

$$V^{\overline{s}_i,\bar{\mu}^i}_i(h) = \max_{s_i \in \overline{\rho}^k_{i,\text{sb}}(S) \cap S_i(h)} V^{s_i,\bar{\mu}^i}_i(h).$$

Specifically, if $S^k_{-i} \cap S_{-i}(h) = \emptyset$, this follows from the fact that $\overline{s}_i \in r_i(\hat{\mu}^i)$ and $\bar{\mu}^i(\cdot|h) = \hat{\mu}^i(\cdot|h)$; if instead $S^k_{-i} \cap S_{-i}(h) \neq \emptyset$, then the above conclusion still holds, because (i) $\overline{\rho}^k_{-i,\text{sb}} \cap S_{-i}(h) \neq \emptyset$ (by the inductive hypothesis), (ii) $\overline{s}_i \in r_i(\mu^i|\overline{\rho}^k_{\text{sb}}(S))$ (by the inductive hypothesis), and (iii) $\bar{\mu}^i(\cdot|h) = \mu^i(\cdot|h)$. Using again the inductive hypothesis, we obtain $\overline{\rho}^k_{i,\text{sb}}(S) = S^k_i$. Moreover, $\bar{\mu}^i \in \Delta^k_i$. Therefore, $r_i(\bar{\mu}^i) \subseteq \overline{\rho}^k_{i,\text{sb}}(S)$. Then, by Lemma 29, $\overline{s}_i \in r_i(\bar{\mu}^i)$. ∎

11.3.2 Iterated Conditional Dominance

The reformulation of strong rationalizability as a reduction procedure opens up the opportunity to characterize it with a notion of iterated dominance. The appropriate notion of dominance is conditional dominance. In Chapter 10 we have shown that a strategy is conditionally undominated if and only if it is weakly sequentially optimal under a CPS, absent any restrictions on players' CPSs given by strategic reasoning. To characterize strong rationalizability with *iterated* conditional dominance, we need to extend the notion of conditional dominance and its relation with weak sequential optimality to a "reduced game" $C = \bigtimes_{i \in I} C_i$, where

[6] The chain rule holds: to see this, pick any $h' \in H$ such that $S^k_{-i} \cap S_{-i}(h') = \emptyset$ and $h \in H$ such that $S^k_{-i} \cap S_{-i}(h) \neq \emptyset$. Then we have $\mu^i(S_{-i}(h')|h) = 0$, as μ^i strongly believes S^k_{-i}.

C_{-i} describes the contingent behavior that each player i expects from the co-players. For each $i \in I$, say that a strategy $\bar{s}_i \in C_i$ is **conditionally dominated in** C if there exist a history $h \in H_i(\bar{s}_i) \cap H(C)$ and a mixed strategy $\sigma_i \in \Delta(C_i)$, with $\sigma_i(S_i(h)) = 1$, such that

$$\forall s_{-i} \in C_{-i} \cap S_{-i}(h), \quad \sum_{s_i \in C_i \cap S_i(h)} \sigma_i(s_i) u_i(\zeta(s_i, s_{-i})) > u_i(\zeta(\bar{s}_i, s_{-i})).$$

We say that $\bar{s}_i \in C_i$ is **conditionally undominated in** C if it is *not* conditionally dominated in C. The set of strategies of player i that are *conditionally undominated* (Not Conditionally Dominated) *in* C is denoted by $\mathrm{NCD}_i(C)$.

Lemma 30. *Fix* $C = \times_{i \in I} C_i \in \mathcal{C}$. *A strategy* $\bar{s}_i \in C_i$ *is conditionally undominated in* C *if and only if there exists* $\mu^i \in \Delta^H_{\mathrm{sb}}(C_{-i})$ *such that* $\bar{s}_i \in r_i(\mu^i | C)$.

Proof. The proof follows the same lines as the proof of Lemma 27 in Chapter 10. Here we adapt the key steps, and the reader can consult the proof of Lemma 27 for further details.

Suppose there exists $\mu^i \in \Delta^H_{\mathrm{sb}}(C_{-i})$ such that $\bar{s}_i \in r_i(\mu^i | C)$. Then, for every $h \in H_i(\bar{s}_i) \cap H(C)$,

$$\bar{s}_i \in \arg \max_{s_i \in C_i \cap S_i(h)} \sum_{s_{-i} \in S_{-i}(h)} u_i(\zeta(s_i, s_{-i})) \mu^i(s_{-i} | h),$$

and since μ^i strongly believes C_{-i},

$$\mu^i(C_{-i} \cap S_{-i}(h) | h) = 1.$$

Then, by Lemma 2,[7] no mixed action $\sigma_i \in \Delta(C_i \cap S_i(h))$ dominates $\bar{s}_i$ in $(C_i \cap S_i(h)) \times (C_{-i} \cap S_{-i}(h))$. Thus, $\bar{s}_i$ is not conditionally dominated in C.

Conversely, suppose now that $\bar{s}_i$ is not conditionally dominated in C. Thus, for each $h \in H_i(\bar{s}_i)$, $\bar{s}_i$ is not dominated in $(C_i \cap S_i(h)) \times (C_{-i} \cap S_{-i}(h))$. Then, by Lemma 2, there exists $\nu^{i,h} \in \Delta(C_{-i} \cap S_{-i}(h))$

$$\bar{s}_i \in \arg \max_{s_i \in C_i \cap S_i(h)} \sum_{s_{-i} \in C_{-i} \cap S_{-i}(h)} u_i(\zeta(s_i, s_{-i})) \nu^{i,h}(s_{-i} | h)$$

[7] Lemma 2 is applied here to the strategic form of the game reduced to C.

For each $h \notin H_i(\bar{s}_i)$, fix $\nu^{i,h} \in \Delta(S_{-i}(h))$ such that, if $h \in H(C_{-i})$, $\nu^{i,h}(C_{-i}) = 1$. Then, construct a CPS $\mu^i \in \Delta_{\text{sb}}^H(C_{-i})$ and check that $\bar{s}_i \in r_i(\mu^i|C)$ as in the proof of Lemma 27. ■

We are now in position to state the equivalence between strong rationalizability and iterated conditional dominance. Let $\text{NCD}(C) = \times_{i\in I} \text{NCD}_i(C) \subseteq S$ denote the set of conditionally undominated strategy profiles in $C = \times_{i\in I} C_i \in \mathcal{C}$. This defines a restriction operator, that is, an operator $\text{NCD} : \mathcal{C} \to \mathcal{C}$ such that $\text{NCD}\,(C) \subseteq C$ for every $C \in \mathcal{C}$. As we did for iterated dominance in Chapter 4, we can represent iterated conditional dominance through the following sequence of subsets $\left(\text{NCD}^k(S)\right)_{k\in\mathbb{N}}$, where—as usual—$\text{NCD}^k = \text{NCD} \circ \text{NCD}^{k-1}$ (with $\text{NCD}^0\,(C) = C$ for every C).

Definition 65. *A profile of strategies $s \in S$ survives iterated conditional dominance if $s \in \text{NCD}^\infty(S) = \bigcap_{k\geq 1} \text{NCD}^k(S)$.*

Theorem 35. *For every $k = 1, 2, ...$, we have $S^k = \text{NCD}^k(S)$. Therefore, a strategy profile is strongly rationalizable if and only if it survives iterated conditional dominance.*

Proof. For each $k \in \mathbb{N} \cup \{\infty\}$, it follows from Proposition 8 that $\overline{\rho}^k_{\text{sb}}(S) = S^k$. By definition, for every $C \in \mathcal{C}$ and $i \in I$, we have $\overline{\rho}_{i,\text{sb}}(C) = r_i(\Delta^H_{\text{sb}}(C_{-i})|C)$. Then, by Lemma 30, $\overline{\rho}^1_{\text{sb}}(S) = \text{NCD}^1(S)$, and by a simple inductive argument, $S^k = \text{NCD}^k(S)$ for every $k \in \mathbb{N}$. ■

In Chapter 10 we have shown that conditional dominance and weak dominance in the strategic form of the game are generically equivalent. In light of this result, strong rationalizability is generically equivalent to **iterated admissibility**, i.e., the iterated deletion of all weakly dominated strategies. To prove this result, we need first to extend Lemma 28 to conditional dominance within a set of strategy profiles C — the proof is in the appendix of this chapter. For any Cartesian set C of strategy profiles, let $\text{NWD}_i\,(C)$ denote the set of strategies of player i that are not weakly dominated in C, that is, are not weakly dominated in the strategic form restricted to C, $\left\langle I, (U_i|_C, C_i)_{i\in I}\right\rangle$, and let $\text{NWD}\,(C) = \times_{i\in I} \text{NWD}_i\,(C)$.

Lemma 31. *Fix arbitrarily a Cartesian subset of strategy profiles $C = \times_{i\in I} C_i$. Generically, $\text{NCD}(C) = \text{NWD}\,(C)$.*

Proposition 9. *Generically,* $S^\infty = \mathrm{NWD}^\infty(S)$*.*

Proof. Fix $k \geq 0$ and suppose by induction that, generically, $S^k = \mathrm{NWD}^k(S)$ (it is trivially true for $n = 0$). By Theorem 35, $S^{k+1} = \mathrm{NCD}^{k+1}(S)$. Let $C = S^k$. Thus, $\mathrm{NCD}^{k+1}(S) = \mathrm{NCD}(C)$. By Lemma 31, we have $\mathrm{NCD}(C) = \mathrm{NWD}(C)$, except for a subset of payoff vectors W with zero-measure closure. By the inductive hypothesis, we also have

$$\mathrm{NCD}(C) = \mathrm{NCD}(\mathrm{NWD}^k(S)) = \mathrm{NWD}^{k+1}(S),$$

except for a subset of payoff vectors W' with zero-measure closure. The union of the closures of W and $W^{'}$ has zero measure. Moreover, it is a closed and contains $W \cup W'$. Hence, it contains the closure of $W \cup W'$. It follows that the closure of $W \cup W'$ has zero measure. We conclude that, generically, $S^{k+1} = \mathrm{NWD}^{k+1}(S)$. Since the game is finite, there exists $M \in \mathbb{N}$ such that $S^\infty = S^M$ and $\mathrm{NWD}^\infty(S) = \mathrm{NWD}^M(S)$, and we have proven that, generically, $S^M = \mathrm{NWD}^M(S)$. ■

11.3.3 Independent Rationalization and Strategic Independence

A natural variant of strong rationalizability arises from the observation that, under the baseline definition, player i is free to believe that *all* his co-players are not rational once just one co-player chooses an action that is not prescribed by any justifiable strategy. To capture the hypothesis that each player i ascribes to each individual co-player j the highest level of strategic sophistication that is consistent with j's behavior, we introduce the following variation of strong rationalizability, which we call **independent strong rationalizability**.

Definition 66. *For each* $i \in I$*, let* $\widehat{S}_i^0 = S_i$*. For each* $n > 0$*, let* $s_i \in \widehat{S}_i^n$ *if and only if there exists* $\mu^i \in \Delta^H(S_{-i})$ *such that:*

1. *for all* $j \neq i$*,* $m = 0, ..., n-1$*, and* $h \in H(\widehat{S}_j^m)$*,*

$$\mu^i\left(\widehat{S}_j^m \times \left(\bigtimes_{k\neq i,j} S_k\right)|h\right) = 1;$$

2. $s_i \in r_i(\mu^i)$*.*

Finally, let $\widehat{S}_i^\infty = \bigcap_{n>0} \widehat{S}_i^n$*.*

Now the question is: is independent strong rationalizability equivalent to strong rationalizability? Yes and no. In terms of strategies, this version of strong rationalizability can yield a different output, but the induced terminal histories do not change. In this case, we say that the two procedures are path-equivalent. (The proof is in the appendix.)

Proposition 10. *For each* $n \in \mathbb{N} \cup \{\infty\}$, $\zeta(\widehat{S}^n) = \zeta(S^n)$.

The intuition behind Proposition 10 is that *independent rationalization* (i.e, the belief that each individual co-player has the highest level of strategic sophistication that is consistent with his own behavior), only has bite at histories that are not compatible with strategic reasoning for one or more co-players (but consistent for at least one co-player).

So far, we allowed for the possibility that players have correlated beliefs about the behavior of different co-players. Sometimes, it makes sense for a player to assume that, according to his subjective beliefs, the strategies of different co-players are mutually independent. This means that what player i believes about co-player j does not depend on the observed behavior of co-player $k \neq j$. This assumption on players' beliefs is called *strategic independence*, and it is naturally complemented by independent rationalization. Here we do not provide the formal definition of strong rationalizability under strategic independence.

11.3.4 Elimination Orders

From an algorithmic viewpoint, it is interesting to understand whether "slowing down" the elimination of conditionally dominated strategies may change the final output of the procedure. For instance, in a two-player game, it could be convenient to alternate the eliminations for the two players in the following way: first compute the justifiable strategies S_1^1 of player 1 only; then compute directly the strategies S_2^2 of player 2 that are justified by a CPS that strongly believes S_1^1; then let $\widetilde{S}_1^3$ be the set of strategies of player 1 that are justified by a CPS that strongly believes S_2^2, and so on. In Chapter 4, we have seen that—in simultaneous-move games—the order of elimination of strictly dominated strategies is immaterial for the final output of iterated strict dominance. Here, the problem is much complicated by the non-monotonicity of strong belief. In the procedure above, we have $\widetilde{S}_1^3 \supseteq S_1^3$, because the set of CPSs that strongly believe S_2^2 are clearly a superset of those that strongly believe

S_2^2 *and* S_2^1. However, the set of CPSs that strongly believe $\widetilde{S}_1^3$ is not a superset of those that strongly believe S_1^3, and at this point $\widetilde{S}_2^4$ need not be a superset of S_2^4. Where does this lead to? Fortunately, although the final output will look different from strong rationalizability in terms of strategies, the induced terminal histories will be exactly the same. We omit the details.

11.4 Appendix

Proof of Lemma 31. Fix a strategy $\bar{s}_i$ and let $W \subset \mathbb{R}^Z$ denote the set of player i's payoff vectors $(y_z)_{z\in Z}$ such that $\bar{s}_i$ is *weakly dominated* in the strategic form with strategy profiles C, *but not conditionally dominated* within C. We show that the closure of W, denoted by $\mathrm{cl}\,(W)$, has zero measure. Since the number of strategies is finite, this yields the result.

Let $\bar{Z}$ be the set of all $z \in Z$ such that $\bar{s}_i \in S_i(z)$. Fix any $\bar{z} \in \bar{Z}$. For each payoff vector $(y_z)_{z\in Z}$, there is a unique vector $(x_z)_{z\in Z}$ such that

$$\begin{cases} y_z = x_z, & \text{if } z \notin \bar{Z}, \\ y_z = x_z + x_{\bar{z}}, & \text{if } z \in \bar{Z}; \end{cases} \tag{11.4.1}$$

that is, vector $(x_z)_{z\in Z}$ such that $x_{\bar{z}} = y_{\bar{z}}/2$, $x_z = y_z - x_{\bar{z}}$ for all $z \in \bar{Z}$, and $x_z = y_z$ for all $z \notin \bar{Z}$. Let $f : \mathbb{R}^Z \to \mathbb{R}^Z$ denote the transformation defined by (11.4.1).

We are going to study for which vectors $(x_z)_{z\in Z}$ the corresponding payoff vector $(y_z)_{z\in Z} = f((x_z)_{z\in Z})$ belongs to $\mathrm{cl}\,(W)$. In particular, we now show that for each $(x_z)_{z\neq\bar{z}} \in \mathbb{R}^{Z\setminus\{\bar{z}\}}$, there is at most one value of $x_{\bar{z}}$ such that $(y_z)_{z\in Z} \in \mathrm{cl}\,(W)$. Fix two values $x < x'$ of $x_{\bar{z}}$, with corresponding payoff vectors $(y_z)_{z\in Z}$ and $(y'_z)_{z\in Z}$. Thus[8]

$$(y'_z)_{z\in\bar{Z}} > (y_z)_{z\in\bar{Z}}, \tag{11.4.2}$$
$$(y'_z)_{z\notin\bar{Z}} = (x_z)_{z\notin\bar{Z}}, \tag{11.4.3}$$
$$(y_z)_{z\notin\bar{Z}} = (x_z)_{z\notin\bar{Z}}. \tag{11.4.4}$$

Suppose that $(y'_z)_{z\in Z} \in \mathrm{cl}\,(W)$. We are going to show that $(y_z)_{z\in Z} \notin \mathrm{cl}\,(W)$, because there are payoff vectors arbitrarily close to $(y_z)_{z\in Z}$ such

[8] Given any two vectors $b = (b_1, ..., b_n), b' = (b'_1, ..., b'_n)$ of real numbers of the same length, we write $b' > b$ when $b'_j > b_j$ for each $j = 1, ..., n$.

that $\bar{s}_i$ is conditionally dominated. Since $(y'_z)_{z\in Z} \in \mathrm{cl}\,(W)$, there exists some $(\widetilde{y}'_z)_{z\in Z} \in W$ that preserves (11.4.2) and (11.4.3):[9]

$$\begin{aligned}(\widetilde{y}'_z)_{z\in\bar{Z}} &> (y_z)_{z\in\bar{Z}}, &(11.4.5)\\ (\widetilde{y}'_z)_{z\notin\bar{Z}} &= (x_z)_{z\notin\bar{Z}}. &(11.4.6)\end{aligned}$$

Since $(\widetilde{y}'_z)_{z\in Z} \in W$, there exists $\bar{\sigma}_i \in \Delta(S_i \backslash \{\bar{s}_i\})$ with $\bar{\sigma}_i(C_i) = 1$ such that

$$\begin{aligned}\forall s_{-i} &\in C_{-i}, \sum_{s_i\in \mathrm{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i)\widetilde{y}'_{\zeta(s_i,s_{-i})} \geq \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}, &(11.4.7)\\ \exists s_{-i} &\in C_{-i}, \sum_{s_i\in \mathrm{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i)\widetilde{y}'_{\zeta(s_i,s_{-i})} > \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}, &(11.4.8)\end{aligned}$$

i.e., $\bar{s}_i$ is weakly dominated in C. Condition (11.4.8) requires that there exist histories $h \in H(\bar{s}_i)$ such that (i) $C_{-i}\cap S_{-i}(h) \neq \emptyset$ and (i) $\hat{s}_i(h) \neq \bar{s}_i(h)$ for some $\hat{s}_i \in \mathrm{supp}\bar{\sigma}_i$. Among these histories, fix one history $\bar{h}$ be such that no $h \prec \bar{h}$ has property (i). This means that, for every $h \prec \bar{h}$ and $\hat{s}_i \in \mathrm{supp}\bar{\sigma}_i$, $\hat{s}_i(h) = \bar{s}_i(\bar{h})$, that is, every $\hat{s}_i$ prescribes the same action. Hence, $\bar{\sigma}_i(S_i(\bar{h})) = 1$. Now fix $s_{-i} \in C_{-i} \cap S_{-i}(\bar{h})$. For each $s_i \in \mathrm{supp}\bar{\sigma}_i$, we have

$$\zeta(s_i, s_{-i}) = \zeta(\bar{s}_i, s_{-i}) \;\Leftrightarrow\; \zeta(s_i, s_{-i}) \in \bar{Z}. \tag{11.4.9}$$

Direction $\Rightarrow$ is obvious; to see $\Leftarrow$, reason by contraposition: $\zeta(s_i, s_{-i}) \neq \zeta(\bar{s}_i, s_{-i})$ requires $\hat{s}_i(h) \neq \bar{s}_i(h)$ for some $h \prec \zeta(s_i, s_{-i})$, and thus $\zeta(s_i, s_{-i})$ cannot be reached with $\bar{s}_i$ under any $s'_{-i} \in S_{-i}$. At this point, we can

[9] Equation (11.4.3) can be preserved because it pertains to terminal histories that are inconsistent with $\bar{s}_i$.

conclude that

$$\begin{aligned}
&\sum_{s_i \in \mathrm{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i)\left(y_{\zeta(s_i,s_{-i})} - y_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.4.9)}{=} \\
&\sum_{s_i \in \mathrm{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i)\left(y_{\zeta(s_i,s_{-i})} - y_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.4.4)}{=} \\
&\sum_{s_i \in \mathrm{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i)\left(x_{\zeta(s_i,s_{-i})} - y_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.4.5)}{>} \\
&\sum_{s_i \in \mathrm{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i)\left(x_{\zeta(s_i,s_{-i})} - \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.4.6)}{=} \\
&\sum_{s_i \in \mathrm{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i)\left(\widetilde{y}'_{\zeta(s_i,s_{-i})} - \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.4.9)}{=} \\
&\sum_{s_i \in \mathrm{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i)\left(\widetilde{y}'_{\zeta(s_i,s_{-i})} - \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.4.7)}{\geq} 0.
\end{aligned}$$

Since this is true for every $s_{-i} \in C_{-i} \cap S_{-i}(\bar{h})$, $\bar{s}_i$ is conditionally dominated by $\bar{\sigma}_i$ in C when $x_{\bar{z}} = x$. Thus, $(y_z)_{z\in Z} \notin W$. Moreover, conditional dominance is preserved in a neighborhood of $(y_z)_{z\in Z}$, because it is defined by strict inequalities. This shows that $(y_z)_{z\in Z} \notin \mathrm{cl}\,(W)$.

Now, the Lebesgue measure of $\mathrm{cl}\,(W)$ is defined as

$$\lambda(\mathrm{cl}\,(W)) = \int \mathbf{1}_{\mathrm{cl}(W)}\left((y_z)_{z\in Z}\right) \prod_{z\in Z} \mathrm{d}y_z.$$

We evaluate this integral after a change of variables, from $(y_z)_{z\in Z}$ to $(x_z)_{z\in Z}$. The change of variables yields

$$\int \mathbf{1}_{\mathrm{cl}(W)}\left((y_z)_{z\in Z}\right) \cdot_{z\in Z} \mathrm{d}y_z = \int \mathbf{1}_{\mathrm{cl}(W)}\left((x_z)_{z\in Z\setminus\bar{Z}}, (x_z + x_{\bar{z}})_{z\in\bar{Z}}\right) k \prod_{z\in Z} \mathrm{d}x_z,$$

where k denotes the determinant of the Jacobian matrix of f. By linearity of f, k is a constant. Using the fact that $\mathbf{1}_{\mathrm{cl}(W)}\left((y_z)_{z\in Z}\right)$ is not zero for at most one value of $x_{\bar{z}}$, we obtain

$$\begin{aligned}\lambda(\mathrm{cl}\,(W)) &= \int \left(\int \mathbf{1}_{\mathrm{cl}(W)} \left((x_z)_{z\in Z\setminus \bar{Z}}\,, (x_z + x_{\bar{z}})_{z\in \bar{Z}} \right) k \mathrm{d}x_{\bar{z}} \right) \prod_{z\neq \bar{z}} \mathrm{d}x_z \\ &= \int 0 \prod_{z\neq\bar{z}} \mathrm{d}x_z = 0.\end{aligned}$$

■

Lemma 28 is a corollary of Lemma 31: just set $C = S$.

Proof of Proposition 10. The proof is by induction, but we need a stronger induction hypothesis than the claim we aim to prove. Recall that $H(C) = \{h \in H : S\,(h) \cap C \neq \emptyset\}$.

Induction hypothesis (n): For each $i \in I$ and $s_i \in S_i^n$ (resp., $\widehat{s}_i \in \widehat{S}_i^n$), there exists $\widehat{s}_i \in \widehat{S}_i^n$ (resp., $s_i \in S_i^n$) such that $s_i(h) = \widehat{s}_i(h)$ for each $h \in H(S^n) = H(\widehat{S}^n)$.

Basis step ($n = 0$): $\widehat{S}_i^0 = S_i^0 = S_i$ for each $i \in I$.

Inductive step ($n + 1$): Let $H^* := H(S^n) = H(\widehat{S}^n)$ and $Z^* := \zeta(S^n) = \zeta(\widehat{S}^n)$, where the equalities follow from the induction hypothesis. We prove that for each $i \in I$ and $s_i \in S_i^{n+1}$, there is $\widehat{s}_i \in \widehat{S}_i^{n+1}$ such that $s_i(h) = \widehat{s}_i(h)$ for each $h \in H^*$; the proof of the converse is identical. Fix any μ^i that strongly believes S_{-i}^n and such that $s_i \in r_i(\mu^i)$. By the induction hypothesis, we can construct a map $\eta : S_{-i}^n \to \widehat{S}_{-i}^n$ that associates each $s_{-i} \in S_{-i}^n$ with some $\widehat{s}_{-i} \in \widehat{S}_{-i}^n$ such that $s_{-i}(h) = \widehat{s}_{-i}(h)$ for every $h \in H^*$. For each $h \in H^*$, since $\mu^i(S_{-i}^n|h) = 1$, we can construct $\hat{\mu}^i(\cdot|h) \in \Delta(\widehat{S}_{-i}^n)$ as $\hat{\mu}^i(s_{-i}|h) = \mu^i(\eta^{-1}(s_{-i})|h)$ for each $s_{-i} \in \widehat{S}_{-i}^n$. It is easy to see that, by construction, $\hat{\mu}^i(S_{-i}(h)|h) = 1$. Moreover, for each $h' \in H^*$ with $S_{-i}(h') \subseteq S_{-i}(h)$, and for each $\overline{S}_{-i} \subseteq S_{-i}(h')$,

$$\begin{aligned}\hat{\mu}^i(\overline{S}_{-i}|S_{-i}(h)) &= \mu^i(\eta^{-1}(\overline{S}_{-i})|S_{-i}(h)) \\ &= \mu^i(\eta^{-1}(\overline{S}_{-i})|S_{-i}(h')) \cdot \mu^i(S_{-i}(h')|S_{-i}(h)) \\ &= \hat{\mu}^i(\overline{S}_{-i}|S_{-i}(h')) \cdot \hat{\mu}^i(S_{-i}(h')|S_{-i}(h)),\end{aligned}$$

where the second equality follows from the fact that $\eta^{-1}(\overline{S}_{-i}) \subseteq S_{-i}(h')$, while the third equality holds because, for each $s_{-i} \in \widehat{S}_{-i}^n$, we have $s_{-i} \in S_{-i}(h')$ if and only if $\eta^{-1}(s_{-i}) \in S_{-i}(h')$. For each $h' \notin H^*$ such that

$\hat{\mu}^i(S_{-i}(h')|h) > 0$ for some $h \in H^*$, derive $\hat{\mu}^i(\cdot|h')$ by conditioning. For every other $h' \notin H^*$, let $\hat{\mu}^i(\cdot|h') = \widetilde{\mu}^i(\cdot|h')$ for any CPS $\widetilde{\mu}^i$ that satisfies condition 1 in Definition 66. Thus, $\hat{\mu}^i$ satisfies the chain rule and condition 1 in Definition 66. We are going to show that

$$\forall h \in H^* \cap H_i(s_i), \quad V_i^{s_i,\mu^i}(h) = \max_{s_i' \in S_i(h)} V_i^{s_i',\hat{\mu}^i}(h);$$

this implies the existence of some $\widehat{s}_i \in r_i(\hat{\mu}^i) \subseteq \widehat{S}_i^{n+1}$ such that $\widehat{s}_i(h) = s_i(h)$ for all $h \in H^*$, as desired. Fix $h \in H^* \cap H_i(s_i)$. Since $h \in H^*$, we have $\widehat{S}_i^n \cap S_i(h) \neq \emptyset$. So, there is at least one $\widehat{s}_i \in \widehat{S}_i^n$ such that

$$V_i^{\widehat{s}_i,\mu^i}(h) = \max_{s_i' \in S_i(h)} V_i^{s_i',\hat{\mu}^i}(h).$$

Since $s_i \in S_i^n$, $\widehat{s}_i \in \widehat{S}_i^n$, and $\mu^i(S_{-i}^n|h) = \hat{\mu}^i(\widehat{S}_{-i}^n|h) = 1$, we have $\zeta(s_i, s_{-i}) \in Z^*$ and $\zeta(\widehat{s}_i, s_{-i}) \in Z^*$ for each s_{-i} with $\mu^i(s_{-i}|h) > 0$ or $\hat{\mu}^i(s_{-i}|h) > 0$. By construction, we have $\mu^i(S_{-i}(z)|h) = \hat{\mu}^i(S_{-i}(z)|h)$ for each $z \in Z^*$. Hence, both s_i and $\widehat{s}_i$ induce the same distribution over terminal histories under $\mu^i(\cdot|h)$ and $\hat{\mu}^i(\cdot|h)$. So,

$$V_i^{s_i,\mu^i}(h) = \max_{s_i' \in S_i(h)} V_i^{s_i',\hat{\mu}^i}(h)$$

as well. ■

12

Equilibrium in Multistage Games

In this chapter we analyze subgame perfect equilibrium, a solution concept that—unlike Nash equilibrium, or iterated admissibility—is based on the details of the multistage game Γ rather than just its strategic form $\mathcal{N}(\Gamma)$. Somewhat informally, a **subgame perfect equilibrium** is a profile of strategies $s^* = (s_i^*)_{i \in I}$ such that each strategy s_i^* is sequentially optimal given the conjecture that other players behave as specified by s_{-i}^*. In other words, the *ex ante* optimality condition of Nash equilibrium is replaced by the more demanding *sequential* optimality condition. Indeed, as in the analysis of rationalizability of Chapter 11, it is *not* assumed that strategies are implemented by a machine or a trustworthy agent. Rather, it is informally assumed that each player makes a plan in advance and, as the play unfolds, he is free to either choose actions as planned or to deviate from the plan; therefore, players must not have incentives to deviate at any history, given that—as is typically assumed in equilibrium analysis—they have correct conjectures about each other. We will show that in finite games where only one player is active at each stage and payoffs are "generic," there is only one subgame perfect equilibrium that can be computed with the so called **backward induction** algorithm, a kind of "inter-personal folding-back" procedure. Such backward induction may seem to have a "common-belief-in-rationality flavor;" therefore, we will discuss the differences and relationships between backward induction and rationalizability. In more general games with finite horizon, one can use a "case-by-case backward induction" method to find the subgame perfect equilibria. Finally, we extend the analysis to account for randomized strategies and chance moves.

12.1 Subgame Perfect Equilibrium

In Section 9.3.1 we noticed that every solution concept for games with simultaneous moves can be applied to the strategic form $\mathcal{N}(\Gamma)$ of a multistage game Γ. This may be a useful first step. Indeed, a strategy is sequentially optimal given a conjecture only if it is an ex ante best reply to this conjecture (see Remark 49 and Proposition 2 in Chapter 10). This implies that strategies that are not rationalizable in $\mathcal{N}(\Gamma)$ cannot be rationalizable in Γ for whatever extension of the rationalizability concept from simultaneous to multistage games. Yet, we noticed in Chapter 11 that, in games with at least two stages, rationality entails a stronger form of justifiability than just being an ex ante best reply to some conjecture and this yields notions of rationalizability in Γ stronger than mere rationalizability in the strategic form $\mathcal{N}(\Gamma)$. Similar considerations apply to traditional equilibrium analysis: there may be Nash equilibria s^* of $\mathcal{N}(\Gamma)$ such that the strategy s_i^* of some player i is not sequentially optimal given that i "correctly" expects the co-players to behave as prescribed by strategy profile s_{-i}^*. The strategy of fighting entry in the Entry Game of Figure 9.7 is a case in point (see Section 9.3.1). The credibility of such strategies is questionable. Such considerations induced traditional game theorists to endorse a refinement[1] of the Nash equilibrium concept called "subgame perfect (Nash) equilibrium" for reasons that will soon be made clear. As in the case of the Nash equilibrium concept, no formal and general justification is offered for the "correct-conjectures" assumption underlying this equilibrium concept. We first focus on equilibria in pure strategies, then move to randomized equilibria.

[1] A "refinement" of a solution concept is a different, more demanding solution concept that, for some games, allows for a strictly smaller set of strategy profiles. For example, initial rationalizability is a refinement of rationalizability in the strategic form, and strong rationalizability is a refinement of initial rationalizability.

12.1.1 Pure Strategies

For any player i and conjecture $\beta^i \in \times_{h\in H} \Delta(\mathcal{A}_{-i}(h))$, let $\hat{r}_i\left(\beta^i\right)$ denote the set of sequentially optimal strategies given β^i, or **sequential best replies** to β^i.[2] Suppose that conjecture β^i is deterministic because it corresponds to a pure strategy profile s_{-i}, that is, $\beta^i(s_{-i}(h)|h) = 1$ for all $h \in H$; then, with an abuse of notation similar to the one adopted for static games, we write $\hat{r}_i(s_{-i})$.

Definition 67. *A pure strategy profile $s^* = (s_i^*)_{i\in I}$ is a* ***subgame perfect (Nash) equilibrium (SPE)*** *if, for each $i \in I$, s_i^* is sequentially optimal given the deterministic conjecture corresponding to s_{-i}^*, that is,*

$$\forall i \in I, s_i^* \in \hat{r}_i\left(s_{-i}^*\right).$$

For every strategy profile $s \in S$ and history $h \in H$, let $\zeta(s|h)$ denote the terminal history $z = (h, s(h), ...)$ induced by s *starting from* h. Note, if s induces h, that is $h \prec \zeta(s)$, then $\zeta(s|h) = \zeta(s)$. But $\zeta(s|h)$ is well defined even if s does not induce h, because $\zeta(s|h)$ depends only on the actions prescribed by s at h and the non-terminal histories following h (if any). For example, in the BoS with an Outside Option of Figure 9.1, $\zeta((\text{Out.B}_1, \text{B}_2)\,|\,(\text{In})) = (\text{In}, (\text{B}_1, \text{B}_2))$. The following remark explains the name "subgame perfect (Nash) equilibrium":

Remark 53. *By inspection of Definition 67, it follows that a strategy profile $s^* = (s_i^*)_{i\in I}$ is an SPE if and only if*

$$\forall i \in I, \forall h \in H, \forall s_i \in S_i,\ u_i(\zeta(s^*|h)) \geq u_i(\zeta(s_i, s_{-i}^*|h)).$$

Thus, for every $h \in H$, an SPE yields a Nash equilibrium in the subgame $\Gamma(h)$ starting at h.

Example 53. In the BoS with an Outside Option there are two (pure) SPEs. Indeed, the BoS subgame starting at (In) has two pure equilibria, (B_1, B_2) and (S_1, S_2). Thus, a SPE must have the form $(a_1.\text{B}_1, \text{B}_2)$ or $(a_1.\text{S}_1, \text{S}_2)$ with $a_1 \in \{\text{In}, \text{Out}\}$. Furthermore, $a_1.\text{X}_1$ must be a best reply to X_2 (with X= B or X = S) at the root, because the strategy pair must

[2] Recall that $r_i\left(\beta^i\right)$ instead denotes the set of *weakly* sequential optimal strategies given β^i. Thus, $\hat{r}_i\left(\beta^i\right) \subseteq r_i\left(\beta^i\right)$ and the inclusion may be strict if player i moves (is active) at leat twice along some given path.

be a Nash equilibrium. Only $\mathrm{In.B_1}$ is a best reply to $\mathrm{B_2}$, while $\mathrm{Out.S_1}$ is the only best reply to $\mathrm{S_2}$ or the required form (also $\mathrm{Out.B_1}$ is a best reply to $\mathrm{S_2}$ at the root, but it does not have the required form, hence, it is not a sequential best reply to $\mathrm{S_2}$). Thus, we obtain two SPEs: $(\mathrm{In.B_1}, \mathrm{B_2})$ and $(\mathrm{Out.S_1}, \mathrm{S_2})$. ▲

We will present a method to find all the SPEs of (finite) two-stage games. Most methods to compute the SPEs of a game are based on an equilibrium version of the OD principle. Say that **strategy profile** $s^* = (s_i^*)_{i \in I}$ satisfies **one-step optimality with correct conjectures** if, for each $i \in I$, strategy s_i^* is one-step optimal given the deterministic conjecture corresponding to s_{-i}^*. With this, Theorem 34 in Section 10.5 yields:

Corollary 7. (OD principle for pure SPE) *Suppose that* Γ *is continuous at infinity. Then a strategy profile is an SPE if and only if it satisfies one-step optimality with correct conjectures.*

By Remark 51 in Section 10.5, it follows that, for every *compact-continuous* game Γ (hence, in particular, for every finite game) a strategy profile is an SPE if and only if it satisfies one-step optimality with correct conjectures.

Backward Induction

We now introduce a class of games with perfect information with a unique SPE that can be computed by a backward calculation called "backward induction."

Recall that a game Γ has **perfect information (PI)** if, for each $h \in H$, only one player—denoted $\iota(h)$—is active at h:

$$\forall h \in H, \exists! i \in I, |\mathcal{A}_i(h)| > 1.$$

Thus, for each $h \in H$,

$$|\mathcal{A}_i(h)| > 1 \Leftrightarrow i = \iota(h),$$

and we can write *histories* as *sequences* of *actions of active players.* Fix two terminal histories $z', z'' \in Z$; if $z' \neq z''$, there is a player who is decisive, or "pivotal," for reaching z' rather than z'', that is, the player who is active at the longest prefix (last common predecessor) of z' and z''. We let $\pi(z', z'')$ denote this player; formally,

$$\pi\left(z', z''\right) = \iota\left(\arg \max_{h \in H : h \prec z', h \prec z''} \ell(h)\right).$$

Definition 68. *A game with perfect information* Γ *has* ***no relevant ties*** *if the active player is never indifferent between distinct continuation paths; formally,*

$$\forall z', z'' \in Z,\ z' \neq z'' \Rightarrow u_{\pi(z',z'')}\left(z'\right) \neq u_{\pi(z',z'')}\left(z''\right).$$

Now, fix a *finite PI game with no relevant ties*. For each $h \in H$ one can compute the SPE-action $s^*_{\iota(h)}(h)$ and the SPE-value $V^*_j(h)$ for each $j \in I$ of reaching h. This is done by induction on $L(\Gamma(h))$, the height of the subgame starting at h.[3] Of course, the value for a player of reaching a terminal history is just the payoff of that history; therefore, $V^*_j(z) = u_j(z)$ for all $j \in I$ and $z \in Z$. With this, **backward induction** (**BI**) is the following recursive procedure:

- If $L(\Gamma(h)) = 1$, then $\left(h, a_{\iota(h)}\right) \in Z$ for every $a_{\iota(h)} \in \mathcal{A}_{\iota(h)}(h)$; thus,
  - $i = \iota(h) \Rightarrow s^*_i(h) = \arg\max_{a_i \in \mathcal{A}_i(h)} V^*_i(h, a_i) = \arg\max_{a_i \in \mathcal{A}_i(h)} u_i(h, a_i)$,
  - $\forall j \in I$, $V^*_j(h) = V^*_j\left(h, s^*_{\iota(h)}(h)\right)$.
- Let $s^*_{\iota(h)}(h)$ and $V^*_j(h)$ be defined for every $h \in \bar{H}$ such that $L(\Gamma(h)) \leq k$ and every $j \in I$; if $L(\Gamma(h)) = k+1$, then $L\left(\Gamma\left(h, a_{\iota(h)}\right)\right) \leq k$ for every $a_{\iota(h)} \in \mathcal{A}_{\iota(h)}(h)$; thus,
  - $i = \iota(h) \Rightarrow s^*_i(h) = \arg\max_{a_i \in \mathcal{A}_i(h)} V^*_i(h, a_i)$,
  - $\forall j \in I$, $V^*_j(h) = V^*_j\left(h, s^*_{\iota(h)}(h)\right)$.

Note that $s^*_{\iota(h)}(h)$ is well defined because the given perfect information game is assumed to be finite and with no relevant ties, which implies that—at each step of the backward procedure—the active player has *only one*

[3] Recall that the height of h is $L(\Gamma(h)) := \max_{z : h \prec z}[\ell(z) - \ell(h)]$.

maximizing action. The procedure can be generalized to other games, e.g., some games with more than one active player at each stage, such as the finitely Repeated Prisoners' Dilemma.

In words, we start from the *last* stage of the game: h is such that all feasible actions terminate the game, that is $L(\Gamma(h)) = 1$.[4] According to the algorithm, the active player selects the payoff maximizing action. This determines a profile of payoffs for all players, denoted $(V_i^*(h))_{i\in I}$. Then we go backward to the second-to-last stage, or—more precisely—we consider histories of height 2: $L(\Gamma(h)) = 2$. The value $V_i^*\left(h, a_{\iota(h)}\right)$ has already been computed for all histories $\left(h, a_{\iota(h)}\right) \in \bar{H}$, because such histories correspond to the last stage of the game, or are terminal. According to the algorithm, the active player $\iota(h)$ chooses the feasible action $a_{\iota(h)} = s^*_{\iota(h)}(h)$ that maximizes $V^*_{\iota(h)}\left(h, a_{\iota(h)}\right)$, the reason is that the active player expects that every following player (possibly himself) would maximize his own payoff in the last stage. The algorithm continues to go backward in this way until it reaches the first stage ($h = \varnothing$).

Example 54. In the ToL4 of Figure 9.2, let T_k (L_k) denote the action of taking (leaving) k euros. It can be verified that there are no relevant ties. We obtain the strategy pair $s^* = (\mathrm{T}_1.\mathrm{T}_3, \mathrm{T}_2.\mathrm{T}_4)$ by backward induction as follows.

▶ $L(\Gamma(h)) = 1$: The only history h with $L(\Gamma(h)) = 1$ is $h = (\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3)$ and $\iota(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3) = 2$. Player 2 maximizes by taking 4 euros; thus, $s_2^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3) = \mathrm{T}_4$ and $V_1^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3) = 0$.

▶ $L(\Gamma(h)) = 2$: The only history h with $L(\Gamma(h)) = 2$ is $h = (\mathrm{L}_1, \mathrm{L}_2)$ and $\iota(\mathrm{L}_1, \mathrm{L}_2) = 1$. Player 1 maximizes by taking 3 euros, because $V_1^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{T}_3) = 3 > 0 = V_1^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3)$; thus, $s_1^*(\mathrm{L}_1, \mathrm{L}_2) = \mathrm{T}_3$ and $V_2^*(\mathrm{L}_1, \mathrm{L}_2) = 0$.

▶ $L(\Gamma(h)) = 3$: The only history h with $L(\Gamma(h)) = 3$ is $h = (\mathrm{L}_1)$ and $\iota(\mathrm{L}_1) = 2$. Player 2 maximizes by taking 2 euros, because $V_2^*(\mathrm{L}_1, \mathrm{T}_2) = 2 > 0 = V_2^*(\mathrm{L}_1, \mathrm{L}_2)$; thus, $s_2^*(\mathrm{L}_1) = \mathrm{T}_2$ and $V_1^*(\mathrm{L}_1) = 0$.

▶ $L(\Gamma(h)) = 4$: The only history h with $L(\Gamma(h)) = 4$ is the root $\varnothing$ and $\iota(\varnothing) = 1$. Player 1 maximizes by taking 1 euros, because $V_1^*(\mathrm{T}_1) = 1 > 0 = V_2^*(\mathrm{L}_1)$; thus, $s_1^*(\varnothing) = \mathrm{T}_1$. ▲

[4]Recall that which stage is the last, second-to-last and so on may be endogenous. For example, we may have a game that lasts for two stages if the first mover chooses Left and three stages if the first mover chooses Right.

By the OD principle, the BI-strategy profile is the unique SPE of Γ:

Remark 54. *For every finite PI game Γ with no relevant ties, the strategy profile s^* obtained with the backward induction algorithm is the unique SPE of Γ.*

"Case-by-Case" Backward Induction

For finite games that do not have a unique SPE computable by backward induction, Corollary 7 still offers a method to compute all the SPEs. We describe it for *two-stage games*. Consider a *two-stage* game Γ. Fix any first-stage, non-terminal (hence, preterminal) action profile $a \in \mathcal{A}(\varnothing) \setminus Z$.[5] Consider the second-stage subgame

$$G^2(a) = \left\langle I, \left(\mathcal{A}_i(a), u_i(a, \cdot)\right)_{i \in I} \right\rangle .$$

In other words, the set of feasible actions of player i in $G^2(a)$ is $\mathcal{A}_i(a)$ and his payoff function in $G^2(a)$ is

$$a' \mapsto u_{i,a}\left(a'\right) = u_i\left(a, a'\right) .$$

For example, in the BoS with an Outside Option $\mathcal{A}(\varnothing) \setminus Z = \{\text{In}\}$ and $G^2(\text{In})$ is the BoS subgame.

Stage 2. Let $NE^2(a)$ denote the set of (pure) Nash equilibria of $G^2(a)$. If $NE^2(a) = \emptyset$ for some $a \in \mathcal{A}(\varnothing) \setminus Z$, then Γ cannot have any (pure) SPE. Suppose that $NE^2(a) \neq \emptyset$ for all $a \in \mathcal{A}(\varnothing) \setminus Z$, and consider all possible selections from the second-stage equilibrium correspondence $a \mapsto NE^2(a)$, that is, all the functions

$$s^2 : (\mathcal{A}(\varnothing) \setminus Z) \to \bigcup_{a \in \mathcal{A}(\varnothing) \setminus Z} \mathcal{A}(a)$$

such that $s^2(a) \in NE^2(a)$ for each $a \in \mathcal{A}(\varnothing) \setminus Z$. Note, the set of such selections is the Cartesian product $\times_{a \in \mathcal{A}(\varnothing) \setminus Z} NE^2(a)$, which has cardinality $\prod_{a \in \mathcal{A}(\varnothing) \setminus Z} \left|NE^2(a)\right|$.

[5] With a small abuse of notation, we do not distinguish between an action profile a and the sequence of length one (a). Therefore, if $(a) \in Z$, we write $a \in Z$.

Stage 1. Each selection $s^2 \in \times_{a\in\mathcal{A}(\varnothing)\setminus Z} NE^2(a)$ is a "**case**" to which we apply a backward induction calculation. Specifically, for each selection s^2, define the *auxiliary* simultaneous-move *game*

$$G^1\left(s^2\right) = \left\langle I, \left(\mathcal{A}_i(\varnothing), u_i^1\left(\cdot, s^2\right)\right)_{i\in I}\right\rangle,$$

where

$$u_i^1\left(a, s^2\right) = \begin{cases} u_i\left(a, s^2(a)\right), & \text{if } a \in \mathcal{A}(\varnothing)\setminus Z, \\ u_i(a), & \text{if } a \in Z. \end{cases}$$

In words, $G^1\left(s^2\right)$ specifies the payoffs of each first-stage action profile a under the hypothesis (commonly believed by the players) that, if a is non-terminal, the following second-stage profile will be $s^2(a)$. Since $s^2(a)$ is a Nash equilibrium of $G^2(a)$ for each $a \in \mathcal{A}(\varnothing)\setminus Z$, every Nash equilibrium $s^1 \in \mathcal{A}(\varnothing)$ of $G^1\left(s^2\right)$ yields an SPE $s = \left(s^1, s^2\right)$: specifically, for each $i \in I$, $s_i(\varnothing) = s_i^1$ and $s_i(a) = s_i^2(a)$ for all $a \in \mathcal{A}(\varnothing)\setminus Z$.

Let $NE^1\left(s^2\right)$ denote the set of (pure) equilibria of the auxiliary $G^1\left(s^2\right)$. Then the number of SPEs of Γ is the summation over "cases" s^2 of number of equilibria in the auxiliary games $G^1\left(s^2\right)$:

$$\sum_{s^2\in\times_{a\in\mathcal{A}(\varnothing)\setminus Z} NE^2(a)} \left|NE^1\left(s^2\right)\right|.$$

Hence, there is only one SPE of Γ *if* (i) the second-stage equilibrium correspondence $a \mapsto NE^2(a)$ is actually a function s^2 (hence $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} NE^2(a)$ is a singleton) and (ii) $G^1\left(s^2\right)$ has a unique equilibrium.

Example 55. The BoS with a dissipative move of Example 39 has two strategically equivalent second-stage BoS subgames, $G^2(B)$ and $G^2(N)$. The second-stage equilibrium correspondence is $B \mapsto \{(u,l),(d,r)\}$ and $N \mapsto \{(U,L),(D,R)\}$. Therefore we obtain $2\times 2 = 4$ cases/selections, i.e., 4 auxiliary games:

$G^1((U,L),(u,l))$	N	B
	4, **1**	4, -1

$G^1((U,L),(d,r))$	N	B
	4, 1	1, **2**

$G^1((D,R),(d,r))$	N	B
	1, **4**	1, 2

$G^1((D,R),(u,l))$	N	B
	1, **4**	4, -1

By inspection of the equilibria of these auxiliary games, we obtain the set of SPEs (see the payoffs in bold):

$$SPE = \{(U.u, N.L.l), (U.d, B.L.r), (D.d, N.R.r), (D.u, N.R.l)\}.$$

We may interpret SPE $(U.d, B.L.r)$ as follows: the dissipative action B is regarded as a signal that Bob is going for his favorite equilibrium in subgame $G^2(B)$, whereas action N is interpreted as a signal that Bob "concedes" and just aims at equilibrium (U, L) of $G^2(N)$. ▲

Properties (i)-(ii) give a sufficient, but not necessary condition for uniqueness, as the following example shows.

Example 56. Ann and Bob play the following two-stage game Γ, where every first-stage action pair but (D, R) is terminal, whereas (D, R) leads to a sort of Battle-of-the-Sexes subgame G^2:

$a\backslash b$	L	R
U	2, 1	1, 2
D	1, 2	G^2

with

G^2	l	r
u	**2, 1**	0, 0
d	0, 0	**1/2, 2**

According to the subgame equilibrium of G^2 we obtain two auxiliary games:

$G^1(u,l)$	L	R
U	2, 1	1, 2
D	1, 2	2, 1

$G^1(d,r)$	L	R
U	2, 1	**1, 2**
D	1, 2	1/2, 2

Auxiliary game $G^1(u,l)$ is a kind of "Matching Pennies" and has no equilibrium; thus, there is no SPE where (u, l) is selected in subgame G^2. Auxiliary game $G^1(d, r)$ instead is dominance solvable and its unique equilibrium is (U, R). Therefore, Γ has a *unique* SPE, $(U.d, R.r)$. In this equilibrium, Ann does not deviate in the first stage because she expects to be "punished" by the subgame equilibrium of G^2—in which she gets only 1/2—if she chooses D. ▲

12.1.2 Randomized Strategies

The definition of the continuation value of strategies and actions can be extended to locally randomized strategies, i.e., behavior strategies. To simplify the probabilistic analysis we focus throughout on *finite* games.

For any behavior strategy profile $\beta = (\beta_i)_{i\in I}$ and histories $h = (a^1, ..., a^{\ell(h)})$, $h' = (h, a^{\ell(h)+1}, ..., a^{\ell(h')})$ let

$$\mathbb{P}^{\beta}(h'|h) = \prod_{t=\ell(h)+1}^{\ell(h')} \prod_{i\in I} \beta_i(a_i^t|h, ..., a^{t-1})$$

denote the probability of reaching h' from h. Then

$$V_i^{\beta}(h) = \sum_{z\in Z(h)} u_i(z)\mathbb{P}^{\beta}(z|h),$$

$$V_i^{\beta}(h, a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} V_i^{\beta}(h, (a_i, a_{-i}))\beta_{-i}(a_{-i}|h),$$

where $\beta_{-i}(a_{-i}|h) := \prod_{j\neq i} \beta_j(a_j|h)$. Obviously $V_i^{\beta}(h)$ and $V_i^{\beta}(h, a_i)$ depend only on the behavior of β in the sub-tree with root h.

Definition 69. *A* ***behavior strategy profile*** $\beta^* = (\beta_i^*)_{i\in I}$ *is an SPE if, for each* $i \in I$*,*

$$\forall h \in H,\ V_i^{\beta^*}(h) \geq \max_{\beta_i} V_i^{\beta_i, \beta^*_{-i}}(h);$$

β^* *satisfies* ***one-step optimality with correct conjectures*** *if, for each* $i \in I$*,*

$$\forall h \in H,\ \mathrm{supp}\beta_i^*(\cdot|h) \subseteq \arg\max_{a_i\in\mathcal{A}_i(h)} V_i^{\beta^*}(h, a_i).$$

The results about dynamic optimality can be extended to behavior strategies, and in particular they imply the OD Principle for randomized SPE:

OD Principle for randomized SPE: *A behavior strategy profile is an SPE if and only if it satisfies one-step optimality with correct conjectures.*

As with pure strategies, there is a "case-by-case" backward induction algorithm to compute the set of SPEs in behavior strategies, which can be easily spelled out for two-stage games and is very similar to the algorithm explained above: Just replace the stage-2 pure equilibrium selection s^2 with a stage-2 mixed equilibrium selection $\beta^2 \in \bigtimes_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a)$, where $MNE^2(a)$ denotes the set of mixed Nash equilibria of $G^2(a)$. There are two important differences with respect to pure strategy equilibria. First, since every finite simultaneous-move game has always at least one mixed Nash equilibrium, we have $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a) \neq \emptyset$ and $MNE\left(G^1\left(\beta^2\right)\right) \neq \emptyset$ for each β^2; therefore, there is a unique SPE in behavior strategies *if and only if* (i) $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a) = \{\beta^2\}$ is a singleton and (ii) $G^1\left(\beta^2\right)$ has a unique mixed equilibrium (that is, (i)-(ii) give a *necessary* and sufficient condition for uniqueness of randomized equilibria). Second, in some rare "nongeneric" games $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a)$ has the power of the continuum and there is a continuum of "cases" to deal with in the stage-1 analysis.

Example 57. The BoS subgames of Example 39 have a completely mixed equilibrium on top of the two pure equilibria, that is, $\left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D, \frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right)$ in $G^2(N)$ and $\left(\frac{4}{5}\delta_u + \frac{1}{5}\delta_d, \frac{1}{5}\delta_l + \frac{4}{5}\delta_r\right)$ in $G^2(B)$. Therefore there are $3\times 3 = 9$ second-stage selections β^2 with 9 corresponding auxiliary games $G^1\left(\beta^2\right)$. We have already computed the 4 pure selections and corresponding pure SPE's in Example 55. Without any actual computation, we can claim that there are at least 5 partially or totally *randomized* additional equilibria, that is, at least one for each non pure selection $\beta^2 = s^2$.[6] For example, the selection given by $\beta^2(N) = \left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D, \frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right)$ in $G^2(N)$ and $\beta^2(B) = (d,r)$ in $G^2(B)$ yields the auxiliary game

$G^1\left(\beta^2\right)$	N	B
	$\frac{4}{5}$, $\frac{4}{5}$	1, **2**

and the partially randomized equilibrium

$$\left(\left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D\right).d, B.\left(\frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right).r\right),$$

where $\beta_a = \left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D\right).d$ is the partially randomized strategy of Ann that would play $\left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D\right)$ upon observing N and plays d after B, and $\beta_b = B.\left(\frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right).r$ is the partially randomized strategy of Bob that plays B in the first stage, would play $\left(\frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right)$ after the counterfactual choice N, and plays r after B. ▲

[6] If there is a selection β^2 that makes Bob (the only first-mover) indifferent between B and N, then there is a corresponding continuum of randomized SPE's parameterized by the probability of burning $\beta_b(B|\varnothing)$.

The OD Principle allows a relatively simple proof of the existence of randomized equilibria in finite games:

Theorem 36. *(Kuhn) Every finite game has at least one subgame perfect equilibrium in behavior strategies.*

Proof. We provide a recursive construction of a profile β. We start from histories of height 1, i.e., those immediately preceding the last stage, and going backward until the first stage.

Basis step. For every history h with height 1, the corresponding last-stage simultaneous game has at least one mixed equilibrium. Let $\beta^*(\cdot|h) \in \times_{i\in I} \Delta(\mathcal{A}_i(h))$ be one of the mixed equilibria, and consider the corresponding expected payoff profile $(v_i^*(h))_{i\in I}$, where $v_i^*(h) := \sum_{a\in\mathcal{A}(h)} \left(\prod_{j\in I} \beta_j^*(a_j|h)\right) u_i(h,a)$.

Inductive step. Suppose that a mixed action profile $\beta^*(\cdot|h)$ and an expected payoff profile $(v_i^*(h))_{i\in I}$ have been assigned to each history with height k or less. Then one can assign a mixed action profile $\beta^*(\cdot|h)$ and a payoff profile $(v_i^*(h))_{i\in I}$ to each history h with height $k+1$: just pick a mixed equilibrium of the game $\left\langle I, (\mathcal{A}_i(h), v_i)_{i\in I}\right\rangle$ such that, for every $i \in I$ and for every $a \in \mathcal{A}(h)$, $v_i(a) = v_i^*(h,a)$ (note that (h,a) has height k or less, so $v_i^*(h,a)$ is well defined).

By construction, the behavior strategy profile β^* satisfies one-step optimality with correct conjectures. Thus, by the OD Principle, $\beta^*$is a subgame perfect equilibrium. ■

12.2 Strategic Reasoning and Backward Induction

Let us go back to games that can be uniquely solved by backward induction, such as the finite PI games with no relevant ties. We call "**outcome**" either a terminal history (path) or the corresponding payoff vector. It is sometimes claimed that the "compelling logic" of the backward induction solution rules out all the non-backward-induction outcomes as inconsistent with rationality (R) and common belief of rationality (CBR).

In Chapter 11 we (informally) considered two versions of the CBR idea: common *initial* belief in rationality (CIBR) and common *strong* belief in rationality (CSBR). We argued that the behavioral implications of RCIBR are characterized by the **initial rationalizability** solution concept, and the behavioral implications of RCSBR are characterized by the **strong rationalizability** solution concept. Therefore, we assess the aforementioned claim by comparing backward induction with initial and strong rationalizability in *finite PI games with no relevant ties.*

Let us first restrict our attention to PI games with two stages, such as the Entry Game of Figure 9.7. The backward induction outcome and strategies coincide with those obtained with both versions of rationalizability, which in this class of games are equivalent. Thus there is a strong similarity between backward induction and (initial, or strong) rationalizability in two-stage PI games.[7]

Now consider games with more than two stages. For such games, initial rationalizability is weaker than backward induction and strong rationalizability, that is, it may allow strategies that are ruled out by either backward induction, or strong rationalizability, or both.

Example 58. Go back to the ToL4 of Figure 9.2. It can be verified in this game that strong rationalizability "mimics" the steps of the backward induction procedure: it eliminates the conditionally dominated strategy $L_2.L_4$ in step 1, strategy $L_1.L_3$ in step 2, strategy $L_2.T_4$ in step 3, and strategy $L_1.T_3$ in step 4. Iterated admissibility yields the same order of eliminations, illustrating Proposition 9. Initial rationalizability coincides with strong rationalizability for the first two steps. This is true by definition for the first step (see Chapter 11). As for the second, note that $L_1.L_3$ is dominated by the reduced strategy T_1 once $L_2.L_4$ has been eliminated. But it can be verified that initial rationalizability cannot eliminate any other strategy in the third step. Indeed, no strategy is dominated in the residual strategic form without $L_2.T_4$ and $L_1.T_3$ (cf. Proposition 7). Specifically, according to the third step of initial rationalizability, player 2 may be initially certain that player 1 immediately takes 1 euro, and upon being surprised by action L_1, player 2 may stop believing that player 1 is rational and (initially) believes in 2's rationality. Thus, player 2 can leave 2 euros hoping that player 1 will leave again. ▲

This shows that backward induction is not based on (rationality and)

[7] Backward induction may be "slower" than (initial or strong) rationalizability if the first mover also moves in the second stage, for at least one first stage action, or if the first mover has an outside option that strictly dominates all other actions. But the end result is the same.

common *initial* belief in rationality. Yet, the example may engender the intuition that strong rationalizability and backward induction are essentially equivalent in this class of games. This intuition is only partially correct. On one hand, it can be shown that strong rationalizability and backward induction yield the *same outcome* (the quite complex proof is omitted).

Theorem 37. *In every finite game with perfect information and no relevant ties, all the strongly rationalizable strategy profiles yield the same terminal history as the unique subgame perfect equilibrium:* $\zeta(S^{\infty}) = \{\zeta(s^{*})\}$.

On the other hand, the following example shows that strongly rationalizable and backward induction *strategies* may be *disjoint*.

Example 59. Consider the PI game between Ann (a) and Bob (b) depicted in Figure 12.1. It can be verified that there are no relevant ties.

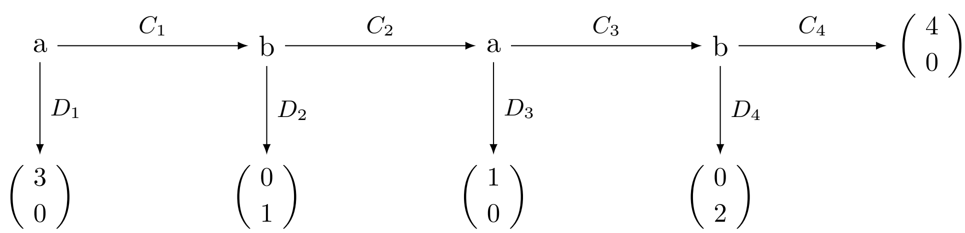


Figure 12.1: A PI game with no relevant ties.

The *subgame perfect equilibrium* is $s^{*} = (D_1.D_3, D_2.D_4)$. Strategies $C_1.D_3$ and $C_2.C_4$ are conditionally dominated. Hence,

$$S^1 = \{D_1.D_3, D_1.C_3, C_1.C_3\} \times \{C_2.D_4, D_2.D_4, D_2.C_4\}$$

(recall that strong rationalizability does not distinguish between behaviorally equivalent strategies). The *strongly rationalizable set* is $S^{\infty} = S^2 = \{D_1.D_3, D_1.C_3\} \times \{C_2.D_4\}$. Thus, $s^{*} \notin S^{\infty}$, because $s_2^{*} \neq C_2.D_4$. Indeed, strong belief in rationality requires that, upon observing C_1, Bob would believe that Ann is implementing the only justifiable (i.e., conditionally undominated) strategy that selects C_1 at the root, that is, $C_1.C_3$. The weakly sequentially optimal strategy given conjecture $C_1.C_3$ is $C_2.D_4$. ▲

The example shows that backward induction strategies may be inconsistent with rationality and strong belief in rationality, which makes Theorem 37 all the more noteworthy.

So, what is the relationship between backward induction and strategic reasoning based on some form of rationality and common belief in rationality? We argued that backward induction is not tightly related to two well understood extensions from simultaneous to sequential games of the RCBR assumptions, although it yields the same *outcome* as rationality and common *strong* belief in rationality.

We note that there are versions of RCBR for sequential games that yield a notion of "backwards rationalizability" more tightly connected to backward induction (see Perea [56] and Battigalli and De Vito [10]). The analysis of such solution concept is beyond the scope of this textbook.

12.3 Subgame Perfection with Chance Moves

The definition of multistage game can be extended to allow for the possibility that some events in the game depend on chance. Chance is modeled as a fictitious player, denoted by 0, that chooses actions with exogenously given conditional probabilities. This means that the probability that a particular chance action a_0 is selected immediately after history h is a number $\beta_0(a_0|h)$ which is part of the description of the game.

We maintain the convention that at each stage there are simultaneous moves by all players, including chance. Since the set of active players (those with more than one feasible action) may depend on time and on the previous history, this assumption does not entail any loss of generality. For example, one may model games where chance moves occur before or after the moves of real players. To minimize on notational changes, we also ascribe a utility function u_0 to the chance player 0, but we assume that it is constant. This is just another notational trick.

A **multistage game with observable actions and chance moves** is a structure

$$\Gamma = \langle (A_i, \mathcal{A}_i(\cdot), u_i)_{i=0}^n , \beta_0 \rangle .$$

The symbols A_i, $\mathcal{A}_i(\cdot)$, u_i have the same meaning as before. From the feasibility correspondences $\mathcal{A}_i(\cdot)$ $(i = 0, ..., n)$ we can derive the set of feasible histories $\bar{H}$, the set of terminal histories Z and the set of non-terminal histories $H = \bar{H} \backslash Z$ (see section 9.1). Functions u_i are real-valued and have domain Z. The only novelty is player 0, chance: for every $z \in Z$, $u_0(z) = 0$, and β_0 describes how the probabilities of chance moves depend on the previous history, that is,

$$\beta_0 = (\beta_0(\cdot|h))_{h\in H} \in \bigtimes_{h\in H} \Delta^o(\mathcal{A}_0(h)),$$

where $\Delta^o(\mathcal{A}_0(h))$ is the set of full-support probability measures on $\mathcal{A}_0(h)$.[8]

Next we define the strategic form. To avoid measure-theoretic technicalities we will henceforth assume that all the sets $\mathcal{A}_0(h)$ $(h \in H)$ are *finite*. Let I be the set of the "real" players, i.e., $I = \{1, ..., n\}$. The set S_i of strategies of $i \in I$ is defined as before. Although "strategies of chance" are well defined mathematical objects, we disregard them, as chance is not a strategic player. Thus, the set of strategy profiles is $S = \bigtimes_{i\in I} S_i$ and similarly $S_{-i} = \bigtimes_{j\in I\setminus\{i\}} S_j$. The presence of chance moves implies that more than one terminal history may be possible when the players follow a particular strategy profile s; let $Z(s)$ denote the set of such terminal histories. The probability of each $z \in Z(s)$ depends on the probability of the chance moves contained in z and is denoted $\hat{\zeta}(z|s)$ (if a player, or an external observer were uncertain about the true strategy profiles, $\hat{\zeta}(z|s)$ would represent the probability of z conditional on the players following the strategy profile s, which explains the notation). Therefore the path function when there are chance moves has the form $\hat{\zeta} : S \to \Delta(Z)$. The formal definition of $Z(s)$ and $\hat{\zeta}(\cdot|s)$ is as follows. Let $\mathbf{a}_i^t(z)$ denote the action of player i at stage t in history z, let $\mathbf{h}^{t-1}(z)$ denote the prefix (initial sub-history) of z of length t, and recall that $\ell(z)$ is the length of z. Then

$$\forall s \in S,\ Z(s) = \{z \in Z : \forall t \in \{1, ..., \ell(z)\}, \forall i \in I,\ \mathbf{a}_i^t(z) = s_i(\mathbf{h}^{t-1}(z))\},$$

$$\forall z \in Z,\ \hat{\zeta}(z|s) = \begin{cases} \prod_{t=1}^{\ell(z)} \beta_0(\mathbf{a}_0^t(z)|\mathbf{h}^{t-1}(z)), & \text{if } z \in Z(s), \\ 0, & \text{if } z \notin Z(s). \end{cases}$$

In words, $\hat{\zeta}(z|s)$ is the product of the probabilities of the actions taken by the chance player in history z.

[8]Recall that $\Delta^o(X) = \{\mu \in \Delta(X) : \text{supp}\mu = X\}$, where the superscript o stands for (relatively) "open" in the hyperplane of vectors whose elements sum to 1.

A **Nash equilibrium** of Γ is a strategy profile s^* such that for every $i \in I$ and for every $s_i \in S_i$,

$$\sum_z \hat{\zeta}(z|s^*)u_i(z) \geq \sum_z \hat{\zeta}(z|s_i, s^*_{-i})u_i(z).$$

In other words, a Nash equilibrium of Γ is an equilibrium of the static game where players simultaneously choose strategies and have payoff functions $U_i(s) = \sum_z \hat{\zeta}(z|s)u_i(z)$ $(i \in I,\, s \in S)$.

In order to define subgame perfect equilibria, we first define the conditional outcome function $\hat{\zeta}(\cdot|h, s)$: for each $h \in H$, the set of terminal histories that can be reached when s is followed starting from h (where h may be *in*consistent with s) is

$$Z(h, s) = \{z \in Z : h \prec z, \forall t \in \{\ell(h){+}1, ..., \ell(z)\}, \forall i \in I, \mathbf{a}_i^t(z) = s_i(\mathbf{h}^{t-1}(z))\};$$

then

$$\forall z \in Z,\, \hat{\zeta}(z|h, s) = \begin{cases} \prod_{t=\ell(h)+1}^{\ell(z)} \beta_0(\mathbf{a}_0^t(z)|\mathbf{h}^{t-1}(z)), & \text{if } z \in Z(h, s), \\ 0, & \text{if } z \notin Z(h, s). \end{cases}$$

A strategy profile s^* is a **subgame perfect equilibrium** if for all $h \in H$, $i \in I$, and $s_i \in S_i$,

$$\sum_z \hat{\zeta}(z|h, s^*)u_i(z) \geq \sum_z \hat{\zeta}(z|h, s_i, s^*_{-i})u_i(z).$$

Let $\hat{\zeta}(\cdot|h, a_i, s) \in \Delta(Z)$ denote the probability measure on Z that results if s is followed starting from h, except that player i chooses $a_i \in \mathcal{A}_i(h)$ at h. A strategy profile s satisfies **one-step optimality with correct conjectures** if for all $h \in H$, $i \in I$, and $a_i \in \mathcal{A}_i(h)$,

$$\sum_z \hat{\zeta}(z|h, s)u_i(z) \geq \sum_z \hat{\zeta}(z|h, a_i, s)u_i(z).$$

It can be shown that in every game with finite horizon or with continuity at infinity (hence in every game with discounting) a strategy profile is a subgame perfect equilibrium if and only if it satisfies one-step optimality with correct conjectures.

13

Repeated Games

A subset of multistage games that has received particular attention in the literature is the one of *repeated games*. Loosely speaking, a repeated game is a strategic environment in which the same players play multiple times the same static game—the **stage game**. More generally, we could think of the repetition of the same static or sequential one-period game; but our analysis, like most of the literature on repeated games, will be restricted to the repetition of the same static (i.e., simultaneous-move) games, so that stages will coincide with periods. See our comments in Section 9.2.1.

Many real-life situations can be modelled as repeated games. Think of a couple which has to decide every weekend what to do. Or consider the daily relationship between two colleagues, who must independently decide whether to exert effort on a common project or to shirk. Or again, suppose that firms are involved in a cartel and must decide whether to set their individual price as suggested by the cartel or to undercut such level in order to reap a larger market share.

In addition to their practical relevance, repeated games are interesting also from a theoretical point of view. Indeed, given the dynamic nature of the strategic interaction, players can adjust their future behavior in order to incentive other players to behave in a certain way. Although this phenomenon may arise in any multistage game, the specific structure of repeated games (namely, the fact that the same stage game is repeated over and over again) makes them the ideal environment to study these reward and punishment strategies.

Before we delve into the formal analysis, it is useful to introduce some taxonomy. First of all, the class of repeated games can be divided in *finitely repeated games* and *infinitely repeated games* depending on whether the stage game is repeated a finite number of times or infinitely many times. Furthermore, repeated games can be classified according to the information feedback that players receive about the behavior of their opponents at previous stages. *Repeated games with perfect monitoring* are games in which each player perfectly observes (and recalls) the actions of all players in previous rounds. Thus, this class of games is a special case of multistage games with observable actions. *Repeated games with imperfect public monitoring* are games in which players do not directly observe what their opponents chose in the past. Instead, in each round, they observe (and recall) a common public stochastic signal, whose realization is affected by the action profile chosen by players in the previous round. *Repeated games with imperfect private monitoring* are games in which players do not directly observe what their opponents chose in the past, nor do they observe a common public signal. Instead, each of them observes (and recalls) a private stochastic signal, whose realization depends on the action taken by players in the previous round.

One of the main questions addressed by the literature on repeated games is the following: "To what extent can the repeated strategic interaction enlarge the set of equilibrium payoffs with respect to those of the stage game?" An answer to this question is provided by so called "folk theorems," namely results with the following flavor.

Informal Folk Theorem. Consider an infinitely repeated game. If players are very patient and they can statistically detect the behavior of their opponents, then any payoff profile compatible with individual optimality (namely, each payoff profile in which players get at least as much as the worst possible payoff consistent with utility-maximizing behavior) is achievable in equilibrium.

In this chapter we follow a different approach and we stress the strategic nature of the repeated interaction. As such, we highlight how dynamic incentives come at play in repeated games and may enable players to punish or reward their opponents. As a by-product of this approach, we introduce the tools needed to prove folk theorems.

Specifically, we will only analyze repeated games with *perfect monitoring*. First, we will introduce some basic notation concerning the one-period game and the associated repeated games. Then, we will prove some simple results relating the Nash equilibria of the one-period game to the subgame perfect equilibria of the corresponding repeated game for a fixed "degree of patience" (discount factor). Finally, we will provide a characterization of the set of (pure) SPE payoff profiles in infinitely repeated games, again, for a fixed "degree of patience." Since it is implausible to assume that agents are arbitrarily patient, the latter kind of results (rather than the "folk theorems") are those most used in the applications of the theory of repeated games.

13.1 Repeated Games with Perfect Monitoring

For any static *compact-continuous* game $G = \langle I, (A_i, v_i)_{i\in I}\rangle$, we let $\Gamma^{\delta,T}(G)$ denote the T-stage game with observable actions obtained by repeating G for $T \in \mathbb{N}\cup\{\infty\}$ times and computing payoffs $u_i : Z \to \mathbb{R}$ (where $Z = A^T$) as discounted time averages with common discount factor $\delta \in (0,1)$;[1] that is, for every terminal history $(a^t)_{t=1}^T \in Z$ and for every player i,

$$u_i(a^1, a^2, ...) = \sum_{t=1}^{T} \delta^{t-1} v_i(a^t). \tag{13.1.1}$$

Note that u_i is well defined even if $T = \infty$, because G is compact-continuous, and so v_i is bounded; therefore, the series of discounted payoffs has a finite sum. In the analysis of finitely repeated games ($T \in \mathbb{N}$) one can also allow for $\delta = 1$, i.e., no discounting. Also note that the most important insights of the analysis can be extended to the more general case where G has *bounded* payoff functions, e.g., when G is an oligopoly game with price setting, homogeneous products, and competition *à la* Bertrand.

In the analysis of infinitely repeated games, it may be convenient to use an equivalent representation of the payoff associated with an infinite history:

$$\bar{u}_i(a^1, a^2, ...) = (1-\delta)\sum_{t=1}^{T} \delta^{t-1} v_i(a^t), \tag{13.1.2}$$

which is a **discounted time average** of the stream of payoffs $\left(v_i(a^t)\right)_{t=1}^{\infty}$. When the stream of payoffs is constant, e.g., $v_i(a^t) = v_i(\bar{a})$ for each t, then

[1] The analysis can be easily extended to allow for heterogeneous discount factors.

the average (13.1.2) is precisely this constant:

$$\bar{u}_i(\bar{a}, \bar{a}, ...) = (1-\delta) \sum_{t=1}^{T} \delta^{t-1} v_i(\bar{a}) = v_i(\bar{a}).$$

In this strategic environment, every history h of length $t \in \mathbb{N}$ is an element of A^t. By convention, $A^0 = \{\varnothing\}$ is the set containing the empty sequence of action profiles.

Example 60. Consider the Prisoners' Dilemma in which the set of players is given by $I = \{1, 2\}$ and the set of actions available to each player i is given by "*C*ooperation" and "*D*efection," that is, $A_i = \{C, D\}$. Utilities are represented in the following table.

	C	D
C	$2, 2$	$0, 3$
D	$3, 0$	$1, 1$

The repeated game with perfect monitoring $\Gamma^{\delta,2}(G)$ is the multistage game obtained by repeating the Prisoners' Dilemma twice. In this case, the set of non-terminal histories is

$$H = \{\varnothing\} \cup A$$

and the payoff of player $i \in I$ at terminal history $\left(a^1, a^2\right)$ is given by

$$u_i\left(a^1, a^2\right) = v_i\left(a^1\right) + \delta v_i\left(a^2\right).$$

▲

Example 61. There are n firms which produce and sell a homogenous good: I is the set of firms, $q_i \in [0, \bar{q}_i]$ is the output of firm i and $\bar{q}_i$ is its capacity. The market (inverse) demand function $P : [0, \sum_i \bar{q}_i] \to \mathbb{R}_+$ is given by

$$P\left(\sum_{j\in I} q_j\right) = \max\{0, \bar{p} - \beta \sum_{j \in I} q_j\}.$$

Firms compete *à la* Cournot choosing the quantity to produce in order to maximize their total profits and keeping into account that the production of the good entails a unit cost equal to c. Thus, each firm i maximizes $v_i(q_1, ...q_n) = q_i P\left(q_i + \sum_{j \neq i} q_j\right) - cq_i$. Game $\Gamma^{\delta,\infty}(G)$ is the infinitely repeated game in which the set of histories is given by

$$\bar{H} = A^{\leq \mathbb{N}_0} = \{\varnothing\} \cup \left(\bigcup_{t \in \mathbb{N}} \left(\bigtimes_{i \in I} [0, \bar{q}_i]\right)^t\right) \cup \left(\bigtimes_{i \in I} [0, \bar{q}_i]\right)^{\mathbb{N}},$$

and the payoff function of firm i is given by

$$\bar{u}_i(\left(q_1^1, ..., q_n^1\right), \left(q_1^2, ..., q_n^2\right), ...) = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \left(q_i^t \max\{0, \bar{p} - \beta \sum_{j \in I} q_j^t\} - cq_i^t\right).$$

▲

In the analysis of repeated games with perfect monitoring, we use subgame perfect equilibrium (SPE) as a solution concept. Notice that since G is compact-continuous and $\delta < 1$, the repeated game is continuous at infinity (see Section 10.5) and, consequently, we can use the OD Principle to characterize SPEs.

13.1.1 Automaton Representation of Strategy Profiles

The equilibrium analysis of repeated games often requires the construction of strategy profiles that attain a certain payoff. Since the set of histories in a repeated game may include many different elements, this task can lead to significant notational complexity. To simplify the analysis, it is often useful to represent strategy profiles by collecting histories in equivalence classes with respect to the strategic behavior of players. This approach underlies the so-called *automaton* representation of a strategy profile.

Definition 70. *An automaton is a profile* $(\Psi, \psi_0, \gamma, \varphi)$ *where:*

- Ψ *is a set of states;*
- $\psi_0 \in \Psi$ *is the initial state;*
- $\gamma : \Psi \to A$ *is a behavioral rule that specifies an action profile for each state;*

- *$\varphi : \Psi \times A \to \Psi$ is a transition rule that specifies how the automaton moves across states based on the current state and by the action profile chosen by players.*

The automaton representation of a strategy profile in a repeated game with perfect monitoring requires to specify a set of states. Each state corresponds to a set of histories in the repeated game at which the behavior of players (under the strategy profile) is constant. Put differently, the behavior of players under a certain strategy profile is measurable with respect to the partition of non-terminal histories represented by the states.

Every strategy profile $s \in S$ can be represented as an automaton. To see this, let $\Psi = H$, $\psi_0 = \varnothing$ and $\gamma(h) = (s_i(h))_{i \in I}$ for every $h \in H$. Finally, let $\varphi(h, a) = (h, a) \in \Psi$. In general, the automaton representation we just provided is not the only possible one; indeed it is often possible to reduce the set of states by grouping histories together—namely, by coarsening the equivalence classes.

Conversely, given an automaton $(\Psi, \psi_0, \gamma, \varphi)$, we can derive a profile of strategies as follows. For every $\ell \geq 1$ and for every $h^\ell \in A^\ell$, define iteratively function $\tilde{\varphi} : \Psi \times H \backslash \{\varnothing\} \to \Psi$ as follows. For $\ell = 1$, let $\tilde{\varphi}\left(\psi_0, h^1\right) = \varphi\left(\psi_0, a^1\right)$. Then, suppose that the map $\tilde{\varphi}\left(\psi_0, h^{\ell-1}\right)$ has been defined for every $h^{\ell-1} \in A^{\ell-1}$; hence let $\tilde{\varphi}\left(\psi_0, h^\ell\right) = \varphi\left(\tilde{\varphi}\left(\psi_0, h^{\ell-1}\right), a^\ell\right)$. So, for every player $i \in I$, we can specify the following strategy: $s_i(\varnothing) = \gamma(\psi_0)$ and, for every $\ell \geq 1$ and for every $h^\ell \in A^\ell$, let $s_i\left(h^\ell\right) = \gamma\left(\tilde{\varphi}\left(\psi_0, h^\ell\right)\right)$.

13.2 Simple Results about the Multiplicity of Equilibria

We prove three simple results relating the Nash equilibria of G to the subgame perfect equilibria of $\Gamma^{\delta,T}(G)$. We first apply the OD principle to show a rather obvious, but important result: playing repeatedly a Nash equilibrium of G is subgame perfect. Actually, playing any fixed sequence of Nash equilibria of G is subgame perfect (for example, alternating between the equilibrium preferred by Rowena and the equilibrium preferred by Colin in the repeated Battle of the Sexes is subgame perfect).

Theorem 38. *Let $\bar{z} = (\bar{a}^1, \bar{a}^2, ...) \in A^T$ be a sequence of Nash equilibria of G; then the strategy profile $\overline{s}$ that prescribes to play $\bar{a}^t$ in each stage*

t (irrespective of previous actions) is a subgame perfect equilibrium of $\Gamma^{\delta,T}(G)$.

Proof. Let us show that there are no incentives to make one-shot deviations from $\overline{s}$. By the OD principle, this implies the thesis.

Fix an arbitrary stage t and a player i. The continuation payoff of i if he chooses a_i in stage t and sticks to $\overline{s}_i$ afterward (assuming that everybody else sticks to $\overline{s}_{-i}$) is

$$v_i(a_i, \bar{a}^t_{-i}) + \sum_{\tau=t+1}^{T} \delta^{\tau-t} v_i(\bar{a}^\tau).$$

Note that the second term in this expression *does not depend on* a_i, because the equilibria of G played in the future may depend on calendar time, but do not depend on past actions. Since $\bar{a}^t$ is an equilibrium of G, for every a_i, $v_i(a_i, \bar{a}^t_{-i}) \leq v_i(\bar{a}^t_i, \bar{a}^t_{-i})$. Therefore i has no incentive to deviate from $\bar{a}^t_i$ in stage t. ■

By Theorem 38, any set of assumptions that guarantees the existence of a pure equilibrium of G also guarantees the existence of a subgame perfect equilibrium (in pure strategies) of $\Gamma^{\delta,T}(G)$.

The next result identifies situations in which a subgame perfect equilibrium *must* be the repetition of the Nash equilibrium of G.

Theorem 39. *Suppose that G is* finitely *repeated ($T < \infty$) and has a* unique *Nash equilibrium a°, then $\Gamma^{\delta,T}(G)$ has a* unique *subgame perfect equilibrium s° that prescribes to play a° always (for every $h \in H$ and for every $i \in I$, ${s_i}^\circ(h) = {a_i}^\circ$).*

For example, the finitely repeated Prisoners' Dilemma has a unique subgame perfect equilibrium: always defect.[2]

Proof. By Theorem 38, s° is a subgame perfect equilibrium. Let s be any subgame perfect equilibrium. We will show that $s = s^\circ$, that is,

[2]The finitely repeated Prisoners' Dilemma is a very simple game. Since the Prisoners' Dilemma has a dominant action equilibrium, the SPE can be obtained by backward induction. Another result that holds for the finitely repeated Prisoners' Dilemma is that, although it has many Nash equilibrium strategy profiles, they are all equivalent, and hence they all induce the permanent defection path.

for every $t \in \{0, ..., T-1\}$, for every $h \in A^t$, $s(h) = a^\circ$ (by convention, $A^0 = \{\varnothing\}$ is the set containing the empty sequence of action profiles). The result will be proved by induction on the number k of stages that still have to be played.

Basis step. Suppose that a history $h \in A^{T-1}$ has just occurred, i.e., the game reached the last stage. Since s is subgame perfect it must prescribe an equilibrium of G at the last stage. But there is only one equilibrium of G, a°; thus $s(h) = a^\circ$.

Inductive step. Suppose, by way of induction, that s prescribes a° in the last k periods ($k \geq 1$), that is, for every period $t \in \{T-k+1, ..., T\}$ and every history h' of length $t-1$ (that is, $h' \in A^{t-1}$), $s(h') = a^\circ$. We show that s prescribes a° in period $T-k$ for every history h of length $T-k-1$ ($h \in A^{T-k-1}$), $s(h) = a^\circ$. Since s is subgame perfect, it must be immune to deviations, in particular to one-shot deviations, in period $T-k$. Therefore, by the inductive hypothesis, for every $h \in A^{T-k-1}$, for every $i \in I$ and for every $a_i \in A_i$,

$$v_i(s(h)) + \sum_{\tau=1}^{k} \delta^\tau v_i(a^\circ) \geq v_i(a_i, s_{-i}(h)) + \sum_{\tau=1}^{k} \delta^\tau v_i(a^\circ).$$

Note that the summation term is the same on both sides of the inequality because the inductive hypothesis implies that if s is followed in the last k periods, future payoffs are independent of current actions. Therefore the action profile $s(h)$ is such that for every $i \in I$, for every $a_i \in A_i$, $v_i(s(h)) \geq v_i(a_i, s_{-i}(h))$, implying that $s(h)$ is a Nash equilibrium of G. But the only Nash equilibrium of G is a°; thus, $s(h) = a^\circ$. ■

Comments. First, there is no need to invoke the OD principle to prove the result above. The principle states that immunity to one-shot deviations implies immunity to all deviations, which holds under some (fairly general) assumptions. The proof above used the converse, immunity to all deviations implies immunity to one-shot deviations, which is trivially true by definition. Second, when a° is a profile of strictly *dominant* actions in G (or, more generally, when a° is the unique rationalizable profile of G), then a similar uniqueness result holds for strong rationalizability: by a result similar to Theorem 37, every profile of strongly rationalizable strategies of $\Gamma^{\delta,T}(G)$ induces path $z^\circ = (a^\circ, ..., a^\circ)$, or $\zeta(S^\infty) = \{z^\circ\}$.

If one of the hypotheses of the previous theorem is removed, i.e., if either $T = \infty$ or G has multiple equilibria, then the thesis may fail. We illustrate this with an example and a general result.

Suppose that $T < \infty$ but G has multiple equilibria. Then the Basis Step of the proof above does not work anymore, because *the last stage equilibrium may depend on past actions.* Consider, for example, the following variation of the Prisoners' Dilemma game:

$1\backslash 2$	C	D	P
C	$4,4$	$0,5$	$-1,0$
D	$5,0$	$2,2$	$-1,0$
P	$0,-1$	$0,-1$	$0,0$

Figure 13.1: A modified Prisoners' Dilemma.

It can be checked that, for δ sufficiently large, the following strategies support the $(4,4)$ outcome in all periods but the last one:

- start playing C;
- if you are in stage $t < T$ and no deviation from (C,C) has occurred, play C;
- if you are in stage $t = T$ and no deviation from (C,C) has occurred, play D;
- if a deviation from (C,C) has occurred, play P in every stage until the last one.

The key observation is that in the second-to-last stage players do not defect because the immediate gain from defection is more than compensated by the future loss caused by a switch to a worse last-stage equilibrium.

Next suppose that $T = \infty$ and δ is "sufficiently high," then the repeated game may have many other equilibria beside the repetition of the stage-game equilibria. We show this for the case in which a stage-game equilibrium a° is strictly Pareto dominated by some other profile a^*.

Theorem 40. (Nash reversion, trigger strategy equilibria) *Let* $G = \langle I, (A_i, v_i)_{i\in I}\rangle$ *be a static game such that there are an equilibrium* a° *and an action profile* a^* *that strictly Pareto-dominates* a°*. Consider the repeated game* $\Gamma^{\delta,\infty}(G)$ *and the strategy profile* $s^* = (s_i^*)_{i\in I}$ *defined as follows:*

$$s_i^*(h) = \begin{cases} a_i^*, & \text{if } h = \varnothing \text{ or } h = (a^*, ..., a^*), \\ a_i{}^\circ, & \text{otherwise.} \end{cases}$$

Then s^* *is a subgame perfect equilibrium if and only if*

$$v_i(a^*) \geq (1-\delta) \sup_{a_i\in A_i} v_i(a_i, a_{-i}^*) + \delta v_i(a^\circ), \tag{IC}$$

that is,

$$\delta \geq \frac{\sup_{a_i\in A_i} v_i(a_i, a_{-i}^*) - v_i(a^*)}{\sup_{a_i\in A_i} v_i(a_i, a_{-i}^*) - v_i(a^\circ)}.$$

Note that by definition $\sup_{a_i\in A_i} v_i(a_i, a_{-i}^*) \geq v_i(a^*)$ and by assumption $v_i(a^*) > v_i(a^\circ)$. Therefore

$$0 \leq \frac{\sup_{a_i\in A_i} v_i(a_i, a_{-i}^*) - v_i(a^*)}{\sup_{a_i\in A_i} v_i(a_i, a_{-i}^*) - v_i(a^\circ)} < 1.$$

Strategies such as those defined in the theorem are called **trigger strategies** because a deviation from the equilibrium path "triggers" some sort of punishment.

Proof of Theorem 40. We show that there are no incentives to make one-shot deviations if and only if condition (IC) holds. By the OD principle, this implies the result.

There are two types of histories: those for which a deviation from a^* has not yet occurred and the others. For the latter histories s^* prescribes to play a° forever after. Since a° is an equilibrium of G, Theorem 38 implies that there cannot be incentives to make one-shot deviations. Suppose now that no deviation from a^* has yet occurred. Then s^* prescribes to play a^*. Consider the incentives of player i. His continuation payoff if he sticks to s_i^* is $v_i(a_i^*)$, because i expects to get $v_i(a_i^*)$ forever (recall that the continuation payoff is a discounted time average). His continuation payoff if he deviates and plays $a_i \neq a_i^*$ is $(1-\delta)v_i(a_i, a_{-i}^*) + \delta v_i(a^\circ)$, because i expects that his opponents stick to a_{-i}^* in the current stage, and that his deviation triggers a switch to a° from the following stage (as prescribed by s^*). Therefore i does not have an incentive to make a one-shot deviation if and only if condition (IC) holds. ∎

As an application of Theorem 40 consider, for example, an infinitely repeated Cournot oligopoly game: a is an output profile, $v_i(a)$ is the one-period profit of firm i, a° is the Cournot equilibrium profile, a^* is the joint-profit maximizing profile, i.e., the profile that would be chosen if the firms were allowed to merge and if the merged firm were not regulated. The trigger-strategy profile s^* may be interpreted as a collusive agreement to keep the price at the monopoly level. If δ is above the threshold the collusive agreement is self-enforcing. "Self-enforcing" in this context does not only mean that there are no incentive to deviate from the equilibrium path (all Nash equilibria satisfy this property), but also that the "punishment" or "war" that is supposed to take place after a deviation is itself self-enforcing, hence "credible," and this is indeed the case because playing the Cournot profile in each stage is an equilibrium (Theorem 38).

13.3 Characterization of the Equilibrium Set

In the modified Prisoners' Dilemma of Figure 13.1, players follow a quite intuitive punishment (or rewarding) code: if both players coordinate on (C, C) in the first $T-1$ stages, such behavior is rewarded by having players coordinating on the Pareto superior Nash equilibrium. Instead, if one of the players (or both) plays something other than C in some stage $t \leq T-1$, then players coordinate on the Pareto inferior Nash equilibrium (P, P) in every subsequent stage. Notice that the punishment code specified in the example is self-enforcing as it is based on players playing in the last stage a Nash equilibrium of the stage game. However, more generally, rewards and punishments can be supported as, in turn, players adjust their strategic behavior in subsequent rounds. We now provide a formal discussion of this issue in the context of infinitely repeated games.

We start making a simple observation that turns out to be extremely powerful: an infinitely repeated game exhibits a recursive structure. Indeed, once we rescale payoffs dividing by δ^ℓ, the subgame that starts after any history $h^\ell \in A^\ell$ is identical to the original repeated game. Therefore, a strategy profile of the whole game induces an SPE in the subgame with root h^ℓ if and only if it is an SPE in the original game. As a result, the set of SPE payoffs in the subgame with root h^ℓ is identical to the one of the original game.

Definition 71. *Let $X \subseteq \mathbb{R}^I$. Action profile $a^* = (a_i^*)_{i \in I}$ is* ***enforceable*** *on X, if there exists a function $\eta : A \to X$ such that for every $i \in I$ and every $a_i \in A_i$,*

$$(1-\delta) v_i(a^*) + \delta \eta_i(a^*) \geq (1-\delta) v_i\left(a_i, a_{-i}^*\right) + \delta \eta_i\left(a_i, a_{-i}^*\right),$$

where $\eta_i(\cdot)$ is the projection of function $\eta(\cdot)$ on the i-th dimension. With this, we say that function $\eta(\cdot)$ enforces a^* (on X) and we refer to $\eta(\cdot)$ as to the enforcing function.

In words, action profile a^* is enforceable on X if we can define "continuation payoff" within subset X in a way that makes action a_i^* incentive compatible for every player i. Furthermore, we say that a payoff profile is pure-action decomposable on X if the payoff of each player can be decomposed in a (weighted) sum of the payoff associated with a profile a^* enforceable on X and of the associated enforcing function.

Definition 72. *Payoff profile $v^* \in \mathbb{R}^I$ is* ***pure-action decomposable*** *on $X \subseteq \mathbb{R}^I$ if there exist a profile a^* enforceable on X such that for every $i \in I$,*

$$v_i^* = (1-\delta) v_i(a^*) + \delta \eta_i(a^*),$$

where $\eta : A \to X$ enforces a^* on X.

If payoff profile v^* is pure-action decomposable, then it can be regarded as the equilibrium payoff profile of the static game $\hat{G} = \left\langle I, (A_i, \hat{v}_i)_{i \in I} \right\rangle$ obtained by modifying the payoff function of $G = \left\langle I, (A_i, v_i)_{i \in I} \right\rangle$ as follows:

$$\forall i,\ \forall a_i,\ \forall a_{-i},\ \hat{v}_i(a_i, a_{-i}) = (1-\delta) v_i(a_i, a_{-i}) + \delta \eta_i(a_i, a_{-i}).$$

The following definition introduces the important concept of self-generating set.

Definition 73. *A set $X \subseteq \mathbb{R}^I$ is* ***pure-action self-generating*** *if every payoff profile in X is pure-action decomposable on X.*

Thus, a set of payoff profiles X is pure-action self-generating if each payoff profile in such a set can be achieved by playing an action profile in the stage game and specifying continuation payoffs that (i) enforce the action profile, and (ii) belong to the set X itself. Note that this last requirement exploits the recursive structure of the repeated game with perfect monitoring. Indeed, since continuation payoffs belong to X, they can be pure-action decomposed as well with continuation values in X. In turn, these second-degree continuation payoffs can be further decomposed and so on.

Proposition 11. *Let $\Gamma^{\delta,\infty}(G)$ be a repeated game with perfect monitoring. Then the set W^* of pure SPE payoffs is the maximal (under set inclusion) pure-action self-generating set.*

Proof. The set of pure SPE payoff profiles is self-generating. Indeed, by definition, each payoff profile in such a set can be decomposed in a payoff profile $v(a^*)$ of the stage game and in a payoff profile $\eta(a^*)$ that emerges in the subgame that starts after a^*. By definition of subgame perfection, this payoff profile is also a pure SPE payoff profile.

Thus, we need to show that any pure-action self-generating set is a subset of W^*. Consider an arbitrary pure-action self-generating set X, and pick any $\hat{v} \in X$. We show that $\hat{v}$ is a pure SPE payoff profile, that is, $\hat{v} \in W^*$. Since $\hat{v}$ is pure-action decomposable on X, we can find an action profile $a^{\hat{v}}$ and a function $\eta^{\hat{v}} : A \to X$ such that $\hat{v} = (1-\delta)\,v\left(a^{\hat{v}}\right)+\delta\eta^{\hat{v}}\left(a^{\hat{v}}\right)$. Since $\eta^{\hat{v}}(a) \in X$ for every $a \in A$, we can further conclude that there exists an action profile $a^{\eta^{\hat{v}}(a)} \in A$ and a continuation payoff function $\eta^{\eta^{\hat{v}}(a)} : A \to X$ such that

$$\eta^{\hat{v}}(a) = (1-\delta)\,v\left(a^{\eta^{\hat{v}}(a)}\right) + \delta\eta^{\eta^{\hat{v}}(a)}\left(a^{\eta^{\hat{v}}(a)}\right).$$

Iterating this procedure, we can define for each $t \in \mathbb{N}$ the following maps $\bar{\eta}^{\hat{v},t} : A^t \to X$. For every history $h^t = \left(a^1, a^2, ..., a^t\right) = \left(h^{t-1}, a^t\right) \in A^t$, define $\bar{\eta}^{\hat{v},1}\left(a^1\right) = \eta^{\hat{v}}\left(a^1\right)$. Next suppose that the map $\bar{\eta}^{\hat{v},t-1} : A^{t-1} \to X$ has been defined for $t \geq 2$; then define $\bar{\eta}^{\hat{v},t}\left(h^t\right) = \eta^{\bar{\eta}^{\hat{v},t-1}\left(h^{t-1}\right)}\left(a^t\right)$. Since the set X is self-generating, we have $\bar{\eta}^{\hat{v},t}\left(h^t\right) \in X$. Hence, for every $i \in I$, we can construct a strategy $s_i^{\hat{v}}$ for player i as follows: define $s_i^{\hat{v}}(\varnothing) = a_i^{\hat{v}}$ and $s_i^{\hat{v}}\left(h^t\right) = a_i^{\bar{\eta}^{\hat{v},t}\left(h^t\right)}$ for every $h^t \in A^t$ and $t \in \mathbb{N}$. By construction, strategy profile $\left(s_1^{\hat{v}}, ..., s_n^{\hat{v}}\right)$ induces payoff $\hat{v}$. Furthermore, after each history $h \in H$, the action profile specified by the strategy profile $\left(s_i^{\hat{v}}\right)_{i\in I}$ is enforceable on X. Thus the behavior of each player is incentive compatible given the enforcing function. We conclude that $\left(s_i^{\hat{v}}\right)_{i\in I}$ is an SPE profile. Thus, $\hat{v}$ is a pure SPE payoff profile. ■

Proposition 11, due to Abreu, Pearce and Stacchetti [2], provides a characterization of the set of pure SPE payoff profiles in terms of self-generating sets. Although such characterization sheds light on some key properties of SPE payoffs, it does not provide any guidance on how to build a strategy profile that induces a certain equilibrium-payoff profile.

An important result due to Abreu [1] helps in this respect. Loosely speaking, this result says that to check whether a certain payoff profile is compatible with equilibrium, we can restrict attention to strategy profiles that specify an "optimal penal code" for each player. This penal code is implemented whenever a player unilaterally deviates from the equilibrium behavior *independently of the specific history at which this deviation occurs* (joint deviations are instead ignored). Put differently, the scheme used to incentivize player i to abide equilibrium behavior is specified independently of the circumstances in which such deviation occurred. In particular, the same punishment is implemented regardless on whether the player's deviation takes place on or off the equilibrium path.

We start by defining simple strategy profiles.

Definition 74. *Fix a path* $\mathbf{a}(0) \in A^{\mathbb{N}}$ *and a profile of paths* $\mathbf{a}_I = (\mathbf{a}(i))_{i\in I} \in \left(A^{\mathbb{N}}\right)^I$. *The* ***simple strategy profile*** $s^{\mathbf{a}(0),\mathbf{a}_I}$ *is the strategy profile represented by the following automaton* $(\Psi, \psi_0, \gamma, \varphi)$*:*

$$
\begin{aligned}
\Psi &= (\{0\} \cup I) \times \mathbb{N}_0, \\
\psi_0 &= (0,0), \\
\gamma(i,t) &= a^t(i), \\
\varphi((j,t),a) &= \begin{cases} (i,0), & \text{if } a_i \neq a_i^t(j) \text{ and } a_{-i} = a_{-i}^t(j), \\ (j,t+1), & \text{otherwise,} \end{cases}
\end{aligned}
$$

where $a^t(i)$ *denotes the* t*-th element of* $\mathbf{a}(i)$*.*

A simple strategy profile can be understood as follows. Players start playing the action profiles according to $\mathbf{a}(0)$ and keep doing so as long as no unilateral deviation is observed. In particular, joint deviations by more than one player are ignored. If player i unilaterally deviates in some round t, namely if $a_i^t \neq a_i^t(0)$ whereas $a_{-i}^t = a_{-i}^t(0)$, then players "punish" player i by beginning to play the profiles in $\mathbf{a}(i)$, the "punishment protocol" for player i. Once players start playing a punishment protocol against player i, namely once they are playing according to $\mathbf{a}(i)$, they keep doing so unless a unilateral deviation from $\mathbf{a}(i)$ is observed (once more, joint deviations are ignored). If for some t', player $j \neq i$ unilaterally deviates from the protocol $\mathbf{a}(i)$—namely, if $a_j^{t'} \neq a_j^{t'}(i)$ whereas $a_{-j}^{t'} = a_{-j}^{t'}(i)$—then players exit punishment protocol $\mathbf{a}(i)$ and start playing the action profiles in protocol $\mathbf{a}(j)$. If the deviator is agent i, then the punishment protocol $\mathbf{a}(i)$ starts again from the beginning.

Profile $\mathbf{a}_I$ can thus be regarded as a *penal code*: it specifies, for each player $i \in I$, a punishment protocol $\mathbf{a}(i)$ to be implemented each time player i unilaterally deviates. In particular, the punishment protocol against a player is independent of the time at which the deviation occurs and of its nature.

Two important examples of simple strategy profiles are the *Nash-reversion trigger strategy* and the *Nash-reversion grim trigger strategy.*

Definition 75. *Fix an infinitely repeated game $\Gamma^{\delta,\infty}(G)$ and a path $\mathbf{a}(0) \in A^{\mathbb{N}}$. For every $i \in I$, let $\mathbf{a}^*(i)$ denote the constant path playing $a^*(i)$ in every period, where $a^*(i)$ is a Nash equilibrium of G, and consider the profile of paths $\mathbf{a}_I^* = (\mathbf{a}^*(i))_{i \in I}$.The simple strategy profile $s^{\mathbf{a}(0),\mathbf{a}_I^*}$ is called the* ***Nash-reversion trigger strategy*** *profile. Furthermore, if, for every $i \in I$, the Nash equilibrium $a^*(i)$ satisfies $v_i(a^*(i)) = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} v_i(a_i, a_{-i})$, then the simple strategy profile $s^{\mathbf{a}(0),\mathbf{a}_I^*}$ is called the* ***Nash-reversion grim trigger strategy*** *profile.*

In words, a Nash-reversion trigger strategy profile is a simple strategy profile in which if player $i \in I$ deviates, the punishment protocol recommends to play a Nash equilibrium of the stage game from there onwards. Moreover, if the Nash equilibrium also "minmaxes" player i (namely, it minimizes the utility of the deviator once we take into account that he is always able to react to the punishment by maximizing his own payoff), then we have a Nash-reversion grim trigger strategy profile.

Example 62. Let $\Gamma^{\delta,\infty}(G)$ be the infinitely repeated game with stage game G given by the Prisoners' Dilemma of Example 60. Consider strategy profile (s_1, s_2) such that, for every $i \in \{1, 2\}$,

$$s_i(h) = \begin{cases} C, & \text{if } h \in \{\varnothing\} \cup \{(a_1, a_2, ...) : \forall t \geq 1, a^t \in \{(C, C), (D, D)\}\}, \\ D, & \text{otherwise.} \end{cases}$$

This strategy profile is Nash-reversion grim trigger strategy (hence a trigger strategy) profile. ▲

Although trigger strategy profiles are easy to describe, they are not, in general, the best way to incentivize players. Indeed, harder punishments are possible if we give up the requirement that (i) the punishment must specify a constant action profile, and (ii) such action profile must be a Nash equilibrium of the stage game. Once we give up these two requirements, we can obtain other relevant concepts, such as *optimal penal codes*. Their importance relies in the fact that they represent the optimal way to incentivize players in a repeated game. As such, whenever a payoff profile is attainable in a pure SPE, it is also attainable with a simple strategy profile in which deviators are punished according to the optimal penal code.

Before providing a formal definition of optimal penal codes, it is useful to prove a preliminary result that has its own relevance.

Lemma 32. *Let $\Gamma^{\delta,\infty}(G)$ be a repeated game with perfect monitoring. Then the set W^* of pure SPE payoffs is compact.*

Proof. We provide the proof for the case in which the stage game G is finite. In this proof we use the concepts of enforceability and self-generation introduced above. The proof for the case where G is compact-continuous can be found in Osborne and Rubinstein [53, Lemma 153.3]. Since the utility of the players in the stage game is bounded (by finiteness of G) and the discount factor δ is strictly smaller than 1, the set W^* is bounded. So we have to show that W^* is closed. To this end, let cl(W^*) denote the closure of W^*. By definition of closure of a set, we have $W^* \subseteq$cl(W^*). We will show that cl(W^*) is pure-action self-generating; by Proposition 11, this will imply cl$(W^*) \subseteq W^*$, and so $W^* =$cl(W^*), as required. Pick any $\hat{v} \in$cl(W^*). We prove that $\hat{v}$ is pure-action decomposable with action profile a^* and enforcing function η^*. First note that, since cl(W^*) is a compact subset of a metric space, there exists a sequence $(\hat{v}_n)_{n\in\mathbb{N}} \in (W^*)^{\mathbb{N}}$ of payoff profiles converging to $\hat{v}$. Proposition 11 implies that each $\hat{v}_n$ is a pure-strategy SPE profile and it is pure-action decomposable on W^*. Thus, we can define a sequence of action profiles $(a_n)_{n\in\mathbb{N}}$ and a sequence of (continuous) functions $(\eta_n : A \to W^*)_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$,

$$\hat{v}_n = (1-\delta)\, v(a_n) + \delta\eta_n(a_n).$$

Next note that $A = \times_{i\in I} A_i$ is finite, hence compact. Moreover $(\mathrm{cl}(W^*))^A$ is a compact metrizable space, and $\eta_n \in (\mathrm{cl}(W^*))^A$ for every $n \in \mathbb{N}$. It follows that $A \times (\mathrm{cl}(W^*))^A$ is compact, and $(a_n, \eta_n) \in A \times (\mathrm{cl}(W^*))^A$ for every $n \in \mathbb{N}$. Therefore we can find a subsequence $(a_{n_k}, \eta_{n_k})_{k\in\mathbb{N}}$ of $(a_n, \eta_n)_{n\in\mathbb{N}}$ such that, as $k \to \infty$, a_{n_k} converges to $a^* \in A$ and η_{n_k} converges to a (continuous) function $\eta^* \in (\mathrm{cl}(W^*))^A$. Furthermore, $v(\cdot)$ is a continuous function, and so $\lim_{k\to\infty} v(a_{n_k}) = v(a^*)$. Hence

$$\begin{aligned}\hat{v} &= \lim_{n\to\infty} \hat{v}_n = \lim_{k\to\infty} \hat{v}_{n_k} = \lim_{k\to\infty} \left((1-\delta)\, v(a_{n_k}) + \delta\eta_{n_k}(a_{n_k})\right) \\ &= (1-\delta)\, v(a^*) + \delta\eta^*(a^*).\end{aligned}$$

Therefore, $\hat{v}$ is pure-action decomposable with action profile a^* and enforcing function η^*. ■

We are now ready to define optimal penal codes and to prove that they exist.

Definition 76. *An* ***optimal penal code*** *is a profile of strategy profiles* $\mathbf{s} \in S^I$ *such that* $\mathbf{s} = \left(s^{\mathbf{a}(i),\mathbf{a}_I}\right)_{i\in I}$ *for some profile of paths* $\mathbf{a}_I = (\mathbf{a}(i))_{i\in I}$*, where each simple strategy profile* $s^{\mathbf{a}(i),\mathbf{a}_I}$ *(*$i \in I$*) is an SPE.*

To understand this definition, interpret each $\mathbf{a}(i)$ as paths "punishing" player i. Then, the i-th element of an optimal penal code is the strategy profile that starts "punishing" player i and, if any player j unilaterally deviates from such "punishment path," then "punishes" j with path $\mathbf{a}(j)$.

Proposition 12. *If repeated game* $\Gamma^{\delta,\infty}(G)$ *has a (pure) subgame perfect equilibrium, then it has an optimal penal code.*

Proof. If there is an SPE, then the set W^* of SPE payoff profiles is non-empty. It follows from Lemma 32 that, for every $i \in I$, the set W_i^* is non-empty and compact. Define $\underline{\mathbf{v}}_i^* = \min W_i^*$ and let $\mathbf{a}_I = (\mathbf{a}(i))_{i\in I}$ be a profile of infinite sequences of actions such that $u_i(\mathbf{a}(i)) = \underline{\mathbf{v}}_i^*$ for every

$i \in I$. Lemma 32 implies that for every $i \in I$ there exists an SPE $s[i]$ that induces paths $\mathbf{a}(i)$, with $u_i(\mathbf{a}(i)) = \underline{\mathbf{v}}_i^*$. We now show that, for every $i \in I$, the simple strategy $s^{\mathbf{a}(i),\mathbf{a}_I}$ is an SPE. To this end, recall that for all $i \in I$ and $t \in \mathbb{N}$,

$$u_i^t\left(a^1, a^2, ...,\right) = (1-\delta)\left(\sum_{m=t}^{\infty} \delta^{m-t} v_i\left(a^m\right)\right)$$

is the payoff of player i if path $\left(a^1, a^2, ...,\right)$ is played. So we have to show that for all $i, j \in I$, $t \in \mathbb{N}$ and $a_j \in A_j$,

$$u_j^t(\mathbf{a}(i)) \geq (1-\delta) v_j\left(a_j, a_{-j}^t(i)\right) + \delta u_j^0(\mathbf{a}(j)), \tag{13.3.1}$$

where $a_{-j}^t(i)$ is the profile of actions of j's co-players at time t in sequence $\mathbf{a}(i)$. The SPE conditions imply that for all $j \in I$, $t \in \mathbb{N}$ and $a_j \in A_j$,

$$u_j^t(\mathbf{a}(i)) \geq (1-\delta) v_j\left(a_j, a_{-j}^t(i)\right) + \delta u_j^{d,t}\left(a_j, a_{-j}^t(i)\right),$$

where $u_j^{d,t}\left(a_j, a_{-j}^t(i)\right)$ is the continuation payoff of player j if he unilaterally deviates from $\mathbf{a}(i)$ at stage t and plays action a_j. Since $s[i]$ is an SPE, we have that $u_j^{d,t}\left(a_j, a_{-j}^t(i)\right) \in W^*$ and consequently $u_j^{d,t}\left(a_j, a_{-j}^t(i)\right) \geq u_j^0(\mathbf{a}(j)) = \underline{\mathbf{v}}_j^*$. We conclude that (13.3.1) holds. ■

An immediate corollary of Proposition 12 is the following: Consider a simple strategy profile such that the punishment protocols are given by an optimal penal code and the initial path is given by $\mathbf{a}(0)$; then such simple strategy profile is self-supporting as long as $\mathbf{a}(0)$ can be supported when the punishment protocols are prescribed by the optimal penal code. The following result states that a sequence of action profiles in the repeated game can be supported in a pure SPE if and only if it can be supported through simple strategies in which an optimal penal code is used.

Proposition 13. *Fix a path* $\mathbf{a}(0) \in A^{\mathbb{N}}$ *and a profile of paths* $\mathbf{a}_I \in \left(A^{\mathbb{N}}\right)^I$ *such that* $\mathbf{s} = \left(s^{\mathbf{a}(i),\mathbf{a}_I}\right)_{i \in I}$ *is an optimal penal code. There exists an SPE* $s[0]$ *that induces* $\mathbf{a}(0)$ *if and only if the simple strategy profile* $s^{\mathbf{a}(0),\mathbf{a}_I^*}$ *is an SPE.*

Proof. The "if" direction is immediate. To prove the "only if" direction, observe that, by assumption, there exist an SPE $s[0]$ that induces $\mathbf{a}(0)$. Then, we can replicate the proof of Proposition 12 with $s[0]$ replacing $s[i]$ and prove the result. ■

14

Bargaining

Bargaining is a pervasive economic phenomenon. Many relevant economic settings involve the confrontation between two or more agents with conflicting interests. Examples abound. A seller and a buyer may bargain over the terms of trade; unions can bargain with entrepreneurs to attain a more favorable split of the total surplus; different political parties may bargain on the allocation of the national budget.

The goal of this Chapter is to introduce the workhorse model used in economics to analyze bargaining environments, namely the *Rubinstein's Bargaining Game* (Rubinstein [57]).[1] In line with this goal, we first analyze a simpler setting that, strictly speaking, does not involve any actual bargaining, the *Ultimatum Game*. The analysis of the Ultimatum Game will highlight some insights which hold also in more complicated settings. Then, we will introduce some flavor of bargaining by considering a twice repeated Ultimatum Game. Finally, we will proceed with the analysis of the Rubinstein's game and we will characterize the set of its equilibria.

14.1 The Ultimatum Game

Two agents, Ann (A) and Bob (B), have to agree on the split of \$1, or a "pie" (surplus) of size 1. Ann moves first and proposes a feasible payoff

[1]See also chapter 7 in Osborne and Rubinstein [53].

split, namely a pair of non-negative numbers summing up to 1.[2] Formally, define

$$X = \left\{ (x_{\mathrm{A}}, x_{\mathrm{B}}) \in [0,1]^{\{\mathrm{A},\mathrm{B}\}} : x_{\mathrm{A}} + x_{\mathrm{B}} = 1 \right\}.$$

Ann proposes $x = (x_{\mathrm{A}}, x_{\mathrm{B}}) \in X$, where x_{B} is the share offered to Bob and $x_{\mathrm{A}} = 1 - x_{\mathrm{B}}$ is the share that Ann demands for herself. Upon observing the offer, Bob can either accept (action y) or reject (action n). If Bob accepts, the proposal is implemented and the game ends. If instead Bob rejects, a default exogenous split $(\bar{x}_{\mathrm{A}}, \bar{x}_{\mathrm{B}}) \in X$ is implemented with a *delay* of one period and the game ends.[3] Delay captures the idea that disagreement is costly and Pareto inefficient. We also assume that each default share is strictly positive, that is, $0 < \bar{x}_{\mathrm{A}} < 1$. This describes the rules of the game, also called *bargaining protocol*. As regards immediate agreements, players are expected utility maximizers with linear vNM utility indexes over the total amount of money they receive. But players are impatient in the following sense: When comparing x in the current period to $\bar{x}$ in the following period, each player i weakly (respectively, strictly) prefers x immediately if and only if $x_i \geq \delta \bar{x}_i$ (respectively, $x_i > \delta \bar{x}_i$), where $\delta \in (0,1)$ is the discount factor.[4] The bargaining protocol and players' preferences over (lotteries of) consequences are common knowledge.

The environment we just described is a two-stage game with complete and perfect information. Using our general notation for multistage games, we have $\Gamma = \langle I, Y, g, (A_i, \mathcal{A}_i(\cdot), v_i)_{i \in I} \rangle$ with the following elements:

- $A_{\mathrm{A}} = X \cup \{\mathrm{wait}\}$, $A_{\mathrm{B}} = \{\mathrm{y}, \mathrm{n}, \mathrm{wait}\}$,
- $\mathcal{A}_{\mathrm{A}}(\varnothing) = X$, $\mathcal{A}_{\mathrm{B}}(\varnothing) = \{\mathrm{wait}\}$, $\mathcal{A}_{\mathrm{A}}((x, \mathrm{wait})) = \{\mathrm{wait}\}$, $\mathcal{A}_{\mathrm{B}}((x, \mathrm{wait})) = \{\mathrm{y}, \mathrm{n}\}$,[5]

[2] More generally, the sum has to be at most 1. Here we apply the resource constraint with equality to simplify the analysis, as this does not affect the equilibria.

[3] More generally, the default shares satisfy $\bar{x}_{\mathrm{A}} \geq 0$, $\bar{x}_{\mathrm{B}} \geq 0$, and $\bar{x}_{\mathrm{A}} + \bar{x}_{\mathrm{B}} \leq 1$.

[4] Under risk neutrality, one can equivalently assume that there is no pure time-discounting, but in case of rejection of the proposal the pie is destroyed with probability $1-\delta$. Furthermore, all the statements and arguments given here extend seamlessly to the more general case of players with different discount factors. We omit this generalization for notational simplicity.

[5] Recall that $\varnothing$ denotes the sequence of length zero, i.e., the root of the game, or empty history.

- $Y = X \times \{1,2\}$ is the set of dated splits, where (x,t) represents a split x implemented at date t,[6]

- $g\left(\left(x,\text{wait}\right),\left(\text{wait},\text{y}\right)\right) = (x,1)$, $g\left(\left(x,\text{wait}\right),\left(\text{wait},\text{n}\right)\right) = (\bar{x},2)$ for all $x \in X$,

- $v_i(x,t) = \delta^{t-1}x_i$ for all $(x,t) \in Y$.

The game features complete information because all of the above is common knowledge, and it also features perfect information because players perfectly observe past histories and there is only one active player at each non-terminal history. We maintain such assumptions about the information structure hold for all the bargaining games analyzed here. Henceforth, we ease notation and identify histories with the sequence of actions taken by the active players, e.g., we write (x,n) instead of $\left(\left(x,\text{wait}\right),\left(\text{wait},\text{n}\right)\right)$.

Before we analyze the Ultimatum Game as defined above, it is useful to consider an *approximation with finite action spaces*. For any $k \in \mathbb{N}$, let

$$X_k = \left\{x \in X : x_{\text{A}} \in \left\{0, \frac{1}{k}, ..., \frac{k-1}{k}, 1\right\}\right\}.$$

For example, if $k = 100$, the proposer can only offer to the respondent 0%, 1%, 2%, ..., 99%, or 100% of the pie. Let the corresponding game be denoted by $\Gamma^{(k)}$. In what follows, we assume k to be "large enough," that is, we assume the finite *grid of possible offers* to be *sufficiently fine* to obtain equilibria that approximate the equilibrium of the game with a continuum of offers.

To analyze the game, we have to identify the **minimally acceptable offer** for Bob, considering that a rejection is worth $\delta\bar{x}_{\text{B}}$ to him:

$$x_{\text{B}}^{(k)} := \min_{x \in X_k} \left\{x_{\text{B}} : x_{\text{B}} \geq \delta\bar{x}_{\text{B}}\right\}.$$

By definition, $\delta\bar{x}_{\text{B}} \leq x_{\text{B}}^{(k)} < \delta\bar{x}_{\text{B}} + 1/k$, with $\delta\bar{x}_{\text{B}} = x_{\text{B}}^{(k)}$ if $\delta\bar{x}_{\text{B}} \in \left\{0, \frac{1}{k}, ..., \frac{k-1}{k}, 1\right\}$. Thus, $\lim_{k\to\infty} x_{\text{B}}^{(k)} = \delta\bar{x}_{\text{B}}$, which implies that $x_{\text{B}}^{(k)} - \delta\bar{x}_{\text{B}} < 1 - \delta$ for k large enough. There are two possibilities, either

[6] Recall that Y denotes the space of possible outcomes, or consequences. Do not confuse Y with "yes."

$\delta\bar{x}_{\mathrm{B}} \in \left\{0, \frac{1}{k}, ..., \frac{k-1}{k}, 1\right\}$, or not; in the latter case, $x_{\mathrm{B}}^{(k)} > \delta\bar{x}_{\mathrm{B}}$. If $x_{\mathrm{B}}^{(k)} > \delta\bar{x}_{\mathrm{B}}$, Bob is strictly better off accepting offer $x^{(k)} = \left(1 - x_{\mathrm{B}}^{(k)}, x_{\mathrm{B}}^{(k)}\right)$, and we can use backward induction to find the unique subgame perfect equilibrium (SPE hereafter): If $x_{\mathrm{A}}^{(k)} > \delta\bar{x}_{\mathrm{A}}$ (i.e., $x_{\mathrm{B}}^{(k)} - \delta\bar{x}_{\mathrm{B}} < 1 - \delta$, which holds if $k > 1/(1-\delta)$), Ann offers $x^{(k)}$ and Bob accepts an offer $x \in X_k$ if and only if $x_{\mathrm{B}} \geq x_{\mathrm{B}}^{(k)}$ (Ann is better off making Bob accept this offer if k is sufficiently large). If instead $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}}$, Bob is indifferent between accepting and rejecting the minimally acceptable offer, because the cost of delay exactly offsets the immediate consumption of a smaller share. Therefore, we have to use "case-by-case" backward induction to find the SPEs:

Case 1 When indifferent, Bob accepts; anticipating this,[7] Ann offers $x^{(k)} = \left(1 - x_{\mathrm{B}}^{(k)}, x_{\mathrm{B}}^{(k)}\right)$; Bob accepts offer $x = (x_{\mathrm{A}}, x_{\mathrm{B}})$ if and only if $x_{\mathrm{B}} \geq x_{\mathrm{B}}^{(k)}$.

Case 2 When indifferent, Bob rejects; anticipating this, Ann has to offer slightly more to make him accept. Assuming that $k > 1/(1-\delta)$ (k large) Ann is better off by offering $\left(1 - x_{\mathrm{B}}^{(k)} - \frac{1}{k}, x_{\mathrm{B}}^{(k)} + \frac{1}{k}\right)$ and make Bob accept. Indeed, $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}} = \delta(1 - \bar{x}_{\mathrm{A}})$ and $1 - \delta(1 - \bar{x}_{\mathrm{A}}) - \frac{1}{k} > \delta\bar{x}_{\mathrm{A}}$ if $k > 1/(1-\delta)$; thus, the minimal offer that Bob is willing to accept gives Ann more than the present values of her default share. To summarize, we have:

Proposition 14. *Suppose that k is large enough, that is, $k > 1/(1-\delta)$. Then, the following is a pure strategy SPE of $\Gamma^{(k)}$: $s_{\mathrm{A}}(\varnothing) = \left(1 - x_{\mathrm{B}}^{(k)}, x_{\mathrm{B}}^{(k)}\right)$ and $s_{\mathrm{B}}(x) = y$ if and only if $x_{\mathrm{B}} \geq x_{\mathrm{B}}^{(k)}$; this is the* unique *SPE if $x_{\mathrm{B}}^{(k)} > \delta\bar{x}_{\mathrm{B}}$. If $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}}$, there is another pure strategy SPE of $\Gamma^{(k)}$: $s_{\mathrm{A}}(\varnothing) = \left(1 - x_{\mathrm{B}}^{(k)} - \frac{1}{k}, x_{\mathrm{B}}^{(k)} + \frac{1}{k}\right)$ and $s_{\mathrm{B}}(x) = y$ if and only if $x_{\mathrm{B}} > x_{\mathrm{B}}^{(k)}$.*

If $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}}$, there is also a continuum of SPEs in partially randomized behavior strategies, we omit the details. As $k \to \infty$ the grid of possible offers approximates the continuum ever more finely, $x_{\mathrm{B}}^{(k)} \to \delta\bar{x}_{\mathrm{B}}$ and all the strategy pairs in the SPE set converge to the pure equilibrium whereby

[7] Recall that, in a subgame perfect equilibrium, Ann's conjecture about Bob is correct.

Ann makes the least acceptable offer and Bob says Yes. It turns out that this is the unique subgame perfect equilibrium of Γ, the Ultimatum Game with a continuum of offers, even if we allow for randomized strategies.[8]

Proposition 15. *The Ultimatum Game has a unique SPE* $(s_{\mathrm{A}}, s_{\mathrm{B}})$*, specifically:*

$$s_{\mathrm{A}}(\varnothing) = (1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}}),\ s_{\mathrm{B}}(x_{\mathrm{A}}, x_{\mathrm{B}}) = \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{B}} \geq \delta\bar{x}_{\mathrm{B}}, \\ \mathrm{n}, & \text{if } x_{\mathrm{B}} < \delta\bar{x}_{\mathrm{B}}. \end{cases}$$

Proof. Let $(\beta_{\mathrm{A}}, \beta_{\mathrm{B}})$ be an SPE in behavior strategies. We will proceed backwards. Faced with an offer $(x_{\mathrm{A}}, x_{\mathrm{B}})$, if Bob accepts this offer, he gets x_{B} immediately. If he rejects, he gets $\bar{x}_{\mathrm{B}}$ one period later, which is worth $\delta\bar{x}_{\mathrm{B}}$. Thus, rationality implies that he replies y if $x_{\mathrm{B}} > \delta\bar{x}_{\mathrm{B}}$, n if $x_{\mathrm{B}} < \delta\bar{x}_{\mathrm{B}}$ and either y or n if $x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}$. As a result, the rejection probability $\beta_{\mathrm{B}}(\mathrm{n}|x)$ may be different from 0 and 1 only if $x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}$. To ease notation, let $\rho = \beta_{\mathrm{B}}(\mathrm{n}|(1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}}))$ denote this rejection probability. With this, the expected payoff of Ann when she proposes $(x_{\mathrm{A}}, x_{\mathrm{B}})$ is

$$\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}})) = \begin{cases} \delta\bar{x}_{\mathrm{A}}, & \text{if } x_{\mathrm{B}} < \delta\bar{x}_{\mathrm{B}}, \\ \rho\delta\bar{x}_{\mathrm{A}} + (1-\rho)(1 - \delta\bar{x}_{\mathrm{B}}), & \text{if } x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}, \\ 1 - x_{\mathrm{B}}, & \text{if } x_{\mathrm{B}} > \delta\bar{x}_{\mathrm{B}}. \end{cases}$$

Now note that

$$1 - \delta\bar{x}_{\mathrm{B}} > \delta(1 - \bar{x}_{\mathrm{B}}) = \delta\bar{x}_{\mathrm{A}},$$

which implies that $\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}}))$ is discontinuous at $x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}$, with

$$\lim_{x_{\mathrm{B}} \searrow \delta\bar{x}_{\mathrm{B}}} \mathbb{E}_{\beta_{\mathrm{B}}}[v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}})] = 1 - \delta\bar{x}_{\mathrm{B}} > \delta\bar{x}_{\mathrm{A}} = \lim_{x_{\mathrm{B}} \nearrow \delta\bar{x}_{\mathrm{B}}} \mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}})).$$

If $\rho > 0$ then

$$1 - \delta\bar{x}_{\mathrm{B}} > \rho\delta\bar{x}_{\mathrm{A}} + (1-\rho)(1 - \delta\bar{x}_{\mathrm{B}}).$$

If $\rho = 0$, then $\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}}))$ attains the supremum $1 - \delta\bar{x}_{\mathrm{B}}$, which is therefore a maximum, at offer $(1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}})$. Hence, the candidate equilibrium of the statement is indeed a (pure) SPE. If instead $\rho > 0$, then

[8] From now, when we claim that a bargaining game has a unique equilibrium, we mean unique among all the SPEs in behavior strategies, which include those in pure strategies. It will turn out that the unique equilibrium is always in pure strategies.

payoff $1 - \delta \bar{x}_{\mathrm{B}}$ can be approached arbitrarily closely, but not attained by $\mathbb{E}_{\beta_{\mathrm{B}}} \left(v_{\mathrm{A}} \middle| \left(1 - x_{\mathrm{B}}, x_{\mathrm{B}}\right)\right)$, and Ann does not have a best reply. This shows that the candidate equilibrium is the unique SPE. ■

For future reference, we want to stress a few key implications of Proposition 15. First, the equilibrium is *unique*. In particular, in such equilibrium the respondent is kept at the present value of his disagreement outcome, $\delta \bar{x}_{\mathrm{B}}$, while the proposer extracts all the remaining surplus, $1 - \delta \bar{x}_{\mathrm{B}}$. Second, the equilibrium payoff is determined by the disagreement share of the (last-stage) respondent or equivalently by the value he can obtain by rejecting the offer coming from the proposer. Third, the equilibrium outcome is Pareto efficient, namely the disagreement outcome $\left(\left(1 - \delta \bar{x}_{\mathrm{B}}, \delta \bar{x}_{\mathrm{B}}\right), 2\right)$, which is inefficient because consumption is delayed, never arises. In the rest of this note we show that these insights hold true even if Bob can make counteroffers and bargaining lasts for (infinitely) many periods.

14.2 2-Period Alternating Offer Game

In the Ultimatum Game, Bob can only accept or reject the offer coming from Ann and he is prevented from proposing counteroffers. To allow for the latter, we extend the Ultimatum Game by adding a subgame following the rejection of Bob. In this subgame, the interaction between the two agents is a new Ultimatum Game in which the roles of the two agents are switched. Thus, Bob becomes the proposer and Ann the respondent. We call this game **2-Period Alternating Offer Game**.

As before, suppose that Ann and Bob have to split \$1. The bargaining protocol is described as follows:[9]

1.P at the beginning of the game Ann makes a proposal $x^1 = \left(x^1_{\mathrm{A}}, x^1_{\mathrm{B}}\right) \in X$;

1.R Bob can either accept (y) or reject (n) offer x^1; if Bob accepts, the proposed split x^1 is implemented in period 1 and the game ends; if Bob rejects, the game moves to period 2;

[9] We leave to the reader the exercise of representing this bargaining protocol with the general mathematical notation for dynamic games.

2.P if the play moves to period 2, Bob makes a proposal $x^2 = \left(x^2_{\mathrm{A}}, x^2_{\mathrm{B}}\right) \in X$;

2.R Ann can either accept (y) or reject (n) offer x^2; if Ann accepts, the proposed split x^2 is implemented in period 2; if Ann rejects the disagreement split $(\bar{x}_{\mathrm{A}}, \bar{x}_{\mathrm{B}})$ is implemented in period 3 and the game ends.

The next proposition states that also the 2-Period Alternating Offer Game has a unique subgame perfect equilibrium and provides its characterization.

Proposition 16. *The 2-Period Alternating Offer Game has a unique SPE* $(s_{\mathrm{A}}, s_{\mathrm{B}})$ *determined as follows: for every* $(x^1, x^2) \in X \times X$

$$s_{\mathrm{A}}(\varnothing) = (1 - \delta(1 - \delta\bar{x}_{\mathrm{A}}), \delta(1 - \delta\bar{x}_{\mathrm{A}})), \ s_{\mathrm{B}}\left(x^1\right) = \begin{cases} \mathrm{y}, & \text{if } x^1_{\mathrm{B}} \geq \delta(1 - \delta\bar{x}_{\mathrm{A}}), \\ \mathrm{n}, & \text{if } x^1_{\mathrm{B}} < \delta(1 - \delta\bar{x}_{\mathrm{A}}), \end{cases}$$

and

$$s_{\mathrm{B}}\left(x^1, \mathrm{n}\right) = (\delta\bar{x}_{\mathrm{A}}, 1 - \delta\bar{x}_{\mathrm{A}}), \ s_{\mathrm{A}}\left(x^1, \mathrm{n}, x^2\right) = \begin{cases} \mathrm{y}, & \text{if } x^2_{\mathrm{A}} \geq \delta\bar{x}_{\mathrm{A}}, \\ \mathrm{n}, & \text{if } x^2_{\mathrm{A}} < \delta\bar{x}_{\mathrm{A}}. \end{cases}$$

Proof. As in the Ultimatum Game with a continuum of offers, if the proposer anticipates that—when indifferent—the respondent says No (with positive probability), then the reduced-form payoff function of the proposer cannot be maximized. Thus, the proposer has a best response to the rational strategy of the respondent only if the respondent accepts when indifferent. This allows to solve for the unique SPE by a kind of modified backward induction procedure, whereby the respondent breaks ties in favour of the proposer.

In particular, the analysis of the second-period game, is the same as for the Ultimatum Game with reversed roles for Bob (proposer) and Ann (respondent). This explains the second-period strategies. With this, we obtain a reduced-form Ultimatum Game where the disagreement split implemented with one-period delay in case of rejection is the equilibrium offer of the second period $(\hat{x}_{\mathrm{A}}, \hat{x}_{\mathrm{B}}) = (\delta\bar{x}_{\mathrm{A}}, 1 - \delta\bar{x}_{\mathrm{A}})$. The solution of the Ultimatum Game with (delayed) disagreement split $(\hat{x}_{\mathrm{A}}, \hat{x}_{\mathrm{B}})$ is that Bob's response strategy is

$$\begin{aligned} s_{\mathrm{B}}\left(x^1\right) &= \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{B}}^1 \geq \delta \hat{x}_{\mathrm{B}}, \\ \mathrm{n}, & \text{if } x_{\mathrm{B}}^1 < \delta \hat{x}_{\mathrm{B}}, \end{cases} \\ &= \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{B}}^1 \geq \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \\ \mathrm{n}, & \text{if } x_{\mathrm{B}}^1 < \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \end{cases} \end{aligned}$$

and Ann offers

$$s_{\mathrm{A}}(\varnothing) = (1 - \delta \hat{x}_{\mathrm{B}}, \delta \hat{x}_{\mathrm{B}}) = \left(1 - \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right)\right).$$

■

Proposition 16 and its proof show that the equilibrium implications of the 2-period model are similar to those of the 1-period model (Ultimatum Game) stated in Proposition 15. First, the equilibrium is unique. Second, the equilibrium outcome[10] is determined by the disagreement share of the first-round respondent, $\hat{x}_{\mathrm{B}}$. However, differently from the Ultimatum Game, the disagreement share $\hat{x}_{\mathrm{B}}$ is endogenous, as it takes into account that if Bob rejects the offer, he becomes the new proposer and, in equilibrium, he "forces" Ann to accept share $\delta \bar{x}_{\mathrm{A}}$. Third, the equilibrium outcome is Pareto efficient, because costly disagreement (rejection and delay) does not occur. Finally, we observe that the strategy of the second-period respondent depends only on the last offer, i.e., it is independent of what happened in the first round.

14.3 Bargaining with Infinite Horizon

The analysis developed in the previous Sections can be easily extended to n-Periods Alternating Offer Games. For this reason, in this section we will jump to the case in which the bargaining game has an infinite horizon and agents keep alternating in the role of proposer and respondent until an agreement is reached. This setting has been studied in an extremely influential paper by Ariel Rubinstein (Rubinstein [57]) and it has become a building block of several applied works.

This is the bargaining protocol:

[10] That is, split $\left(1 - \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right)\right)$ in period 1.

- (*First-Step Rule*) In period 1 Ann can offer any split $x = (x_{\mathrm{A}}, x_{\mathrm{B}}) \in X$. Bob can accept (y) or reject (n). If Bob accepts offer x, the agreement is immediately implemented, i.e., the outcome (consequence) is $(x, 1)$; if Bob rejects, the play moves to the following period.

- (*Inductive-Step Rule*) If no agreement was reached before period $t > 1$, the player who rejected the offer in period $t-1$ becomes the proposer in period t. The proposer can offer any split $x \in X$. If offer x is accepted, the agreement is immediately implemented and the outcome is (x, t); if it is rejected the play moves to period $t+1$.

Thus, in the Rubinstein's bargaining game, Ann and Bob keep switching roles until an agreement is reached. In particular, Ann plays in the role of the proposer in period $t \in \mathbb{N}$ if and only if t is odd, and Bob plays in the role of proposer in period $t \in \mathbb{N}$ if and only if t is even.

Formally, this strategic environment can be represented as an infinite-horizon game with perfect information. In particular,[11]

- The set of **non-terminal histories** is the union $H = H_P \cup H_R$ of the set of histories where a proposer is active,

$$H_P := \bigcup_{t \in \mathbb{N}} (X \times \{\mathrm{n}\})^{t-1},$$

  and the set of histories where a respondent is active

$$H_R := \bigcup_{t \in \mathbb{N}} \left((X \times \{\mathrm{n}\})^{t-1} \times X \right).$$

- For all $t \in \mathbb{N}$, $h \in (X \times \{\mathrm{n}\})^{t-1}$ and $x \in X$, $\mathcal{A}_{\mathrm{A}}(h) = X$, $\mathcal{A}_{\mathrm{B}}(h) = \{\mathrm{wait}\} = \mathcal{A}_{\mathrm{A}}(h, x)$ and $\mathcal{A}_{\mathrm{B}}(h, x) = \{\mathrm{y,n}\}$ if t is odd, whereas $\mathcal{A}_{\mathrm{B}}(h) = X$, $\mathcal{A}_{\mathrm{A}}(h) = \{\mathrm{wait}\} = \mathcal{A}_{\mathrm{B}}(h, x)$ and $\mathcal{A}_{\mathrm{A}}(h, x) = \{\mathrm{y,n}\}$ if t is even.

- The set of **finite terminal histories**—or agreement histories—is

$$Z_{\mathrm{y}} := \bigcup_{t \in \mathbb{N}} \left((X \times \{\mathrm{n}\})^{t-1} \times (X \times \{\mathrm{y}\}) \right).$$

[11] Recall that $\mathbb{N} = \{1, 2, ...\}$ and that we adopt the convention $(X \times \{\mathrm{n}\})^0 = \{\varnothing\}$, where $\varnothing$ is the empty history.

- The set of **infinite terminal histories**—or permanent disagreement histories—is

$$Z_{\text{n}} := (X \times \{\text{n}\})^{\infty},$$

  the set of terminal histories is therefore $Z := Z_{\text{y}} \cup Z_{\text{n}}$.

- The **outcome** (or consequence) **space** is $Y = (X \times \mathbb{N}) \cup \{D\}$, where (x, t) means that agreement on x is reached in period t, and D conveniently denotes the **permanent disagreement** outcome; the **outcome** (or consequence) **function** $g : Z \to Y$ is

$$g(z) = \begin{cases} (x, t), & \text{if } z = (h, x, \text{y}) \in (X \times \{\text{n}\})^{t-1} \times (X \times \{\text{y}\}), \\ D, & \text{if } z \in Z_{\text{n}}. \end{cases}$$

- The **utility function** $v_i : Y \to \mathbb{R}$ of each player i is

$$\begin{aligned} \forall (x, t) \in (X \times \mathbb{N}), \quad & v_i(x, t) = \delta^{t-1} x_i, \\ & v_i(D) = 0. \end{aligned}$$

The implied **payoff function** $u_i = v_i \circ g : Z \to \mathbb{R}$ of each player $i \in \{\text{A}, \text{B}\}$ is given by

$$u_i(z) = \begin{cases} \delta^{t-1} x_i, & \text{if } z = (h, x, \text{y}) \in (X \times \{\text{n}\})^{t-1} \times (X \times \{\text{y}\}), \\ 0, & \text{if } z \in Z_{\text{n}}. \end{cases}$$

Note that u_i satisfies the property of **continuity at infinity**: for every $\varepsilon > 0$ there exists a sufficiently large t, such that if two terminal histories z' and z'' have a t-period common prefix, then $|u_i(z') - u_i(z'')| \leq \varepsilon$. Indeed, for every $t \in \mathbb{N}$, every non-terminal history $h \in (X \times \{\text{n}\})^t$, and every pair of terminal continuations of h, viz. $z' = (h, ...) \in Z$ and $z'' = (h, ...) \in Z$, we have $|u_i(z') - u_i(z'')| \leq \delta^t$. This implies that the *OD principle* holds: a strategy s_i is sequentially optimal given conjecture β^i (possibly a deterministic conjecture s_{-i}) if and only if s_i is one-step optimal given β^i.[12]

In the following analysis, it is convenient to partition H into four subsets according to two binary criteria: whether A or B is the active player, and whether the active player is a proposer or a respondent. We

[12] See Section 10.5.

let $H_{i,P}$ (respectively $H_{i,R}$) denote the set of histories where $i \in \{\mathrm{A}, \mathrm{B}\}$ is a proposer (respectively, a respondent). Thus

$$H = H_{\mathrm{A},P} \cup H_{\mathrm{B},R} \cup H_{\mathrm{B},P} \cup H_{\mathrm{A},R},$$

with

$$\begin{aligned}
H_{\mathrm{A},P} &= \bigcup_{t\in\{1,3,5,\ldots\}} (X \times \{\mathrm{n}\})^{t-1}, \\
H_{\mathrm{B},R} &= \bigcup_{t\in\{1,3,5,\ldots\}} (X \times \{\mathrm{n}\})^{t-1} \times X, \\
H_{\mathrm{B},P} &= \bigcup_{t\in\{2,4,6,\ldots\}} (X \times \{\mathrm{n}\})^{t-1}, \\
H_{\mathrm{A},R} &= \bigcup_{t\in\{2,4,6,\ldots\}} (X \times \{\mathrm{n}\})^{t-1} \times X.
\end{aligned}$$

The set of strategies in the Rubinstein's bargaining game allows rather complex behavior and history-dependence. For instance, an agent may react to a low offer by rejecting it and by offering an even lower share to his opponent once he becomes proposer in the following period. However, we can define a simple subset of strategies, namely stationary strategies. A strategy for player i is stationary if:

(i) the proposal made by i when he plays in the role of the proposer is always the same, hence it is independent of the previous history;

(ii) the reply of agent i when he plays in the role of the respondent depends only on the last offer, hence it is independent of the history that precedes such an offer.

Formally, a strategy s_i ($i \in \{\mathrm{A,B}\}$) is **stationary** if $s_i(h) = s_i(h')$ for all $h, h' \in H_{i,P}$, and $s_i(h, x) = s_i(h', x)$ for all $h, h' \in H_{-i,P}$, and $x \in X$.[13] An SPE is stationary if both players use stationary strategies. Notice that if each player follows a stationary strategy, each bargaining round is independent of the past: the proposer does not condition his behavior on the past history, while the reply of the respondent depends only on the offer that is currently on the table. Moreover, by their very nature,

[13] Note that, if $h \in H_{-i,P}$, then $(h, x) \in H_{i,R}$.

stationary strategies can be easily described by the proposal that an agent makes when he plays in the role of the proposer and by a reply function that specifies how the respondent reacts to any offer $x \in X$.

Since Rubinstein's game has infinite horizon, its SPE cannot be computed by backward induction. Nevertheless, we can use a "conjecture-and-verify" approach that exploits our analysis of the 2-Period Alternating Offer Game to characterize an SPE.[14]

In the finite-horizon equilibrium, offers are accepted and reply functions are "monotone," that is, every offer that gives the respondent less than the equilibrium offer is rejected and every offer that gives more is accepted. However, the sequence of equilibrium offers is not stationary, because as the deadline comes closer equilibrium offers are more influenced by the default split.[15]

In the infinite horizon game there is no deadline and this implies that the game has a **stationary structure**. Specifically, for every $i \in \{\mathrm{A}, \mathrm{B}\}$ and every pair of histories $h, h' \in H_{i,P}$, the subgames starting after h and h' look exactly the same: i is the first proposer of an infinite-horizon alternating-offer game; furthermore, for every offer $x \in X$, the subgames starting after (h, x) and (h', x) again look exactly the same: $-i$ has to respond and—in case of rejection—he will be the first proposer in an infinite-horizon alternating-offer game. Therefore we conjecture that there is a stationary SPE whereby equilibrium offers are always accepted and reply functions are "monotone" in the sense explained above, with symmetry when the roles of the players are switched, that is, the proposer always demand the same for herself and the acceptance threshold of the respondent is what is left for himself by such proposal. Let x_P^* denote the **share that the proposer demands for herself**. Then we conjecture that there is a stationary SPE $(s_\mathrm{A}, s_\mathrm{B})$ of the following form: for all $h \in H_{\mathrm{A},P}$, $h' \in H_{\mathrm{B},P}$, $x = (x_\mathrm{A}, x_\mathrm{B}) \in X$:

$$
\begin{aligned}
s_\mathrm{A}(h) &= (x_P^*, 1 - x_P^*)\,, \; s_\mathrm{A}(h', x) = \begin{cases} \text{y}, & \text{if } x_\mathrm{A} \geq 1 - x_P^*, \\ \text{n}, & \text{if } x_\mathrm{A} < 1 - x_P^*, \end{cases} \\
s_\mathrm{B}(h') &= (1 - x_P^*, x_P^*)\,, \; s_\mathrm{B}(h, x) = \begin{cases} \text{y}, & \text{if } x_\mathrm{B} \geq 1 - x_P^*, \\ \text{n}, & \text{if } x_\mathrm{B} < 1 - x_P^*. \end{cases}
\end{aligned}
$$

[14] This approach is borrowed from Gibbons [36].

[15] We proved this for two periods of barganing, but it is intuitive that the same holds for finitely many periods.

With this, we can find the implied value of x_P^* relying on the analysis of the 2-Period Alternating Offer Game: according to the candidate equilibrium, if the play moves to period $t = 3$ the split $(x_P^*, 1 - x_P^*)$ is implemented in that period. Therefore, the first two periods strategies must be determined by backward induction (breaking ties in favor of the proposer) as in the 2-period model with default split $(\bar{x}_{\mathrm{A}}, \bar{x}_{\mathrm{B}}) = (x_P^*, 1 - x_P^*)$. We have seen that Ann's equilibrium offer in such game is

$$(1 - \delta (1 - \delta \bar{x}_{\mathrm{A}}), \delta (1 - \delta \bar{x}_{\mathrm{A}})) = (1 - \delta (1 - \delta x_P^*), \delta (1 - \delta x_P^*)).$$

Therefore, to have a stationary equilibrium of the infinite-horizon game we must have

$$x_P^* = 1 - \delta (1 - \delta x_P^*).$$

Solving this equation we find

$$x_P^* = \frac{1 - \delta}{1 - \delta^2} = \frac{1 - \delta}{(1 - \delta)(1 + \delta)} = \frac{1}{1 + \delta}.$$

To check that the stationary strategies specified above with $x_P^* = 1/(1 + \delta)$ indeed form an SPE it is enough to verify that each one is one-step optimal given the conjecture that the opponent follows his strategy in the candidate equilibrium. We leave this as an exercise.

Proposition 17. *Rubinstein's bargaining game has an SPE in stationary strategies* $(s_{\mathrm{A}}, s_{\mathrm{B}})$ *determined as follows: for all* $h \in H_{\mathrm{A},P}$, $h' \in H_{\mathrm{B},P}$, *and* $x = (x_{\mathrm{A}}, x_{\mathrm{B}}) \in X$*:*

$$\begin{aligned} s_{\mathrm{A}}(h) &= \left(\frac{1}{1+\delta}, \frac{\delta}{1+\delta}\right), \ s_{\mathrm{A}}(h', x) = \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{A}} \geq \frac{\delta}{1+\delta}, \\ \mathrm{n}, & \text{if } x_{\mathrm{A}} < \frac{\delta}{1+\delta}, \end{cases} \\ s_{\mathrm{B}}(h') &= \left(\frac{\delta}{1+\delta}, \frac{1}{1+\delta}\right), \ s_{\mathrm{B}}(h, x) = \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{B}} \geq \frac{\delta}{1+\delta}, \\ \mathrm{n}, & \text{if } x_{\mathrm{B}} < \frac{\delta}{1+\delta}. \end{cases} \end{aligned}$$

Proposition 18 below establishes that this is the unique SPE of the Rubinstein's game.

Proposition 18. *The strategy pair of Proposition 17 is the unique SPE of Rubinstein's bargaining game.*

Proof. Consider any subgame that starts with player i being active as the proposer. All these subgames have the same structure up to a normalization of utilities. In particular, for every $h, h' \in H_{i,P}$, the set of continuation strategies available to agent j ($j \in \{\text{A,B}\}$) in the subgame with initial node h is isomorphic to the set of continuation strategies available to agent j in the subgame with initial node h'. The only difference among these subgames is that the respective payoff functions can be obtained one from the other by a positive linear transformation, which does not affect incentives.

To prove our result, it is useful to introduce some additional notation. Take any $h \in H_{i,P}$. Let Γ_i^h be the subgame with initial node h in which we normalize payoffs so that immediate agreement on split $x = (x_\text{A}, x_\text{B})$ yields payoff x_A to Ann (x_B to Bob).[16] Given this normalization, the sum of the players' payoffs at any terminal history of the subgame Γ_i^h is at most 1. Denote with $E\left(\Gamma_i^h\right)$ the set of SPE equilibria in subgame Γ_i^h and with $V_j\left(\Gamma_i^h\right)$ ($j \in \{\text{A,B}\}$) the set of SPE payoffs of agent j. Since all subgames where i is the proposer are isomorphic, $E\left(\Gamma_i^h\right)$ and $V_j\left(\Gamma_i^h\right)$ are independent of $h \in H_{i,P}$, and it makes sense to write $E\left(\Gamma_i\right)$ and $V_j\left(\Gamma_i\right)$. Finally, denote with $\underline{v}_j\left(\Gamma_i\right)$ and $\overline{v}_j\left(\Gamma_i\right)$ the infimum and the supremum of $V_j\left(\Gamma_i\right)$. By Proposition 17, we know that $E\left(\Gamma_i\right)$ and $V_j\left(\Gamma_i\right)$ are nonempty and that $\overline{v}_j\left(\Gamma_i\right) > 0$.

Consider any subgame that starts at period-t history $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{A,P}$ (t odd). At such history Ann is the proposer. Rationality and the expectation of an equilibrium continuation imply that Bob accepts any offer x such that $x_\text{B} > \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right)$ and rejects any offer x such that $x_\text{B} < \delta\underline{v}_\text{B}\left(\Gamma_\text{B}\right)$. In words, Bob accepts (respectively, rejects) any offer that gives him more than the maximum (respectively, less than the minimum) equilibrium payoff he can get in the subgame that would start after his rejection. Thus if Ann offers $x = (x_\text{A}, x_\text{B})$ with $x_\text{B} > \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right)$, the behavior of players would yield consequence (x, t) and Ann would get a payoff equal to $1 - x_\text{B}$. As a result, $\lim_{x_\text{B} \searrow \delta\overline{v}_\text{B}(\Gamma_\text{B})} (1 - x) = 1 - \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right)$ is a lower bound on the equilibrium payoff that Ann can get in subgame Γ_A. Formally:

$$\underline{v}_\text{A}\left(\Gamma_\text{A}\right) \geq 1 - \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right). \tag{14.3.1}$$

Moreover, we have already argued that in equilibrium Bob rejects any offer

[16] That is, if $h \in (X \times \{\text{n}\})^{t-1}$, the payoff function of i in Γ_i^h satisfies $u_i^h(h, x, \text{y}) = x_i = \delta^{1-t} u_i(h, x, \text{y})$.

$x = (x_A, x_B)$ with $x_B < \delta \underline{v}_B(\Gamma_B)$. Thus, if Ann offers $x = (x_A, x_B)$ with $x_B < \delta \underline{v}_B(\Gamma_B)$, the maximum she can get is smaller than the maximum sum of payoffs attainable in the subgame starting after the rejection (which, from the perspective of period t, is equal to $\delta \cdot 1 = \delta$) minus the minimum Bob can get (which, from the perspective of period t, is equal to $\delta \underline{v}_B(\Gamma_B)$). Instead, if Ann offers $x = (x_A, x_B)$ with $x_B \geq \delta \underline{v}_B(\Gamma_B)$, the maximum she can obtain is at most $1 - \delta \underline{v}_B(\Gamma_B)$. Indeed, if Bob accepts the offer, Ann would get $1 - x_B$, while if he rejects, the maximum Ann can get is once more lower than $\delta(1 - \underline{v}_B(\Gamma_B))$. Since $\delta < 1$ and $\lim_{x_B \searrow \delta \underline{v}_B(\Gamma_B)} (1 - x) = 1 - \underline{v}_B(\Gamma_B)$, we conclude that :

$$\overline{v}_A(\Gamma_A) \leq 1 - \delta \underline{v}_B(\Gamma_B) \tag{14.3.2}$$

By replicating the previous analysis for subgame Γ_B, we get:

$$\underline{v}_B(\Gamma_B) \geq 1 - \delta \overline{v}_A(\Gamma_A) \tag{14.3.3}$$

$$\overline{v}_B(\Gamma_B) \leq 1 - \delta \underline{v}_A(\Gamma_A) \tag{14.3.4}$$

Combining (14.3.1) with (14.3.4) and (14.3.2) with (14.3.3), we get:

$$\begin{aligned}\underline{v}_A(\Gamma_A) &\geq 1 - \delta(1 - \delta \underline{v}_A(\Gamma_A)) \Rightarrow \underline{v}_A(\Gamma_A) \geq \frac{1}{1+\delta}, \\ \overline{v}_A(\Gamma_A) &\leq 1 - \delta(1 - \delta \overline{v}_A(\Gamma_A)) \Rightarrow \overline{v}_A(\Gamma_A) \leq \frac{1}{1+\delta}.\end{aligned}$$

By definition, $\underline{v}_A(\Gamma_A) \leq \overline{v}_A(\Gamma_A)$. Thus, $\underline{v}_A(\Gamma_A) = \overline{v}_A(\Gamma_A) = v_A(\Gamma_A) = \frac{1}{1+\delta}$. Therefore, $V_A(\Gamma_A) = \left\{\frac{1}{1+\delta}\right\}$. Following similar steps, we can get $\underline{v}_B(\Gamma_B) = \overline{v}_B(\Gamma_B) = v_B(\Gamma_B) = \frac{1}{1+\delta}$, so that $V_B(\Gamma_B) = \left\{\frac{1}{1+\delta}\right\}$. Given the normalization in the payoffs, we conclude that in every subgame of the Rubinstein's game that starts at period-t history $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{i,P}$ (t odd for $i = \text{A}$, even for $i = \text{B}$) player i (the proposer) gets a payoff equal to $\delta^{t-1} \frac{1}{1+\delta}$.

Now, consider subgame Γ_A from the perspective of Bob. We have just showed that by rejecting the offer, Bob will obtain $\delta v_B(\Gamma_B) = \frac{\delta}{1+\delta}$. Obviously, Bob can follow a strategy that prescribes to reject all offers made in period t. Thus, $\underline{v}_B(\Gamma_A) \geq \delta v_B(\Gamma_B) = \frac{\delta}{1+\delta}$. Moreover, the maximum payoff Bob can get in subgame Γ_A is smaller or equal than the maximum sum of payoffs available (namely, 1) minus the minimum payoff

Ann can get (by our previous reasoning, we know that $\underline{v}_{\text{A}}(\Gamma_{\text{A}}) = v_{\text{A}}(\Gamma_{\text{A}}) = \frac{1}{1+\delta}$). Thus, $\overline{v}_{\text{B}}(\Gamma_{\text{A}}) \leq 1 - v_{\text{A}}(\Gamma_{\text{A}}) = \frac{\delta}{1+\delta}$. By definition $\underline{v}_{\text{B}}(\Gamma_{\text{A}}) \leq \overline{v}_{\text{B}}(\Gamma_{\text{A}})$. We conclude that $\overline{v}_{\text{B}}(\Gamma_{\text{A}}) = \underline{v}_{\text{B}}(\Gamma_{\text{A}}) = v_{\text{B}}(\Gamma_{\text{A}}) = \frac{\delta}{1+\delta}$ and $V_{\text{B}}(\Gamma_{\text{A}}) = \left\{\frac{\delta}{1+\delta}\right\}$. A similar reasoning implies that $\overline{v}_{\text{A}}(\Gamma_{\text{B}}) = \underline{v}_{\text{A}}(\Gamma_{\text{B}}) = v_{\text{A}}(\Gamma_{\text{B}}) = \frac{\delta}{1+\delta}$ and $V_{\text{A}}(\Gamma_{\text{B}}) = \left\{\frac{\delta}{1+\delta}\right\}$. Thus, in every subgame of the Rubinstein's game that starts at period-t history $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{i,P}$ (t even for $i = \text{A}$, odd for $i = \text{B}$) player j (the receiver) gets a payoff equal to $\delta^{t-1}\frac{\delta}{1+\delta}$.

Now, pick $i \in \{\text{A,B}\}$ and consider any subgame with initial node $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{i,P}$. By the previous argument, we know that the sum of players' equilibrium payoffs in this subgame must be equal to $\delta^t\frac{1}{1+\delta} + \delta^t\frac{\delta}{1+\delta} = \delta^t$. Therefore, the only consequence compatible with equilibrium payoffs is (x^*, t), where $x^* = (x_P^*, x_R^*) = \left(\frac{1}{1+\delta}, \frac{\delta}{1+\delta}\right)$ and x_P^* (x_R^*) is the share the proposer demands for himself (offers to the respondent). Indeed, any consequence (x, t) with $x \neq x^*$ would yield payoffs different from the unique equilibrium payoffs we identified above. Similarly, any consequence (x, t') with $t' > t$ would yield payoffs whose sum could be at most $\delta^{t+1} < \delta^t$. Obviously, in the subgame that starts with initial node $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{i,P}$, consequence (x, t) is reached if and only if i proposes x^* and j accepts at history (h, x^*). Finally, consider any history (h, x) with $x = (x_P, x_R) \neq x^*$. Rationality and the expectation of an equilibrium continuation on the respondent's side imply that in equilibrium Bob accepts at any history $h \in (X \times \{\text{n}\})^{t-1} \times \{x\} \subseteq H_{i,R}$ with $x_R > \delta v_{\text{B}}(\Gamma_{\text{B}}) = \frac{\delta}{1+\delta}$, and rejects at any history $h \in (X \times \{\text{n}\})^{t-1} \times \{x\} \subseteq H_{i,R}$ with $x_R < \delta v_{\text{B}}(\Gamma_{\text{B}}) = \frac{\delta}{1+\delta}$. ■

Proposition 18 shows that infinite-horizon alternating-offers game shares the same features that we highlighted in a finitely repeated Ultimatum Game. In particular, although this game has infinite horizon and we cannot apply backward induction, the game has a unique subgame perfect equilibrium, which is stationary and rather simple to describe.[17] In this equilibrium, the proposer always asks for himself a fraction $\frac{1}{1+\delta}$ and

[17] This is in stark contrast with other infinite-horizon games that admit multiple equilibria and equilibrium outcomes. Such uniqueness is one of the reasons for the success of Rubinstein's model for applied work.

offers to the respondent $\frac{\delta}{1+\delta}$. The respondent, in turn, accepts any offer that gives him at least $\frac{\delta}{1+\delta}$ and rejects everything else.

The difference $\frac{1}{1+\delta} - \frac{\delta}{1+\delta} = \frac{1-\delta}{1+\delta}$ can be interpreted as the proposer's advantage and it is decreasing in δ. This is not surprising. When players are impatient, i.e., when δ is low, disagreement is costly and the respondent will be reluctant to reject the offer. As a result, the proposer will be able to extract a higher share of the total surplus.

Furthermore, the equilibrium characterization implies that an agreement is reached in the first period. To put it differently, costly disagreement never happens. This is a consequence of the fact that disagreement is costly for both players and that such cost is common knowledge between them.

15

Multistage Games with Incomplete Information

In this chapter we extend the analysis of static games with incomplete information to game forms with a multistage structure. As in Chapter 8, we first analyze rationalizability in games with payoff uncertainty and then we move on to the analysis of equilibria in Bayesian games. We observed in Chapters 9 and 12 that some Nash equilibria of multistage games are un-intuitive because the strategies of one or more players are "irrational" in some subgame. This led most game theorists to adopt subgame perfect equilibrium, a refinement of Nash equilibrium, as the standard solution concept. Similarly, there are Bayesian equilibria of multistage Bayesian games that prescribe "irrational" behavior at histories that cannot be reached in equilibrium. The reaction of most game theorists was to extend the subgame perfect equilibrium concept to multistage Bayesian games to refine Bayesian equilibrium and obtain a notion of "perfect Bayesian equilibrium." Unlike the complete-information case, however, they could not agree on a canonical definition, except for simple cases like leader-follower games. Here we provide the most general definition that satisfies some basic consistency requirements. Then we focus on a special class of leader-follower games, the signaling games, where all the definitions in the literature coincide with ours.

15.1 Multistage Games with Payoff Uncertainty

Consider a **multistage game tree** with observable actions (see Section 9.2 of Chapter 9), i.e., a structure

$$\left\langle I, (A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle .$$

The sets of histories H, Z and $\bar{H} = H \cup Z$, where H (respectively, Z) are non-terminal (respectively, terminal), are defined in the usual way.

A multistage environment with incomplete information is obtained by adding to the previous structure a set of states of nature and payoff functions that depend on the state of nature and the terminal history (cf. Section 8.1 of Chapter 8). A **multistage game with payoff uncertainty** and observable actions is given by the following mathematical structure

$$\hat{\Gamma} = \left\langle I, \Theta_0, (\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i)_{i\in I}\right\rangle ,$$

where $\left\langle I, (A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle$ is a multistage game tree, all the sets Θ_j ($j \in I \cup \{0\}$) are nonempty, Θ_0 is the space of residual uncertainty, and Θ_i the set of information-types of player i. As usual, we let $\Theta = \bigtimes_{i\in I\cup\{0\}} \Theta_i$. The outcome function $g : \Theta \times Z \to Y$ and the utility functions $v_i : \Theta_i \times Y \to \mathbb{R}$ depend on the vector of parameters $\theta \in \Theta$ which is not commonly known. For each $i \in I$, payoff function u_i is written in a parameterized form

$$\begin{array}{rccl} u_i : & \Theta \times Z & \to & \mathbb{R}, \\ & (\theta, z) & \mapsto & v_i\left(\theta_i, g\left(\theta, z\right)\right). \end{array}$$

If the set Θ is a singleton (or, more generally, if each payoff function u_i is independent of $\theta \in \Theta$), then the game is said to have **complete information**.

Note that the definition of $\hat{\Gamma}$ requires that each feasible set $\mathcal{A}_i(h)$ be independent of the vector of parameters $\theta \in \Theta$. In some applications it makes sense to allow the set of feasible actions of some player i at some history h to depend on θ_i. The above definition of $\hat{\Gamma}$ can be generalized as follows: for each $i \in I$, the feasibility correspondence $\mathcal{A}_i(\cdot,\cdot) : \Theta_i \times A^{<\mathbb{N}_0} \rightrightarrows A_i$ associates with each information-type θ_i and finite sequence of action profiles $h = (a^t)_{t=1}^{\ell}$ a set of actions $\mathcal{A}_i(\theta_i, h)$ that are feasible for θ_i immediately after h. Note, since player i knows θ_i he always knows his feasible actions. Such generalization would not change the substance of the results. Furthermore, to simplify the probabilistic analysis to come, we assume that the game is finite, i.e., $\bar{H}$ *and* Θ *are finite*.

In line with the genuine incomplete-information interpretation, we assume that there is *no ex ante stage*: players are "born" with their private information. Thus, we do not define strategies as choice rules that depend on private information and history, but rather as choice rules that only depend on history. As usual, we let $S_i = \times_{h \in H} \mathcal{A}_i(h)$ denote the set of pure strategies of player $i \in I$, and we let $S = \times_{i \in I} S_i$ denote the set of strategy profiles.

We obtained multistage games with payoff uncertainty by putting together ideas and concepts related to static games with incomplete information and multistage games with (observable actions and) complete information. Both are special cases of this larger class of games. This transition may appear so seamless that one risks neglecting an important new issue: if there is incomplete information, some player i does not know (θ_0 or) the information-types θ_{-i} of his co-players. If there were only one stage, it would be enough to posit *exogenous* (type-dependent) beliefs about such exogenous unknowns. But, if there are at least two stages, one has to consider i's beliefs about co-players' information-types conditional on their observed actions. The key issue is that such beliefs are necessarily *endogenous*, because—besides i's initial exogenous beliefs—they depend on what i thinks about the relationship between co-players' information-types and their behavior. In other words, player i has to update his beliefs about the information-types of others. To understand this issue, one first has to go back to probability theory and recall how it models beliefs about unknown parameters as they are updated upon observing new evidence.

15.2 Intermezzo: Bayes Rule

Consider an agent i who is initially uncertain about two things: a parameter $\theta \in \Theta$ and a variable (or a vector of variables) with realizations $x \in X$, where Θ and X are assumed to be *finite*. In a later stage i observes x. Even though i cannot observe θ, he is able to assess the probability of each realization $x \in X$ conditional on each $\theta \in \Theta$, which is denoted by $\mathrm{P}(x|\theta)$.

For example, consider an urn of unknown composition. It is only known that the urn contains between 1 and 10 balls, which may differ only in the color, *Black* or *White*. A first ball is drawn, its color is observed and then it is put back into the urn. Then a second ball is drawn (possibly the same as before) and its color is observed. Let θ denote the proportion of *White* balls in the urn. The set Θ of possible values of this parameter is finite:

$$\Theta = \bigcup_{n=1}^{10} \left\{ \theta : \exists k \in \{0, ..., n\}, \theta = \frac{k}{n} \right\}.$$

Let *Black* draws correspond to number 0 (since black is the absence of light) and *White* draws correspond to number 1. Then $X = \{0,1\}^2$ and the probability of $x = (x_1, x_2)$ conditional on the proportion of white balls being θ is

$$\mathrm{P}(x|\theta) := \theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2}.$$

For example, $\mathrm{P}((1,0)|\theta) = \theta(1-\theta)$.

Agent i also assigns a *subjective probability* $\mathrm{P}(\theta)$ to all the possible values of θ. Note that $\mathrm{P}(x|\theta)$ is well defined even if i assigns probability 0 to θ. For example, for some reason i may be certain that the urn does not contain more than 5 balls, and hence $\mathrm{P}\big(\theta = \frac{1}{10}\big) = 0$. Yet i thinks that if θ *were* $\frac{1}{10}$ then the probability of $x = (0,0)$ *would be* $\frac{9}{10} \cdot \frac{9}{10}$, that is, $\mathrm{P}\big((0,0)|\frac{1}{10}\big) = \frac{81}{100}$.

Fix any finite uncertainty space Ω (later, we will relate Ω to Θ and X). The law of conditional probabilities says that for every pair of events $E, F \subseteq \Omega$,

$$\mathrm{P}(F) > 0 \Rightarrow \mathrm{P}(E|F) = \frac{\mathrm{P}(E \cap F)}{\mathrm{P}(F)}.$$

Note that the same condition can be written more compactly as

$$\mathrm{P}(E \cap F) = \mathrm{P}(E|F)\mathrm{P}(F).$$

Here we may assume that the space of uncertainty is $\Omega = \Theta \times X$. Each realization x corresponds to event $E = \Theta \times \{x\}$, each parameter value θ corresponds to event $F = \{\theta\} \times X$, each pair (θ, x) is the singleton $\{(\theta, x)\}$. Then the joint probability of the pair (θ, x) is

$$\mathrm{P}(\theta, x) = \mathrm{P}(x|\theta)\mathrm{P}(\theta).$$

Note that the collection of events $\{F \subseteq \Omega : \exists \theta \in \Theta, F = \{\theta\} \times X\}$ forms a partition of $\Omega = \Theta \times X$. This partition of Ω induces a corresponding partition of every event $E \subseteq \Omega$, and in particular of events of the form $E = \Theta \times \{x\}$. Then, we can compute the probability of every x (that is, event $\Theta \times \{x\}$) starting from the probabilities of the "cells" $\{(\theta, x)\} = (\Theta \times \{x\}) \cap (\{\theta\} \times X)$, where $\theta \in \Theta$. In turn, these probabilities can be obtained from the conditional and prior probabilities $\mathrm{P}(x|\theta)$ and $\mathrm{P}(\theta)$, $\theta \in \Theta$. Thus, we obtain the formula expressing the marginal, or **predictive probability** of x as

$$\mathrm{P}(x) = \sum_{\theta' \in \Theta} \mathrm{P}(x, \theta') = \sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta'). \tag{15.2.1}$$

The problem is to derive from these elements $\mathrm{P}(\theta|x)$, the probability that i would assign to each $\theta \in \Theta$ upon observing evidence $x \in X$. There are two possibilities, either $\mathrm{P}(x) = 0$ or $\mathrm{P}(x) > 0$.

If $\mathrm{P}(x) = 0$, then $\mathrm{P}(\theta|x)$ cannot be derived from the previous data. *This does not mean that i is unable to assess the conditional probability* $\mathrm{P}(\theta|x)$, it only means that $\mathrm{P}(\theta|x)$ is not determined by the other probabilistic assessments expressed above.

If $\mathrm{P}(x) > 0$, then $\mathrm{P}(\theta|x) = \frac{\mathrm{P}((\theta,x))}{\mathrm{P}(x)}$. Substituting $\mathrm{P}(x)$ with the expression given by (15.2.1) we obtain

$$\mathrm{P}(\theta|x) = \frac{\mathrm{P}(x|\theta)\mathrm{P}(\theta)}{\sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta')}. \tag{15.2.2}$$

Eq. (15.2.2) is known as **Bayes Formula,** that is, an equation expressing $\mathrm{P}(\theta|x)$ as a function of the conditional probabilities $\mathrm{P}(x|\theta')$ and the prior probabilities $\mathrm{P}(\theta')$, with $\theta' \in \Theta$.

We call **Bayes Rule** the rule that says that whenever (15.2.2) *can* be applied, then it *must* be applied. Since (15.2.2) can be applied if and only if $\mathrm{P}(x) > 0$, we may write Bayes Rule in the following compact form:

$$\forall x \in X, \forall \theta \in \Theta,\ \mathrm{P}(\theta|x) \left(\sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta') \right) = \mathrm{P}(x|\theta)\mathrm{P}(\theta). \tag{15.2.3}$$

Bayes Rule is *not violated* if either $\mathrm{P}(x) = 0$ (both sides of (15.2.3) are zero), or $\mathrm{P}(x) > 0$ and $\mathrm{P}(\theta|x)$ is computed with (15.2.2). Therefore, whenever an assessment of prior and conditional probabilities satisfies (15.2.3), we say that it is **consistent with Bayes rule**. In particular, Bayes rule "holds" if $\mathrm{P}(x) = 0$: another way to express this point is that if the antecedent in the material implication

$$\mathrm{P}(x) > 0 \Rightarrow \mathrm{P}(\theta|x) = \frac{\mathrm{P}((\theta, x))}{\sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta')}$$

is false, then the material implication holds.[1]

15.3 Rational Planning with Incomplete Information

Before we move on to the analysis of solution concepts for multistage games with incomplete information, we have to extend the analysis of rational planning of Chapter 10. Specifically, we need to take into account that player i, given his (information-)type θ_i, is uncertain about the types θ_{-i} of the co-players. To simplify the notation, we assume that *there is no residual uncertainty*, that is, Θ_0 is a *singleton*, and we remove θ_0 from formulas.[2]

Definition 77. *Fix player i and type θ_i in a multistage game with payoff uncertainty $\hat{\Gamma}$. A* ***conjecture*** *is an array of probability measures*

$$\beta^i = (\beta^i(\cdot|\theta_{-i}, h))_{\theta_{-i} \in \Theta_{-i}, h \in H} \in \left(\bigtimes_{h \in H} \Delta(\mathcal{A}_{-i}(h)) \right)^{\Theta_{-i}}.$$

A ***personal system of beliefs*** *of type θ_i is an array of probability measures*

$$\mu_i\left(\cdot|\theta_i, \cdot\right) = (\mu_i(\cdot|\theta_i, h))_{h \in H} \in \Delta(\Theta_{-i})^H.$$

A pair $\left(\beta^i, \mu_i\left(\cdot|\theta_i, \cdot\right)\right)$ is called ***personal assessment*** *of type θ_i.*

As with complete-information games, to ease intuition, it is better to start thinking about two-person games, so that $-i$ is the co-player other than i. In this case, $\beta^i(\cdot|\theta_{-i}, \cdot) \in \bigtimes_{h \in H} \Delta(\mathcal{A}_{-i}(h))$ is like a behavior strategy of $-i$. Thus, this is a generalization of the notion of conjecture in

[1] The material implication $p \Rightarrow q$ is verified if either p is false, or both p and q are true.

[2] Alternatively, we could interpret θ_{-i} as including θ_0.

two-person static games with complete information, where we represented conjectures about $-i$ and mixed actions of $-i$ with the same mathematical object, a probability measure over A_{-i}. With more than two players, for some h, $\beta^i(\cdot|\theta_{-i}, h) \in \Delta(\mathcal{A}_{-i}(h))$ may be a correlated proability measure over $\times_{j \neq i} \mathcal{A}_j(h)$. In the analysis of perfect Bayesian equilibrium of Section 15.7, we will assume that each $\beta^i(\cdot|\theta_{-i}, h)$ is the product of the marginal measures (mixed actions) $\beta_j(\cdot|\theta_j, h) \in \Delta(\mathcal{A}_j(h))$, where $\beta_j(\cdot|\theta_j, \cdot)$ is the equilibrium behavior strategy of j. Thus, it will make sense to write $\beta^i = \beta_{-i}$.

As a matter of interpretation, conjecture β^i is like a family of statistical models that type θ_i deems possible. This family is parameterized by θ_{-i}, the unknown type profile of the co-players (cf. Section 15.2). Thus, $\beta^i(a_{-i}|\theta_{-i}, h)$ is the probability of a_{-i} given θ_{-i} and conditional on "evidence" h, $\mu_i(\theta_{-i}|\theta_i, h)$ is the probability assigned by type θ_i of i to θ_{-i} conditional on observing h, and $\mu_i(\cdot|\theta_i, \varnothing)$ is the exogenous initial belief of type θ_i.[3]

Given $\hat{\Gamma}$ and type θ_i (hence, payoff function $u_{i,\theta_i} : \Theta_{-i} \times Z \to \mathbb{R}$), personal assessment $\left(\beta^i, \mu_i(\cdot|\theta_i, \cdot)\right)$ yields a **subjective decision tree**. The relevant conditional probabilities are derived as follows.

Suppose that the plan of type θ_i is described by behavior strategy $\beta_i(\cdot|\theta_i, \cdot)$ and fix a personal assessment $\left(\beta^i, \mu_i(\cdot|\theta_i, \cdot)\right)$. Let

$$\begin{aligned} \mathbb{P}^{\beta}(a|\theta_i, \theta_{-i}, h) &= \beta_i(a_i|\theta_i, h)\beta^i(a_{-i}|\theta_{-i}, h), \qquad (15.3.1) \\ \mathbb{P}^{\beta^i,\mu_i}(a_{-i}|\theta_i, h) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta^i(a_{-i}|\theta_{-i}, h)\mu_i(\theta_{-i}|\theta_i, h), \\ \mathbb{P}^{\beta,\mu_i}(a|\theta_i, h) &= \sum_{\theta_{-i}\in\Theta_{-i}} \mathbb{P}^{\beta}(a|\theta_i, \theta_{-i}, h)\mu_i(\theta_{-i}|\theta_i, h) \end{aligned}$$

for all $h \in H$, $a = (a_i, a_{-i}) \in \mathcal{A}(h)$, and $\theta_{-i} \in \Theta_{-i}$. We can use the chain rule to obtain the probability of history $h' = \left(a^1, ..., a^{\ell(h')}\right)$ conditional on prefix $h = \left(a^1, ..., a^{\ell(h)}\right)$ (with $\ell(h) < \ell(h')$) given type profile $\theta = (\theta_i, \theta_{-i})$:

$$\mathbb{P}^{\beta}\left(h'|\theta, h\right) = \prod_{t=\ell(h)+1}^{\ell(h')} \mathbb{P}^{\beta}\left(a^t|\theta, \left(h, ..., a^{t-1}\right)\right), \qquad (15.3.2)$$

[3]According to the interpretation given above, $\mu_i(\cdot|\theta_i, \varnothing)$ is the "prior" of type θ_i in the sense of Bayesian statistics, and each $\mu_i(\cdot|\theta_i, h)$ (with $h \neq \varnothing$) is the "posterior" of θ_i given evidence h.

with the convention that $\mathbb{P}^{\beta}\left(a^t|\theta,\left(h,...,a^{t-1}\right)\right) = \mathbb{P}^{\beta}\left(a^t|\theta,h\right)$ if $t-1 = \ell(h)$. Consistency with Bayes rule requires that, for all type profiles $\theta_{-i} \in \Theta_{-i}$ and histories $(h,a) \in H$,

$$\mathbb{P}^{\beta,\mu_i}(a|\theta_i,h) > 0 \Rightarrow \mu_i(\theta_{-i}|\theta_i,(h,a)) = \frac{\mathbb{P}^{\beta}(a|\theta,h)\mu_i(\theta_{-i}|\theta_i,h)}{\mathbb{P}^{\beta,\mu_i}(a|\theta_i,h)}. \quad (15.3.3)$$

(Eq. (15.3.3) is a "one-step" Bayes rule; given eq.s (15.3.1) and the chain rule (15.3.2), eq. (15.3.3) implies the "multi-step" version of Bayes rule.) Eq.s (15.3.1) and eq. (15.3.3) yield

$$\mathbb{P}^{\beta,\mu_i}(a|\theta_i,h) = \beta_i(a_i|\theta_i,h)\mathbb{P}^{\beta^i,\mu_i}(a_{-i}|\theta_i,h).$$

Hence, for every $a_i \in \mathrm{supp}\beta_i(\cdot|\theta_i,h)$, eq. (15.3.3) yields a version of Bayes rule whereby the belief about the co-players' types depends on what co-players just did:

$$\mathbb{P}^{\beta^i,\mu_i}(a_{-i}|\theta_i,h) > 0 \Rightarrow \mu_i(\theta_{-i}|\theta_i,(h,(a_i,a_{-i}))) = \frac{\beta^i(a_{-i}|\theta_{-i},h)\mu_i(\theta_{-i}|\theta_i,h)}{\mathbb{P}^{\beta^i,\mu_i}(a_{-i}|\theta_i,h)},$$

or, equivalently,

$$\mu_i(\theta_{-i}|\theta_i,(h,(a_i,a_{-i})))\mathbb{P}^{\beta^i,\mu_i}(a_{-i}|\theta_i,h) = \beta^i(a_{-i}|\theta_{-i},h)\mu_i(\theta_{-i}|\theta_i,h). \quad (15.3.4)$$

Thus, $\mu_i(\theta_{-i}|\theta_i,(h,(a_i,a_{-i})))$ is *independent* of own action a_i within subset $\mathrm{supp}\beta_i(\cdot|\theta_i,h) \subseteq \mathcal{A}_i(h)$. The following definition of Bayes consistency requires that eq. (15.3.4) holds within the whole set $\mathcal{A}_i(h)$, not just the subset $\mathrm{supp}\beta_i(\cdot|\theta_i,h)$.

Definition 78. *Personal assessment $\left(\beta^i,\mu_i\left(\cdot|\theta_i,\cdot\right)\right)$ is **Bayes consistent** if (15.3.4) holds for all $\theta_{-i} \in \Theta_{-i}$ and $(h,(a_i,a_{-i})) \in H$.*

Under personal Bayes consistency, *known results about rational planning—such as the OD principle—extend to the incomplete information case.* A detailed and formal analysis is provided in the appendix of this chapter. Here we only specify the elements to be used in the equilibrium analysis below.

To better connect with the analysis of perfect Bayesian equilibria of Section 15.7, we consider the possibility that i's plan, which depends on his type θ_i, is a behavior strategy $\beta_i\left(\cdot|\theta_i,\cdot\right) \in B_i = \times_{h\in H}\Delta(\mathcal{A}_i(h))$. To ease notation, in the remainder of this section we let $\beta_i = \beta_i(\cdot|\theta_i, \cdot)$ and $\mu_i = \mu_i(\cdot|\theta_i, \cdot) \in \Delta(\Theta_{-i})^H$ respectively denote the behavior strategy and personal system of beliefs *of a given type* θ_i of player i.[4] Thus, we fix a type θ_i who plans according to behavior strategy β_i and has personal assessment (β^i, μ_i), and we first define the value for θ_i of using β_i starting from history h *given* a type profile θ_{-i}:

$$V_{\theta_i}^{\beta_i,\beta^i}(\theta_{-i}, h) = \sum_{z \in Z(h)} \mathbb{P}^{\beta_i,\beta^i}(z|\theta_i, \theta_{-i}, h)\, u_i(\theta_i, \theta_{-i}, z).$$

Since type θ_i does not know θ_{-i}, to obtain the value of using β_i from h he must use the system of beliefs μ_i:

$$V_{\theta_i}^{\beta_i,\beta^i,\mu_i}(h) = \sum_{\theta_{-i} \in \Theta_{-i}} \mu_i(\theta_{-i}|\theta_i, h)\, V_{\theta_i}^{\beta_i,\beta^i}(\theta_{-i}, h).$$

Similarly, the value of choosing action a_i at h with future behavior given by β_i is

$$\begin{aligned} & V_{\theta_i}^{\beta_i,\beta^i,\mu_i}(h, a_i) \\ = & \sum_{\theta_{-i} \in \Theta_{-i}} \mu_i(\theta_{-i}|\theta_i, h) \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h)\, V_{\theta_i}^{\beta_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))). \end{aligned}$$

Note that these values depend only on the specification of β_i and β^i at histories that weakly follow h.

Definition 79. *Fix a type θ_i with personal assessment (β^i, μ_i) and a behavior strategy $\bar{\beta}_i \in B_i$. We say that $\bar{\beta}_i$ is*

- ***one-step optimal given** (β^i, μ_i) if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{\beta}_i,\beta^i,\mu_i}(h) = \max_{a_i \in \mathcal{A}_i(h)} V_{\theta_i}^{\bar{\beta}_i,\beta^i,\mu_i}(h, a_i)\,; \qquad (15.3.5)$$

- ***sequentially optimal given** (β^i, μ_i) if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{\beta}_i,\beta^i,\mu_i}(h) = \max_{\beta_i \in B_i} V_{\theta_i}^{\beta_i,\beta^i,\mu_i}(h). \qquad (15.3.6)$$

[4] This makes sense because feasibility constraints and hard information about co-players are type-independent. Hence, each type of i has the same set of behavior strategies B_i and the same exogenous uncertainty space Θ_{-i}.

The following extension of the OD principle is proved in the appendix.

Proposition 19. (OD principle with incomplete information) *Fix a type* θ_i *with personal assessment* (β^i, μ_i)*. If* (β^i, μ_i) *is Bayes consistent, then one-step optimality given* (β^i, μ_i) *is equivalent to sequential optimality given* (β^i, μ_i).

The following example illustrates how the equivalence between one-step and sequential optimality depends on Bayes consistency.

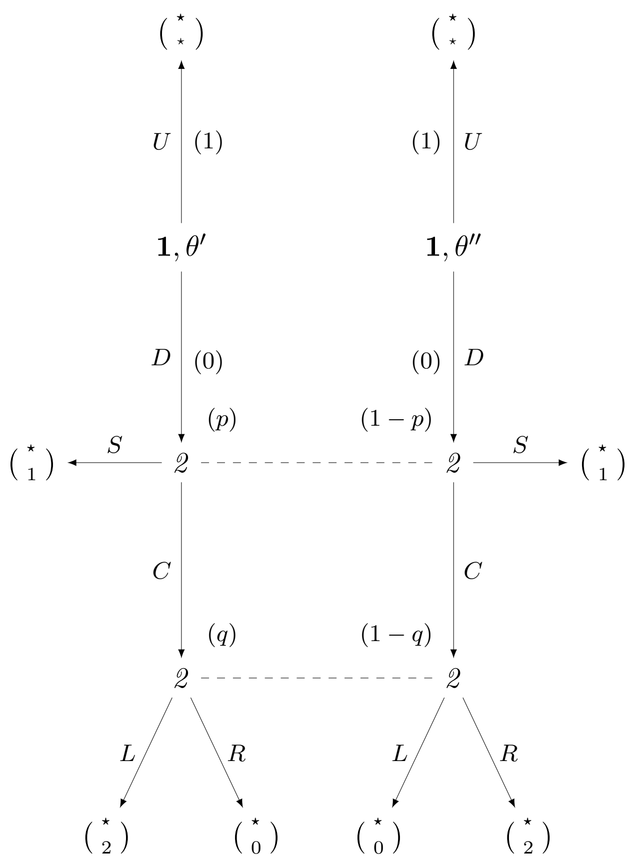


Figure 15.1: A subjective decision tree for player 2.

Example 63. Consider the subjective decision tree for player 2 depicted in Figure 15.1. Player 1 knows the true state of nature (his type) and moves first, choosing between U (up) and D (down). Player 2 has no private information, and his assessment is represented by the numbers in parentheses. The conjecture of 2 is given by $\beta^2(U|\theta') = 1 = \beta^2(U|\theta'')$. Since he initially expects player 1 to go up (U) with probability 1, his initial belief about θ cannot affect his conditional beliefs about θ given (D) and (D,C), because they cannot be pinned down by Bayes rule. Hence, the initial belief is not shown. The system of beliefs is given by $\mu_2(\theta'|D) = p$ and $\mu_2(\theta'|D,C) = q$. This assessment of player 2 is Bayes consistent if and only if $p = q$, because eq. (15.3.4) implies that his belief conditional on (D,C) cannot depend on the fact the he chose C instead of S. Note that the expected value of action L (respectively R) given (D,C) is $2q$ (respectively $2(1-q)$). Given Bayes consistency, if $p > 1/2$ (respectively, $p < 1/2$) the unique sequentially optimal strategy of 2 is also the only one that satisfies one-step optimality, $C.L$ (respectively, $C.R$). Now suppose that Bayes consistency is violated because $p < 1/2$ and $q > 1/2$. Then the unique strategy that satisfies one-step optimality is $S.L$: indeed, $2q > 2(1-q)$ which justifies going left (L) after (D,C); furthermore, player 2 *predicts* that he would go left after (D,C), thus the expected value of taking action C given D is $2p < 1$, which justifies S after D. With this inconsistent assessment player 2 changes his beliefs about θ and therefore his preferences over continuation strategies. This makes it impossible to satisfy sequential optimality, which requires to continue and go right $(C.R)$ after D and to go left (L) after (D,C). ▲

15.3.1 Justifiability and Dominance

In Chapter 8 we modeled conjectures as probability measures $\mu^i \in \Delta(\Theta_{-i} \times A_{-i})$, we considered the sets $r_i(\mu^i, \theta_i)$ of best replies to conjectures μ^i for types θ_i, we called an action a_i "justifiable" for a type θ_i if $a_i \in r_i(\mu^i, \theta_i)$ for some $\mu^i \in \Delta(\Theta_{-i} \times A_{-i})$, and we related justifiability and dominance. In Chapter 10 we defined conditional probability systems (CPSs) $\mu^i \in \Delta^H(S_{-i})$, we considered the sets of weakly sequential best replies $r_i(\mu^i)$ to such CPSs, we called a strategy s_i "justifiable" if $s_i \in r_i(\mu^i)$ for some $\mu^i \in \Delta^H(S_{-i})$, and we related justifiability to conditional dominance. The reason to consider *weak* sequential optimality in the definition of justifiability was that it is invariant to behavioral equivalence, and we took the perspective of an external observer, or a co-player, who wonders whether i may exhibit some pattern of behavior s_i

and who cannot distinguish, nor cares to distinguish between behaviorally equivalent strategies $s_i' \approx_i s_i''$, because they are necessarily realization-equivalent (Lemma 24). Merging these ideas, here we represent how player i would update or revise his beliefs about the co-players as the play unfolds with CPSs on $\Theta_{-i} \times S_{-i}$, we define justifiability for a type by means of such CPSs, and we relate it to conditional dominance for a type.

Let

$$\mathcal{H}_{-i} = \{C_{-i} \subseteq \Theta_{-i} \times S_{-i} : \exists h \in H, C_{-i} = \Theta_{-i} \times S_{-i}(h)\}.$$

A CPS on $\Theta_{-i} \times S_{-i}$ is an array of conditional probability measures

$$\bar{\mu}^i = \left(\bar{\mu}^i(\cdot|C_{-i})\right)_{C_{-i}\in\mathcal{H}_{-i}} \in \bigtimes_{C_{-i}\in\mathcal{H}_{-i}} \Delta(C_{-i})$$

such that the chain rule holds, that is, for all $C_{-i}, D_{-i} \in \mathcal{H}_{-i}$, and $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$,

$$E_{-i} \subseteq D_{-i} \subseteq C_{-i} \Rightarrow \bar{\mu}^i(E_{-i}|C_{-i}) = \bar{\mu}^i(E_{-i}|D_{-i})\,\bar{\mu}^i(D_{-i}|C_{-i}).$$

As in Section 10.4 of Chapter 10, whenever convenient we ease notation by writing $\bar{\mu}^i(\cdot|h) = \bar{\mu}^i(\cdot|\Theta_{-i} \times S_{-i}(h))$, and we let $\Delta^H(\Theta_{-i} \times S_{-i})$ denote the set of CPSs on $\Theta_{-i} \times S_{-i}$. With this, it should be always kept in mind that if h' and h'' differ only because of actions taken by player i, so that $S_{-i}(h') = S_{-i}(h'')$, then $\bar{\mu}^i(\cdot|h') = \bar{\mu}^i(\cdot|h'')$.

Note that CPSs can be related to Bayes consistent personal assessments. Recall that

$$S_{-i}(h, a_{-i}) = \{s_{-i} \in S_{-i}(h) : s_{-i}(h) = a_{-i}\}$$

is the set of strategies of others allowing h and selecting a_{-i} given h.

Definition 80. *Personal assessment* $\left(\beta^i, \mu_i\right)$ *is* ***consistent with*** *CPS* $\bar{\mu}^i$ *if, for all* $h \in H$, $\theta_{-i} \in \Theta_{-i}$, *and* $a_{-i} \in \mathcal{A}_{-i}(h)$,

$$\beta^i(a_{-i}|\theta_{-i}, h)\,\mu_i(\theta_{-i}|h) = \bar{\mu}^i(\{\theta_{-i}\} \times S_{-i}(h, a_{-i})\,|h). \tag{15.3.7}$$

Note that, taking the summation with respect to co-players' actions in eq. (15.3.7), we obtain

$$\begin{aligned}\mu_i(\theta_{-i}|h) &= \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h)\,\mu_i(\theta_{-i}|h)\\ &= \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h,a_{-i})\,|h)\\ &= \bar{\mu}^i\left(\{\theta_{-i}\}\times\left(\cup_{a_{-i}\in\mathcal{A}_{-i}(h)}S_{-i}(h,a_{-i})\right)|h\right)\\ &= \bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h)\,|h),\end{aligned}$$

because $\sum_{a_{-i}\in\mathcal{A}_{-i}(h)}\beta^i(a_{-i}|\theta_{-i},h) = 1$ and $\cup_{a_{-i}\in\mathcal{A}_{-i}(h)}S_{-i}(h,a_{-i}) = S_{-i}(h)$. Furthermore, this implies that if $\left(\beta^i,\mu_i\right)$ is consistent with $\bar{\mu}^i$, then

$$\bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h)\,|h) > 0 \Rightarrow \beta^i(a_{-i}|\theta_{-i},h) = \frac{\bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h,a_{-i})\,|h)}{\bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h)\,|h)}.$$

Remark 55. *For each CPS $\bar{\mu}^i \in \Delta^H(\Theta_{-i}\times S_{-i})$, there is a* Bayes consistent *personal assessment that is consistent with $\bar{\mu}^i$.*

This observation is important because it shows that the only parts of a personal assessment $\left(\beta^i,\mu_i\right)$ consistent with a CPS $\bar{\mu}^i$ that are not determined by eq. (15.3.7) are the conditional probabilities of actions $\beta^i(a_{-i}|\theta_{-i},h)$ given types deemed negligible conditional on h, and such probabilities do not determine the conditional values of carrying out strategies. Thus, we may as well define such values by means of CPSs: for all $\theta_i\in\Theta_i$, $s_i\in S_i$, $\bar{\mu}^i\in\Delta^H(\Theta_{-i}\times S_{-i})$, $h\in H_i(s_i)$, and $z\in Z(h)$, let

$$\mathbb{P}^{s_i,\bar{\mu}^i}(\theta_{-i},z|h) = \begin{cases} 0, & \text{if } s_i\notin S_i(z),\\ \bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(z)\,|h), & \text{if } s_i\in S_i(z),\end{cases}$$

$$V_{\theta_i}^{s_i,\bar{\mu}^i}(h) = \sum_{(\theta_{-i},z')\in\Theta_{-i}\times Z(h)} u_i\left(\theta_i,\theta_{-i},z'\right)\mathbb{P}^{s_i,\bar{\mu}^i}(\theta_{-i},z'|h).$$

Definition 81. *Strategy s_i^* is **weakly sequentially optimal for type** θ_i given CPS $\bar{\mu}^i$, written $s_i^*\in r_i\left(\bar{\mu}^i,\theta_i\right)$, if*

$$\forall h\in H_i(s_i^*),\ V_{\theta_i}^{s_i^*,\bar{\mu}^i}(h) = \max_{s_i\in S_i(h)} V_{\theta_i}^{s_i,\bar{\mu}^i}(h).$$

*Strategy s_i^* is **justifiable for type** θ_i if there exists some CPS $\bar{\mu}^i\in\Delta^H(\Theta_{-i}\times S_{-i})$ such that $s_i^*\in r_i\left(\bar{\mu}^i,\theta_i\right)$.*

As in Chapter 10, we can give an analogous definition of weak sequential optimality given a Bayes consistent personal assessment and use Remark 55 to show that it is equivalent to the foregoing definition with CPSs (cf. Proposition 5).

For any fixed type θ_i, it is convenient to consider the section $r_{i,\theta_i} : \Delta^H(\Theta_{-i} \times S_{-i}) \rightrightarrows S_i$ at θ_i of the weak sequential optimality correspondence $r_i : \Delta^H(\Theta_{-i} \times S_{-i}) \times \Theta_i \rightrightarrows S_i$, that is, $r_{i,\theta_i}(\bar{\mu}^i) = r_i(\bar{\mu}^i, \theta_i)$ for all $\bar{\mu}^i$. With this, the set of weakly sequentially optimal strategies for type θ_i is $r_{i,\theta_i}(\Delta^H(\Theta_{-i} \times S_{-i}))$.

Using essentially the same arguments as in Chapter 10, we can relate justifiability and conditional dominance. To ease notation, let us define the parameterized normal-form (or strategic-form) payoff function

$$\begin{array}{rccl} \hat{U}_i : & \Theta \times S & \to & \mathbb{R} \\ & (\theta, s) & \mapsto & u_i(\theta, \zeta(s)). \end{array}$$

For any mixed strategy $\sigma_i \in \Delta(S_i)$, let

$$\hat{U}_i(\theta, \sigma_i, s_{-i}) = \mathbb{E}_{\sigma_i}\left(\hat{U}_{i,\theta,s_{-i}}\right) = \sum_{s_i \in S_i} \hat{U}_i(\theta, s_i, s_{-i})\, \sigma_i(s_i).$$

Definition 82. *Strategy $\bar{s}_i$ is* ***conditionally dominated for type*** *θ_i if there are $h \in H_i(\bar{s}_i)$ and $\sigma_i \in \Delta(S_i(h))$ such that*

$$\forall (\theta_{-i}, s_{-i}) \in \Theta_{-i} \times S_{-i}(h),\ \hat{U}_i(\theta_i, \theta_{-i}, \bar{s}_i, s_{-i}) < \hat{U}_i(\theta_i, \theta_{-i}, \sigma_i, s_{-i}).$$

The set of pairs (s_i, θ_i) such that s_i is not conditionally dominated for θ_i is denoted NCD_i, and we let $NCD = \times_{i \in I} NCD_i$.

Lemma 33. *A strategy is not conditionally dominated for a type if and only if it is justifiable for that type, that is, for each $i \in I$,*

$$NCD_i = \bigcup_{\theta_i \in \Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H(\Theta_{-i} \times S_{-i})\right).$$

15.4 Rationalizability

In Section 8.2 of Chapter 8 we analyzed rationalizability in static games with payoff uncertainty. In Chapter 11 we analyzed initial and strong rationalizability in multistage games with complete information, two legitimate extensions of complete-information rationalizability characterizing the behavioral implications of a weaker and a stronger version of the idea of rationality and common belief in rationality. Here we can take stock of those analyses and merge them to obtain an analysis of rationalizability in multistage games with payoff uncertainty. Similarly to Chapter 11, the behavioral implication of mere rationality is justifiability. As in Section 15.3, to ease notation we assume that there is no residual uncertainty, that is, Θ_0 is a singleton and can be neglected.[5]

15.4.1 Initial Rationalizability

Initial rationalizability characterizes the behavioral implications of Rationality and Common Initial Belief in Rationality. Following the same template of Chapters 8 and 11, we first extend the rationalization operator to multistage games with payoff uncertainty.

Let $\mathcal{C}$ denote the collection of Cartesian subsets $C = \times_{i \in I} C_i$ with $C_i \subseteq \Theta_i \times S_i$ for each $i \in I$. We interpret a set $C \in \mathcal{C}$ as type-dependent restrictions on behavior implied by strategic reasoning and we assume that each player *initially believes* that such restrictions hold for the co-players.[6] Let

$$\Delta_{\varnothing}^{H}\left(C_{-i}\right) = \left\{\bar{\mu}^{i} \in \Delta^{H}\left(\Theta_{-i} \times S_{-i}\right) : \bar{\mu}^{i}\left(C_{-i} | \varnothing\right) = 1\right\}$$

denote the set of CPSs that initially assign probability 1 to C_{-i}. The initial rationalization operator $\rho : \mathcal{C} \to \mathcal{C}$ is the self-map defined as follows: for every $C \in \mathcal{C}$,

$$\begin{aligned} \rho\left(C\right) &= \mathop{\times}_{i \in I} \left\{\left(\theta_i, s_i\right) : \exists \bar{\mu}^{i} \in \Delta_{\varnothing}^{H}\left(C_{-i}\right), s_i \in r_i\left(\bar{\mu}^{i}, \theta_i\right)\right\} \\ &= \mathop{\times}_{i \in I} \left(\bigcup_{\theta_i \in \Theta_i} \left\{\theta_i\right\} \times r_{i,\theta_i}\left(\Delta_{\varnothing}^{H}\left(C_{-i}\right)\right)\right). \end{aligned}$$

It can be verified that ρ is *monotone*, because $C'_{-i} \subseteq C_{-i}$ implies $\Delta_{\varnothing}^{H}\left(C'_{-i}\right) \subseteq \Delta_{\varnothing}^{H}\left(C_{-i}\right)$. Thus, iterating ρ starting from the whole set of

[5] Analyzing the more general case does not create any conceptual difficulty.

[6] The restrictions given by C_i may be type-dependent because the sections of C_i at different types θ'_i and θ''_i (C_{i,θ'_i} and C_{i,θ''_i}) may be different. Furthermore, it makes sense to require that $\mathrm{proj}_{\Theta_i} C_i = \Theta_i$, because types are exogenous and cannot be deleted as such. But this is not strictly necessary for our analysis.

profiles of types and strategies we obtain a (weakly) decreasing sequence of subsets. To ease notation, we write $\Theta \times S$ instead of $\bigtimes_{i\in I}(\Theta_i \times S_i)$, given the obvious isomorphism between the two sets. Similarly, for games with more than two players, we identify $\Theta_{-i} \times S_{-i}$ with $\bigtimes_{j\neq i}(\Theta_j \times S_j)$. With this, we consider the sequence $(\rho^n(\Theta \times S))_{n\in\mathbb{N}}$, where

$$\rho^1(\Theta \times S) = \bigtimes_{i\in I}\left(\bigcup_{\theta_i\in\Theta_i}\{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H(\Theta_{-i} \times S_{-i})\right)\right)$$

contains all the profiles $(\theta_i, s_i)_{i\in I}$ such that s_i is justifiable for θ_i, thus representing the type-dependent behavioral implications of rationality. Since $(\rho^n(\Theta \times S))_{n\in\mathbb{N}}$ is decreasing, we can define

$$\rho^\infty(\Theta \times S) = \bigcap_{n\in\mathbb{N}}\rho^n(\Theta \times S).$$

By finiteness of $\Theta \times S$, the limit $\rho^\infty(\Theta \times S)$ is attained after finitely many iterations.

Definition 83. *A profile of types and strategies $(\theta_i, s_i)_{i\in I}$ is **initially rationalizable** if $(\theta_i, s_i)_{i\in I} \in \rho^\infty(\Theta \times S)$.*

Leveraging on the monotonicity of ρ, one can prove an analog of Theorem 23. We postpone a discussion of this to Section 15.4.4, where we present Directed Rationalizability. As in Proposition 7 of Chapter 11, one can characterize initial rationalizability by means of iterated dominance. For each Cartesian set $C \in \mathcal{C}$, let $\mathrm{ND}(C) \subseteq C$ denote the set of profiles $(\theta_i, s_i)_{i\in I} \in C$ such that s_i is not dominated within C for θ_i by any mixed strategy $\sigma_i \in \Delta(C_{i,\theta_i})$, where $C_{i,\theta_i} = \{s'_i \in S_i : (\theta_i, s'_i) \in C_i\}$ is the section at θ_i of C_i. In other words, $\mathrm{ND}(C)$ is the set of non-dominated profiles in the restricted strategic form $\left\langle I, (C_i, U_i|_C)_{i\in I}\right\rangle$. The restriction operator ND: $\mathcal{C} \to \mathcal{C}$ is a self-map that can be iterated starting from any $C \in \mathcal{C}$. Note that, by Lemma 33, $\rho(\Theta \times S) = NCD$. Thus, it makes sense to start the iteration from the set NCD of profiles of types and strategies such that the latter are not conditionally dominated for the given types. The proof of the following result is similar to the proof of Proposition 7.

Proposition 20. *$\rho^\infty(\Theta \times S) = \mathrm{ND}^\infty(NCD)$.*

In words, initial rationalizability can be computed by first eliminating (for every player i) each (θ_i, s_i) such that s_i is conditionally dominated for θ_i, and then iteratively eliminating pairs (θ_i, s_i) such that s_i is dominated for θ_i in the restricted strategic form obtained from the previous eliminations. The reason why conditional dominance is used only in the first step is the same as in Chapter 11: initial rationalizability allows "surprised" players to abandon their initial belief in the rationality of the co-players even if the co-players' observed behavior is consistent with rationality.

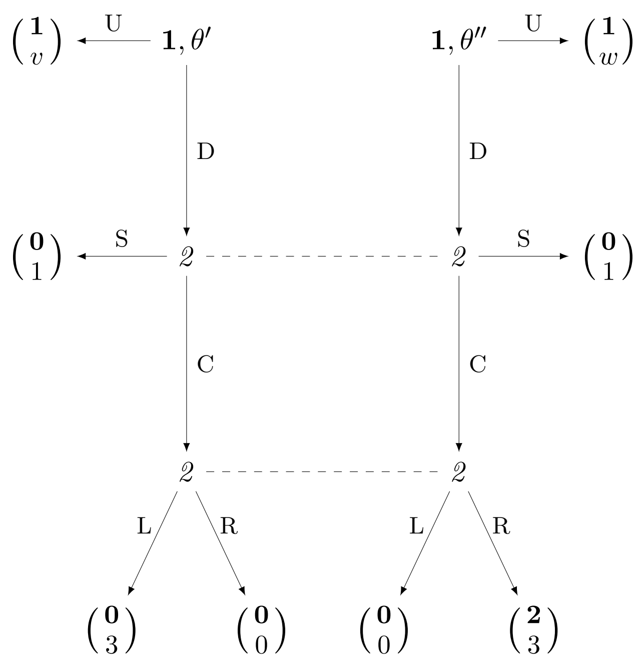


Figure 15.2: A 3-stage game with payoff uncertainty.

Example 64. In the multistage game with payoff uncertainty depicted in Figure 15.2, D is dominated by U for type θ' of player 1, but not for type θ''. The reduced strategy S is dominated for (the unique type of) player 2 by mixed strategy $\frac{1}{2}\delta_{\text{C.L}} + \frac{1}{2}\delta_{\text{C.R}}$ conditional on D. The remaining two strategies C.L and C.R are justifiable. For example, any $\bar{\mu}^2$ such that

$\bar{\mu}^2(\theta', \mathrm{D}|\mathrm{D}) = \frac{1}{2}$ justifies both. Thus,

$$\rho^1(\Theta \times S) = NCD = \left\{ (\theta', \mathrm{U}), (\theta'', \mathrm{U}), (\theta'', \mathrm{D}) \right\} \times \{\mathrm{C.L}, \mathrm{C.R}\}.$$

Initial rationalizability stops here. If player 2 initially believes that both types of player 1 would choose U, which is consistent with initial belief in rationality, then no revised belief about player 1's type conditional upon observing D violates the chain rule (or Bayes consistency). In particular, player 2 may deem θ' at least as likely as θ'' conditional on D despite the fact that D is dominated for θ' but not for θ'', thus justifying C.L. If, instead, player 2 strongly believed in the rationality of 1, then he should assign probability 0 to θ' conditional on D. Only C.R is optimal given such beliefs. Yet, initial rationalizability does not rely on strong belief in rationality. ▲

15.4.2 Strong Rationalizability

Strong rationalizability characterizes the behavioral implications of Rationality and Common Strong Belief in Rationality, which in turn captures the best rationalization principle: rational players strongly believe in the rationality of their co-players, and more generally ascribe to them the highest degree of strategic sophistication consistent with their observed behavior.

For any nonempty subset $C_{-i} \subseteq \Theta_{-i} \times S_{-i}$, let $\Delta^H_{\mathrm{sb}}(C_{-i})$ denote the set of CPSs of player i that **strongly believe** $C_{-i} \subseteq \Theta_{-i} \times S_{-i}$, that is,

$$\Delta^H_{\mathrm{sb}}(C_{-i}) = \left\{ \bar{\mu}^i \in \Delta^H(\Theta_{-i} \times S_{-i}) : \begin{array}{l} \forall h \in H, \\ (\Theta_{-i} \times S_{-i}(h)) \cap C_{-i} \neq \emptyset \Rightarrow \bar{\mu}^i(C_{-i}|h) = 1 \end{array} \right\}.$$

Note that, letting $h = \varnothing$, we obtain $\Theta_{-i} \times S_{-i}(\varnothing) = \Theta_{-i} \times S_{-i}$ and $(\Theta_{-i} \times S_{-i}(\varnothing)) \cap C_{-i} = C_{-i}$. Thus, for every nonempty C_{-i}, strong belief in C_{-i} implies initial belief in C_{-i}: $\Delta^H_{\mathrm{sb}}(C_{-i}) \subseteq \Delta^H_{\varnothing}(C_{-i})$.

With this, we define a (weakly) decreasing sequence of Cartesian subsets $(C^n_{\mathrm{sb}})_{n\in\mathbb{N}}$ representing the type-dependent behavioral implications of rationality and the strategic reasoning steps reflecting the best rationalization principle.

Definition 84. *Consider the following elimination procedure:*
(Step $n = 0$*) For each* $i \in I$*, let* $C^0_{i,\mathrm{sb}} = \Theta_i \times S_i$*,* $C^0_{-i,\mathrm{sb}} = \Theta_{-i} \times S_{-i}$*, and*

$C^0_{\text{sb}} = \Theta \times S$.
*(**Step** $n > 0$) For each $i \in I$, let*

$$\begin{aligned}
\Delta_i^n &= \bigcap_{m=0}^{n-1} \Delta_{\text{sb}}^H \left(C^m_{-i,\text{sb}}\right), \\
C^n_{i,\text{sb}} &= \left\{(\theta_i, s_i) \in \Theta_i \times S_i : \exists \bar{\mu}^i \in \Delta_i^n, s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\} \\
&= \bigcup_{\theta_i \in \Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta_i^n\right),
\end{aligned}$$

$C^n_{-i,\text{sb}} = \times_{i \neq i} C^n_{j,\text{sb}}$, $C^n_{\text{sb}} = \times_{i \in I} C^n_{i,\text{sb}}$.
*A profile of types and strategies $(\theta_i, s_i)_{i \in I}$ is **strongly rationalizable** if $(\theta_i, s_i)_{i\in I} \in C^\infty_{\text{sb}} = \bigcap_{n>0} C^n_{\text{sb}}$, that is, if $(\theta_i, s_i) \in C^\infty_{i,\text{sb}} = \bigcap_{n>0} C^n_{i,\text{sb}}$ for each player $i \in I$.*

Note that the sequence $(C^n_{\text{sb}})_{n\in\mathbb{N}}$ is indeed (weakly) decreasing because, by definition, $\Delta_i^n \subseteq \Delta_i^{n-1}$ and thus $r_{i,\theta_i}\left(\Delta_i^n\right) \subseteq r_{i,\theta_i}\left(\Delta_i^{n-1}\right)$ for every $i \in I$ and $\theta_i \in \Theta_i$.

Let us compare strong and initial rationalizability. To ease the comparison, write the sequence of subsets obtained with the initial rationalizability procedure as $\left(C^n_\varnothing\right)_{n\in\mathbb{N}} = \left(\rho^n\left(\Theta \times S\right)\right)_{n\in\mathbb{N}}$ with $C^n_\varnothing = \times_{i\in I} C^n_{i,\varnothing}$. Since $C^0_{-i,\text{sb}} = \Theta_{-i} \times S_{-i}$, $\Delta_i^1 = \Delta^H\left(\Theta_{-i} \times S_{-i}\right)$ and $C^1_{i,\text{sb}}$ is the set of pairs (θ_i, s_i) such that s_i is justifiable for θ_i. Thus, $C^1_{\text{sb}} = \rho\left(\Theta \times S\right) = C^1_\varnothing$. For $n = 2$,

$$\Delta_i^2 = \Delta^H_{\text{sb}}\left(C^1_{-i,\text{sb}}\right) \subseteq \Delta^H_\varnothing\left(C^1_{-i,\text{sb}}\right) = \Delta^H_\varnothing\left(C^1_{-i,\varnothing}\right)$$

$(i \in I)$, where the inclusion follows from the fact that strong belief implies initial belief; thus, $C^2_{\text{sb}} \subseteq C^2_\varnothing$. In the game of Example 64 (which will be analyzed in detail below), $\Delta_i^2 = \left\{\bar{\mu}^2 : \bar{\mu}^2\left(\theta', \text{D}|\text{D}\right) = 0\right\}$ and $C^2_{2,\text{sb}} = \{\text{C.R}\}$, while $C^2_{2,\varnothing} = C^1_{2,\varnothing} = \{\text{C.L}, \text{C.R}\}$. The inclusion indeed holds for all steps $n > 1$. The proof of the following result is a rather straightforward extension of the proof of Remark 52 in Chapter 11.

Remark 56. *Initial rationalizability is weaker than strong rationalizability, that is,*

$$C^n_{\text{sb}} \subseteq C^n_\varnothing := \rho^n\left(\Theta \times S\right)$$

for all $n \in \mathbb{N} \cup \{\infty\}$.

Extending the analysis of Chapter 11 to allow for incomplete information, we could define a restriction operator $\bar{\rho}_{\text{sb}} : \mathcal{C} \to \mathcal{C}$ based on constrained optimization so that $C^n_{\text{sb}} = \bar{\rho}^n_{\text{sb}} (\Theta \times S)$ for every $n \in \mathbb{N} \cup \{\infty\}$; we skip the details. This in turn yields a characterization of strong rationalizability by means of iterated conditional dominance, which we report below without proof.

Recall that, for any subset $C_i \subseteq \Theta_i \times S_i$ and type θ_i, we let $C_{i,\theta_i} \subseteq S_i$ denote the section of $C_i \subseteq \Theta_i \times S_i$ at θ_i. Similarly, we let $C_{i,\theta_i} (h) \subseteq S_i (h)$ denote the section of $C_i \cap (\Theta_i \times S_i (h))$ at θ_i, that is,

$$C_{i,\theta_i} (h) = \{s_i \in S_i (h) : (\theta_i, s_i) \in C_i\} .$$

Also, for any nonempty subset $C \subseteq \Theta \times S$, let

$$H (C) = \{h \in H : \exists (\theta, s) \in C, h \prec \zeta (s)\}$$

denote the set of non-terminal histories consistent with C.

Definition 85. *Fix a nonempty Cartesian subset $C \in \mathcal{C}$, a player $i \in I$, and a pair $(\theta_i, \bar{s}_i) \in C_i$. Strategy $\bar{s}_i$ is* ***conditionally dominated in*** C ***for type*** θ_i *if there are $h \in H_i(\bar{s}_i) \cap H(C)$ and $\sigma_i \in \Delta(C_{i,\theta_i} (h))$, such that*

$$\forall (\theta_{-i}, s_{-i}) \in C_{-i} \cap (\Theta_{-i} \times S_{-i}(h)) , \hat{U}_i (\theta_i, \theta_{-i}, \bar{s}_i, s_{-i}) < \hat{U}_i (\theta_i, \theta_{-i}, \sigma_i, s_{-i}) .$$

We say that $\bar{s}_i \in C_{i.\theta_i}$ is ***conditionally undominated in*** C ***for type*** θ_i *if it is* not *conditionally dominated in C for type θ_i. With this, we let* NCD(C) *denote the set of profiles $(\theta_i, s_i)_{i \in I}$ such that s_i is* conditionally undominated in *C for θ_i, for every $i \in I$.*

Theorem 41. $C^n_{\text{sb}} = \text{NCD}^n (\Theta \times S)$ *for all $n \in \mathbb{N} \cup \{\infty\}$.*

Example 65. Go back to the multistage game with payoff uncertainty of Example 64. Strong rationalizability yields a unique solution. We already know that

$$C^1_{\text{sb}} = \text{NCD} (\Theta \times S) = C^1_{\varnothing} = \left\{ \left(\theta', \text{U}\right), \left(\theta'', \text{U}\right), \left(\theta'', \text{D}\right) \right\} \times \{\text{C.L}, \text{C.R}\} .$$

Next, note that $C^1_{1,\text{sb}} \cap (\Theta_1 \times \{\text{D}\}) = \{(\theta'', \text{D})\}$. Thus, C.L is conditionally dominated in C^1_{sb}. No other elimination is possible in Step 2, because U is a best reply for θ'' to C.L, D is a best reply for θ'' to C.R, and C.R is sequentially optimal given any $\bar{\mu}^2$ such that $\bar{\mu}_2(\theta', \mathrm{D}|\mathrm{D}) = 0$. Therefore,

$$C^2_{\mathrm{sb}} = \mathrm{NCD}^2(\Theta \times S) = \left\{ \left(\theta', \mathrm{U}\right), \left(\theta'', \mathrm{U}\right), \left(\theta'', \mathrm{D}\right) \right\} \times \{\mathrm{C.R}\}.$$

Finally, in Step 3 we eliminate (θ'', U), because D is the unique best reply to C.R for type θ'':

$$C^3_{\mathrm{sb}} = \mathrm{NCD}^3(\Theta \times S) = \left\{ \left(\theta', \mathrm{U}\right), \left(\theta'', \mathrm{D}\right) \right\} \times \{\mathrm{C.R}\}.$$

This leaves only one strategy for each type. Therefore, the elimination procedure stops at Step 3: $C^3_{\mathrm{sb}} = C^\infty_{\mathrm{sb}}$. ▲

15.4.3 Rationalizability and Iterated Admissibility

As in Chapter 11, one can establish a tight connection (a generic equivalence) between conditional dominance and weak dominance in the strategic form, and move from there to establish a connection between strong rationalizability and a form of iterated admissibility. Recall that, for any subset $C_i \subseteq \Theta_i \times S_i$, we let $C_{i,\theta_i} \subseteq S_i$ denote its section at θ_i.

Definition 86. *Fix a nonempty Cartesian subset $C \in \mathcal{C}$, a player $i \in I$, and a pair $(\theta_i, \bar{s}_i) \in C_i$. Strategy $\bar{s}_i$ is **weakly dominated for type** θ_i **in** C if there is a mixed strategy $\sigma_i \in \Delta(C_{i,\theta_i})$ such that*

$$\begin{aligned} \forall (\theta_{-i}, s_{-i}) &\in C_{-i}, U_i(\theta_i, \theta_{-i}, \bar{s}_i, s_{-i}) \leq U_i(\theta_i, \theta_{-i}, \sigma_i, s_{-i}) \text{ and} \\ \exists \left(\bar{\theta}_{-i}, \bar{s}_{-i}\right) &\in C_{-i}, U_i\left(\theta_i, \bar{\theta}_{-i}, \bar{s}_i, \bar{s}_{-i}\right) < U_i\left(\theta_i, \bar{\theta}_{-i}, \sigma_i, \bar{s}_{-i}\right). \end{aligned}$$

*Strategy $\bar{s}_i$ is **admissible for** θ_i **in** C if it is* not *weakly dominated for θ_i in C. We let $\mathrm{NWD}(C)$ denote the set of profiles $(\theta_i, s_i)_{i \in I} \in C$ such that s_i is admissible in C for θ_i for every player $i \in I$.*

It is relatively straightforward to show that if a strategy $\bar{s}_i$ is conditionally dominated in C for a type θ_i then it is weakly dominated for θ_i in C. By essentially the same arguments used in Chapter 11, one can show that the converse is true for almost all parameterized payoff functions. To make this precise, note that appending to structure $\left\langle I, (\Theta_i, A_i, \mathcal{A}_i(\cdot))_{i \in I} \right\rangle$ a profile of parameterized payoff functions $(u_i)_{i \in I} \in \mathbb{R}^{\Theta \times Z \times I}$, we obtain a game with payoff uncertainty. Recall that a subset of a Euclidean space is deemed **negligible** if its closure has 0 Lebesgue measure.

Lemma 34. *Fix a finite structure* $\langle I, (\Theta_i, A_i, \mathcal{A}_i(\cdot))_{i\in I}\rangle$ *and a nonempty* $C \in \mathcal{C}$. *For all profiles of parameterized payoff functions* $(u_i)_{i\in I} \in \mathbb{R}^{\Theta\times Z\times I}$ *except at most a negligible set,* $\mathrm{NCD}(C) = \mathrm{NWD}(C)$.

Using Lemma 34, one can prove that strong rationalizability is generically equivalent to iterated admissibility, that is, the maximal iterated removal of pairs (θ_i, s_i) such that s_i is weakly dominated for θ_i in the restricted set of profiles not eliminated in previous steps. Furthermore, initial rationalizability is generically equivalent to the removal of all pairs (θ_i, s_i) such that s_i is weakly dominated for θ_i (in $\Theta \times S$) followed by the iterated elimination of pairs (θ_i, s_i) such that s_i is dominated for θ_i.

Theorem 42. *Fix a finite structure* $\langle I, (\Theta_i, A_i, \mathcal{A}_i(\cdot))_{i\in I}\rangle$. *For all profiles* $u \in \mathbb{R}^{\Theta\times Z\times I}$ *except at most a negligible set,*

$$C^n_{\mathrm{sb}} = \mathrm{NWD}^n(\Theta\times S),\ \rho^n(\Theta\times S) = \left(\mathrm{ND}^{n-1}\circ \mathrm{NWD}\right)(\Theta\times S)$$

for all $n \in \mathbb{N}\cup\{\infty\}$.

Example 66. Go back again to the multistage game with payoff uncertainty of Examples 64 and 65. We can verify that its parameterized payoffs are "generic" in the sense of Theorem 42, that is, they do not belong to the negligible set for which strong rationalizability differs from iterated admissibility. Its (reduced) strategic form can be represented by a pair of matrices, one for each θ, where player 2 chooses the rows and only player 1 knows the true matrix (payoffs of player 1 are in **bold**).

θ'	U	D
S	**1**, v	**0**, 1
C.L	**1**, v	**0**, 3
C.R	**1**, v	**0**, 0

θ''	U	D
S	**1**, w	**0**, 1
C.L	**1**, w	**0**, 0
C.R	**1**, w	**2**, 3

Note, in particular, that S is weakly dominated by $\frac{1}{2}\delta_{\mathrm{C.L}} + \frac{1}{2}\delta_{\mathrm{C.R}}$. Then we obtain

$$\begin{aligned}
\mathrm{NWD}^1(\Theta\times S) &= \left\{(\theta',\mathrm{U}),(\theta'',\mathrm{U}),(\theta'',\mathrm{D})\right\}\times\{\mathrm{C.L},\mathrm{C.R}\} = C^1_{\mathrm{sb}},\\
\mathrm{NWD}^2(\Theta\times S) &= \left\{(\theta',\mathrm{U}),(\theta'',\mathrm{U}),(\theta'',\mathrm{D})\right\}\times\{\mathrm{C.R}\} = C^2_{\mathrm{sb}},\\
\mathrm{NWD}^3(\Theta\times S) &= \left\{(\theta',\mathrm{U}),(\theta'',\mathrm{D})\right\}\times\{\mathrm{C.R}\} = C^3_{\mathrm{sb}},\\
\mathrm{NWD}^\infty(\Theta\times S) &= \mathrm{NWD}^3(\Theta\times S) = C^3_{\mathrm{sb}} = C^\infty_{\mathrm{sb}}.
\end{aligned}$$

▲

15.4.4 Directed Rationalizability

Following up on the analysis of Section 8.2 of Chapter 8, we provide notions of "directed rationalizability" for multistage games with payoff uncertainty. That is, we provide extensions of the notions of initial and strong rationalizability to situations where some contextual restrictions on players' beliefs are transparent. Let us first clarify the restrictions we consider and what we mean by "transparent" in the analysis of multistage games, where conditional beliefs as the game unfolds play a crucial role.

It is natural to consider restrictions concerning *initial exogenous beliefs*, that is, assumptions about what each player i would believe about θ_{-i} at the beginning of the game if he were of type θ_i (cf. Section 8.3 of Chapter 8). Such assumptions can be represented by restricted belief sets $\bar{\Delta}_{\theta_i} \subseteq \Delta(\Theta_{-i})$. However, it may be plausible and interesting to consider also restrictions on initial beliefs about both types and behavior, and/or restrictions on beliefs conditional on observing some histories of play. Therefore, we will be more flexible and posit, for each type θ_i of each player i, a (possibly) restricted set of CPSs $\Delta_{i,\theta_i} \subseteq \Delta^H(\Theta_{-i} \times S_{-i})$.

Given this, how should we extend the meaning of "transparency" to the present environment? One candidate is a notion of "initial transparency," according to which an event is initially transparent if it is true and *commonly believed at the beginning of the game*. This is the simplest extension of the transparency idea from static games to multistage games, and it is sufficient to motivate and interpret the notion of initial directed rationalizability presented below. However, initial transparency is not germane to the idea of strong rationalizability, which captures forward-induction reasoning. Thus, we consider a strengthening based on the notion of "common full belief." We say that a player **fully believes** an event E if he assigns probability 1 to E conditional on *every* history $h \in H$. We say that event E is **transparent** if E is true and there is common full belief of E.

Note that some events cannot be fully believed, hence they cannot be transparent, because they are contradicted by some non-terminal history. Here, however, we only consider the transparency of events concerning what players believe about the state of nature θ and (possibly) behavior. Events that exclusively concern what players believe cannot be contradicted by the observation of players' behavior. Therefore, it is always possible to assume that some features of players' beliefs are transparent.[7]

Let us start with initial directed rationalizability. This solution concept posits restrictions on players' beliefs and characterizes the information-dependent behavioral implications of (a) Rationality, (b) Common Initial Belief in Rationality, and (c) (initial) transparency of the posited restrictions. For each $i \in I$ and $\theta_i \in \Theta_i$, let $\Delta_{i,\theta_i} \subseteq \Delta^H(\Theta_{-i} \times S_{-i})$ denote the restricted, nonempty set of CPSs for information-type θ_i of player i. For a given profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$, define the **initial Δ-rationalization operator** $\rho_\Delta : \mathcal{C} \to \mathcal{C}$ as follows: for every $C \in \mathcal{C}$,

$$\begin{aligned}\rho_\Delta(C) &= \bigtimes_{i\in I}\left\{(\theta_i, s_i) : \exists\bar{\mu}^i \in \Delta^H_\varnothing(C_{-i}) \cap \Delta_{i,\theta_i}, s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\} \\ &= \bigtimes_{i\in I}\left(\bigcup_{\theta_i\in\Theta_i}\{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H_\varnothing(C_{-i}) \cap \Delta_{i,\theta_i}\right)\right).\end{aligned}$$

(Recall that $\Delta^H_\varnothing(C_{-i})$ denotes the set of CPSs that initially assign probability 1 to C_{-i}.)

For a fixed Δ, self-map ρ_Δ is *monotone*, because $C'_{-i} \subseteq C_{-i}$ implies $\Delta^H_\varnothing\left(C'_{-i}\right) \subseteq \Delta^H_\varnothing(C_{-i})$. Thus, as we iterate ρ_Δ starting from the whole set of profiles of types and strategies, we obtain the (weakly) decreasing sequence of subsets $(\rho^n_\Delta(\Theta \times S))_{n\in\mathbb{N}}$. In particular, $\rho^1_\Delta(\Theta \times S) = \rho_\Delta(\Theta \times S)$ contains all the profiles $(\theta_i, s_i)_{i\in I}$ such that, for each i, s_i is justified for θ_i by some CPS $\bar{\mu}^i \in \Delta_{i,\theta_i}$. Let

$$\rho^\infty_\Delta(\Theta \times S) = \bigcap_{n\in\mathbb{N}} \rho^n_\Delta(\Theta \times S).$$

Definition 87. *Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of restricted sets of CPSs. A profile of types and strategies $(\theta_i, s_i)_{i\in I}$ is **initially Δ-rationalizable** if $(\theta_i, s_i)_{i\in I} \in \rho^\infty_\Delta(\Theta \times S)$.*

As a matter of fact, whether the belief restrictions are assumed to be transparent or only initially transparent does not make any difference here,

[7] Someone might object that a player's behavior may reveal something about his beliefs. But this objection would be based on a misunderstanding. Any inference about beliefs based on observed behavior relies on an hypothesized link between beliefs and behavior, such as that chosen actions must maximize subjective expected utility. Thus, observations about behavior can only contradict the *joint hypothesis* that a player's beliefs have some features *and* such link holds.

because initial rationalizability is silent on what players would believe if they were surprised by the realization of an unexpected history.

Similarly to Section 8.2 of Chapter 8, for each $C \in \mathcal{C}$, map $\Delta \mapsto \rho_\Delta(C)$ is monotone in the following sense: if $\Delta_{i,\theta_i} \subseteq \Delta'_{i,\theta_i}$ for every i and θ_i, written $\Delta \subseteq \Delta'$, then $\rho_\Delta(C) \subseteq \rho_{\Delta'}(C)$. By monotonicity of ρ_Δ, this implies that the initial directed rationalizability map $\Delta \mapsto \rho^\infty_\Delta(\Theta \times A)$ is monotone.

Remark 57. *Fix two profiles of restrictions, Δ and Δ'. If $\Delta \subseteq \Delta'$, then $\rho^n_\Delta(\Theta \times S) \subseteq \rho^n_{\Delta'}(\Theta \times S)$ for every $n \in \mathbb{N} \cup \{\infty\}$.*

As argued in Section 8.2 of Chapter 8, the special case of restrictions that only concern exogenous beliefs deserves attention. We say that $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ is a profile of **restrictions on exogenous beliefs** if, for each $i \in I$ and each $\theta_i \in \Theta_i$, there is a nonempty set $\bar{\Delta}_{i,\theta_i} \subseteq \Delta(\Theta_{-i})$ such that

$$\Delta_{i,\theta_i} = \left\{\bar{\mu}^i \in \Delta^H(\Theta_{-i} \times S_{-i}) : \mathrm{marg}_{\Theta_{-i}}\bar{\mu}^i(\cdot|\varnothing) \in \bar{\Delta}_{i,\theta_i}\right\},$$

where $\mathrm{marg}_{\Theta_{-i}}\bar{\mu}^i(\cdot|\varnothing)$ denotes the initial marginal belief about θ_{-i} derived from $\bar{\mu}^i$, that is, $\mathrm{marg}_{\Theta_{-i}}\bar{\mu}^i(\theta_{-i}|\varnothing) = \bar{\mu}^i(\{\theta_{-i}\} \times S_{-i}|\varnothing)$ for each $\theta_{-i} \in \Theta_{-i}$.

Given the monotonicity of ρ_Δ, one can prove the following extension of Theorem 26 in Chapter 8 to the present multistage framework.

Theorem 43. *Consider a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of restrictions on* exogenous *beliefs. Then*

$$\mathrm{proj}_\Theta \rho^n_\Delta(\Theta \times S) = \Theta$$

for all $n \in \mathbb{N} \cup \{\infty\}$, and

$$\rho_\Delta(\rho^\infty_\Delta(\Theta \times S)) = \rho^\infty_\Delta(\Theta \times S).$$

Furthermore, for every $C \in \mathcal{C}$,

$$C \subseteq \rho_\Delta(C) \Rightarrow C \subseteq \rho^\infty_\Delta(\Theta \times S).$$

As noticed in Chapter 8, $\mathrm{proj}_\Theta \rho^\infty_\Delta(\Theta \times S) = \Theta$ implies that, for every type θ_i of each player i, the set of initially Δ-rationalizable strategies for type θ_i is nonempty.

The following example illustrates initial directed rationalizability. Although we neglected the residual uncertainty Θ_0 to simplify our abstract notation, we reintroduce it in this particular case.

Example 67. Consider a leader-follower game, where player 1 (she) is the leader and player 2 (he) is the follower. Each player has to choose between two restaurants, one on the left-hand side of the main street[8] and one on the right-hand side. Thus, $A_i = \{\ell_i, r_i\}$, $S_1 = A_1$, and $S_2 = \{\ell_2, r_2\}^{\{\ell_1, r_1\}}$. The unknown state $\theta_0 \in \Theta_0 = \left\{\theta_0^\ell, \theta_0^r\right\}$ determines whether the best restaurant is on the left or on the right. Each player i has private information $\theta_i \in \Theta_i = \left\{\theta_i^\ell, \theta_i^r\right\}$, which is believed to be correlated to the quality of restaurants, but the payoff depends only on the chosen restaurant and on θ_0:

$$u_i\left(\theta, (a_1, a_2)\right) = \begin{cases} 1, & \text{if } a_i = \ell_i \text{ and } \theta_0 = \theta_0^\ell, \text{ or } a_i = r_i \text{ and } \theta_0 = \theta_0^r, \\ 0, & \text{otherwise.} \end{cases}$$

It is transparent that player 1 thinks that her private information signals quality, and that player 2 thinks that player 1 is an expert who has more accurate information in the following sense: if he knew θ_1 he would deem the state θ_0 "signalled" by θ_1 more likely than the other *independently of his private information* θ_2. Formally, we consider the following sets of CPSs featuring restrictions on exogenous initial beliefs: using obvious abbreviations for initial, marginal, and conditional beliefs,[9] for each "street side" $x \in \{\ell, r\}$,

$$\Delta_{1,\theta_1^x} = \left\{\bar{\mu}^1 : \bar{\mu}^1\left(\theta_0^x\right) > \frac{1}{2}\right\},$$

and for each $\theta_2 \in \left\{\theta_2^\ell, \theta_2^r\right\}$,

$$\Delta_{2,\theta_2} = \left\{\bar{\mu}^2 : \forall x \in \{\ell, r\}, \bar{\mu}^2\left(\theta_0^x, \theta_1^x\right) > 0, \bar{\mu}^2\left(\theta_0^x | \theta_1^x\right) > \frac{1}{2}\right\}.$$

Then, initial Δ-rationalizability is computed in two steps: it implies that player 1 follows her private information and player 2 follows the choice of player 1 independently of his own private information:

$$\rho_\Delta^2\left(\Theta \times S\right) = \Theta_0 \times \left\{\left(\theta_1^\ell, \ell_1\right), \left(\theta_1^r, r_1\right)\right\} \times \left\{\left(\theta_2, s_2\right) : s_2\left(\ell_1\right) = \ell_2, s_2\left(r_1\right) = r_2\right\}.$$

[8] Given a "canonical direction."

[9] For example, $\bar{\mu}^1\left(\theta_0\right) = \bar{\mu}^1\left(\{\theta_0\} \times \Theta_2 \times S_2 | \varnothing\right)$ and $\bar{\mu}^2\left(\theta_0^x | \theta_1^x\right) = \frac{\bar{\mu}^2\left(\left\{\theta_0^x, \theta_1^x\right\} \times S_1 | \varnothing\right)}{\bar{\mu}^2\left(\Theta_0 \times \left\{\theta_1^x\right\} \times S_1 | \varnothing\right)}$.

We leave the formal proof as an exercise. The key insight is that, given the belief restrictions, if the follower is initially certain of the rationality of the leader, then he initially assigns strictly positive probability to each action of the leader; hence, he cannot be "surprised" and, for each action of the leader, he believes that this action reveals her private information. Since he "trusts" the private information of the leader more than his own, he follows suit. ▲

We now consider strong directed rationalizability. This solution concept posits restrictions on players' beliefs and characterizes the behavioral implications of (a) Rationality, (b) transparency of the posited restrictions, and (c) Common Strong Belief in (a)-(b). Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$. For each $i \in I$ and $\theta_i \in \Theta_i$, and for any nonempty subset $C_{-i} \subseteq \Theta_{-i} \times S_{-i}$, we let $\Delta^H_{\mathrm{sb},\theta_i}(C_{-i})$ denote the set of CPSs in Δ_{i,θ_i} that strongly believe $C_{-i} \subseteq \Theta_{-i} \times S_{-i}$, that is,

$$\Delta^H_{\mathrm{sb},\theta_i}(C_{-i}) = \Delta^H_{\mathrm{sb}}(C_{-i}) \cap \Delta_{i,\theta_i}.$$

Note that $\Delta^H_{\mathrm{sb},\theta_i}(C_{-i}) = \Delta_{i,\theta_i}$ if $C_{-i} = \Theta_{-i} \times S_{-i}$.

Definition 88. *Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of restricted sets of CPSs. Consider the following elimination procedure:*
*(**Step** $n = 0$) For each $i \in I$, let $C^{\Delta,0}_{i,\mathrm{sb}} = \Theta_i \times S_i$, $C^{\Delta,0}_{-i,\mathrm{sb}} = \Theta_{-i} \times S_{-i}$, and $C^{\Delta,0}_{\mathrm{sb}} = \Theta \times S$.*
*(**Step** $n > 0$) For each $i \in I$ and $\theta_i \in \Theta_i$, let*

$$\Delta^n_{i,\theta_i} = \bigcap_{m=0}^{n-1} \Delta^H_{\mathrm{sb},\theta_i}\left(C^{\Delta,m}_{-i}\right),$$

and for each $i \in I$ let

$$\begin{aligned} C^{\Delta,n}_{i,\mathrm{sb}} &= \left\{(\theta_i, s_i) \in \Theta_i \times S_i : \exists \bar{\mu}^i \in \Delta^n_{i,\theta_i}, s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\} \\ &= \bigcup_{\theta_i\in\Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta^n_{i,\theta_i}\right). \end{aligned}$$

Also, let $C^{\Delta,n}_{-i,\mathrm{sb}} = \times_{i\neq i} C^{\Delta,n}_{j,\mathrm{sb}}$ and $C^{\Delta,n}_{\mathrm{sb}} = \times_{i\in I} C^{\Delta,n}_{i,\mathrm{sb}}$.
*A profile of types and strategies $(\theta_i, s_i)_{i\in I}$ is **strongly Δ-rationalizable** if $(\theta_i, s_i)_{i\in I} \in C^{\Delta,\infty}_{\mathrm{sb}} = \bigcap_{n>0} C^{\Delta,n}_{\mathrm{sb}}$, that is, if $(\theta_i, s_i) \in C^{\Delta,\infty}_{i,\mathrm{sb}} = \bigcap_{n>0} C^{\Delta,n}_{i,\mathrm{sb}}$ for each player $i \in I$.*

Note that the sequence $\left(C_{\mathrm{sb}}^{\Delta,n}\right)_{n\in\mathbb{N}}$ is (weakly) decreasing because, by definition, $\Delta_{i,\theta_i}^{n} \subseteq \Delta_{i,\theta_i}^{n-1}$ and thus $r_{i,\theta_i}\left(\Delta_{i,\theta_i}^{n}\right) \subseteq r_{i,\theta_i}\left(\Delta_{i,\theta_i}^{n-1}\right)$ for every $i \in I$ and $\theta_i \in \Theta_i$. The proof of the following result is left to the reader as an exercise.

Remark 58. *The first steps of the strong and initial directed rationalizability procedures coincide, but strong directed rationalizability refines initial directed rationalizability:* $C_{\mathrm{sb}}^{\Delta,1} = \rho_{\Delta}^{1}\left(\Theta \times S\right)$ *and* $C_{\mathrm{sb}}^{\Delta,n} \subseteq \rho_{\Delta}^{n}\left(\Theta \times S\right)$ *for every* $n \geq 2$*. If* $\Delta = \left(\Delta_{i,\theta_i}\right)_{i\in I,\theta_i\in\Theta_i}$ *is a profile of restrictions on* exogenous *beliefs, then*

$$\mathrm{proj}_{\Theta} C_{\mathrm{sb}}^{\Delta,n} = \Theta$$

for all $n \in \mathbb{N}\cup\{\infty\}$*, which implies that, for every type* θ_i *of each player* i*, the set of strongly* Δ*-rationalizable strategies is nonempty.*

The following example illustrates strong directed rationalizability.

Example 68. Player 1 is a young unemployed person who has to decide whether to get an education (e) or not (n). There is only one kind of job, with a fixed wage W, and education is *not* necessary to apply for the job. Furthermore, education does *not* enhance productivity. Player 2 is an employer who has to decide whether to hire player 1 or not, after observing his decision.

Player 1 can be a high type θ^H or a low type θ^L. Higher types are more productive, yield more additional revenue if hired ($R^H > R^L$), and incur a lower subjective cost to get an education ($c^H < c^L$). Player 1 knows his type. We assume that $0 < c^H < W$ and $R^H = R > W > R^L = 0$. This game can be represented as in Figure 15.3.

We let $x.y$ denote the strategy of the employer who selects action $x \in \{j, n\}$ at history (n) and action $y \in \{J, N\}$ at history (e). Let us first consider strong rationalizability. Note that, for type θ^{L}, action n is dominant. Strong rationalizability yields

$$\left\{\left(\theta^{\mathrm{H}},\mathrm{e}\right),\left(\theta^{\mathrm{H}},\mathrm{n}\right),\left(\theta^{\mathrm{L}},\mathrm{n}\right)\right\} \times \{j.J, n.J\}.$$

Indeed, at the second step strategies $j.N$ and $n.N$ are eliminated: strong belief in rationality implies that, if player 1 gets an education, then player 2 must conclude that his type is high; this entails that J is optimal.

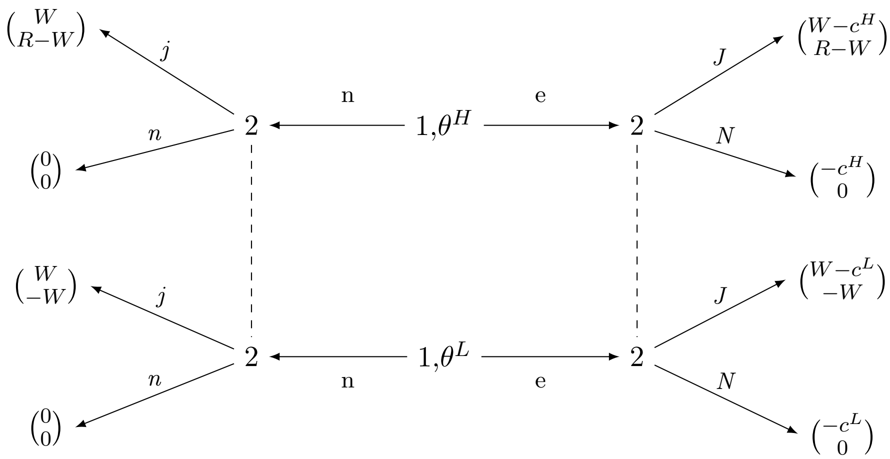


Figure 15.3: A “job market” game.

Assume now that (it is transparent that) at the beginning of the game player 2 assigns a low probability to the high type. Specifically, every CPS $\bar{\mu}^2 \in \Delta_2$ is such that

$$\bar{\mu}^2\left(\left\{\theta^{\mathrm{H}}\right\} \times \{\mathrm{n}, \mathrm{e}\} \,|\varnothing\right) < \frac{W}{R}.$$

(Note that $\frac{W}{R} \in (0,1)$.) With this restriction on the beliefs of player 2 and no restriction on the beliefs of any type of player 1, directed rationalizability is computed in three steps:

$$C_{\mathrm{sb}}^{\Delta,3} = \left\{\left(\theta^{\mathrm{H}}, \mathrm{e}\right), \left(\theta^{\mathrm{L}}, \mathrm{n}\right)\right\} \times \{n.J\}.$$

At the first step, $\left(\theta^{\mathrm{L}}, \mathrm{e}\right)$ is eliminated because, as noted above, action n is dominant for type θ^{L}. Strategy $j.J$ is also eliminated at the first step, because at least one of the two beliefs after observing e and n must be derived from $\bar{\mu}^2(\cdot|\varnothing)$ by conditioning; hence at least one of the two conditional beliefs cannot assign probability higher than $\frac{W}{R}$ to θ^{H}. The same argument as the one used for strong rationalizability shows that strategies $j.N$ and $n.N$ are eliminated at the second step. Since both $j.J$ and $j.N$ were eliminated, at the third step type θ^{H} chooses action e. We leave the formal proof as an exercise. ▲

We also note a rather subtle point: unlike initial directed rationalizability, *strong directed rationalizability is not monotone in the beliefs restrictions*. In particular, strong Δ-rationalizability may not refine strong rationalizability (which is obtained by letting $\Delta_{i,\theta_i} = \Delta^H(\Theta_{-i} \times S_{-i})$ for all i and θ_i). This follows from the fact that the strong belief set $\Delta^H_{\text{sb}}(C_{-i})$ is not monotone in C_{-i} (see the discussion in Section 11.3 of Chapter 11). The following example illustrates this point.

Example 69. Consider the following game. Player 1 moves first, choosing between In and Out. After Out, the game ends. After In, player 2 moves, choosing between $left$, *center*, and *right*, then the game ends. Player 1 is privately informed of a payoff-relevant parameter $\theta_1 \in \{0, 1\}$. Players' payoffs depend on the terminal history and on θ_1 as follows (see Figure 15.4).

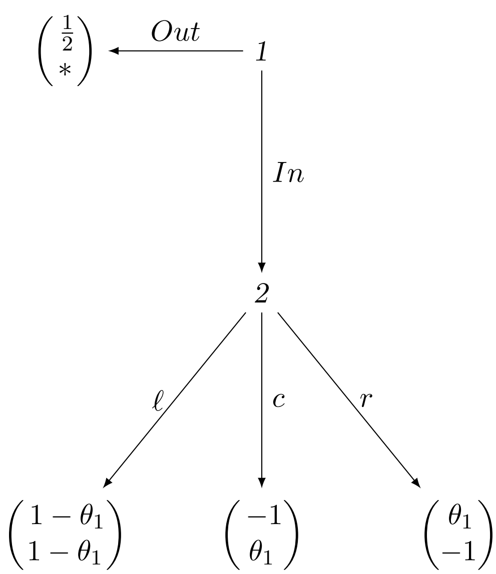


Figure 15.4: A leader-follower game with asymmetric information.

We first analyze the game with strong rationalizability (that is, without belief restrictions), which can be computed by iterated conditional dominance. Since r is conditionally dominated, it is eliminated in Step 1. Given this, in Step 2, In is eliminated for type $\theta_1 = 1$. In Step 3, player 2 rationalizes In assuming that it was chosen by type $\theta_1 = 0$ (forward induction), therefore c is eliminated. Finally, in Step 4, Out is eliminated for type $\theta_1 = 0$. To conclude, Out is the only strongly rationalizable strategy for type $\theta_1 = 1$, In is the only strongly rationalizable strategy for type $\theta_1 = 0$, and ℓ is the only strongly rationalizable strategy for player 2:

$$C_{\text{sb}}^{\infty} = C_{\text{sb}}^{4} = \{(0, In), (1, Out)\} \times \{\ell\}.$$

Next we consider directed rationalizability assuming that (only) the following is transparent: player 2 becomes certain of type $\theta_1 = 1$ upon observing In, that is,

$$\Delta_2 = \left\{\bar{\mu}^2 \in \Delta^H(\Theta_1 \times S_1) : \mu_2((1, In) | (In)) = 1\right\}.$$

Given this, both ℓ and r are eliminated in Step 1 of directed rationalizability. Then, in Step 2 In is eliminated for *both* types of player 1. *This makes it impossible to rationalize In* in Step 3. Hence, the only strongly Δ-rationalizable strategy of both types of player 1 is Out, and the only strongly Δ-rationalizable strategy of player 2 is c:

$$C_{\text{sb}}^{\Delta,\infty} = C_{\text{sb}}^{\Delta,3} = \{(0, Out), (1, Out)\} \times \{c\}.$$

▲

Note that in Example 69 we considered a restriction on beliefs conditional on observed behavior, which are *endogenous*. The reason is that, if the restrictions Δ only concern *exogenous* beliefs, then the set of strong Δ-rationalizable paths (not the strategies) is monotone in the posited restrictions Δ.[10]

15.5 Multistage Bayesian Games

Structure $\left\langle I, \Theta_0, (\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i)_{i \in I}\right\rangle$ does not specify the exogenous interactive beliefs of the players, i.e., their beliefs about each other's types. Therefore, it is not rich enough to define traditional notions of equilibrium, e.g. Bayesian equilibrium and refinements thereof.

As in the case of static games, we should add exogenous interactive beliefs structures *à la* Harsanyi to carry out a traditional analysis of

[10] A proof is available by request.

games with incomplete information. Since the dynamics involve additional complications of strategic analysis, we *simplify* the *interactive beliefs* aspect: we assume that the set of types *à la* Harsanyi T_i coincides with (more precisely, is isomorphic to) Θ_i. To emphasize this assumption we use the phrase "simple Bayesian game" (cf. Section 8.4 of Chapter 8). Compared to the previous analysis of rationalizability, we allow for the presence of residual uncertainty, that is, we allow for the possibility that $|\Theta_0| > 1$.[11]

A multistage Bayesian game with information-types, or **simple multistage Bayesian game** (with observable actions), is a (*finite*) structure

$$\Gamma = \left\langle I, \Theta_0, \left(\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i, (p_i(\cdot|\theta_i))_{\theta_i \in \Theta_i}\right)_{i \in I}\right\rangle$$

where

- $\left\langle I, \Theta_0, (\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i)_{i \in I}\right\rangle$ is a (*finite*) game with payoff uncertainty

- for every $i \in I$ and $\theta_i \in \Theta_i$, $p_i(\cdot|\theta_i) \in \Delta(\Theta_0 \times \Theta_{-i})$ is the *initial* exogenous belief of information-type θ_i.

Alternatively, we could specify an ex ante belief, or "prior," $P_i \in \Delta(\Theta)$ for every player i such that $P_i(\theta_i) = P_i(\Theta_0 \times \{\theta_i\} \times \Theta_{-i}) > 0$ for each $\theta_i \in \Theta_i$, and derive the initial belief of each type θ_i as a conditional probability, that is,

$$p_i(\theta_0, \theta_{-i}|\theta_i) = \frac{P_i(\theta_0, \theta_i, \theta_{-i})}{P_i(\theta_i)} \tag{15.5.1}$$

for all $(\theta_0, \theta_{-i}) \in \Theta_0 \times \Theta_{-i}$.[12]

---

[11]In the analysis of rationalizability, we assumed distributed knowledge of the θ parameter (that is, $|\Theta_0| = 1$) and neglected Θ_0 merely for notational simplicity.

[12]Game theorists use the term "prior" somewhat liberally to mean any kind of subjective probability measure assigned before receiving any kind of information. In Bayesian inferential statistics, instead, a **prior** is a subjective probability measure over statistical models, e.g., urn models. In the present context, the priors of Bayesian statistics are the initial beliefs of types $p_i(\cdot|\theta_i)$, because they are the starting points of the Bayesian updating process.

Note also that for every profile of initial beliefs of types $(p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}$ we can find a class of ex ante beliefs that generate them according to (15.5.1): just pick any strictly positive measure $\lambda_i \in \Delta^\circ(\Theta_i)$ and compute the ex ante probability of each $(\theta_0,\theta_i,\theta_{-i})$ as follows:

$$P_{\lambda_i}(\theta_0,\theta_i,\theta_{-i}) = p_i(\theta_0,\theta_{-i}|\theta_i)\lambda_i(\theta_i).$$

Since λ_i is strictly positive,

$$\begin{aligned} P_{\lambda_i}(\theta_i) &= \sum_{(\theta_0,\theta_{-i})\in\Theta_0\times\Theta_{-i}} p_i(\theta_0,\theta_{-i}|\theta_i)\lambda_i(\theta_i) \\ &= \lambda_i(\theta_i) \sum_{(\theta_0,\theta_{-i})\in\Theta_0\times\Theta_{-i}} p_i(\theta_0,\theta_{-i}|\theta_i) = \lambda_i(\theta_i) > 0. \end{aligned}$$

Therefore P_{λ_i} is an ex ante belief that generates $(p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}$, that is,

$$\frac{P_{\lambda_i}(\theta_0,\theta_i,\theta_{-i})}{P_{\lambda_i}(\theta_i)} = \frac{p_i(\theta_0,\theta_{-i}|\theta_i)\lambda_i(\theta_i)}{\lambda_i(\theta_i)} = p_i(\theta_0,\theta_{-i}|\theta_i)$$

for all $(\theta_0,\theta_i,\theta_{-i})$. If there are at least two information-types for player i, then there is a continuum of ex ante beliefs that generate $(p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}$. If instead there is only one type of player i, say $\bar{\theta}_i$, then there is no essential difference between i's ex ante belief and i's belief given his (unique) type $\bar{\theta}_i$, because $\Theta = \Theta_0 \times \left\{\bar{\theta}_i\right\} \times \Theta_{-i}$ is isomorphic to $\Theta_0 \times \Theta_{-i}$.

We keep writing $p_i(\cdot|\theta_i)$ for the initial exogenous belief of type θ_i—as if there were an ex ante stage—for several reasons. First, it is traditional in much of the applied-theory literature. Second, in some applications, it does make sense to think of an ex ante stage in which all player have the same information (i.e., they just know that the set of possible states of nature is Θ). Third, many formulas involving conditional probabilities are more easily readable under the "ex-ante interpretation" that θ is determined at random and each i observes the realization θ_i (cf. Section 8.6 in Chapter 8).

Simple multistage Bayesian games are traditionally analyzed by means of a refined notion of Bayesian equilibrium, just like multistage games with observable actions and complete information are traditionally analyzed by computing subgame perfect Nash equilibria. The following sections are devoted to such traditional analysis. However, we point out that traditional equilibrium analysis is not a must. For example, we can use directed rationalizability assuming that there are transparent restrictions on exogenous beliefs determined by the belief maps $\theta_i \mapsto p_i(\cdot|\theta_i)$. Alternatively, one can determine the ex ante strategic form of the Bayesian game and apply iterated admissibility. Both approaches can be very insightful.

15.6 Bayesian Equilibrium

We now show how the notion of Bayesian equilibrium given for the static case (see Section 8.4 of Chapter 8) can be naturally extended to simple multistage Bayesian games.

Fix a simple multistage Bayesian game Γ. A **choice rule**, or (extended) decision function of player i is an element of the set $S_i^{\Theta_i}$ that can be interpreted as the deterministic conjecture of player $j \neq i$ about i's contingent behavior as a function of his information-type θ_i. For each $i \in I$, we let

$$\Sigma_i = S_i^{\Theta_i} = \left(\bigtimes_{h \in H} \mathcal{A}_i(h) \right)^{\Theta_i}$$

denote the set of all choice rules of player i. A generic element of Σ_i is denoted by σ_i (not to be confused with a mixed strategy). Going back to the definition of Bayesian equilibrium of Chapter 8 and replacing actions a_i with strategies s_i, we obtain the following:

Definition 89. *A (pure)* ***Bayesian equilibrium*** *of* Γ *is a profile of choice rules* $(\sigma_i)_{i \in I} \in \bigtimes S_i^{\Theta_i}$ *such that, for all* $i \in I$, $\theta_i \in \Theta_i$, *and* $s_i \in S_i$,

$$\begin{aligned} & \sum_{(\theta_0, \theta_{-i}) \in \Theta_0 \times \Theta_{-i}} p_i(\theta_0, \theta_{-i}|\theta_i)\, u_i(\theta_0, \theta_i, \theta_{-i}, \zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i}))) \\ \geq & \sum_{(\theta_0, \theta_{-i}) \in \Theta_0 \times \Theta_{-i}} p_i(\theta_0, \theta_{-i}|\theta_i)\, u_i(\theta_0, \theta_i, \theta_{-i}, \zeta(s_i, \sigma_{-i}(\theta_{-i}))), \end{aligned}$$

where $\sigma_{-i}(\theta_{-i}) = (\sigma_j(\theta_j))_{j \neq i}$.

The definition can be extended to allow for randomization, which requires a suitable notion of randomized choice rules. As in multistage games with complete information, we can think of players choosing pure strategies at random (given their type), or players who randomize over actions as the play unfolds. The latter notion of randomization is more convenient. A behavioral (randomized) choice rule, or **extended behavior strategy** of player i is an element of the set $B_i^{\Theta_i}$, where $B_i = \times_{h\in H} \Delta(\mathcal{A}_i(h))$ denotes the set of behavior strategies. A generic element of the set $B_i^{\Theta_i}$ is denoted by

$$(\beta_i(\cdot|\theta_i,\cdot))_{\theta_i\in\Theta_i} = (\beta_i(\cdot|\theta_i,h))_{\theta_i\in\Theta_i,h\in H}\,,$$

where $\beta_i(a_i|\theta_i,h)$ is the probability of choosing a_i given θ_i and h.

If we define Γ using "priors" $P_i \in \Delta(\Theta)$, we can characterize the set of Bayesian equilibria of Γ as the Nash equilibria of the *ex ante* strategic form. For each $i\in I$, define $U_i : \times_{j\in I}\Sigma_j \to \mathbb{R}$ by

$$U_i(\sigma_i,\sigma_{-i}) = \sum_{\theta\in\Theta} P_i(\theta)u_i\left(\theta,\zeta\left(\sigma_i\left(\theta_i\right),\sigma_{-i}\left(\theta_{-i}\right)\right)\right),$$

where $P_i \in \Delta(\Theta)$ is any ex ante belief generating $(p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}$. The **ex ante strategic form** of Γ is the static game

$$\mathcal{AS}(\Gamma) = \langle I, (\Sigma_i, U_i)_{i\in I}\rangle\,.$$

It can be checked that *a profile σ is a Bayesian equilibrium of Γ if and only if it is a Nash equilibrium of $\mathcal{AS}(\Gamma)$*. An analogous result holds for randomized Bayesian equilibria.

As anticipated above, it is convenient to represent randomized Bayesian equilibria of multistage Bayesian games as profiles of extended behavior strategies $\beta \in \times_{i\in I} B_i^{\Theta_i}$ rather than profiles of (type-dependent) mixed strategies.[13] We can interpret extended behavior strategies as conditional distributions in heterogeneous populations of agents. To see this, suppose that $\pi_i(\theta_i,s_i)$ is the fraction of agents in population i of type θ_i who play s_i.[14] For every $X_i \subseteq S_i$, set $\pi_i(\theta_i,X_i) = \sum_{s_i\in X_i}\pi_i(\theta_i,s_i)$. Next recall that $S_i(h)$ denotes the set of strategies of (population) i that allow (do not prevent) h, and $S_i(h,a_i)$ denotes the set of strategies of i that allow h *and* select a_i given h. Then, let

$$\beta_i(a_i|\theta_i,h) = \frac{\pi_i(\theta_i,S_i(h,a_i))}{\pi_i(\theta_i,S_i(h))} \text{ if } \pi_i(\theta_i,S_i(h)) > 0.$$

[13] By an incomplete-information version of Theorem 32, the two approaches are equivalent.

[14] Milgrom and Weber [47] call π_i a "distributional strategy."

With this, $\beta_i(a_i|\theta_i, h)$ is the fraction of agents in population i taking action a_i at h, among those whose type is θ_i and whose strategy does not prevent h.

We can also characterize the Bayesian equilibria of Γ as the Nash equilibria of the *interim* strategic form. Assume for notational simplicity that $\Theta_i \cap \Theta_j = \emptyset$ for each $i, j \in I$, $i \neq j$ (this is just a matter of labelling). The **interim strategic form** of Γ is the static $\left(\sum_{i\in I} |\Theta_i|\right)$-player game

$$\mathcal{IS}(\Gamma) = \left\langle \bigcup_{i\in I} \Theta_i, (S_{\theta_i}, U_{\theta_i})_{i\in I, \theta_i \in \Theta_i} \right\rangle$$

where

- $S_{\theta_i} = S_i$ for each θ_i,
- the expected payoff function $U_{\theta_i} : \times_{j\in I} S_j^{\Theta_j} \to \mathbb{R}$ is given by

  $$U_{\theta_i}((s_{\theta_j})_{j\in I, \theta_j\in\Theta_j}) = \sum_{(\theta_0,\theta_{-i})\in\Theta_0\times\Theta_{-i}} p_i(\theta_0, \theta_{-i}|\theta_i) u_i\left(\theta_0, \theta_i, \theta_{-i}, \zeta(s_{\theta_i}, s_{\theta_{-i}})\right),$$

  with $s_{\theta_{-i}} = (s_{\theta_j})_{j\in I\setminus\{i\}} \in S_{-i}$.

It can be checked that *a profile σ is a Bayesian equilibrium of Γ if and only if it is a Nash equilibrium of $\mathcal{IS}(\Gamma)$*. An analogous result holds for randomized Bayesian equilibria.

15.7 Perfect Bayesian Equilibrium

Like Nash equilibrium, Bayesian equilibrium allows for non maximizing choices at histories that are not supposed to occur in equilibrium. For multistage games with complete information, this problem was addressed using the notion of subgame perfect equilibrium. Thus, traditional game theorists find it natural to try to extend this notion to multistage Bayesian games and define some kind of "perfect Bayesian equilibrium" (PBE) concept. Here we define the weakest (i.e., the most general) among the meaningful versions of the PBE concept.[15]

[15] Some versions of PBE found in textbooks are either ambiguously defined, or are serioulsy flawed because they do not assume that Bayes consistency holds starting from every history (cf. Definition 78).

As we did in Section 15.3, to ease notation we neglect residual uncertainty, that is, we consider a simple multistage Bayesian game Γ where Θ_0 is a *singleton* and remove θ_0 from formulas. It is convenient to look at *randomized* equilibria as this makes it more transparent how the analysis relies on Bayes rule. To connect to the analysis of rational planning, a candidate equilibrium profile $\beta = (\beta_i)_{i \in I}$ of extended behavior strategies yields a "correct conjecture" for (each type of) each player i, with

$$\beta^i (a_{-i}|\theta_{-i}, h) = \prod_{j \neq i} \beta_j (a_j|\theta_j, h) \tag{15.7.1}$$

for all $h \in H$, $\theta_{-i} \in \Theta_{-i}$, and $a_{-i} \in \mathcal{A}_{-i}(h)$. Furthermore, we need to specify a profile $(\mu_i(\cdot|\theta_i, h))_{i \in I, \theta_i \in \Theta_i}$ of personal systems of beliefs, one for each type of each player, otherwise we cannot compute conditional expected payoffs. Of course, for each player $i \in I$, the initial belief of each type $\theta_i \in \Theta_i$ about the types of the other players must be the belief $p_i(\cdot|\theta_i)$ exogenously specified in Bayesian game Γ.

Definition 90. *A* ***system of beliefs*** *in* Γ *is a profile* $\mu = (\mu_i)_{i \in I} \in \left(\Delta(\Theta_{-i})^{\Theta_i \times H}\right)^I$ *where each* $\mu_i(\cdot|\theta_i, \cdot) \in \Delta(\Theta_{-i})^H$ *(*$i \in I$*,* $\theta_i \in \Theta_i$*) is a personal system of beliefs such that* $\mu_i(\cdot|\theta_i, \varnothing) = p_i(\cdot|\theta_i)$*. A pair* (β, μ) *given by a profile of extended behavior strategies and a system of beliefs is called* ***assessment****.*

Clearly, beliefs μ must be related to β and are therefore *endogenous.* Hence, a candidate equilibrium cannot be just an extended behavior strategy profile, it has to be an *assessment* (β, μ). With this, fix a candidate equilibrium assessment (β, μ); when can we say that (β, μ) is a PBE? First note that, for each $h \in H$, beliefs $(\mu_i(\cdot|\theta_i, h))_{i \in I, \theta_i \in \Theta_i}$ define a (h, μ)-**continuation** (Bayesian) **game** $\Gamma(h, \mu)$: consider the set of feasible continuations of h, $\{h' \in A^{\leq \mathbb{N}} : (h, h') \in \bar{H}\}$, the resulting θ-dependent payoffs and the interactive beliefs $(\mu_i(\cdot|\theta_i, h))_{i \in I, \theta_i \in \Theta_i}$. (Of course, the $(\varnothing, \mu)$-"continuation game" is the Bayesian game itself, $\Gamma(\varnothing, \mu) = \Gamma$, because $\mu_i(\cdot|\theta_i, \varnothing) = p_i(\cdot|\theta_i)$ for all $i \in I$, and $\theta_i \in \Theta_i$.) Intuitively, (β, μ) is a PBE if it satisfies two conditions:

- (**Interpersonal**) **Bayes consistency:** the system of beliefs μ and the behavior strategies β must be related to each other *via Bayes rule* (thus, the players use the candidate equilibrium strategies of the co-players as conjectures in order to update beliefs *via* Bayes rule).

- **Continuation equilibrium** (often called "sequential rationality"): for each $h \in H$, β induces a Bayesian equilibrium of the (h, μ)-continuation game $\Gamma(h, \mu)$.

However, *it is not obvious how to specify Bayes consistency.* Game theorists have proposed different definitions. The reason is that, on top of mere consistency with Bayes rule (see Definition 78), they wanted to incorporate additional assumptions in the spirit of traditional equilibrium analysis, such as that players "update in the same way" (differences in beliefs are only due to differences in information), and that beliefs satisfy independence across opponents. Unfortunately, it was not very clear which additional assumptions one was trying to incorporate in the PBE concept, nor how to exactly express those assumptions in a mathematical language. Appeals to intuition and references to particular examples dominated the analysis. Hence the mess: *there is no universally accepted, or "canonical," notion of PBE* despite the fact that many applied papers refer to the PBE concept as if such a universally accepted notion existed. We consider below a minimalistic (i.e., general) notion of PBE based *only* on Bayes consistency as explained in Section 15.3.

Each player i, given his type θ_i, is uncertain about the types θ_{-i} of the co-players. As in the SPE analysis of complete information games, the candidate equilibrium behavior strategies β are commonly believed to be complied with, starting from any history h and whatever the number of deviations from β implied by h.[16] Thus, the probabilities of simultaneous and future actions are always believed to be determined by β given the players' types. A new action profile a taken at h is the evidence upon which player i updates his subjective belief about θ_{-i}.

Definition 91. *Assessment (β, μ) is a* ***perfect Bayesian equilibrium (PBE)*** *of Γ if, for every player $i \in I$ and type $\theta_i \in \Theta_i$, (1) personal assessment $\left(\beta^i, \mu_i(\cdot|\theta_i, \cdot)\right)$ (with conjecture β^i given by eq. (15.7.1)) is Bayes consistent and (2) behavior strategy $\beta_i(\cdot|\theta_i, \cdot)$ is sequentially optimal given $\left(\beta^i, \mu_i(\cdot|\theta_i, \cdot)\right)$.*

[16] A deviation from extended behavior strategy β_j is *detected* at history h if h contains an action a_j taken at some earlier history $h' \prec h$ such that $\beta_j(a_j|\theta_j, h') = 0$ *for every* θ_j.

Note that assessment (β, μ) satisfies condition (2) of Definition 91 if and only if the continuation of β from each $h \in H$ induces a Bayesian equilibrium in the (h, μ)-continuation game $\Gamma(h, \mu)$. Letting $h = \varnothing$ (empty history), we obtain the intuitive observation that, *if* (β, μ) *is a PBE of* Γ, *then* β *is a Bayesian equilibrium of* Γ.[17]

Let $\mathbb{E}_{\beta,\mu_i}(u_i|\theta_i, h, a_i) = V_{\theta_i}^{\beta,\mu_i}(h, a_i)$ denote the conditional expected payoff for type θ_i given h and a_i, under assessment (β, μ). By an application of the OD Principle (Proposition 19), we obtain the following result:

Corollary 8. *An assessment* (β, μ) *such that each type of each player satisfies Bayes consistency is a PBE if and only if, for all* $i \in I$, $\theta_i \in \Theta_i$, *and* $h \in H$,

$$\text{supp}\beta_i(\cdot|\theta_i, h) \subseteq \arg\max_{a_i \in \mathcal{A}_i(h)} \mathbb{E}_{\beta,\mu}(u_i|\theta_i, h, a_i).$$

The PBE concept of Definition 91 is quite weak. For example, it allows players who start with the same information and beliefs in a given stage to update in different ways, although they observe exactly the same evidence, that is, the actions played at that stage. Also, it allows player i to change his beliefs about the type of j upon observing the action of some other player k, as if k had some private signal about the type of j. The reason is that the *interpersonal* Bayes consistency requirements bite at histories with positive probability, but not at histories with 0 probability.

Example 70. Consider the assessment partially shown in Figure 15.5, where p denotes the belief of player 2 conditional on D and q denotes the belief of player 3 conditional on (D, C). First of all, since D has zero probability, even if players 2 and 3 have the same (uniform) prior and observe the same evidence (action D is taken by player 1), they may have different beliefs conditional on D, i.e., it is possible that $\mu_3(\theta'|D) \neq p =: \mu_2(\theta'|D)$. In a sense, they can come up with different explanations of the unexpected move by player 1. Now suppose for the sake of the discussion that instead they have the same belief given D: $\mu_3(\theta'|D) = p$. If the candidate equilibrium behavior strategy of 2 assigns strictly positive probability to C—that is, $\gamma =: \beta_2(C|D) > 0$—then Bayes consistency

[17] We will consider the relationship between PBE and Bayesian equilibrium more in detail in the analysis of a special case; see Section 15.8.

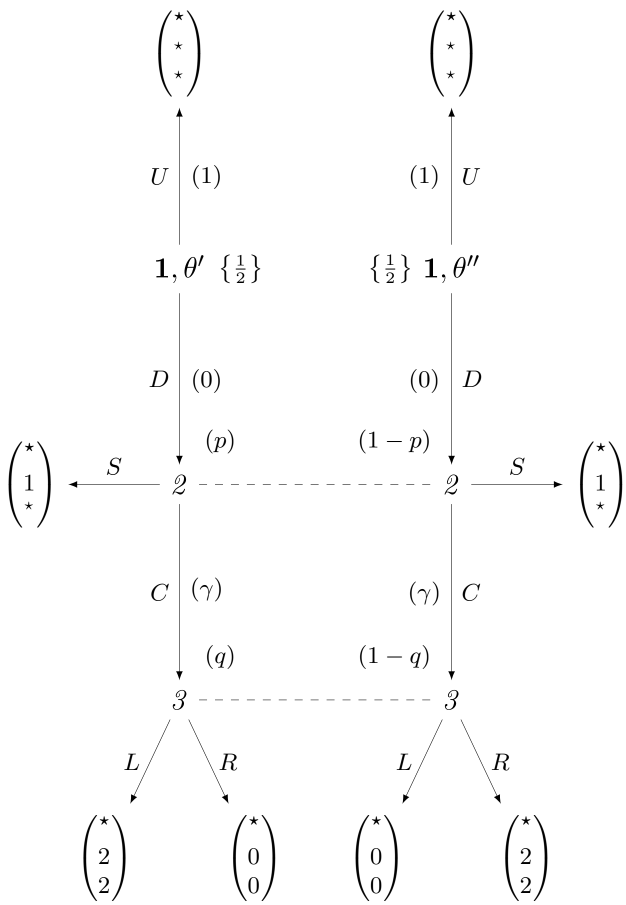


Figure 15.5: Assessment in a three-person game.

implies that 2 and 3 hold the same belief conditional on (D, C), i.e., $p = q$. If instead $\gamma = 0$ then Bayes consistency only implies $\mu_2(\theta'|D, C) = p$, but it does *not* imply $p = q$; in other words, 3 may change his belief about 1's type just because he is surprised by 2. Stronger notions of PBE require, in this example, that $\mu_3(\theta'|D) = p = q$. More generally, they require that players with the same initial beliefs and information update in the same way and that updated beliefs about a co-player's type depend only on that co-player's actions. ▲

15.8 Equilibrium in Signaling Games

We now study an easy case widely considered in applications, for which everybody agrees on the meaning of PBE:

Definition 92. *A* ***leader-follower game*** *is a two-person, two-stage game with observable actions where only one player is active at each stage and the second mover is different from the first mover, for each non-terminating action of the first mover. A* ***signaling game*** *(or* ***sender-receiver game****) is a leader-follower game with payoff uncertainty $\hat{\Gamma}$ where the second mover has no private information.*

So, a **signaling game** is a two-stage game with payoff uncertainty and perfect information (only one active player at stage 1, at most one active player at stage 2) where only the first mover has private information.[18] Thus, we assume that θ is known to player 1, whose action[19] $a_1 \in A_1$ is observed by player 2 before he chooses $a_2 \in \mathcal{A}_2(a_1)$.[20] The possible parameter values $\theta \in \Theta$ are the **types** of player 1, the informed player. The actions of the informed player are also called **signals** or **messages** because they may be thought to reveal her private information. Thus, the first player is also called "**sender**" and the second player is also called "**receiver.**" Let $A_2 = \bigcup_{a_1 \in A_1} \mathcal{A}_2(a_1)$. The payoffs of sender and receiver

[18] This is not a crucial assumption, it just allows to simplify the notation. The extension of the analysis to signaling games with two-sided incomplete information is conceptually straightforward.

[19] In some models the set of feasible actions of player 1 depends on θ and is denoted by $\mathcal{A}_1(\theta)$. The set of potentially feasible actions of player 1 is $A_1 = \bigcup_{\theta \in \Theta} \mathcal{A}_1(\theta)$.

[20] If player 2 is not active after action a_1, then $\mathcal{A}_2(a_1)$ is a singleton.

are given by the functions

$$\begin{aligned} u_1 &: \Theta \times A_1 \times A_2 \to \mathbb{R}, \\ u_2 &: \Theta \times A_1 \times A_2 \to \mathbb{R}. \end{aligned}$$

As we notice more generally for multistage games with incomplete information, a signaling game can be analyzed by some versions of rationalizability. We have already done this in previous examples of signaling games, such as those in Examples 68 and 69. Yet, in order to define the traditional notion of equilibrium we need to add a description of players' exogenous beliefs. In particular, to obtain a **Bayesian signaling game** Γ, we just have to append to signaling game (with payoff uncertainty) $\hat{\Gamma}$ an initial belief $p \in \Delta(\Theta)$ of the uninformed player, i.e., of the receiver. We assume for simplicity that $p(\theta) > 0$ for every $\theta \in \Theta$. This *prior* belief is an *exogenously given* element of the model, whereas (behavior) strategies have to be determined through equilibrium analysis. Furthermore, consistently with our interpretation of Bayesian games, it should be informally assumed that the prior p of the receiver is "transparent," that is, the sender is certain that p is the prior of the receiver, the receiver is certain that the sender is certain of this, etc.

A(n extended) **behavior strategy for the sender** is an array of probability measures $\beta_1 = (\beta_1(\cdot|\theta))_{\theta\in\Theta} \in (\Delta(A_1))^{\Theta}$: under the interpretation that the sender is "born" of a given type θ, β_1 is not a strategy in the intuitive sense, but rather a randomized decision rule expressing how the behavior of the sender depends on his type. A **behavior strategy for the receiver** is an array of probability measures $\beta_2 = (\beta_2(\cdot|a_1))_{a_1\in A_1} \in \times_{a_1\in A_1} \Delta(\mathcal{A}_2(a_1))$.

Here the receiver has the role of agent i in Section 15.2 on Bayes rule, with $X = A_1$. The receiver is initially uncertain about (θ, a_1) and has a prior probability measure on $\Theta \times A$ with marginal $p \in \Delta(\Theta)$. *In an equilibrium assessment* (β, μ), β_1 represents the conjecture of the receiver about the probability of each action of the sender conditional on each possible type (parameter value) θ. Thus, the probabilistic assessment of the receiver is such that

$$\begin{aligned} \forall\theta &\in \Theta,\, \mathbb{P}^{\beta,\mu}(\theta) = \mu(\theta|\varnothing) = p(\theta), \\ \forall\theta &\in \Theta, \forall a_1 \in A_1,\, \mathbb{P}^{\beta,\mu}(a_1|\theta) = \beta_1(a_1|\theta),\, \mathbb{P}^{\beta,\mu}(\theta, a_1) = \beta_1(a_1|\theta)p(\theta), \\ \forall a_1 &\in A_1,\, \mathbb{P}^{\beta,\mu}(a_1) = \sum_{\theta'\in\Theta} \beta_1(a_1|\theta')p(\theta'). \end{aligned}$$

Here $\mu(\theta|a_1)$ denotes the probability that the receiver would assign to θ upon observing action a_1 of the sender. Since the receiver chooses a_2 after he has observed a_1 so as to maximize the expectation of $u_2(\theta, a_1, a_2)$, the system of conditional probabilities $\mu = (\mu(\cdot|a_1))_{a_1 \in A_1}$ is an essential ingredient of equilibrium analysis. Applying the terminology of Section 15.7 to the special case of signaling games, μ is called "system of beliefs."

If the predictive probability of action a_1 is positive, $\mathbb{P}^{\beta,\mu}(a_1) = \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta') > 0$, then Bayes formula applies and we must have

$$\mu(\theta|a_1) = \frac{\mathbb{P}^{\beta,\mu}(\theta, a_1)}{\mathbb{P}^{\beta,\mu}(a_1)} = \frac{\beta_1(a_1|\theta)p(\theta)}{\sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')}.$$

Since β_1 is endogenous, also the system of beliefs μ *is endogenous*. Thus, we have to determine through equilibrium analysis the triple (β_1, β_2, μ). As we have explained in Section 15.7, (β_1, β_2, μ) (a profile of behavior strategies plus a system of beliefs) is called "assessment."

For any given assessment (β_1, β_2, μ) we use the following notation to abbreviate conditional expected payoff formulas:

$$\begin{aligned} \mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1, \cdot\right)\right) &= \sum_{a_2 \in \mathcal{A}_2(a_1)} \beta_2(a_2|a_1) u_1(\theta, a_1, a_2), \\ \mathbb{E}_{\mu}\left(u_2\left(\cdot, a_1, a_2\right)\right) &= \sum_{\theta \in \Theta} \mu(\theta|a_1) u_2(\theta, a_1, a_2). \end{aligned}$$

Thus, $\mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1, \cdot\right)\right)$ is the expected payoff for the sender of choosing action a_1 given that his type is θ and assuming that his conjecture about the behavior of the receiver is given by β_2.[21] Similarly, $\mathbb{E}_{\mu}\left(u_2\left(\cdot, a_1, a_2\right)\right)$ is the expected payoff for the receiver of choosing action a_2 given that he has observed a_1 and assuming that his conditional beliefs about θ are determined by μ.

In equilibrium, an action of player i can have positive conditional probability only if it maximizes the conditional expected payoff of i. Furthermore, the equilibrium assessment must be consistent with Bayes

[21] There is no loss of generality in representing a conjecture of player 1 as a behavior strategy of player 2. If player 1 had a conjecture of the form $\sigma_2 \in \Delta(S_2)$, where S_2 is the set of pure strategies of player 2, then we could derive from σ_2 a realization-equivalent behavior strategy. A similar argument holds for the conjecture of player 2 about player 1.

rule. Therefore we obtain three equilibrium conditions for the three components of (β_1, β_2, μ):

Definition 93. *Assessment* (β_1, β_2, μ) *is a* ***perfect Bayesian equilibrium*** (PBE) *of the Bayesian signaling game* Γ *if it satisfies the following conditions:*

$$\forall \theta \in \Theta, \text{supp}\beta_1(\cdot|\theta) \subseteq \arg\max_{a_1 \in A_1} \mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1, \cdot\right)\right) \tag{BR$_1$}$$

$$\forall a_1 \in A_1, \text{supp}\beta_2(\cdot|a_1) \subseteq \arg\max_{a_2 \in \mathcal{A}_2(a_1)} \mathbb{E}_{\mu}\left(u_2\left(\cdot, a_1, a_2\right)\right) \tag{BR$_2$}$$

$$\forall a_1 \in A_1, \forall \theta \in \Theta, \mu(\theta|a_1) \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta') = \beta_1(a_1|\theta)p(\theta). \tag{CONS}$$

Thus, from the perspective of an analyst who wants to compute the set of equilibrium assessments (β_1, β_2, μ), there are three systems of weak inequalities and equalities, one for each group of "unknowns" (endogenous variables) β_1, β_2, and μ. Note that (CONS)—consistency with Bayes rule—can also be expressed as

$$\begin{aligned} \forall a_1 &\in A_1, \forall \theta \in \Theta, \\ \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta') &> 0 \Rightarrow \mu(\theta|a_1) = \frac{\beta_1(a_1|\theta)p(\theta)}{\sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')}. \end{aligned}$$

Each equilibrium condition involves two out of the three groups of endogenous variables β_1, β_2 and μ: (BR$_1$) says that each mixed action $\beta_1(\cdot|\theta) \in \Delta(A_1)$ is a best reply to β_2 for type θ of the sender (for each $\theta \in \Theta$), (BR$_2$) says that each mixed action $\beta_2(\cdot|a_1) \in \Delta(\mathcal{A}_2(a_1))$ is a best reply for the receiver to the conditional belief $\mu(\cdot|a_1) \in \Delta(\Theta)$ (for each $a_1 \in A_1$), and (CONS) says that the triple (p, β_1, μ) is consistent with Bayes rule.

As we noticed more generally for multistage Bayesian games, PBE refines Bayesian equilibrium. Here we provide an explicit proof of this observation for the special case of signaling games.

Remark 59. *Each PBE* (β_1, β_2, μ) *of a Bayesian signaling game* Γ *is such that the (extended) behavior strategy pair* (β_1, β_2) *is a Bayesian equilibrium of* Γ*, that is, a (randomized) Nash equilibrium of the ex ante strategic form of* Γ*.*

Proof. Fix a PBE (β_1, β_2, μ). By inspection of the definition of PBE, the best reply condition for each type of the sender is satisfied. To see that also the strategic-form best reply condition of the receiver is satisfied note that the ex ante expected utility of the receiver can be expressed as

$$\begin{aligned}
\mathbb{E}_{\beta_1,\beta_2}(u_2) &= \sum_{\theta\in\Theta} p(\theta) \sum_{a_1\in A_1} \beta_1(a_1|\theta) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, u_2(\theta, a_1, a_2) \\
&= \sum_{a_1\in A_1} \sum_{\theta\in\Theta} \beta_1(a_1|\theta)\, p(\theta) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, u_2(\theta, a_1, a_2) \\
&= \sum_{a_1\in A_1} \sum_{\theta\in\Theta} \mu(\theta|a_1)\mathbb{P}(a_1) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, u_2(\theta, a_1, a_2) \\
&= \sum_{a_1\in A_1} \mathbb{P}(a_1) \sum_{a_2\in A_2} \beta_2(a_2|a_1) \sum_{\theta\in\Theta} \mu(\theta|a_1) u_2(\theta, a_1, a_2) \\
&= \sum_{a_1\in A_1} \mathbb{P}(a_1) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, \mathbb{E}_\mu(u_2(\cdot, a_1, a_2))
\end{aligned}$$

where the third equality follows from Bayes rule. Thus, β_2 maximizes $\mathbb{E}_{\beta_1,\beta_2}(u_2)$ if and only if each term $\sum_{a_2\in A_2} \beta_2(a_2|a_1)\, \mathbb{E}_\mu(u_2(\cdot, a_1, a_2))$ with $\mathbb{P}(a_1) > 0$ is maximized. Since condition (BR_2) requires that *each* term $\sum_{a_2\in A_2} \beta_2(a_2|a_1)\, \mathbb{E}_\mu(u_2(\cdot, a_1, a_2))$ is maximized (whether or not $\mathbb{P}(a_1) > 0$), it follows that also the best reply condition of the receiver holds, and (β_1, β_2) is a Bayesian equilibrium of the signaling game. ■

It is worth noting that the equilibrium probability

$$\mathbb{P}(a_1) = \sum_{\theta'\in\Theta} \beta_1(a_1|\theta')p(\theta')$$

of some action a_1 *may be zero*. Fix a_1 and suppose, for example, that for each $\theta \in \Theta$, there is some action a_1^θ such that $\mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1^\theta, \cdot\right)\right) > \mathbb{E}_{\beta_2}(u_1(\theta, a_1, \cdot))$. Then action/signal a_1 must have zero probability because it is not a best reply to β_2 for any type θ of the sender. *Yet we assume that the belief* $\mu(\cdot|a_1)$ *is well defined* and the receiver takes a best reply to this belief. This is a *perfection* requirement analogous to the subgame perfection condition for games with observable actions and complete information. A perfect Bayesian equilibrium is a Bayesian equilibrium that satisfies perfection and consistency with Bayes rule.

Furthermore, even if $\mu(\cdot|a_1)$ cannot be computed with Bayes formula, it may still be the case that the equilibrium conditions put incentive constraints on the possible values of $\mu(\cdot|a_1)$, because mixed action $\beta_2\,(\cdot|a_1)$ has to be a best reply to $\mu(\cdot|a_1)$. The following example illustrates this point.

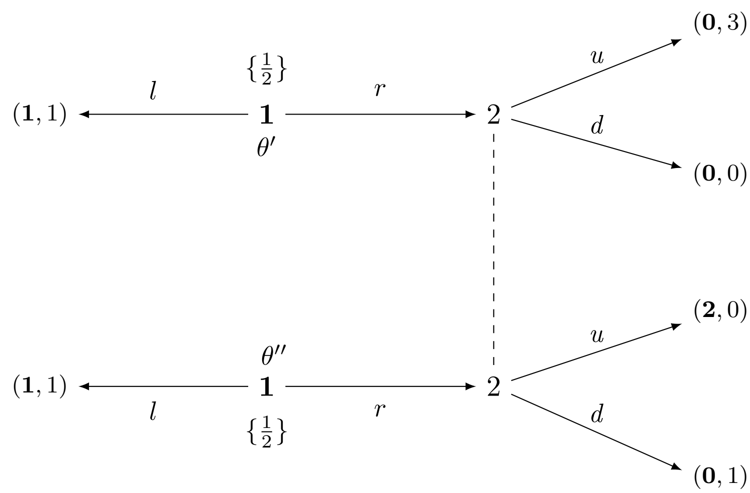


Figure 15.6: A signaling game

Example 71. Consider the signaling game depicted in Figure 15.6. The payoffs of the sender are in **bold.** The set $\mathcal{A}_2(l)$ is a singleton and therefore the action of player 2 after l is not shown. If the sender goes right (r) then the receiver can go up (u) or down (d), i.e., $\mathcal{A}_2(r) = \{u, d\}$.

Note that action r is dominated for type θ' of the sender. Therefore $\beta_1(r|\theta') = 0$ in every PBE. Now we show that in equilibrium we also have $\beta_1(r|\theta'') = 0$. Suppose, by way of contradiction, that $\beta_1(r|\theta'') > 0$ in a PBE. Then Bayes formula applies and $\mu(\theta''|r) = 1$. But then the best reply of the receiver is down, i.e., $\beta_2(d|r) = 1$, and the best reply of type θ'' is left, i.e., $\beta_1(l|\theta'') = 1 - \beta_1(r|\theta'') = 1$, contradicting our initial assumption. We conclude that, in every PBE, r is chosen with probability zero and $\mu(\cdot|r)$ cannot be determined with Bayes formula. Yet the equilibrium conditions put a constraint on $\mu(\cdot|r)$: in equilibrium d must be (weakly) preferred to u (if the receiver chooses u after r then type θ'' chooses r and we have just shown that this cannot happen in equilibrium). Therefore

$$\begin{aligned} \mu(\theta''|r) &\geq 3\mu(\theta'|r) \\ \text{or } \mu(\theta''|r) &\geq \frac{3}{4}. \end{aligned}$$

The set of equilibrium assessments is

$$\left\{(\beta_1, \beta_2, \mu) : \beta_1(l|\theta') = \beta_1(l|\theta'') = 1, \beta_2(d|r) = 1, \mu(\theta''|r) > \frac{3}{4}\right\}$$
$$\cup \left\{(\beta_1, \beta_2, \mu) : \beta_1(l|\theta') = \beta_1(l|\theta'') = 1, \beta_2(d|r) \geq \frac{1}{2}, \mu(\theta''|r) = \frac{3}{4}\right\}.$$

▲

The PBE assessments in Example 71 are examples of "pooling" equilibria. A **pooling equilibrium** is a PBE assessment where all types of the sender choose the same pure action with probability one, i.e., there exists $a_1^* \in A_1$ such that for every $\theta \in \Theta$, $\beta_1(a_1^*|\theta) = 1$. In this case Bayes rule implies that the posterior on θ conditional on the equilibrium action a_1^* is the same as the prior: $\mu(\cdot|a_1^*) = p(\cdot)$.

The polar case is when different types choose different pure actions: a **separating equilibrium** is a PBE assessment such that each type θ of player 1 chooses some action $a_1(\theta)$ with probability one ($\beta_1(a_1(\theta)|\theta) = 1$) and $a_1(\theta') \neq a_1(\theta'')$ for all θ' and θ'' with $\theta' \neq \theta''$. A separating equilibrium may exist only if A_1 has at least as many elements as Θ. *If A_1 and Θ have the same number of elements* (cardinality) then in a separating equilibrium each action is chosen with ex ante positive probability (because $p(\theta) > 0$ for each $\theta \in \Theta$) and the action of player 1 *perfectly reveals* her private information (if A_1 has more elements than Θ then the actions that in equilibrium are chosen by some type are perfectly revealing, the others need not be revealing).

We provide an example of a signaling game with a separating equilibrium.

Example 72. Consider the signaling game depicted in Figure 15.7. The game can be interpreted as follows: a truck driver (sender) enters in a pub where an aggressive customer (receiver) has to decide whether to start a fight (*f*, upward) or acquiesce (a, downward). There are two types of truck drivers: 90% of them are surly (θ^s) and like to eat sausages (s) for

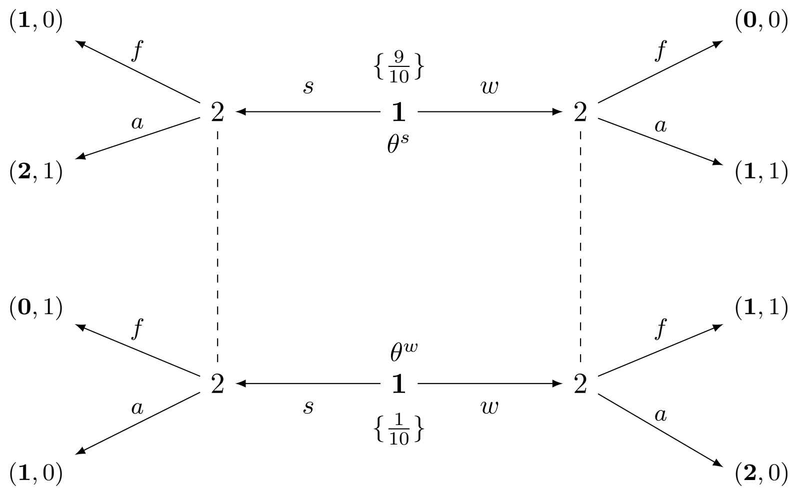


Figure 15.7: The "sausage-whipped cream" game

breakfast; the remaining 10% are wimps (θ^w) and prefer a dessert with whipped cream (w). Each type of truck driver receives an incremental utility equal to 2 from his favorite breakfast and an incremental utility equal to 1 from the other breakfast. Furthermore both types incur a loss of 1 util if they have to fight. The receiver prefers to fight with a wimp and to avoid the fight with a surly driver.

This game has only one "reasonable" PBE and it is separating.[22] Indeed, the payoffs are such that for each type of driver it is weakly dominant to have his preferred breakfast; thus, iterated deletion of weakly dominated actions yields the equilibrium $\beta_1(s|\theta^s) = 1 = \beta_1(w|\theta^w)$, $\beta_2(a|s) = 1 = \beta_2(f|w)$, $\mu(\theta^s|s) = 1 = \mu(\theta^w|w)$.

The game has also two sets of pooling equilibria (meaning that one type of player 1 chooses a weakly dominated action). In the first set of assessments each type has sausages for breakfast and player 2 would fight if and only he observed a whipped-cream breakfast: $\beta_1(s|\theta^s) = 1 = \beta_1(s|\theta^w)$, $\beta_2(a|s) = 1 = \beta_2(f|w)$, $\mu(\theta^s|s) = \frac{9}{10}$, $\mu(\theta^w|w) \geq \frac{1}{2}$. In the second set of assessments each type has whipped cream for breakfast and player 2

[22] Actually, this example is a modification of a well-known game where the cost of a fight is *larger* than the marginal benefit from having the preferred breakfast and all equilibria are pooling.

would fight if and only if he observed a sausage breakfast: $\beta_1(w|\theta^s) = 1 = \beta_1(w|\theta^w)$, $\beta_2(a|w) = 1 = \beta_2(f|s)$, $\mu(\theta^s|w) = \frac{9}{10}$, $\mu(\theta^w|s) \geq \frac{1}{2}$. ▲

15.8.1 An Algorithm to Compute Pure PBE's.

As for other finite-horizon games, also the equilibria of signaling games can be computed with a "case-by-case backward induction" algorithm. For simplicity, we focus on pure equilibria, that is, any PBE (β_1, β_2, μ) such that β_1 and β_2 are pure ($|\mathrm{supp}\beta_1(\cdot|\theta)| = 1$ for each θ, and $|\mathrm{supp}\beta_2(\cdot|a_1)| = 1$ for each a_1). We denote by $\mathbf{a}_1 \in A_1^\Theta$ the decision function of the sender, whereas $s_2 \in \bigtimes_{a_1 \in A_1} \mathcal{A}_2(a_1)$ denotes the pure strategy of the receiver (i.e., $\beta_1(\mathbf{a}_1(\theta)|\theta) = 1$ for each θ, and $\beta_2(s_2(a_1)|a_1) = 1$ for each a_1). Working backward, we start from the analysis of stage 2 (receiver) and then we move back to the analysis of stage 1 (sender). The analysis of stage 2, in turn, is broken down into a preliminary step and a main step.

Stage 2, preliminary step: deletion of dominated actions. For any a_1, an action $a_2 \in \mathcal{A}_2(a_1)$ is **conditionally dominated** given a_1 if there exists some $\alpha_2 \in \Delta(\mathcal{A}_2(a_1))$ such that

$$\forall\theta \in \Theta, \sum_{a_2' \in \mathcal{A}_2(a_1)} u_2(\theta, a_1, a_2')\, \alpha_2(a_2') > u_2(\theta, a_1, a_2);$$

otherwise, a_2 is **conditionally undominated** given a_1. Denote by $\mathcal{ND}(a_1) \subseteq \mathcal{A}_2(a_1)$ the set of conditionally undominated actions given a_1. An obvious adaptation of Lemma 2 (the *Wald-Pearce Lemma*) implies that, for every $a_1 \in A_1$ and $a_2 \in \mathcal{A}(a_1)$,

$$a_2 \in \mathcal{ND}(a_1) \iff \left(\exists\nu \in \Delta(\Theta), a_2 \in \arg\max_{a_2' \in \mathcal{A}_2(a_1)} \mathbb{E}_\nu\left(u_2(\cdot, a_1, a_2')\right)\right)$$

(see Lemma 33). In other words, a second-mover action is conditionally undominated if and only if it is a best reply to some conditional belief about the types of the first mover given the observed action of the first mover. With this, it is obvious that the conditionally dominated actions of player 2 are unjustifiable and must have zero probability in equilibrium. Therefore, the preliminary step requires to delete all the conditionally dominated actions of the receiver.[23]

[23] A similar preliminary step could be applied to player 1, but we do not explicitly

Stage 2, main step. Pick a strategy $s_2 \in \bigtimes_{a_1 \in A_1} \mathcal{ND}(a_1)$, that is, pick one action $s_2(a_1) \in \mathcal{ND}(a_1)$ among the undominated actions given each a_1. For each $a_2 \in \mathcal{ND}(a_1)$, let

$$J(a_2|a_1) = \left\{ \nu \in \Delta(\Theta) : a_2 \in \arg \max_{a_2' \in \mathcal{A}_2(a_1)} \mathbb{E}_\nu \left(u_2 \left(\cdot, a_1, a_2' \right) \right) \right\}$$

denote the nonempty set of beliefs $\nu \in \Delta(\Theta)$ that justify a_2 as a best reply given a_1. A strategy $s_2 \in \bigtimes_{a_1 \in A_1} \mathcal{ND}(a_1)$ (i.e., a selection from the undominated action correspondence $a_1 \mapsto \mathcal{ND}(a_1)$) is a "**case.**" Clearly, there are $\prod_{a_1 \in A_1} |\mathcal{ND}(a_1)|$ cases to be considered. For every such case s_2 (by Lemma 2 and Lemma 33), there is a system of beliefs $\mu \in \bigtimes_{a_1 \in A_1} J(s_2(a_1)|a_1)$ that makes s_2 a kind of "stage-2 equilibrium."[24]

Stage 1. For each case $s_2 \in \bigtimes_{a_1 \in A_1} \mathcal{ND}(a_1)$, pick a decision function $\mathbf{a}_1 \in \bigtimes_{\theta \in \Theta} \arg\max_{a_1 \in A_1} u_1(\theta, a_1, s_2(a_1))$, that is, pick an action

$$\mathbf{a}_1(\theta) \in \arg \max_{a_1 \in A_1} u_1(\theta, a_1, s_2(a_1))$$

for each type $\theta \in \Theta$. Finally, verify whether there exists a system of beliefs $\mu \in \bigtimes_{a_1 \in A_1} J(s_2(a_1)|a_1)$ that is consistent with Bayes rule given decision function $\mathbf{a}_1$, i.e., such that, for each $a_1 \in A_1$,

$$\sum_{\theta' : \mathbf{a}_1(\theta') = a_1} p(\theta') > 0 \Rightarrow \left(\forall \theta \in \mathbf{a}_1^{-1}(a_1), \mu(\theta|a_1) = \frac{p(\theta)}{\sum_{\theta' : \mathbf{a}_1(\theta') = a_1} p(\theta')} \right),$$

where

$$\sum_{\theta' : \mathbf{a}_1(\theta') = a_1} p(\theta') = p\left(\mathbf{a}_1^{-1}(a_1)\right) = \mathbb{P}^{\mathbf{a}_1}(a_1)$$

is the probability of action/signal a_1 determined by decision function $\mathbf{a}_1$. If such μ exists, then assessment $(\mathbf{a}_1, s_2, \mu)$ is a pure PBE; otherwise, there

include such step in the algorithm for the following reason: Suppose that a_1 is dominated for type θ of player 1. Clearly, $\beta_1(a_1|\theta) = 0$ in any equilibrium. But the "deletion" of such a_1 for type θ may induce a naive user of the algorithm to illegitimately infer that $\mu(\theta|a_1) = 0$. Such inference is correct and follows from Bayes rule if $\beta_1(a_1|\theta') > 0$ for some θ', but this cannot be presumed. It may be the case that, in equilibrium, $\beta_1(a_1|\theta') = 0$ for every θ'; hence, Bayes formula cannot determine the "off-path" conditional belief $\mu(\cdot|a_1)$ and there may be an equilibrium assessment such that $\mu(\theta|a_1) > 0$.

[24] An element of $\bigtimes_{a_1 \in A_1} J(s_2(a_1)|a_1) \subseteq \Delta(\Theta)^{A_1}$ is indeed a system of beliefs because it associates a probability measure on Θ with each action a_1 of the sender.

is no μ such that $(\mathbf{a}_1, s_2, \mu)$ is a PBE. Note that here we are applying Bayes rule in the special case where the behavior strategy of the sender is deterministic, that is, $\beta_1(a_1|\theta) = 1$ if $\mathbf{a}_1(\theta) = a_1$, and $\beta_1(a_1|\theta) = 0$ if $\mathbf{a}_1(\theta) \neq a_1$.

We illustrate the algorithm computing the equilibria of the previous "sausage-whipped cream" example (Figure 15.7). First note the preliminary stage-2 step is vacuous because the receiver has no conditionally dominated action. Since the receiver weakly prefers to fight if and only if the conditional probability of the surly type is not larger than 1/2, the sets of justifying beliefs are as follows:

$$\begin{aligned}
J(f|s) &= \left\{\mu(\cdot|s) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|s) \leq \frac{1}{2}\right\}, \\
J(a|s) &= \left\{\mu(\cdot|s) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|s) \geq \frac{1}{2}\right\}, \\
J(f|w) &= \left\{\mu(\cdot|w) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|w) \leq \frac{1}{2}\right\}, \\
J(a|w) &= \left\{\mu(\cdot|w) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|w) \geq \frac{1}{2}\right\}.
\end{aligned}$$

Also note that if the expected reply of the receiver does not depend on the signal—the breakfast of the truck driver—, then the sender eats his preferred breakfast. Thus, there are 2×2 possible cases, or undominated strategies of the receiver (we list first the reply to $s$ausage, second the reply to $w$hipped cream):

- **Case** $f.f$: θ^s chooses s, θ^w chooses w, thus $\mu(\theta^s|s) = 1$ and $\mu(\theta^s|w) = 0$. Since $1 = \mu(\theta^s|s) > \frac{1}{2}$, $\mu(\cdot|s) \notin J(f|s)$ and there is no PBE where the receiver plays $f.f$.

- **Case** $a.a$: θ^s chooses s, θ^w chooses w, thus $\mu(\theta^s|s) = 1$ and $\mu(\theta^w|w) = 1$. Since $0 = \mu(\theta^s|w) < \frac{1}{2}$, $\mu(\cdot|w) \notin J(a|w)$ and there is no PBE where the receiver plays $a.a$.

- **Case** $a.f$: θ^s chooses s, θ^w is indifferent because eating his preferred breakfast (whipped cream) would be compensated by the cost of a fight.

- **Subcase** $w|\theta^w$**:** θ^w chooses w. Thus, $\mu(\theta^s|s) = 1 \geq \frac{1}{2}$ and $\mu(\theta^s|w) = 0 \leq \frac{1}{2}$, which implies that $\mu(\cdot|s) \in J(a|s)$ and $\mu(\cdot|w) \in J(f|w)$ as required. We thus obtain the separating PBE (β_1, β_2, μ) with $\beta_1(s|\theta^s) = \beta_1(w|\theta^w) = 1$, $\beta_2(a|s) = \beta_2(f|w) = 1$, $\mu(\theta^s|s) = \mu(\theta^w|w) = 1$.
- **Subcase** $s|\theta^w$**:** also θ^w chooses s. Thus, $\mu(\theta^s|s) = p(\theta^s) = \frac{9}{10} \geq \frac{1}{2}$, which implies that $\mu(\cdot|s) \in J(a|s)$; on the other hand, $\mu(\cdot|w)$ is not determined by Bayes rule and we can pick any $\mu(\cdot|w) \in J(f|w)$, e.g., $\mu(\theta^s|w) = \frac{1}{3}$, to obtain a set of pure PBE's (β_1, β_2, μ) with $\beta_1(s|\theta^s) = \beta_1(s|\theta^w) = 1$, $\beta_2(a|s) = \beta_2(f|w) = 1$, $\mu(\theta^s|s) = \frac{9}{10}$, $\mu(\theta^s|w) \leq \frac{1}{2}$.

- **Case** $f.a$: θ^w chooses w, θ^s is indifferent because eating his preferred breakfast (sausage) would be compensated by the cost of a fight.

  - **Subcase** $s|\theta^s$**:** θ^s chooses s. Thus, $\mu(\theta^s|w) = 0 < \frac{1}{2}$**,** $\mu(\cdot|w) \notin J(a|w)$ and there is no PBE where the receiver plays $f.a$ and the sender eats his preferred breakfast.
  - **Subcase** $w|\theta^s$**:** also θ^s chooses w. Thus, $\mu(\theta^s|w) = p(\theta^s) = \frac{9}{10} \geq \frac{1}{2}$, which implies that $\mu(\cdot|w) \in J(a|w)$; on the other hand, $\mu(\cdot|s)$ is not determined by Bayes rule and we can pick any $\mu(\cdot|s) \in J(f|s)$, e.g., $\mu(\theta^s|s) = \frac{1}{3}$, to obtain a set of pure PBE's (β_1, β_2, μ) with $\beta_1(w|\theta^s) = \beta_1(w|\theta^w) = 1$, $\beta_2(a|w) = \beta_2(f|s) = 1$, $\mu(\theta^s|w) = \frac{9}{10}$, $\mu(\theta^s|s) \leq \frac{1}{2}$.

15.9 Appendix

Here we study an extension of the analysis of rational planning of Chapter 10 to encompass incomplete information. Compared to Chapter 10, we focus on uncertainty about the true map from terminal histories to "utils" and on Bayesian updating, while we trim some lengthy arguments explained in detail there. In what follows, we keep the same notation as in Section 15.3 and let $\beta^i \in \left(\bigtimes_{h\in H} \Delta(\mathcal{A}_{-i}(h))\right)^{\Theta_{-i}}$ denote a conjecture of i about the type-dependent behavior of the co-players, keeping in mind that in the perfect Bayesian equilibrium analisys such conjectures are determined by profiles β_{-i} of type-dependent behavior strategies.

Fix once and for all a type θ_i of player i. To ease notation we let $\mu_{\theta_i} = \mu_{\theta_i}(\cdot|\cdot) \in \Delta(\Theta_{-i})^H$ denote the personal system of beliefs of type θ_i, that is, with reference to the notation used for Bayesian games we write $\mu_{\theta_i}(\theta_{-i}|h)$ instead of $\mu_i(\theta_{-i}|\theta_i, h)$ for the probability assigned by type θ_i to profile θ_{-i} conditional on h. Hence we write a personal assessment of θ_i as a pair $\left(\beta^i, \mu_{\theta_i}\right)$. Similarly, we write $u_{\theta_i}(\theta_{-i}, z) = u_i(\theta_i, \theta_{-i}, z)$ (that is, u_{θ_i} is the section of u_i at θ_i). With this, we *fix* a *Bayes consistent* personal assessment $\left(\beta^i, \mu_{\theta_i}\right)$ and analyze the resulting subjective decision problem for type θ_i with parameterized payoff function $u_{\theta_i} : \Theta_{-i} \times Z \to \mathbb{R}$.

For a *fixed* θ_{-i}, the probability of action profile $a = (a_i, a_{-i})$ conditional on h given strategy s_i and conjecture β^i is

$$\mathbb{P}^{s_i,\beta^i}(a|\theta_{-i}, h) = \begin{cases} 0, & \text{if } a_i \neq s_i(h), \\ \beta^i(a_{-i}|\theta_{-i}, h), & \text{if } a_i = s_i(h). \end{cases}$$

Using the chain rule of conditional probabilities, we can define the probability of any history $h' \in H$ conditional on a prefix h ($h \prec h'$), given that s_i is played from h onward: let $h = (a^1, ..., a^{\ell(h)})$ and $h' = (a^1, ..., a^{\ell(h)}, ..., a^{\ell(h')})$, then

$$\mathbb{P}^{s_i,\beta^i}(h'|\theta_{-i}, h) = \prod_{t=\ell(h)+1}^{\ell(h')} \mathbb{P}^{s_i,\beta^i}\left(a^t|\theta_{-i}, \left(h, ..., a^{t-1}\right)\right),$$

with the convention that $\mathbb{P}^{s_i,\beta^i}\left(a^t|\theta_{-i}, \left(h, ..., a^{t-1}\right)\right) = \mathbb{P}^{s_i,\beta^i}\left(a^t|\theta_{-i}, h\right)$ if $t - 1 = \ell(h)$. With this, we can determine for all $s_i \in S_i$ and $h \in H$ the value for type θ_i of reaching h *given* θ_{-i} and given that s_i is followed from h onward:

$$V_{\theta_i}^{s_i,\beta^i}(\theta_{-i}, h) = \sum_{z \in Z(h)} \mathbb{P}^{s_i,\beta^i}(z|\theta_{-i}, h)\, u_{\theta_i}(\theta_{-i}, z),$$

which, of course, *depends only on the behavior of s_i at histories that weakly follow h.* Next we define, for all $s_i \in S_i$ and $h \in H$, the **subjective value for θ_i conditional on reaching** h given that s_i is followed from h

onward:[25]

$$V_{\theta_i}^{s_i,\beta^i,\mu_i}(h) = \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h)\, V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},h)$$

$$= \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{z\in Z(h)} \mathbb{P}^{s_i,\beta^i}(z|\theta_{-i},h)u_{\theta_i}(\theta_{-i},z),$$

which—again—depends only on the behavior of s_i at histories that weakly follow h. Similarly, for every $a_i \in \mathcal{A}_i(h)$ we define the **value of taking action** a_i at h given that s_i will be followed from the next stage:

$$V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,a_i) = \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h)V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},(h,(a_i,a_{-i}))),$$

which depends only on the behavior of s_i at histories that follow h after taking action a_i.

Remark 60. *For every $s_i \in S_i$ and for every $h \in H$,*

$$V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,s_i(h)) = V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).$$

[25] To ease notation, here and in the following formulas we write function symbols like V and $\mathbb{P}$ with μ_i instead of μ_{θ_i} in the superscript, and with θ_i as a subscript of the symbol. Consider the following interpretation: $\mu_i : \Theta_i \times H \to \Delta(\Theta_{-i})$ is the system of beliefs of player i before he observes his private information θ_i. Writing θ_i in the subscript of V or $\mathbb{P}$ means that we select from μ_i its section at θ_i, that is, μ_{θ_i}.

Proof. By inspection of the definitions we have

$$
\begin{aligned}
& V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,s_i(h)) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h) V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},(h,(s_i(h),a_{-i}))) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h) \\
& \cdot \left(\sum_{z\in Z(h,(s_i(h),a_{-i}))} u_{\theta_i}(\theta_{-i},z) \mathbb{P}^{s_i,\beta^i}(z|\theta_{-i},(h,(s_i(h),a_{-i}))) \right) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{z\in Z(h)} u_{\theta_i}(\theta_{-i},z) \mathbb{P}^{s_i,\beta^i}(z|\theta_{-i},h) \\
= & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},h) \\
= & V_i^{s_i,\beta^i}(h)
\end{aligned}
$$

where the third equality follows from the chain rule. ■

As we did in Chapter 10, we define recursively the subjective value of reaching a history h and of taking an action a_i at h, under the presumption that the behavior of player i will be subjectively rational in the following stages. We use the symbol $\hat{V}_{\theta_i}^{\beta^i,\mu_i}$ to denote such values to emphasize that they are optimal given personal assessment (β^i,μ_{θ_i}). Similarly, for all $h\in H$ and $a_{-i}\in\mathcal{A}_{-i}(h)$, we let

$$\mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) = \sum_{\theta_{-i}\in\Theta_{-i}} \beta^i(a_{-i}|\theta_{-i},h)\mu_{\theta_i}(\theta_{-i}|h) \tag{15.9.1}$$

denote the probability assigned by type θ_i to a_{-i} conditional on h according to personal assessment (β^i,μ_{θ_i}). With this, the recursion is based on the height of histories $h\in H$, here denoted by $L(h)=\max_{z\in Z(h)}\ell(z)-\ell(h)$.[26]

[26]In Chapter 10, each $h\in H$ determined a subgame $\Gamma(h)$ and the height of h was written as the maximum length of such subgame, $L(\Gamma(h))$. Here we do not have a subgame in the same sense as Chapter 10, because payoffs are unknown, but we still have a subtree, which is all that matters.

- If $L(h) = 1$, then $(h, a) \in Z$ for every $a \in \mathcal{A}(h)$. Taking this into account, for each $a_i \in \mathcal{A}_i(h)$ let

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h) u_{\theta_i}(\theta_{-i}, (h, (a_i, a_{-i}))),$$

and let

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) = \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i).$$

- Suppose $\hat{V}_{\theta_i}^{\beta^i,\mu_i}$ has been defined for every h with $L(h) \leq k$. Then if $L(h) = k + 1$ let

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, (a_i, a_{-i})),$$

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) = \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i).$$

Definition 94. *A strategy $\bar{s}_i \in S_i$ is*

- ***folding-back optimal given*** $\left(\beta^i, \mu_{\theta_i}\right)$ *if*

$$\forall h \in H,\ \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, \bar{s}_i(h)) = \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h);$$

- ***one-step optimal given*** $\left(\beta^i, \mu_{\theta_i}\right)$ *if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, \bar{s}_i(h)) = \max_{a_i\in\mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i);$$

- ***sequentially optimal given*** $\left(\beta^i, \mu_{\theta_i}\right)$ *if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) = \max_{s_i\in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).$$

The following result is immediate from Definition 94:

Remark 61. *If a strategy is sequentially optimal given $\left(\beta^i, \mu_{\theta_i}\right)$, then it is also one-step optimal given $\left(\beta^i, \mu_{\theta_i}\right)$.*

The following results show that, under personal Bayes consistency, folding-back optimality and sequential optimality are equivalent to one-step optimality.

Proposition 21. *(*Folding Back Principle) *Fix a Bayes consistent assessment* $\left(\beta^i, \mu_{\theta_i}\right)$*. Then*
(I) A strategy $\bar{s}_i$ *is one-step optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$ *if and only if*

$$\begin{aligned} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) &= \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i), \\ V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) &= \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) \end{aligned} \tag{15.9.2}$$

for every $h \in H$ *and* $a_i \in \mathcal{A}_i(h)$*.*
(II) A strategy $\bar{s}_i$ *is folding-back optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$ *if and only if* $\bar{s}_i$ *is one-step optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$*.*

Proof. We prove by induction on the height of non-terminal histories that *if* $\bar{s}_i$ is *one-step optimal* given $\left(\beta^i, \mu_{\theta_i}\right)$ then (15.9.2) holds for every $h \in H$ and $a_i \in \mathcal{A}_i(h)$. As in the proof of Proposition 3 in Chapter 10, we omit (and leave to the reader) the proofs that the converse of this statement holds, that (II) holds, and that (I) is equivalent to (II).

Suppose that $\bar{s}_i$ is one-step optimal given $\left(\beta^i, \mu_{\theta_i}\right)$.

Basis step. Fix any $h \in H$ such that $L(h) = 1$. Then, for each $a_i \in \mathcal{A}_i(h)$, by inspection of the definitions of $V_{\theta_i}^{s_i,\beta^i,\mu_i}(h, a_i)$ and $\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i)$ (see the basis step in the recursion) we obtain

$$\begin{aligned} & V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\ = & \sum_{\theta_{-i} \in \Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h) u_i(\theta_i, \theta_{-i}, (h, (a_i, a_{-i}))) \\ = & \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i). \end{aligned}$$

Since $\bar{s}_i$ is one-step optimal,

$$\begin{aligned} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) &= V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\ &= \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h). \end{aligned}$$

Inductive step. Suppose that (15.9.2) holds for each $h \in H$ with $L(h) \le k$. Now, fix any h with $L(h) = k + 1$. Then $L(h, a) \le k$ for each $a \in \mathcal{A}(h)$. Therefore, for every $a_i \in \mathcal{A}_i(h)$,

$$\begin{aligned}
&V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\
&\overset{\text{(def.)}}{=} \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\
&= \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h)\mu_{\theta_i}(\theta_{-i}|h) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\
&\overset{(15.3.4)}{=} \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mu_{\theta_i}(\theta_{-i}|\,(h, (a_i, a_{-i})))\mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\
&= \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|\,(h, (a_i, a_{-i}))) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\
&\overset{\text{(def.)}}{=} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, (a_i, a_{-i})) \\
&\overset{\text{(I.H.)}}{=} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, (a_i, a_{-i})) \overset{\text{(def.)}}{=} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i),
\end{aligned}$$

where the first, fifth, and seventh equalities hold by definition, the third equality follows from eq. (15.3.4) in the definition of Bayes consistency, the sixth equality follows from the inductive hypothesis, and the remaining equalities are immediate. To see that the third equality follows from Bayes consistency, note that eq. (15.3.4) yields

$$\mu_{\theta_i}(\theta_{-i}|\,(h, (a_i, a_{-i})))\mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) = \beta^i(a_{-i}|\theta_{-i}, h)\mu_{\theta_i}(\theta_{-i}|h),$$

where both sides of the equality represent the probability of (θ_{-i}, a_{-i}) given h.

Thus we obtain

$$\begin{aligned}
V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) &= V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, \bar{s}_i(h)) \overset{\text{(loc.opt.)}}{=} \max_{a_i\in\mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\
&= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h),
\end{aligned}$$

where the first equality follows from Remark 60 and the second equality holds because $\bar{s}_i$ is locally optimal (loc.opt.) at h. ■

Theorem 44. (Optimality principle) *Fix a Bayes consistent assessment* $(\beta^i, \mu_{\theta_i})$. *A strategy of player* i *is sequentially optimal given* $(\beta^i, \mu_{\theta_i})$ *if and only if it is folding-back optimal given* $(\beta^i, \mu_{\theta_i})$.

Proof. (If) Let $\bar{s}_i$ be *folding-back optimal* given $(\beta^i, \mu_{\theta_i})$. We will prove by induction on the height of non-terminal histories that

$$\forall h \in H,\ \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) \geq \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).$$

With this, since $\bar{s}_i$ is folding-back optimal, it is one-step optimal (Proposition 21 II), therefore $V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) = \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h)$ for every $h \in H$ (Proposition 21 I); this implies that $\bar{s}_i$ is sequentially optimal given $(\beta^i, \mu_{\theta_i})$.

Basis step. Fix any $h \in H$ such that $L(h) = 1$. By Proposition 21, $\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i)$ for every $a_i \in \mathcal{A}_i(h)$. Since, starting from h, every action profile $(a_i, a_{-i}) \in \mathcal{A}(h)$ terminates the game,

$$\max_{a_i \in \mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) = \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).$$

Therefore

$$\begin{aligned} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) &= \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) \\ &= \max_{a_i \in \mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) = \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h). \end{aligned}$$

Inductive step. Suppose by way of induction that $\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h') \geq \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h')$ for every $h' \in H$ with $L(h') \leq k$. Now fix any h with $L(h) = k+1$. Then $L(h, a) \leq k$ for each $a \in \mathcal{A}(h)$, and the inductive assumption (I.H.) yields

$$\begin{aligned} &\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) \\ &\overset{\text{(def.)}}{\geq} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) \\ &\overset{\text{(def.)}}{=} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, (a_i, a_{-i})) \\ &\overset{\text{(I.H.)}}{\geq} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h, (a_i, a_{-i})) \end{aligned}$$

for every $a_i \in \mathcal{A}_i(h)$. Therefore,

$$
\begin{aligned}
& \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) \\
\geq\ & \max_{a_i \in \mathcal{A}_i(h)} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i})) \\
=\ & \max_{a_i \in \mathcal{A}_i(h)} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{s_i \in S_i} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i})) \\
=\ & \max_{a_i \in \mathcal{A}_i(h)} \max_{s_i \in S_i} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i})) \\
=\ & \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h)
\end{aligned}
$$

where the first equality is obvious, the second holds because each term in the summation can be maximized independently of the others, and the third equality holds because global maximization from h is equivalent to picking at h an action that yields the largest maximal continuation value.

(Only if) If $\bar{s}_i$ is sequentially optimal, then it is one-step optimal (Remark 61), which in turn implies that it is folding-back optimal (Proposition 21 II). ■

Proposition 21 and Theorem 44 yield the OD principle:

Corollary 9. (One-deviation principle) *Fix a Bayes consistent assessment $(\beta^i, \mu_{\theta_i})$. A strategy of player i is sequentially optimal given $(\beta^i, \mu_{\theta_i})$ if and only if it is one-step optimal given $(\beta^i, \mu_{\theta_i})$.*

16

Imperfectly Observable Actions

Bibliography

[1] Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica,* 56, 383-396.

[2] Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica,* 58, 1041-1063.

[3] Aliprantis, C. R. and K. C. Border (2006 [1999]): *Infinite Dimensional Analysis.* Berlin: Springer-Verlag.

[4] Attanasi, G. and R. Nagel (2007): "A Survey of Psychological Games: Theoretical Findings and Experimental Evidence," in *Games, Rationality and Behaviour. Essays on Behavioural Game Theory and Experiments*, ed. by A. Innocenti, and P. Sbriglia. Houndmills: Palgrave McMillan, 204-232.

[5] Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics,* 1, 67-96.

[6] Aumann, R. J. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica,* 55, 1-18.

[7] Battigalli P. (2003): "Rationalizability in Infinite, Dynamic Games of Incomplete Information," *Research in Economics,* 57, 1–38.

[8] Battigalli, P. and G. Bonanno (1999): "Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory," *Research in Economics*, 53, 149-226.

[9] Battigalli, P., R. Corrao, and M. Dufwenberg (2019): "Incorporating Belief-Dependent Motivation in Games," *Journal of Economic Behavior and Organization*, 167, 185-218.

[10] Battigalli, P. and N. De Vito (2021): "Beliefs, Plans, and Perceived Intentions in Dynamic Games," *Journal of Economic Theory*, 195, 105283.

[11] Battigalli, P. and M. Dufwenberg (2009): "Dynamic Psychological Games," *Journal of Economic Theory*, 144, 1-35.

[12] Battigalli, P. and M. Dufwenberg (2021): "Belief-Dependent Motivations and Psychological Game Theory," *Journal of Economic Literature*, 60, 833-882.

[13] Battigalli, P. and M. Siniscalchi (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356-391.

[14] Battigalli, P. and M. Siniscalchi (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3 (1), Article 3.

[15] Battigalli, P., M. Gilli, and M. C. Molinari (1992): "Learning and Convergence to Equilibrium in Repeated Strategic Interaction," *Research in Economics (Ricerche Economiche),* 46, 335-378.

[16] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2022): "Selfconfirming Equilibrium and Model Uncertainty," *Theory and Decisions,* 92, 455-512.

[17] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2016): "A Note on Comparative Ambiguity Aversion and Justifiability," *Econometrica*, 84, 1903-1916.

[18] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and T. Sargent (2016): "A Framework for the Analysis of Self-Confirming Policies," Bocconi Univeristy, IGIER w.p. 573.

[19] Battigalli, P., A. Di Tillio, E. Grillo, and A. Penta (2011): "Interactive Epistemology and Solution Concepts for Games with Asymmetric Information," *The B.E. Journal of Theoretical Economics*, 11 (1) (*Advances*), Article 6.

[20] Ben-Porath, E. and E. Dekel (1992): "Signaling Future Actions and the Potential for Sacrifice," *Journal of Economic Theory*, 57, 36-51.

[21] Bernheim, D. (1984): "Rationalizable Strategic Behavior," *Econometrica*, 52, 1007-1028.

[22] Binmore, K. G., (1982 [1977]): *Mathematical Analysis: A Straightforward Approach*. Cambridge UK: Cambridge University Press.

[23] Brandenburger, A. (2007): "The Power of Paradox," *International Journal of Game Theory,* 35, 465-492.

[24] Brandenburger, A. and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1391-1402.

[25] Boergers, T. and M. Janssen (1995): "On the Dominance Solvability of Large Cournot Games," *Games and Economic Behavior*, 8, 297-321.

[26] Dekel, E. and M. Siniscalchi (2014): "Epistemic Game Theory," in P. Young and S. Zamir (Eds.) *Handbook of Game Theory*, Vol. 4, 619-702.

[27] Dekel, E., D. Fudenberg, and D. Levine (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282-303.

[28] Dekel, E., D. Fudenberg, and S. Morris (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15-40.

[29] Dufwenberg, M. and M. Stegeman (2002): "Existence and Uniqueness of Maximal Reductions Under Iterated Strict Dominance," *Econometrica*, 70, 2007-2023.

[30] Esponda, I. (2013): "Rationalizable Conjectural Equilibrium: A Framework for Robust Predictions," *Theoretical Economics*, 8, 467-501.

[31] Eyster, E. and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623-1672.

[32] Friedman, J.W. (1987): *Game Theory with Applications to Economics,* Oxford: Oxford University Press.

[33] Fudenberg, D. and D. Levine (1998): *The Theory of Learning in Games*, Cambridge MA: MIT Press.

[34] Fudenberg, D. and J. Tirole (1991): *Game Theory,* Cambridge MA: MIT Press.

[35] Geanakoplos, J., D. Pearce, and E. Stacchetti (1989): "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1, 60-79.

[36] Gibbons, R. (1992): *Game Theory for Applied Economists.* Princeton: Princeton University Press.

[37] Gilboa, I. and M. Marinacci (2013): "Ambiguity and the Bayesian Paradigm," in D. Acemoglu, M. Arellano, and E. Dekel (Eds.) *Advances in Economics and Econometrics: Theory and Applications (Vol. 1)*, Cambridge UK: Cambridge University Press, 179-242.

[38] Gilli, M. (1999): "Adaptive Learning in Imperfect Monitoring Games," *Review of Economic Dynamics*, 2, 472-485.

[39] Guesnerie, R. (1992): "An Exploration on the Eductive Justifications of the Rational-Expectations Hypothesis," *American Economic Review*, 82, 1254-1278.

[40] Harsanyi, J. (1967-68): "Games of Incomplete Information Played by Bayesian Players. Parts I, II, III," *Management Science*, 14, 159-182, 320-334, 486-502.

[41] Harsanyi, J. (1995): "Games with Incomplete Information," *American Economic Review*, 85, 291-303.

[42] Hellwig, M., W. Leininger, P. Reny, and A. Robson (1990): "Subgame Perfect Equilibrium in Continuous Games of Perfect Information: An Elementary Approach to Existence and Approximation by Discrete Games," *Journal of the Economic Theory,* 52, 406-422.

[43] Kuhn, H. W. (1953): "Extensive Games and the Problem of Information," in: H.W. Kuhn and A.W. Tucker (Eds.), *Contributions to the Theory of Games II*, Princeton NJ: Princeton University Press, 193-216.

[44] Marinacci, M. (2015): "Model Uncertainty," *Journal of the European Economic Association*, 13, 1022-1100.

[45] Maynard Smith, J. (1982): *Evolution and the Theory of Games,* Cambridge: Cambridge University Press.

[46] Milgrom, P. and J. Roberts (1991): "Adaptive and Sophisticated Learning in Normal-Form Games," *Games and Economic Behavior*, 3, 82-100.

[47] Milgrom, P., and R. Weber (1985): "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operation Research,* 10, 619-632.

[48] Moulin, H. (1984): "Dominance-Solvability and Cournot Stability," *Mathematical Social Sciences*, 7, 83-102.

[49] Nash, J. (1950): "Equilibrium Points in N-Person Games," *Proceedings of the National Academy of Sciences*, 36, 48-49.

[50] Nash, J. (1950): *Non-cooperative Games*. Ph.D. thesis, Princeton University.

[51] Motta, M. (2004): *Competition Policy: Theory and Practice*. Cambridge University Press.

[52] Ok, E. A. (2007): *Real Analysis with Economic Applications*, Princeton NJ: Princeton University Press.

[53] Osborne, M. and A. Rubinstein (1994): *A Course in Game Theory*, Cambridge MA: MIT Press.

[54] PEARCE, D. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1029-1050.

[55] PEREA, A. (2001): *Rationality in Extensive Form Games*, Boston, Dordrecht: Kluwer Academic Publishers.

[56] PEREA, A. (2014): "Belief in the Opponents' Future Rationality," *Games and Economic Behavior,* 83, 231-254.

[57] RUBINSTEIN, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-109.

[58] RUBINSTEIN, A. (1989): "The Electronic Mail Game: Strategic Behavior Under Almost Common Knowledge," *American Economic Review*, 79, 385-391.

[59] RUBINSTEIN, A. (1989): "Comments on the Interpretation of Game Theory," *Econometrica*, 59, 909-924.

[60] SANDHOLM, W. (2010): *Population Games and Evolutionary Dynamics,* Cambridge MA: MIT Press.

[61] SION, M. (1958): "On General Minimax Theorems," *Pacific Journal of Mathematics*, 171-176.

[62] SHIMOJI, M. (2004): "On the Equivalence of Weak Dominance and Sequential Best Response," *Games and Economic Behavior*, 48, 385-402.

[63] SHIMOJI, M. (2004): "On Forward Induction in Money-Burning Games," *Economic Theory*, 19, 637-648.

[64] SHIMOJI, M. AND J. WATSON (1998): "Conditional Dominance, Rationalizability, and Game Forms," *Journal of Economic Theory,* 83, 161-195.

[65] TOPKIS, D. M. (1998): *Supermodularity and Complementarity.* Princeton University Press.

[66] VICKREY, W. (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16, 8-37.

[67] von Neumann, J. (1928): "Zur Theorie der Gesellschaftsspiele," *Math. Annalen*, 100, 295-320.

[68] von Neumann, J. and O. Morgenstern (2004 [1944]): *Theory of Games and Economic Behavior,* Princeton NJ: Princeton University Press.

[69] Wald, A. (1949): "Statistical Decision Functions," *Annals of Mathematical Statistics*, 20, 165-205.

[70] Wald, A. (1950): *Statistical Decision Functions*. New York: Wiley.

[71] Weibull, J. (1995): *Evolutionary Game Theory,* Cambridge MA: MIT Press.

[72] Weibull, J. W. (1996): "The Mass Action Interpretation" in H. Kuhn et al. "The work of John Nash in game theory,"*Journal of Economic Theory*, 69, 153-185.

[73] Weinstein, J. (2016): "The Effect of Changes in Risk Attitude on Strategic Behavior," *Econometrica*, 84, 1881-1902.