

UNIVERSITY OF CALIFORNIA

Los Angeles

## ESSAYS ON REPUTATION

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Economics

by

Marco Celentani

1993

© Copyright by

Marco Celentani

1993

The dissertation of Marco Celentani is approved.

---

Joseph M. Ostroy

---

David Hirshleifer

---

David K. Levine, Committee Chair

University of California, Los Angeles

1993

# Contents

1	Overview	1
2	Reputation with Deterministic Stage Games	8
2.1	Introduction . . . . .	9
2.2	The Model . . . . .	12
2.3	The Result . . . . .	15
2.4	The Quality Game . . . . .	20
2.5	Discussion . . . . .	24
3	Reputation in Repeated Games with Two Long Run Players	27
3.1	Introduction . . . . .	28
3.2	The Model . . . . .	33
3.3	Establishing a Reputation Against a Patient Opponent . . . . .	35
3.4	The Value of Reputation with an Arbitrarily Patient Opponent . . . . .	41
3.5	Special Cases . . . . .	49
3.5.1	Reputation with a Short Run Opponent . . . . .	49
3.5.2	Games of Con°icting Interest . . . . .	50
3.6	Conclusion . . . . .	51
4	Reputation in Dynamic Games	54
4.1	Introduction . . . . .	55

4.2	Description of the Game . . . . .	60
4.2.1	Best Response and Aggregate Best Response . . . . .	62
4.3	The Perturbed Game . . . . .	64
4.4	The Case with No Strategic Externality . . . . .	66
4.5	Including Strategic Externalities . . . . .	69
4.6	Patient Small Players . . . . .	74
4.6.1	The Failure of Reputation in the Durable Goods Monopoly	75
4.6.2	No Irreversible Actions . . . . .	77
4.7	Conclusions . . . . .	80
4.8	Proofs: . . . . .	81
4.8.1	Proof of Lemma 8 . . . . .	81
4.8.2	Proof of Theorem 4 . . . . .	83
4.8.3	Proof of Theorem 5 . . . . .	86
4.8.4	Proof of Theorem 6 . . . . .	89
5	References	92

# List of Figures

2.1	Quality Game . . . . .	26
2.2	Quality Game with unspecified payoff <sup>®</sup> for the short run player . . .	26

## ACKNOWLEDGEMENTS

I wish to acknowledge the frequent and very stimulating conversations I had with Joe Ostroy and Jean-Laurent Rosenthal throughout my academic life at UCLA. Special thanks to Sushil Bikhchandani, Michele Boldrin, Bryan Ellickson, Drew Fudenberg, Roger Farmer, Gary Hansen, David Hirshleifer, Axel Leijonhufvud, John Riley, Manuel Santos, and Bill Zame for useful questions, discussions, and suggestions. I feel indebted to my fellow students at UCLA and in particular to Alex David and Wolfgang Pesendorfer for the countless discussions we had. Finally, I wish to thank my advisor, David Levine, for his endless encouragement and support and most of all for his intellectual stimulus, that I like to think had a strong influence on my formation as an economist.

The financial support of University of California Graduate Fellowship, National Science Foundation, Fondazione Einaudi, Torino, Italy, and Ferruzzi Finanziaria, Italy, is gratefully acknowledged.

## VITA

- 4 June 1963      Born, Naples, Italy
- 1987              Laurea, Economics  
University of Naples Federico II  
Naples, Italy
- 1990              M.A., Economics  
University of California, Los Angeles  
Los Angeles, California
- 1989 - 1992      Teaching Assistant and Associate  
Department of Economics  
University of California, Los Angeles  
Los Angeles, California
- 1992 - 1993      Assistant Professor  
Department of Economics  
Universidad Carlos III de Madrid  
Madrid, Spain

## PUBLICATIONS AND PRESENTATIONS

Celentani, M. (1990). "Spanning by Frequent Trading of Long-Lived Securities", *Studi Economici*, 42(3), pp. 59-82.

Celentani, M. (1991). "Reputation with Deterministic Stage Games", UCLA Working Paper no. 636S.

Celentani, M. (1992). "Breaking Collusive Agreements in the Criminal Sector"

Celentani, M. (1993). "Regulating the Organized Crime Sector"



ABSTRACT OF THE DISSERTATION

ESSAYS ON REPUTATION

by

Marco Celentani

Doctor of Philosophy in Economics

University of California, Los Angeles, 1993

Professor David K. Levine, Chair

The goal of the literature on reputation effects is to provide equilibrium characterizations for games in which there is some uncertainty relative to the utility function of a player. The purpose of this dissertation is to study the conditions under which this kind of reputational argument can be applied to frameworks that are interesting for their economic applications.

Chapter 2 considers a repeated game between a long run player and a sequence of short run players who are alive for one period only and therefore do not care about the future. When the stage game is allowed to be a sequential move one, if a short run player does not expect the long run player to play in a certain way, he might choose an action that does not reveal the strategy chosen by the long run player. Chapter 2 introduces a perturbation on the type of the short run players that guarantees that all informational nodes of the stage game are visited with strictly positive probability on any equilibrium path. This assumption is shown to be sufficient for the long run player to be able to establish a reputation for repeatedly playing any particular stage game pure strategy

Chapter 3 studies reputational arguments in repeated games between a patient player (player 1) who tries to establish a reputation against a patient opponent (player 2). In such a case it is possible that player 2's play might prevent player 1 from establishing an appropriate reputation. The assumption that the action chosen by player 2 is not perfectly observed by player 1 is shown to have very strong implications for equilibrium characterization, in particular in the case in which player 2 is sufficiently patient. In this case it is shown that an arbitrarily patient player 1 can guarantee himself an equilibrium average discounted payoff<sup>®</sup> which is at least equal to the highest payoff<sup>®</sup> from a correlated strategy subject to the constraint that player 2 gets at least his pure strategy minmax payoff<sup>®</sup>.

Chapter 4 turns attention to dynamic games (i.e. to games in which current payoff<sup>®</sup> opportunities depend on the past history of the game through a state variable, such as capital or debt) between a large player and a continuum of small players. When the small players have a fixed discount factor while the large player is arbitrarily patient, it is shown that in any Nash equilibrium the large player is guaranteed at least the optimal commitment payoff<sup>®</sup>. The case in which both the large and the small players are arbitrarily patient is then analyzed and it is shown that the large player will only be able to exploit his reputation if the transition function is reversible, in the sense that players can move from one state to another only if they can also return. An example shows how the failure of this condition in the durable goods monopoly problems prevents player 1 from successfully establishing a desirable reputation.

# Chapter 1

## Overview

The literature on reputation effects studies the implications for equilibrium behavior of a possibly small amount of uncertainty relative to the preferences of one player. The general idea first introduced by Fudenberg and Levine (1989) is that if a player's utility function is not known, he has the option of imitating the play of a player with a utility function different from his own in order to convince his opponents that he will play in a certain way, or in other words, in order to establish a reputation for a desirable behavior.

Most of the literature on reputation effects to date has concentrated on the analysis of repeated simultaneous games between a long run player (a player who lives forever and cares about the future) and a sequence of short run players (who are alive for one period only), in which the long run player can establish a reputation for a particular behavior.

The purpose of this dissertation is to study how this general argument can be applied to frameworks which are interesting for their economic applications

Chapter 2 addresses the problems arising when stage game is not a simultaneous move game but a sequential one. In this case the long run player might be prevented from building a reputation since, if the short run players do not expect him to play in a certain way they might play an action that does not "reveal" the long run player's strategy. A point in case is the Quality Game in which a short run player has to decide whether to buy a product or not and the long run player has to decide whether to produce a high quality or a low quality product, but the quality of the product is revealed only if a purchase occurs. If the short run players do not expect the long run player to produce a high quality product they will not buy and the long run player has no way to establish a reputation for producing high quality products.

In order to avoid this problem Chapter 2 introduces a perturbation also on

the types of the short run players, so as to guarantee that on any equilibrium path all informational nodes of the stage game are visited with strictly positive probability. This allows the long run player to establish a reputation for the so called Stackelberg strategy (the strategy that maximizes his payoff<sup>®</sup> subject to his opponent playing a best response) which guarantees him the Stackelberg payoff<sup>®</sup>. It is finally argued that since the long run player can always establish a reputation for such a strategy and therefore obtain the corresponding payoff<sup>®</sup>, he has to get at least this much in any equilibrium.

Even though Chapter 2 introduces a perturbation on the types of the short run players, the crucial point is to make sure that all informational nodes of the stage game are visited with strictly positive probability on any equilibrium path. In this sense the main point of Chapter 2 is that introducing a perturbation on the informational structure of a game can be beneficial to a long run player who might want to establish a reputation.

This point is a general one and is pursued in Chapter 3 in which reputation effects are studied in a repeated game between two patient (long run) players. Trying to establish a reputation against a patient opponent, a player can run into a problem similar to the one observed for a sequential move stage game. As Schmidt (1993) pointed out, the nature of the problem is twofold: first of all, a patient player cares not only about his opponent's current play but also about his future one; second, a player might believe that if he does play a best response to his opponent's expected play, then he will be punished thereafter.

Schmidt (1993) solved the problem considering a particular class of games, games of conflicting interests with respect to player 1, in which player 2's best response to player 1's static Stackelberg strategy gives player 2 his minmax payoff<sup>®</sup>.

Chapter 3's goal is to provide an equilibrium characterization of a more general

class of games. This is done pursuing the idea introduced in Chapter 2 of introducing a perturbation on the outcomes of player 2's actions. This perturbation makes sure that all finite length histories occur with strictly positive probability. Under the assumption that there exist with strictly positive probability types of player 1 that are committed to strategies that depend only on histories of finite length, player 1 is allowed to establish a reputation for any such strategy.

The implications of the above assumptions are shown to be very strong. First of all Schmidt's (1993) result is shown to be a special case of this more general model. Moreover the result is valid for all finite stage games once the perturbation on the outcome of player 2 is introduced. Finally and more importantly, the equilibrium characterization for games which do not have conflicting interests is stronger than the one proposed by Schmidt (1993), since the optimal strategy to commit to is not a fixed action (the static Stackelberg strategy) but it is a history (and in particular time) dependent strategy. It is shown that, if player 2 is sufficiently patient, then an arbitrarily patient player 1 can guarantee himself the highest payoff from a correlated strategy subject only to the constraint that player 2 gets more than his minmax payoff.

Two interesting applications of the general model discussed in Chapter 3 are a repeated game between two patient players, in which player 2's trembles make sure that the outcome of his action is not perfectly observed, or a repeated principal agent model with the usual assumption of invariance of the support of the distribution over outcomes as a function of the action chosen by the agent. In the latter case, the implications of the model presented are shown to be very strong, since the principal can establish a reputation for making payments which depend not only on the current outcome but also on the past history of the game, thus inducing the first best effort level on the agent's side, while appropriating all the

net surplus.

Many economic problems have the feature that a state variable such as capital, debt or money, provides a link between present actions and future payoff opportunities. As an example, games that describe the strategic interaction between a government and households usually involve state variables. It is in this context that the problem of time consistency of optimal government policy arises: since ex-ante and ex-post optimal policies differ, even a benevolent government may not be able to achieve the optimal commitment outcome.

Chapter 4 turns attention to this kind of problems to consider a general class of dynamic games with one large player and a large number of small players, i.e. to games in which current payoff opportunities may depend on the history of the game through a state variable.

As in the previous two Chapters, the large player has some private information about his type, i.e. the small players are uncertain about the type of large player they are facing. This uncertainty may be very small in the sense that the large player is of one particular type with a probability close to one.

The goal of Chapter 4 is to study the conditions under which the usual reputational arguments can be extended to a dynamic game, and the conditions under which reputational arguments fail.

Since games with a large number (continuum) of small players will be considered, it will be assumed that the individual play of the small players is not observed. In a purely repeated game this assumption would imply that each small player behaves like a short-lived player, since his actions will affect neither his future payoffs nor the public history of the game. In a dynamic game the presence of state variables creates an intertemporal link and introduces a new strategic dimension to the problem. Even though a small player cannot influence

his opponent's future play, he can change the value of his individual state variable, thereby affecting his own future payoff opportunities. Therefore a small player's behavior will depend on the (expected) future actions of the large player

As is clear, the presence of state variables makes it more difficult for the large player to establish a reputation: small players have to become convinced that the large player will follow a particular strategy not only in the current period but also in the future. The more the small players' behavior is affected by play in the distant future the harder it will be for the large player to gain from establishing a reputation.

The first result of Chapter 4 applies to the case where the small players have a fixed discount factor while the large player is arbitrarily patient. If there is a commitment type that plays the strategy to which the large player would want to commit then in any Nash equilibrium the large player is guaranteed at least the optimal commitment payoff.

The case in which both the large and the small players are arbitrarily patient is particularly relevant for policy games, in which, for example, the payoff function of the government is equal to the payoff function of the median voter. Then, if players are very patient, the small players' action may be affected by very distant future outcomes.

In this case it is shown that the large player will only be able to exploit his reputation if the transition function is reversible, in the sense that players can move from one state to another only if they can also return.

This condition is satisfied in capital accumulation games, but is not satisfied, for example, in the standard durable goods monopoly. Once a customer has purchased the durable good, he has reached an irreversible state. An example shows how in the durable goods monopoly reputational arguments fail to guarantee the



large player his optimal commitment payoff<sup>®</sup>.

## Chapter 2

# Reputation with Deterministic Stage Games

## 2.1 Introduction

Kreps and Wilson (1982) and Milgrom and Roberts (1982) have provided an explanation of the chain store paradox, assuming that there is a "chance" that the incumbent is a commitment type: this fact can be exploited by a "sane" incumbent that can therefore build up a reputation for toughness.

Fudenberg and Levine (1989) build up on these results to provide a lower bound on the Nash equilibrium payoffs of the long run player. However their main result (Theorem 1) applies only to games in which the commitment strategy of the long run player is revealed regardless of the strategies the short run players choose: this is true in simultaneous move games, in sequential move games in which the long run player moves first, and in some sequential move games in which the short run player moves first: an example of this last class is the chain store game, in which the strategy the short run players choose before the reputation is established is exactly the one that reveals the strategy by which the long run player builds up his reputation.

Fudenberg and Levine (1989) also provide a generalization of Theorem 1 (Theorem 2) in which the Stackelberg payoff is redefined to keep into account the fact that the outcome of the stage game may not reveal the long run player's strategy; the new bound is computed making use of the fact that the observed outcome of the stage game in general restricts a subset of the strategy space to which the strategy chosen by the long run player must belong.

Unfortunately, in some games this result does not provide a higher lower bound than the minimum payoff for the long run player. For example, consider the quality game with extensive form as in Fig. 1. The short run player moves first and decides whether to buy a product or not; if he decides not to buy, the game ends and both players get 0; if he decides to buy, the long run player decides whether to

produce a low quality product, thus making a larger profit and causing the short run player a loss, or a high quality product, in which case the profit is smaller but the short run player's payoff is positive.

When this stage game is repeated an infinite number of times, the lower bound provided in Theorem 2 in Fudenberg and Levine (1989) is just 0: if the prior probability that the long run player is committed to high quality is less than  $\frac{1}{5}$ , there is an equilibrium in which no short run player ever buys, no information is revealed, and the long run player payoff is 0 (cfr. Fudenberg and Levine (1989), pp. 772-773).

The purpose of this chapter is to provide a different generalization of Theorem 1 in Fudenberg and Levine (1989), one that uses perturbations of the original game with the property that every information set is reached with positive probability in the stage game.

The idea is simply to assume that not only the type of the long run player is uncertain, but also the types of the short run players are, and that for each strategy there exists at least one type of short run player, that is selected with strictly positive probability, that has that strategy as a strictly dominant strategy.

As is shown in Example 1, this might require a substantial increase in the number of periods necessary to build up a reputation with respect to the sequential move game, but may nevertheless provide a significantly higher lower bound on the Nash equilibrium payoffs of the long run player.

As in Fudenberg and Levine (1989) we provide a lower bound on the Nash equilibrium payoffs to the long run player by computing a lower bound on the payoff to the so called Stackelberg strategy to be defined. This strategy need not be the optimal one for the long run player, but since it is always feasible, the optimal strategy has to yield at least as high a payoff.

Our result holds for any stage game (simultaneous move<sup>1</sup> or sequential move) in which the realized actions of the long run player are observed, and the probability distributions over types of long run and short run players have full support.

Fudenberg and Levine (1991) show that there is a lower bound on the Nash equilibrium payoffs to the long run player also when the public outcome of the stage game is a random variable that provides only stochastic information about the strategy the long run player chose.

Our model is a special case of theirs in that in sequential move stage games in which the short run player moves first the public outcome only reveals the action of the long run player and not his strategy. Restricting to this special class of games, however, lets us explicitly compute the lower bound, and thus narrow down the set of equilibrium payoffs to the long run player.

The range of applications of our result is very wide. In the following we just want to mention a few applications of the quality game.

International loan contracts are many times not enforceable or very costly to enforce. They are therefore well described by the quality game: an international lender decides whether to give credit to a foreign agent and the latter then decides whether to repay the loan or renege on his debt. Even though repaying is sub-optimal in the stage game, it is a way of establishing a reputation for repayment that in turn guarantees prolonged access to international loan markets.

Illegal contracts are also not enforceable: nevertheless cocaine dealers or illegal lottery organizers can decide to sell high quality cocaine or to pay the prizes in order to establish a reputation for "honesty".

Importers usually get short term credit from their suppliers. In some less

---

<sup>1</sup>For simultaneous move stage games our result coincides with that of Fudenberg and Levine (1989).

developed countries, however, trading houses do not enforce these contracts, so that the importer has an incentive to renege on it and, by backward induction, foreign traders refuse him credit. Also in this case, however, the importer can guarantee himself a higher discounted payoff in the repeated game by establishing a reputation for repayment.

In Section 2.2 we describe the game and introduce the notation. The result is derived in Section 2.3. Section 2.4 provides two examples of the quality game that give substance to the results of the previous section. Section 2.5 provides a discussion of the result.

## 2.2 The Model

A long run player (player 1) plays a fixed stage game against an infinite sequence of short run players (player 2). The long run player chooses a strategy  $s_1$  from a finite nonempty set  $S_1$  and the short run player chooses an action  $s_2$  from a finite nonempty set  $S_2$ . The corresponding mixed strategy spaces are denoted by  $\mathcal{S}_1$  and  $\mathcal{S}_2$ .

The public outcome of the stage game is given by a mapping  $y : S_1 \times S_2 \rightarrow Y$ , and is to be interpreted as the revealed actions of the long run and the short run player. When the stage game is simultaneous move or sequential move with the long run player moving first, the action reveals the long run player's strategy. But when the stage game is sequential move and the short run player moves first the long run player's revealed action does not reveal what he would have done, had the short run player chosen a different strategy.

The unperturbed stage game is described by the payoffs to the long run and the short run players, a mapping  $u : Y \rightarrow \mathbb{R}^2$ ; with an abuse of notation we let

$u(y(\frac{3}{4})) = (u_1(y(\frac{3}{4}_1; \frac{3}{4}_2)); u_2(y(\frac{3}{4}_1; \frac{3}{4}_2)))$  denote the expected payoffs corresponding to the mixed strategy profile  $\frac{3}{4}$ . In the unperturbed repeated game the long run player maximizes the normalized discounted value of expected payoffs

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t u_1^t$$

Each period's short run player maximizes that period's payoff,  $u_2^t$ .

Both long run and short run players can condition their play on the past history of the game. Let  $H_t = Y^t$  denote the set of possible histories of the game; then mixed strategies are mappings  $\frac{3}{4}_1^t : H_{t-1} \rightarrow S_1$ , and  $\frac{3}{4}_2^t : H_{t-1} \rightarrow S_2$ .

Let  $B : S_1 \rightarrow S_2$  be the correspondence that maps mixed strategies by the long run player in the stage game to the best responses of the short run player. Then we define the Stackelberg payoff  $u_1^s$  as:

$$u_1^s = \max_{s_1 \in S_1} \min_{\frac{3}{4}_2 \in B(s_1)} u_1(s_1; \frac{3}{4}_2)$$

the Stackelberg leader strategy as the  $s_1^s$  that solves

$$\max_{s_1 \in S_1} \min_{\frac{3}{4}_2 \in B(s_1)} u_1(s_1; \frac{3}{4}_2)$$

and the Stackelberg follower strategy as the  $s_2^s$  that solves

$$\min_{\frac{3}{4}_2 \in B(s_1^s)} u_1(s_1^s; \frac{3}{4}_2)$$

Intuitively,  $s_2^s$  is the strategy of the short run player that the long run player wants to induce.

In the perturbed game the payoffs of the long run player, as well as those of the short run player, are made dependent on their types which are assumed to be private knowledge. For simplicity we assume that there are a countable number

of types both of long and short run players:

$$-_1 = f!_0^1; !_1^1; ::; g; -_2 = f!_0^2; !_1^2; ::; g$$

The payoffs are therefore a mapping  $u_i : S_1 \times S_2 \times -_i \rightarrow \mathbb{R}$ , and the mixed strategies are mappings  $\sigma_i^t : H_{t-1} \times -_i \rightarrow S_i$ . We let  $!_1^0$  and  $!_2^0$  be the rational players; in other words we assume that their payoffs are as in the unperturbed game:  $u_i(\sigma_1^t; \sigma_2^t; !_i^0) = u_i(\sigma_1^t; \sigma_2^t)$ ,  $i = 1; 2$ .

The priors on the types are probability distributions  $\pi_1 : -_1 \rightarrow [0; 1]$  and  $\pi_2 : -_2 \rightarrow [0; 1]$  that are assumed to be common knowledge.

For all  $s_1 \in S_1$  let  $!_i(s_1)$  be a type of player  $i = 1; 2$  that has strategy  $s_1$  as a dominant strategy in the repeated game. In the following we will make the following assumptions about the types of long run and short run players:

**Assumption 1** There exists a  $\pi_1 > 0$  such that  $\pi_1(!_1(s_1)) > \pi_1$  for all  $s_1 \in S_1$ .

**Assumption 2** There exists a  $\pi_2 > 0$  such that  $\pi_2(!_2(s_2)) > \pi_2$  for all  $s_2 \in S_2$ .

In the following we will call a Stackelberg leader type a long run player that has  $s_1^a$  as a dominant strategy in the repeated game, and we denote by  $!_1^a$  the event that the long run player is such type, and by  $!_1^{\bar{a}}$  the event that the long run player is not such type. We will denote by  $!_2^j$  the event that the short run player is the type that has  $s_2^j$  as a dominant strategy.<sup>2</sup>

Let  $H^a$  be the set of histories such that the play of the long run player is consistent with the description of the Stackelberg type for all  $t$ , and let  $h^a$  denote the event  $h \in H^a$ . Finally, let  $\mathcal{H}_t^a$  be the random variable  $\Pr(s_1^t = s_1^a | h_{t-1})$  and let  $n(\mathcal{H}_t^a \cdot \mathcal{H})$  be the random variable denoting the number (possibly infinite) of the random variables  $\mathcal{H}_t^a$  for which  $\mathcal{H}_t^a \cdot \mathcal{H}$ .

<sup>2</sup>Less strong assumptions about the types of short run players can be made; see Section 2.4.



## 2.3 The Result

First we show that  $\Pr(\sigma_1^a | h_t)$  is nondecreasing in  $t$  when  $h_t$  is the truncation of a history  $h \in H^a$ .

**Lemma 1** For any infinite history  $h \in H^a$  such that the truncated histories  $h_t$  have positive probability,  $\Pr(\sigma_1^a | h_t)$  is nondecreasing in  $t$ .

**Proof:** We want to show that

$$\begin{aligned} \Pr(\sigma_1^a | h_t) &= \Pr(\sigma_1^a | y(s_1^a; s_2); h_{t-1}) \\ &= \frac{\Pr(\sigma_1^a | h_{t-1}) \Pr(y(s_1^a; s_2) | \sigma_1^a)}{\Pr(\sigma_1^a | h_{t-1}) \Pr(y(s_1^a; s_2) | \sigma_1^a) + (1 - \Pr(\sigma_1^a | h_{t-1})) \Pr(y(s_1^a; s_2) | \sigma_1^a)} \\ &\geq \Pr(\sigma_1^a | h_{t-1}) \end{aligned} \quad (2.1)$$

Inequality (2.1) is equivalent to

$$\frac{\Pr(y(s_1^a; s_2) | \sigma_1^a)}{\Pr(\sigma_1^a | h_{t-1}) \Pr(y(s_1^a; s_2) | \sigma_1^a) + (1 - \Pr(\sigma_1^a | h_{t-1})) \Pr(y(s_1^a; s_2) | \sigma_1^a)} \geq 1 \quad (2.2)$$

which is in turn equivalent to

$$\Pr(y(s_1^a; s_2) | \sigma_1^a) \cdot \Pr(y(s_1^a; s_2) | \sigma_1^a) \quad (2.3)$$

which is trivially satisfied since  $\Pr(y(s_1^a; s_2) | \sigma_1^a) = 1$ .

The following Lemma computes an upper bound on the probability that the probability that the long run player plays  $s_1^a$  is less than a fixed probability  $\epsilon$  when the stage game is repeated a number of times, and is to be used to compute the lower bound on the Nash equilibrium payoffs to the long run player. In the following we will assume that the cardinality of  $S_1$  is  $N + 1$  and will denote by  $[x]$  the operator integral part ( $[x]$  is the greatest integer less than or equal to  $x$ ).

**Lemma 2** Let  $0 < \epsilon < 1$ . Suppose that  $(\epsilon_1^a; \epsilon_2^a)$  are such that  $\Pr(h^a | \sigma_1^a) = 1$ . Let  $K_1 = \lceil \log \frac{1}{1 - (1 - \epsilon)^{N+1}} \rceil + 1$ , and  $\delta > 0$  let

$$K_2(\delta) = \lceil \log \frac{1}{(1 - \delta)^{1 - K_1}} \rceil + 1:$$

Then,  $\delta^2 > 0$ ,  $\Pr(n(\frac{1}{4}_t^\alpha \cdot \frac{1}{4}) > K_1 \downarrow K_2(2)jh^\alpha) \cdot 2$ .

Remark 1. The purpose of Lemma 2 is to provide an upper bound on the probability that the probability that the long run player plays  $s_1^\alpha$  is less than a given  $\frac{1}{4} \in [0; 1)$  after the stage game has been played a given number of times, and to make this upper bound dependent on  $\alpha_1, \alpha_2$  (the lower bounds on  $\alpha_1$  and  $\alpha_2$ ) and  $\frac{1}{4}$  only, and otherwise independent of  $(\alpha_1; \alpha_1)$  and  $(\alpha_2; \alpha_2)$ . To do this we argue that whenever  $\frac{1}{4}_t^\alpha = \Pr(s_1^t = s_1^\alpha | h_{t-1})$  is low, if  $s_1^\alpha$  is played, there is a strictly positive probability that  $\Pr(s_1^t = s_1^\alpha | h_t)$  increases by a nontrivial amount. Since  $\Pr(s_1^t = s_1^\alpha | h_t)$  has to be less than or equal to 1, this cannot happen too often, so that the probability that  $\frac{1}{4}_t^\alpha$  is low in many periods has to be low.

Proof: By Bayes's law we have

$$\begin{aligned} \Pr(s_1^t = s_1^\alpha | h_t) &= \Pr(s_1^t = s_1^\alpha | y(s_1^\alpha; s_2); h_{t-1}) \\ &= \frac{\Pr(s_1^t = s_1^\alpha | h_{t-1}) \Pr(y(s_1^\alpha; s_2) | s_1^\alpha)}{\Pr(s_1^t = s_1^\alpha | h_{t-1}) \Pr(y(s_1^\alpha; s_2) | s_1^\alpha) + (1 - \Pr(s_1^t = s_1^\alpha | h_{t-1})) \Pr(y(s_1^\alpha; s_2) | s_1^\alpha)} \end{aligned}$$

Substituting  $\Pr(y(s_1^\alpha; s_2) | s_1^\alpha) = 1$  in the numerator of the previous fraction and recognizing that the denominator is equal to  $\Pr(y(s_1^\alpha; s_2) | h_{t-1}; s_2^\alpha)$  we have

$$\Pr(s_1^t = s_1^\alpha | h_t) = \frac{\Pr(s_1^t = s_1^\alpha | h_{t-1})}{\Pr(y(s_1^\alpha; s_2) | h_{t-1})} \quad (2.4)$$

where  $\Pr(s_1^t = s_1^\alpha | h_1) = \alpha_1(1 - \alpha_1) \cdot \alpha_1$ .

$\Pr(y(s_1^\alpha; s_2) | h_{t-1})$  is the probability that  $y(s_1^\alpha; s_2)$  is observed, which is equal to the probability that  $s_1^\alpha$  is being played plus the probability that other strategies observationally equivalent to  $s_1^\alpha$  for  $s_2$  are being played. Define  $S_1^\alpha(s_2)$  as the set of strategies of the long run player different from  $s_1^\alpha$  that are observationally equivalent to  $s_1^\alpha$  when the short run player plays  $s_2$ , i.e.  $S_1^\alpha(s_2) = \{s_1 \in S_1^\alpha : y(s_1; s_2) = y(s_1^\alpha; s_2)\}$ . With this notation (2.4) can be rewritten as

$$\Pr(s_1^t = s_1^\alpha | h_t) = \frac{\Pr(s_1^t = s_1^\alpha | h_{t-1})}{\Pr(s_1^t = s_1^\alpha | h_{t-1}) + \sum_{s_1 \in S_1^\alpha(s_2)} \Pr(s_1^t = s_1 | h_{t-1})} \quad (2.5)$$

Saying that  $\Pr(s_1^t = s_1^a | h_{t-1}) > \frac{1}{N}$  is equivalent to saying that

$$\sum_{s_1 \in S_1^a(s_2)} \Pr(s_1^t = s_1 | h_{t-1}) < 1 - \frac{1}{N} \quad (2.6)$$

Given that the cardinality of  $S_1$  is  $N + 1$ , a sufficient condition for (2.6) to be satisfied is  $\Pr(s_1^t = s_1 | h_{t-1}) < \frac{1}{N}$  for all  $s_1 \in S_1^a$ , where  $\frac{1}{N} = (1 - \frac{1}{N}) = N$ .

Now suppose  $\exists s_1 \in S_1^a$  such that  $\Pr(s_1^t = s_1 | h_{t-1}) > \frac{1}{N}$ . Since  $s_1 \in S_1^a$ , there exists an  $s_2$  such that  $s_1$  is not observationally equivalent to  $s_1^a$ ,  $s_1 \notin S_1^a(s_2)$  (in other words,  $y(s_1; s_2) \notin y(s_1^a; s_2)$ ). If the short run player plays such an  $s_2$  (an event that, by Assumption 2, happens with probability at least  $\frac{1}{2}$ ), then by (2.5) we have that

$$\Pr(s_1^t | h_t) \geq \frac{\Pr(s_1^t | h_{t-1})}{1 - \frac{1}{N}}$$

since the denominator of (2.5) is less than or equal to  $1 - \frac{1}{N}$ . In the following we will call such an  $s_2$  an information revealing strategy.

If the stage game is repeated  $K$  times and every time an information revealing  $s_2$  is selected, then

$$\Pr(s_1^t | h_t) \geq \frac{1}{(1 - \frac{1}{N})^K}$$

However, since

$$\Pr(s_1^t | h_t) < 1 \quad (2.7)$$

if

$$\frac{1}{(1 - \frac{1}{N})^K} > 1 \quad (2.8)$$

inequality (2.7) is violated and a contradiction to the hypothesis that  $\Pr(s_1^t = s_1 | h_{t-1}) > \frac{1}{N}$ , any  $s_1 \in S_1^a$ , is obtained. Taking the log of (2.8) and substituting  $\frac{1}{N} = 1 - (1 - \frac{1}{N}) = N$  the condition becomes

$$K > \frac{\log \frac{1}{1 - \frac{1}{N}}}{\log (1 - \frac{1}{N})}$$

Defining  $K_1 = \lceil \log \frac{1}{1 - \frac{1}{N}} / \log (1 - \frac{1}{N}) \rceil + 1$  provides the first part of the result.

Finally, we want to find an upper bound on the number of times the stage game is played and the probability that  $\frac{1}{4} < \frac{1}{2}$  is less than a given  $\epsilon > 0$ , when the long run player plays  $s_1^*$ , i.e. we want to find the smallest integer  $K_2(\epsilon)$  such that

$$\Pr(n(\frac{1}{4} - \frac{1}{2}) > K_1 \epsilon K_2(\epsilon)) < \epsilon \quad (2.9)$$

The probability on the left hand side of inequality (2.9) is less than or equal to the probability that information revealing  $s_2$  are played less than  $K_1$  times when the stage game is repeated  $K_1 \epsilon K_2(\epsilon)$  times.

Suppose that the stage game is played  $K_2(\epsilon)$  times; then the probability that no information revealing  $s_2$  is played is

$$\hat{p} = (1 - \frac{1}{2})^{K_2(\epsilon)} \quad (2.10)$$

and  $1 - \hat{p}$  is the probability that at least one information revealing  $s_2$  is played.

If the stage game is played  $K_1 \epsilon K_2(\epsilon)$  times, i.e. if the experiment of playing the stage game  $K_2(\epsilon)$  times is repeated  $K_1$  times, the probability that at least  $K_1$  information revealing  $s_2$  are played is greater than  $(1 - \hat{p})^{K_1}$ . Therefore a sufficient condition for the probability that less than  $K_1$  information revealing  $s_2$  are played when the stage game is repeated  $K_1 \epsilon K_2(\epsilon)$  times to be less than  $\epsilon$  is

$$(1 - \hat{p})^{K_1} > 1 - \epsilon$$

whence

$$\hat{p} < 1 - (1 - \epsilon)^{1/K_1} \quad (2.11)$$

Substituting (2.10) in (2.11) and rearranging provides

$$K_2(\epsilon) > \frac{\log(1 - (1 - \epsilon)^{1/K_1})}{\log(1 - \frac{1}{2})}$$

Defining  $K_2(\epsilon) = \lceil \log(1 - (1 - \epsilon)^{1/K_1}) / \log(1 - \frac{1}{2}) \rceil + 1$  concludes the proof.  $\square$

Remark 2: The lower bound on  $\alpha_2, \beta_2$ , is to be interpreted as a lower bound on the probability that information revealing  $s_2$  are played. In simultaneous move stage games and in simultaneous move stage games in which the long run player moves first all  $s_2$  are information revealing because the strategy of the long run player is observed. For this class of game our result coincides with the one of Fudenberg and Levine (1989).

We are now ready to state the main result. Let  $V_1(\alpha; \alpha_1; \alpha_2; \beta_1^0)$  be the least Nash equilibrium payoff to a long run player of type  $\beta_1^0$ , with payoffs as in the unperturbed game, when the discount factor is  $\alpha$ . Then

**Theorem 1** Let Assumptions 1 and 2 be satisfied, and let  $\alpha_1, \alpha_2$  be the probability that the short run player is the rational type. Then for all  $\alpha > 0$ , there exists a  $K(\alpha_1; \alpha_2; \alpha) = K^\alpha$  otherwise independent of  $(\alpha_1; \beta_1)$  and  $(\alpha_2; \beta_2)$  such that

$$V_1(\alpha; \alpha_1; \alpha_2; \beta_1^0) \geq (1 - \alpha_2)(1 - \alpha_2)\alpha^{K^\alpha} u_1^\alpha + (1 - \alpha_1)(1 - \alpha_2)(1 - \alpha_2)\alpha^{K^\alpha} \min u_1 \quad (2.12)$$

**Proof:** Suppose the long run player always plays the Stackelberg strategy. Since the best response correspondence<sup>3</sup>  $B(\beta_1^\alpha)$  is upper hemi-continuous, each element of  $B(\beta_1^\alpha)$  is near to an element of  $B(\beta_1^\alpha)$  when  $\beta_1^\alpha$  is sufficiently near to one. Since  $s_2$  is finite, if  $\beta_2$  is near to an element of  $B(\beta_1^\alpha)$ , then it must place probability close to one on  $s_2^\alpha$ . Since the rational short run player has to be indifferent between all strategies that he is willing to assign positive probability, there is a probability  $\eta < 1$  such that  $B(\beta_1^\alpha) \cap B(\beta_1^\alpha)$  whenever  $\beta_1^\alpha > \eta$ .

Set  $K^\alpha = K^\alpha(\alpha) = K^\alpha(\alpha; \alpha_1; \alpha_2; \eta) = K_1 \cup K_2(\alpha)$ . If the long run player always plays  $s_1^\alpha$ , then from Lemma 2 it follows that the probability that there are more

---

<sup>3</sup>Recall that  $B(\cdot)$  is the best response correspondence of the rational short run player.

than  $K^{\pi}(2)$  occasions where the rational short run player plays outside of  $B(s_1^{\pi})$  (corresponding to the events  $\frac{1}{4}^{\pi} \leq \frac{1}{4}$ ) is less than  $\frac{1}{2}$ . In the worst case these events occur at the beginning of the game where the payoffs are discounted the least. Recalling that only a fraction  $1 - \frac{1}{2}$  of short run players is rational provides the right hand side of (2.12). Since the Stackelberg strategy is always feasible for the long run player, the right hand side is a lower bound on any Nash equilibrium payoff.  $\square$

Remark 3: As said in Remark 2, in the case in which the stage game is simultaneous move or sequential move with the long run player moving first,  $\beta_2 = 1$  and the lower bound in Theorem 1 coincides with the lower bound in Theorem 1 in Fudenberg and Levine (1989). The same is true for sequential move stage games in which the short run player moves first and in which the short run players choose an information revealing  $s_2$  when  $\frac{1}{4}^{\pi} \leq \frac{1}{4}$ , such as the chain store game.

## 2.4 The Quality Game

In the following we want to discuss an important application of our results, the quality game. The analysis will turn out to be simpler than in the previous section given the simple structure of the game. In particular  $S_1$  has only 2 elements, therefore  $N = 1$  and  $1 - (1 - \frac{1}{4}) = N = \frac{1}{4}$ .

Example 1. Consider the version of the quality game whose extensive form is described in Fig. 2. When  $a = 1$ ,  $b = \frac{1}{2}$ , and  $c = 0$ , as argued in the introduction, provided that  $\beta_1(\frac{1}{4})$  is not too high, the lower bound for the long run player Nash equilibrium payoffs given by Theorem 2 in Fudenberg and Levine (1989) is just  $\min u_1 = 0$ .

Now suppose that there are two types of short run player, the rational player,

$\theta_2^0$ , with payoffs as given above, and a second one,  $\theta_2^a$ , with payoffs such that he always buys. Suppose that these payoffs are  $a = 1$ ,  $b = 1/2$ , and  $c = 0$ . The rational player  $\theta_2^0$  on the other hand buys only if  $\frac{1}{4} \leq \theta_1 \leq 1/2$ .

In this example we only have one long run player commitment type ( $\theta_1^a$ ) and one short run player commitment type ( $\theta_2^a$ ). In the following we will therefore replace  $\theta_1$  and  $\theta_2$  with  $\theta_1^a = \theta_1(\theta_1^a)$  and  $\theta_2^a = \theta_2(\theta_2^a)$ . Finally, notice that since type  $\theta_2^a$  always buys, we can disregard the term  $(1 - \theta_2)$  in (2.12), since buying is a best response to producing high quality.

Let  $\theta_1^a = 1/4$ . Then  $K_1 = \lceil \log \theta_1^a \rceil + 1 = 4$ :

Suppose  $\theta_2^a = 3/4$ , and let  $\epsilon = 1/99$ . Then we have:

$$V_1(\epsilon; \theta_1^a, \theta_2^a; \theta_1^0) \geq \max_{\theta_2} (1 - \theta_2) \epsilon^{99^{K_1 + K_2(\epsilon)}} = 1/60 > 0$$

which is obtained maximizing with respect to  $\theta_2$  the right hand side of inequality (2.12) in Theorem 1. The  $\theta_2$  that maximizes that expression turns out to be  $1/11$ , which implies that  $K_2(\epsilon) = 10$ .

As claimed above, the introduction of uncertainty on the side of the short run players improves substantially the lower bound on the long run player Nash equilibrium payoffs. <sup>2</sup>

The purpose of the next example is to assess the sharpness of the lower bound on Nash equilibrium payoffs  $V_1$  that is computed using only  $\theta_1^a$  and  $\theta_2^a$ . We will show that, while the use of additional information relative to the distribution of the short run player type does provide a better bound, the induced improvement is far from dramatic.

Example 2. Suppose we introduce another type of short run player,  $\theta_2^1$ , with payoffs  $a = 1$ ,  $b = 1/3$ , and  $c = 0$ , and whose prior is  $\theta_1(\theta_2^1) = 1/2$ ; short run players of type  $\theta_2^1$  buy if  $\frac{1}{4} \leq \theta_1 \leq 1/4$ , thus increasing the probability that the long

run player's action be revealed. In this case, as suggested above,  $V_1$  turns out to be larger.

In such a case  $K^\pi = K^0 + K^{00}$  where  $K^0 = K^\pi(2; 1_1^\pi; 1_2^\pi; \frac{1}{4}^0) = K_1(1_1^\pi; \frac{1}{4}^0) \& K_2(2; 1_2^\pi)$  and  $K^{00} = K^\pi(2; Pr(!_1^{\pi} | h_{K^0}); 1_2^\pi + 1_2^1; \frac{1}{4}) = K_1(Pr(!_1^{\pi} | h_{K^0}); \frac{1}{4}) \& K_2(2; 1_2^\pi + 1_2^1)$ . To see this assume that at least one type  $!_2^\pi$  short run player is selected when the stage game is repeated  $K^0$  times. Then we have that  $Pr(!_1^{\pi} | h_{K^0}) \leq 1_1^\pi = \frac{1}{4} = \frac{1}{4}$ , since  $\frac{1}{4} \cdot \frac{1}{4} = 1/4$ . In this case  $K^{00} = K^\pi(2; 1_1^\pi = \frac{1}{4}; 1_2^\pi + 1_2^1; \frac{1}{4})$ . Since in computing  $K^{00}$  we have assumed that an event had happened whose probability is  $1 - (1 - 1_2^\pi)^{K^0}$ , the probability that  $\frac{1}{4} \leq \frac{1}{4} = \frac{1}{4}$  is equal to  $(1 - (1 - 1_2^\pi)^{K^0})$  and therefore

$$V_1(1; 1_1^\pi; 1_2^\pi; 1_2^1; !_1^0) \geq \max(1 - (1 - (1 - 1_2^\pi)^{K^0}) \cdot \frac{1}{4}, 1 - (1 - 1_2^\pi)^{K^0}) = \frac{1}{4}$$

In the previous examples we have made the assumption that a type of short run player exists with strictly positive probability that had  $s_2^\pi$ , an information revealing strategy as a strictly dominant strategy, which implies that that type of short run player will play  $s_2^\pi$  regardless of the long run player he believes to face.

Another assumption that is perfectly consistent with the structure of the model is the following:

**Assumption 3** A type of short run player exists with strictly positive probability that plays  $s_2^\pi$  provided that the probability that the long run player is the Stackelberg leader type is greater than or equal to  $1_1^\pi$ , the prior probability that he is of that type.

In the quality game studied above Assumption 3 means that

$$1_1^\pi a + (1 - 1_1^\pi) b \geq c; \tag{2.13}$$



whereas Assumption 2 was equivalent to

$$a \geq c; \quad b \geq c: \quad (2.14)$$

As is clear Assumption 2 is stronger than Assumption 3 in that (2.14) implies (2.13) but not vice versa: (2.13) might hold also when  $a \geq c$  but  $b < c$ . In the context of the quality game this means that the short run player commitment types do not prefer purchase to no purchase independently of the quality; it just means that given their preferences they are more willing to take the risk of buying than the rational short run player.

Consider again the game of Fig. 2, and suppose that type  $\theta_2^H$  has payoffs  $a = 1, b = \frac{1}{9}, c = 0$ . If we assume, as in Example 1, that  $\theta_1^H = \frac{1}{3}$ , we then have  $\theta_1^H a + (1 - \theta_1^H)b = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{1}{9} = c = 0$ , Assumption 3 is satisfied and our results follow.

A major difference between Assumptions 2 and 3 however exists. Suppose that in the game of Fig. 2 the payoff to the long run player when he produces low quality is 4 rather than  $\frac{3}{2}$ . If we make Assumption 2, and  $\theta_2^H = \frac{1}{3}$  it turns out that the Stackelberg leader strategy is to produce low quality, since in this case his expected payoff is  $\frac{1}{3} \cdot 4 = \frac{4}{3}$ . In other words if enough short run players exist that always buy and the difference between the payoff to the long run player when he produces low and high quality is large enough, it might be better for him to exploit the short run commitment types rather than building a reputation for honesty.

If we make Assumption 3, on the other hand, and we assume that  $b < c$ , the same result does not hold: after the first time the long run player produces low quality he is revealed to be the rational type ( $\Pr(\theta_1^H) = 0$ ) and no other short run player is guaranteed to ever buy in the future, not even the commitment types.

While we think that the two assumptions we have been discussing can be

appropriate for different games, we also believe that Assumption 2 is interesting in that it highlights that reputation does not always work.

In the examples we have presented so far we have chosen a discount factor that is not too large: if the reference period is one month,  $\delta = .99$  translates to a yearly interest rate of 12.8%. We have chosen to do so to stress the fact that the result doesn't hold only for very patient long run players. However in many economic examples the relevant reference period can be shorter: if the relevant period is for example one week, a weekly discount factor  $\delta = .999$  would translate to a yearly interest rate of 5.3%, and in this case  $V_1$  in Example 1 would be larger than .93.

## 2.5 Discussion

Whenever the strategy of a player is not perfectly observed, that player might be prevented from establishing a reputation for an appropriate behavior. This chapter showed that by introducing types of short run players that are such that all informational nodes of the stage games are reached with strictly positive probability in any equilibrium, a long run player can actually establish a reputation for establishing any particular stage game pure strategy.

In a more general sense the point of this chapter was to show that the negative result pointed at at the beginning of the section is not robust with respect to perturbations of the information structure of the game.

An alternative more general framework is one in which it is assumed that the action of the short run player is not perfectly observed<sup>4</sup> and that the support of the distribution over outcomes is invariant with respect to the action chosen by the short run player.

---

<sup>4</sup>Suppose there is some noise or that the short run player trembles.

This approach lends itself to more general applications and will be pursued further in Chapter 3 to study reputational effects in infinitely repeated games between long run players.

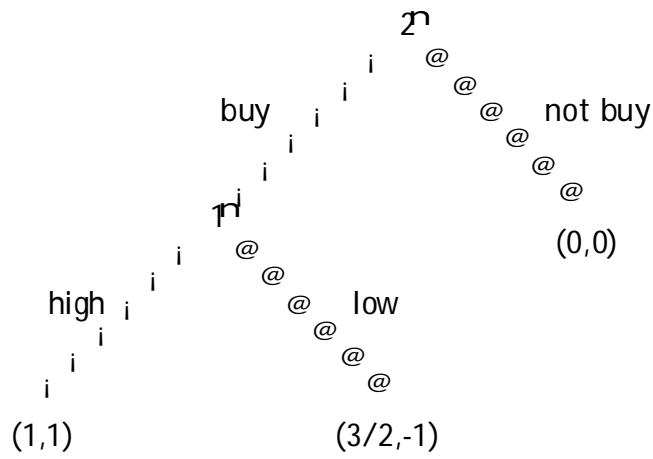


Figure 2.1: Quality Game

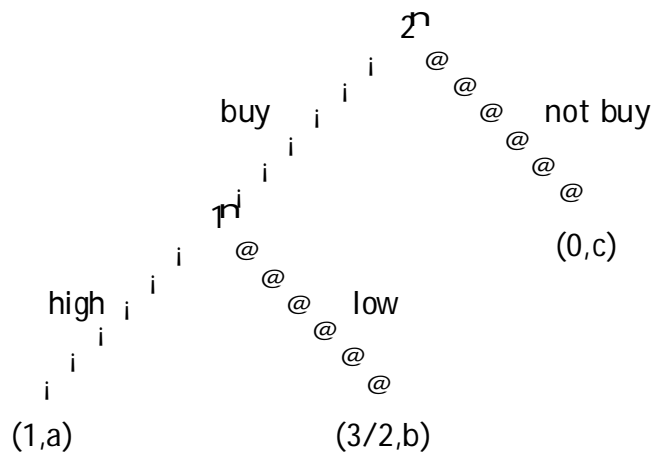


Figure 2.2: Quality Game with unspecified payoffs for the short run player

## Chapter 3

# Reputation in Repeated Games with Two Long Run Players

## 3.1 Introduction

Since the work of Kreps and Wilson (1982) and Milgrom and Roberts (1982), the existence of even a small amount of uncertainty relative to the payoff function of a player has been used to provide predictions of the outcome of repeated strategic interaction between two or more players.

The seminal work of Fudenberg and Levine (1989) considers the case of an infinitely repeated game between a long run player (a player with a positive discount factor who maximizes the present value of his infinitely repeated game payoff) and an infinite sequence of short run players who observe all previous play of the game (or, equivalently of a single player with zero discount factor who, in each period, maximizes his current payoff). In this setting, Fudenberg and Levine (1989) find that if there is strictly positive probability that the long run player is a commitment type who always plays a particular action regardless of the previous play, and if the long run player imitates this type, then there is a finite number of periods in which the opponent may not play a best response to such action.

As a consequence, the long run player can obtain the so called Stackelberg payoff in all but a finite number of repetitions of the stage game, which in turn implies that, if he is sufficiently patient, in any equilibrium his average payoff cannot be lower than a payoff which is arbitrarily close to the Stackelberg payoff.

Chapter 2 extends the results of Fudenberg and Levine (1989) to the case in which a long run and a short run player play repeatedly a sequential game. An interesting example of this class of game is the so called "Quality Game" in which the short run player has to decide whether to buy a product or not, and the long run player has to decide whether to produce a high quality or a low quality product, but the quality of the product is observed and made public only if purchase occurs. In this case the problem is that, if short run players believe

that the product will be a low quality one, then there is an equilibrium in which they will not to and the long run player has no way of establishing a reputation for producing high quality.

In Chapter 2 a perturbation on the types of the short run players was introduced that made sure that all informational nodes of the stage game be visited with strictly positive probability, and it was shown that the probability that the number of periods in which the short run player may believe that the long run player is unlikely to play the Stackelberg strategy exceeds a given finite number becomes arbitrarily small as this finite number increases.

Schmidt (1993) considers the case of a repeated game between two players who both have nonzero discount factors, and mostly deals with the case in which there is uncertainty only relative to player 1's payoff function, so that only player 1 can establish a reputation. His main contribution is that Fudenberg and Levine's (1989) result can be extended to the case of two long run players provided that the stage game belongs to the class of games with conflicting interest with respect to player 1. This means that the strategy of player 1 that maximizes player 1's payoff subject to the constraint that player 2 play a short run best response is also the strategy that minmaxes player 2.

The main problem in the case of two patient players whose actions are observed is that guaranteeing that player 1 is likely to play a certain action today does not imply that player 2 will play a best response to it, since he also cares about future payoffs. This implies that he might believe that the probability that player 1 will keep playing a given action if he currently plays a best response to this action is arbitrarily low, and in particular he might believe that if he does play a best response today, he will get a low payoff in the continuation game. In such a case, since player 2 would never play a best response to the action player 1 is trying

to establish a reputation for, the probability that the continuation play will be unfavorable to him, if he plays a best response, may well stay arbitrarily high since player 1's behavior in this contingency is never observed.

Confining attention to games with conflicting interest with respect to player 1, however, implies that if player 2 does not play a best response to player 1's Stackelberg strategy he gets less than his minmax payoff, while playing a best response today he would get his minmax payoff today and no less than his minmax in any continuation game.

As is clear, the main problem with the case of two long run players and perfect action observability, is that infinite strategies may not be observed if some informational nodes are never reached along the equilibrium path. The purpose of this Chapter is to apply an argument similar to Chapter 2 to make sure that all informational nodes are visited, so that the probability that player 1 is repeatedly playing a history dependent strategy will stay bounded away from 1 with arbitrarily small probability. This will be accomplished assuming that while the action of player 1 is perfectly observed, the action of player 2 is not: players commonly observe only a noisy outcome of the choice of the action of player 2, and there is a sufficient amount of noise so as to guarantee that all fixed length finite histories occur with strictly positive probability.

Making this assumption will be shown to be sufficient to guarantee that player 1 can successfully establish a reputation for an appropriate strategy in all finite stage games with perfect observability of player 1's action and imperfect observability of player 2's action. It will also be argued that introducing this kind of perturbation on the observed outcome of player 2's play, not only allows to generalize Schmidt's (1993) result, but also that a tighter equilibrium characterization is possible. In fact, when we consider the case of two long run player in games



which do not have conflicting interest with respect to player 1, the strategy player 1 would most like to commit to is not necessarily a fixed action but can also be a history dependent strategy.

This intuition can be easily explained considering the classical prisoner's dilemma: it has been argued that player 1 might want to commit to tit-for-tat rather than to the static Stackelberg strategy which gives only the static Nash equilibrium payoff<sup>®</sup> (which in this case is equal to the minmax payoff<sup>®</sup>). If the value of the discount factor of player 2 is sufficiently high, however, player 1 might want to commit to a strategy in which he occasionally plays Cheat and which calls for a sufficiently strong punishment for player 2 if he fails to play Cooperate.

As in the rest of the literature on reputation on repeated games, the goal of this Chapter is to characterize the set of Bayesian Nash equilibrium by finding a lower bound on player 1's Bayesian Nash equilibrium payoff<sup>®</sup>.

In the remainder of the Chapter the assumption will be made that types that are committed to any finitely repeating pure strategy<sup>1</sup> exist with strictly positive probability.

Under this assumption it will be shown that for a fixed discount factor of player 2, the equilibrium payoff<sup>®</sup> to an arbitrarily patient player 1 can be no less than an amount which is arbitrarily close to the best payoff<sup>®</sup> he could obtain by committing to any strategy subject to the condition that player 2 will play the best response to that strategy player 1 likes the least<sup>2</sup>.

The intuition that player 1 can take better advantage of his opportunity to establish a reputation will be shown to be true, in the sense that, if player 2

---

<sup>1</sup>A strategy is finitely repeated if there is an integer  $T$ , such that the strategy at time  $t$  is only determined by the history of the last  $T$  rounds.

<sup>2</sup>The assumption that player 2 plays the best response player 1 likes the least is made since in order to find a lower bound on Bayesian Nash equilibrium payoff<sup>®</sup> it is not possible to assume player 2's cooperation.

is sufficiently patient, then a sufficiently patient player 1 will be able to get an average payoff<sup>®</sup> which cannot be substantially less than the highest payoff<sup>®</sup> from a correlated strategy that gives player 2 more than his pure strategy minmax payoff<sup>®</sup>.

Fudenberg and Levine (1991) deal with the case of a long run player playing against a sequence of short run opponents a  $\bar{\tau}$ -xed stage game with imperfect action observability also for player 1, and analyze the result when the long run player is allowed to establish a reputation for a mixed strategy as well. In the present Chapter we will concentrate on the case in which the action of player 1 is perfectly observed since the only assumption which is necessary for the result is that there is a sufficient amount of noise in the observation of the action of player 2, and the more general result could be obtained at the expense of heavier notation.

Allowing a player to establish a reputation for a mixed strategy, in general increases the payoff<sup>®</sup> he could thus obtain. Since in the case in which player 2 is sufficiently patient a sufficiently patient player 1 is guaranteed to obtain almost the highest expected payoff<sup>®</sup> he could obtain with a correlated strategy subject to player 2 getting more than his pure strategy minmax payoff<sup>®</sup>, however, the introduction of mixed strategies gives a higher bound on player 1's Bayesian Nash equilibrium payoff<sup>®</sup> only if the minmax payoff<sup>®</sup> for player 2 is strictly less than his pure strategy minmax payoff<sup>®</sup>.

The rest of the Chapter is organized as follows. Section 3.2 introduces notation and describes the infinitely repeated game. The general result is given in 3.3. Section 3.4 provides the stronger equilibrium characterization of the case in which player 2 is sufficiently patient. In Section 3.5 the relationship between the results of the previous two Sections and existing literature (in particular Chapter 2 and Schmidt (1993)) is discussed. Section 3.6 provides a discussion of the results as well as directions of further research.

## 3.2 The Model

Consider a repeated game between two players, player 1 and player 2. Let  $A_1$  and  $A_2$  denote the finite (pure) action sets of the two players in the stage game with generic elements  $a_1$  and  $a_2$ , and let  $\Delta_i \subset A_i$  denote respectively mixed actions and mixed action spaces for player  $i = 1, 2$ . Further let  $A = A_1 \times A_2$ , and  $\Delta = \Delta_1 \times \Delta_2$  denote the spaces of pure and mixed strategy profiles.

At the end of each period  $t = 1, 2, \dots$  the action chosen by player 1 is observed by both players, while the action chosen by player 2 is not public knowledge: players commonly observe only a stochastic outcome drawn from a finite set,  $y \in Y$ . The probability distribution over outcomes depends on the action chosen by player 2 and is denoted by  $\mu(y|a_2)$  for a pure action  $a_2$ ; with an abuse of notation we will denote by  $\mu(y|\sigma_2)$  the probability distribution over outcomes  $y \in Y$  for a given mixed action  $\sigma_2$  which is defined in the obvious way from the probability distribution for pure actions.

Player 1 can be one of countably many types  $\theta \in \Theta$ . The types are drawn from a common knowledge prior  $\pi$  assigning positive probability to all points in  $\Theta$ . Player 1's type is private knowledge and is not known to player 2. In the following we will focus on a particular type  $\theta_0$ , which we refer to as the "rational type".

Stage game payoffs are  $u_1(a_1; y)$  for type  $\theta_0$  player 1 and  $u_2(a_1; a_2; y)$  for player 2. Player 1 has discount factor  $\beta_1$  and player 2 has discount factor  $\beta_2$ ; both player 2 and type  $\theta_0$  player 1 maximize the average discounted payoff in the infinitely repeated game. It will also be assumed that both  $u_1$  and  $u_2$  are bounded below from 0, and that  $u_1 \geq \underline{u}_1$  and  $u_2 \geq \underline{u}_2$ .

Types of player 1 other than type  $\theta_0$  have preferences over probability distributions over sequences of player 1 actions and public outcomes, but these are not

necessarily representable in a time separable form.

We will denote by  $h_t \in H_t = (A_1 \in Y)^{t-1}$  the public history of the game up to time  $t$ , and by  $h_t^i \in H_t^i = A_2^{t-1}$  the private history of player  $i$  up to time  $t$ .  $H = H_1$  and  $H^i = H_1^i$  will denote infinite histories.

A type behavior strategy for player 1 is a mapping  $\sigma_1 : H_1 \rightarrow A_1^1$ ,  $\sigma_1 = (\sigma_{11}; \dots; \sigma_{1t}; \dots)$  where  $\sigma_{1t} : H_t \rightarrow A_1$ . A behavior strategy for player 2 is a mapping  $\sigma_2 : H_1 \in H_1^2 \rightarrow A_2^1$ ,  $\sigma_2 = (\sigma_{21}; \dots; \sigma_{2t}; \dots)$  where  $\sigma_{2t} : H_t \in H_t^2 \rightarrow A_2$ . A behavior strategy for player 1 is a mapping  $\sigma_1 : - \rightarrow S_1^1$ , where  $S_1^1$  denotes the set of (infinite) type behavior strategies.

A Bayesian Nash equilibrium is a behavior strategy for player 1, and a behavior strategy for player 2, together with a set of probability beliefs over the set of types of player 1, such that: (i) for each type of player 1 given player 2's behavior strategy, no other type behavior strategy yields a distribution over time sequences of own actions and public outcomes that is preferred to the one obtained under his type behavior strategy; (ii) for player 2 given player 1's behavior strategy and the probability beliefs, no other behavior strategy yields a distribution over time sequences of own actions and public outcomes that is preferred to the one obtained under his behavior strategy; (iii) probability beliefs are updated using Bayes's rule whenever applicable.

With an abuse of notation we will denote by  $u_i(\sigma_{it}(h_{t-1}); \sigma_{2t}(h_{t-1}; h_{t-1}^2))$  the expected payoff to player  $i = 1, 2$  if players 1 and 2 are using respectively strategies  $\sigma_1$  and  $\sigma_2$  the public history of the game at time  $t$  is  $h_{t-1}$  and the private history of player 2 is  $h_{t-1}^2$ .

We will call a type behavior strategy for player 1 repeating if there exists an integer  $T$  such that play at time  $t = T + 1; \dots$  is entirely determined by the history between  $t \leq T$  and  $t \leq 1$ . Notice that for any  $T < \infty$  there are countably

many pure repeating strategies for player 1. A type of player 1 whose preferences are such that playing the type behavior strategy  $\sigma_1$  is strictly dominant is called committed to that strategy, and we will denote the event that player 1 is such a type by  $\mathbb{1}(\sigma_1)$ .

### 3.3 Establishing a Reputation Against a Patient Opponent

The purpose of this section is to study the general conditions under which reputation for any particular behavior can be established by player 1 when player 2 is patient. In the remainder of this section we will use the two following assumptions

**Assumption 4** If  $\sigma_1$  is pure repeating then  $\mathbb{1}(\sigma_1) > 0$ .

**Assumption 5** There exists a  $\epsilon \in (0, 1)$  such that  $\mathbb{1}(y_j^{\theta_2}) > \epsilon$ , for all  $\theta_2 \in \Theta_2$  and all  $y \in Y$ .

Assumption 4 guarantees the existence of "irrational" types to assure that the rational player 1 can hope to build a reputation for punishing player 2.

Assumption 5 is the truly substantive assumption: it says that the support of the distribution over outcomes does not depend on the action chosen by player 2. If the support of the distribution over outcomes depended on the action chosen by player 2 then it would be easy to construct counterexamples to the theorems below. The crucial point is that if player 2's play excludes certain outcomes  $y \in Y$ , then player 2 will not learn how player 1 would have responded to those contingencies, and this can easily prevent player 1 from building a reputation for particular responses to those contingencies.

Before analyzing reputation in our model, we calculate as a benchmark how much the long-run player might hope to get by precommitting.

**Definition 1** For all  $\delta_2 < 1$ ,  $\delta_2 \in B_2(\delta_1; \delta_2)$  if there is no other  $\delta_2 \in S_2^1$  such that after some history  $h_t$

$$\sum_{k=t}^{\infty} \delta_2^{k-t} u_2(\delta_1; \delta_2) + \delta_2 < \sum_{k=t}^{\infty} \delta_2^{k-t} u_2(\delta_1; \delta_2):$$

We will say that  $\delta_2 \in B(\delta_1^N; \delta_2)$ ,  $\delta_1^N \in S_1^N$ , if  $\delta_2 \in B_2(\delta_1; \delta_2)$ , for all  $\delta_1 \in S_1^1$  whose N-truncation coincides with  $\delta_1^N$ . Moreover, we will say that  $\delta_2^N \in B(\delta_1^N; \delta_2)$ , if  $\delta_2^N$  is the N-truncation of a  $\delta_2 \in B_2(\delta_1; \delta_2)$ .

**Definition 2** Let  $\hat{U}_1(\delta_1; \delta_2) = E \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_1(\delta_1; \delta_2) | g$ . Further let

$$\hat{U}_1^N(\delta_2) = \sup_{s_1 \in S_1^N} \inf_{\delta_2 \in B_2(s_1; \delta_2)} \hat{U}_1(s_1; \delta_2);$$

and let  $\hat{U}_1^N(\delta_2) = \hat{U}_1^N(0; \delta_2)$

$\hat{U}_1^N(\delta_2)$  denotes the maximal expected time average that player 1 can achieve by committing to a pure strategy, when the discount factor of player 2 is  $\delta_2$ .

Our goal is to prove the following theorem:

**Theorem 2** Suppose Assumptions 4 and 5 are satisfied. Then  $\lim_{\delta_1 \rightarrow 1} \hat{U}_1(\delta_1; \delta_2) = \hat{U}_1^N(\delta_2)$ .

We will prove this theorem via two Lemmas.

The purpose of the following Lemma is to show that there is a strategy such that if player 2 plays an  $\epsilon$ -best response to a N-truncation of it, and N is sufficiently large, then if player 1 is sufficiently patient, the average discounted payoff he will get in the N-fold repeated game is almost the average discounted payoff he would get in the infinitely repeated game, if player 2 played an  $\epsilon$ -best response to his strategy.

Lemma 3 For all  $\epsilon > 0$ ,  $\delta > 0$ , there exist  $N$ , a pure strategy for player 1  $s_1^N = (s_{11}^N(h_t); \dots; s_{1N}^N(h_{t+N}))$ , and a  $\epsilon_{11} < 1$ , such that for all  $\epsilon_1 > \epsilon_{11}$  if  $\delta_{22} \geq \delta$ , the average discounted payoff to type  $i_0$  player 1 in the  $N$ -fold repeated game is at least  $U_1^a(2^2; \epsilon_2) - \epsilon$ .

Proof: For all  $\delta > 0$ ,  $\epsilon = 4\delta > 0$  there exists a pure strategy  $s_1 \in S_1^1$  such that

$$\inf_{\delta_{22} \geq \delta} U_1(s_1; \delta_{22}) > U_1^a(2^2; \epsilon_2) - \frac{\epsilon}{4} \quad (3.1)$$

Let  $T$  satisfy

$$\frac{1}{T} E \inf_{\delta_{22} \geq \delta} \sum_{t=1}^T u_1(s_{1t}(h_{t-1}); \delta_{2t}(h_{t-1})) > U_1^a(2^2; \epsilon_2) - \frac{\epsilon}{2}$$

for all  $T \geq T$ . Then choose  $N > T$  and  $\epsilon$  such that  $\epsilon_{22} = (1 - \epsilon_2) < \delta_{22}$  and  $\epsilon_{11} \epsilon = N < \epsilon_2$ .

Let  $s_1^N$  be the  $N$ -truncation of a  $s_1$  satisfying (3.1)<sup>3</sup>. Now consider an  $N$ -fold repetition of the stage game in which player 1 plays  $s_1^N$ . If player 2 plays an  $\epsilon_2$  best response to  $s_1^N$ , since  $\epsilon$  is such that  $\epsilon_{22} = (1 - \epsilon_2) < \delta_{22}$ , then in the first  $N - \epsilon$  periods of the  $N$ -fold repeated game he plays a  $\delta_{22}$  best response to  $s_1$ . Since  $\epsilon_{11} \epsilon = N < \epsilon_2$  it follows that

$$\inf_{\delta_{22} \geq \delta} \frac{1 - \epsilon_1}{1 - \epsilon_1^N} \sum_{t=1}^N u_1(s_{1t}(h_{t-1}); \delta_{2t}(h_{t-1})) > U_1^a(2^2; \epsilon_2) - \frac{\epsilon}{2} - \epsilon_{11} \sum_{t=1}^{\epsilon} \frac{1}{N} \frac{1 - \epsilon_1}{1 - \epsilon_1^N} \epsilon$$

Since for all  $\epsilon = 4\delta$  there exists a  $\epsilon_{11} < 1$  such that  $\epsilon_{11} \sum_{t=1}^N \frac{1}{N} \frac{1 - \epsilon_1}{1 - \epsilon_1^N} \epsilon < \epsilon = 4\delta$ , the Lemma follows.  $\square$

The purpose of the next Lemma is to show that if player 1 always plays a given strategy and there exists with strictly positive probability a type of player 1 that is committed to that strategy, then for every integer  $\epsilon$ ,  $\epsilon < 1$ ,  $\delta > 0$ , the probability that the number of periods in which player 2 will expect player 1 to

<sup>3</sup>As is clear  $s_1^N$  is a pure repeating strategy.

play like the commitment type in the following  $\zeta$  periods with probability less than  $\eta$  exceeds a given finite number is less than  $\epsilon$ .

Let  $s_1^\pi$  be the infinite repetition of strategy  $s_1^N$  constructed in the proof of Lemma 3, and let  $H^\pi$  be the set of histories consistent with player 1 playing  $s_1^\pi$ . Let  $\mathbb{1}_t^{\pi, \zeta} = \Pr(a_{1t} = s_{1t}^\pi(h_{t-1}));$  for all  $h_{t-1} \in H_{t-1}^\pi; t^0 = t; \dots; t + \zeta - 1$ ) and let  $n(\mathbb{1}_t^{\pi, \zeta} \cdot \eta)$  denote the random variable indicating the number of periods in which  $\mathbb{1}_t^{\pi, \zeta} \cdot \eta$  in the infinitely repeated game. Finally, let  $\theta^\pi = \theta(s_1^\pi)$  be a type that is committed to  $s_1^\pi$  and let  $\theta^\pi$  denote the event that the type of player 1 is not  $\theta^\pi$  ( $\theta^\pi = 1 - \theta^\pi$ ). Since  $s_1^\pi$  is pure repeating,  $\theta^\pi \in \Theta$ , and therefore  $\theta^\pi = \Pr(\theta^\pi) > 0$ .

**Lemma 4** Let  $0 < \eta < 1$ . Suppose that Assumption 2 is satisfied and that  $(\beta_{1t}, \beta_{2t})$  are such that  $\Pr(\theta^\pi | \beta) = 1$ . Let  $K_1 = \zeta (\log \theta^\pi = \log \eta)$  and for all  $\epsilon > 0$  let  $K_2(\epsilon) = \log(1 - (1 - \epsilon)^{1=K_1}) = \log(1 - \epsilon^{K_1})$ . Then for all  $\epsilon > 0$

$$\Pr(n(\mathbb{1}_t^{\pi, \zeta} \cdot \eta) > K_1 K_2(\epsilon) | \theta^\pi) < \epsilon;$$

**Proof:** By Bayes's law we have

$$\begin{aligned} \Pr(\theta^\pi | h_t) &= \Pr(\theta^\pi | h_{t-1}; a_{1t}; y_t) \\ &= \frac{\Pr(\theta^\pi | h_{t-1}) \Pr(a_{1t}; y_t | \theta^\pi)}{\Pr(\theta^\pi | h_{t-1}) \Pr(a_{1t}; y_t | \theta^\pi) + (1 - \Pr(\theta^\pi | h_{t-1})) \Pr(a_{1t}; y_t | \theta^\pi)} \\ &= \frac{\Pr(\theta^\pi | h_{t-1}) \Pr(a_{1t} | \theta^\pi)}{\Pr(\theta^\pi | h_{t-1}) \Pr(a_{1t} | \theta^\pi) + (1 - \Pr(\theta^\pi | h_{t-1})) \Pr(a_{1t} | \theta^\pi)} \\ &= \frac{\Pr(\theta^\pi | h_{t-1})}{\Pr(a_{1t})} \end{aligned} \tag{3.2}$$

Let  $\mathbb{1}_t^{\pi, \zeta} = \Pr(a_{1t} = s_{1t}^\pi(h_{t-1}); a_{1t} \neq s_{1t}^\pi(h_{t-1})$  for some  $h_{t-1} \in H_{t-1}^\pi; t^0 = t + 1; \dots; t + \zeta - 1$ ). Then (3.2) can be rewritten as

$$\Pr(\theta^\pi | h_t) = \frac{\Pr(\theta^\pi | h_{t-1})}{\mathbb{1}_t^{\pi, \zeta} + \mathbb{1}_t^{\pi, \zeta}};$$



Suppose that  $\frac{1}{4}^{\ell} < \frac{1}{4}$  for some  $\ell$ . This means that there is a history  $h_{t-1}$ , such that  $\Pr(a_{1t} = s_{1b}^{\pi}(h_{t-1})) < \frac{1}{4}$ ,  $t = t + 1; \dots; t + \ell - 1$ . Now suppose that this history actually occurs<sup>4</sup>, and that  $\Pr(!^{\pi} | h_{t-1}) = \Pr(!^{\pi} | h_t)$ , which means that player 1's play until time  $t - 1$  led to no belief updating. From (3.2) we have

$$\Pr(!^{\pi} | h_t) = \frac{\Pr(!^{\pi} | h_{t-1})}{\Pr(a_{1t} = s_{1b}^{\pi}(h_{t-1}))}$$

If player 1 plays  $a_{1t} = s_{1b}^{\pi}(h_{t-1})$ , then the probability that he is type  $!^{\pi}$  has to go up by at least a factor of  $1/\frac{1}{4}$ , since  $\Pr(a_{1t} = s_{1b}^{\pi}(h_{t-1})) < \frac{1}{4}$ . Given that  $\Pr(!^{\pi} | h_0) = \pi$ , if history  $h_{t-1}$  occurs, then  $\Pr(!^{\pi} | h_t) \geq \pi/\frac{1}{4}$ . If a sequence of  $\ell$  stage games is repeated  $K$  times,  $\frac{1}{4}^{\ell} < \frac{1}{4}$  at the beginning of each sequence, and the appropriate history  $h_{t-1}$  occurs in each of the  $K$  repetitions of the sequence, then

$$\Pr(!^{\pi} | h_{K\ell}) \geq \frac{\pi}{\frac{1}{4}^K} \quad (3.3)$$

However, since  $\Pr(!^{\pi} | h_t) \leq 1$ , if  $\pi/\frac{1}{4}^K > 1$  inequality (3.3) is violated and a contradiction to the hypothesis that  $\frac{1}{4}^{\ell} < \frac{1}{4}$  at the beginning of each of the  $K$  repetitions of the  $\ell$ -fold repeated game is obtained. Taking the log of (3.3) we obtain the definition of  $K_1$ .

Suppose now that the stage game is repeated  $K_1 K_2^{(2)\ell}$  times. We want to find the smallest number  $K_2^{(2)}$  such that

$$\Pr(n(\frac{1}{4}^{\ell} \cdot \frac{1}{4}) > K_1 K_2^{(2)\ell} \pi) \leq \epsilon \quad (3.4)$$

is satisfied for a given  $\epsilon > 0$ . Suppose that  $\frac{1}{4}^{\ell} < \frac{1}{4}$ , and let  $h_{t-1}$  be such that  $\Pr(a_{1t} = s_{1b}^{\pi}(h_{t-1})) < \frac{1}{4}$ . Then by Assumption 5, a lower bound on the probability that  $h_{t-1}$  occurs is  $\epsilon^{\ell}$ . Therefore if the stage game is repeated  $K_2^{(2)\ell}$  times, the probability that no appropriate history  $h_{t-1}$  occurs is

$$\bar{A} = (1 - \epsilon^{\ell})^{K_2^{(2)\ell}} \quad (3.5)$$

---

<sup>4</sup>Remember that by Assumption 5 all histories in  $H^{\pi}$  occur with strictly positive probability.

and  $1 - \hat{A}$  is the probability that at least one appropriate history occurs.

If the stage game is repeated  $K_1 K_2(\delta)^{K_1}$  times, the probability that at least  $K_1$  appropriate histories occur is greater than  $(1 - \hat{A})^{K_1}$ . Therefore a sufficient condition for the probability that less than  $K_1$  appropriate histories occur when the stage game is repeated  $K_1 K_2(\delta)^{K_1}$  times to be less than  $\epsilon$  is

$$1 - (1 - \hat{A})^{K_1} > \epsilon \quad (3.6)$$

from which

$$\hat{A} > 1 - (1 - \epsilon)^{1/K_1} \quad (3.7)$$

Substituting (3.5) into (3.7) and rearranging provides

$$(1 - \delta)^{K_2(\delta)} > 1 - (1 - \delta)^{1/K_1}$$

Taking logs gives

$$K_2(\delta) > \frac{\log(1 - (1 - \delta)^{1/K_1})}{\log(1 - \delta)}$$

which concludes the proof.  $\square$

We are now in the position to prove Theorem 2.

**Theorem 2** Suppose Assumptions 4 and 5 are satisfied. Then

$$\lim_{\delta \rightarrow 1} N_1(\delta; \delta) = U_1^*(\delta):$$

**Proof:** From Lemma 3 for all  $\epsilon > 0$ ,  $\delta > 0$ , there exists an  $N$  and a pure repeating strategy for player 1  $s_1^N$  such that if player 2 plays an  $\delta$  best response to it, there exists a  $\epsilon_{11} < 1$  such that for all  $\delta > \epsilon_{11}$  the average discounted payoff to player 1 in the  $N$ -fold repeated game is at least  $U_1^*(\delta; \delta) - \epsilon$ . Since the  $\delta$  best response correspondence of player 2,  $B_2(\delta; \delta)$ , is upper hemi-continuous, there exists a  $\eta < 1$  such that if  $\delta_1^N$  is such that  $\delta_1^N > \eta$ , then  $B_2(\delta_1^N; \delta) \cap B_2(s_1^N; \delta)$ .

Let  $\delta$  in Lemma 4 equal  $N^{-1}$ . Then we know that for all  $\epsilon > 0$  and for all  $\beta < 1$  the probability that the number of periods in which  $\frac{1}{4} \beta^N < \beta$  is larger than  $K_1 K_2(2) \delta$  is less than  $\epsilon$ .

Since  $K_1 K_2(2) \delta$  is finite for all  $\epsilon > 0$ , for all  $\epsilon > 0$  there exists a  $\beta_1$  such that for all  $\beta > \beta_1$  if player 1 always plays  $s_1^*$  then his infinitely repeated game average discounted payoff is at least  $U_1^*(2; \beta_2) - \epsilon$ . Since strategy  $s_1^*$  is always feasible for the rational type in any Bayesian Nash equilibrium, the rational player 1 (type  $\theta_1$ ) has to get at least what he would get by playing  $s_1^*$  and the theorem follows.

2

### 3.4 The Value of Reputation with an Arbitrarily Patient Opponent

Let  $\underline{u}_2$  denote the pure strategy minmax for player 2:

$$\underline{u}_2 = \min_{a_2 \in A_2} \max_{a_1 \in A_1} u_2(a_1; a_2)$$

In addition to the assumptions we made in the previous section, in this section we will also need the following assumption

**Assumption 6** There exists an  $a_2 \in A_2$  such that  $u_2(a) > \underline{u}_2$ .

Assumption 6 says that there is a profile that is better for player 2 than the pure strategy minmax. This is a mild non-degeneracy condition. If it were to fail, the indifference of the player 2 might well make him immune to threats by the long-run player.

Let  $\theta \in \Delta(A)$  be a probability distribution on pure strategy profiles<sup>5</sup>. Then define the set of enforceable pure action profiles  $E$  the set of correlated action

---

<sup>5</sup>This means that  $\theta$  can also be a correlated strategy.

profiles such that the payoff to player 2 is strictly larger than his minmax,  $\underline{u}_2$ :

$$E = \{\theta \in \Theta \mid u_2(\theta) > \underline{u}_2\}$$

Given the definition of  $E$  we will now define  $\bar{U}_1^\alpha$  as follows:

$$\bar{U}_1^\alpha = \sup_{\theta \in E} u_1(\theta) \quad (3.8)$$

Our goal is to show that when player 2 is sufficiently patient, an arbitrarily patient player 1 can get the highest payoff subject to the constraint that player 2 is getting strictly more than his minmax level. In other words we want to prove the following theorem:

**Theorem 3** Suppose Assumptions 4-6 are satisfied. Then

$$\lim_{\alpha \rightarrow 1} \lim_{\beta \rightarrow 1} N_1(\beta; \alpha) = \bar{U}_1^\alpha:$$

We will prove this theorem via several Lemmas.

Let  $a^N = (a_1^N; \dots; a_N^N) \in A^N$  and let

$$u_i(a^N; \beta_i) = \frac{1}{1 - \beta_i} \sum_{t=1}^{\infty} \beta_i^{t-1} u_i(a_t^N)$$

denote the average discounted payoff to player  $i = 1, 2$  in the  $N$ -fold repeated game under action profile  $a^N$ .

**Lemma 5** For all  $\hat{\epsilon}_1 > 0$  there exist  $N, \beta_{11} < 1, \beta_{21} < 1$  such that for all  $\beta_1 > \beta_{11}, \beta_2 > \beta_{21}, u_1(a^N; \beta_1) > \sup_{\theta \in E} u_1(\theta) + \hat{\epsilon}_1$  and  $u_2(a^N; \beta_2) > \underline{u}_2$ :

**Proof:** By continuity of  $u_1$  and  $u_2$ , for all  $\hat{\epsilon}_1 > 0$  we can find  $\theta^\alpha$  such that

$$u_1(\theta^\alpha) > \sup_{\theta \in E} u_1(\theta) + \frac{\hat{\epsilon}_1}{3}$$

and  $u_2(\hat{a}^2) > \underline{u}_2$ . Again by continuity of  $u_1$  and  $u_2$  for all  $\hat{\epsilon}_1 = 3 > 0$  we can find  $\hat{\epsilon}_{11} < 1, \hat{\epsilon}_{21} < 1, a^N \in A^N$  such that for all  $\epsilon_1 > \hat{\epsilon}_{11}, \epsilon_2 > \hat{\epsilon}_{21}$

$$u_1(a^N; \epsilon_1) > u_1(\hat{a}^1) + \frac{\hat{\epsilon}_1}{3}$$

and  $u_2(a^N; \epsilon_2) > \underline{u}_2$  and the proof is complete.  $\square$

Now consider an  $N$ -period pure strategy  $a^N \in A^N$  and let

$$u_1(a^N) = \frac{1}{N} \sum_{t=1}^N \sum_{y \in Y} \frac{1}{2} (y_j a_{2t}^t) u_1(a_{1t}^t; y)$$

denote the average payoff to player 1 from playing  $a^N$  in the  $N$ -fold repeated game.

Now let  $a^{N^1} \in A^{N^1}$  be the  $N^1$ -fold repetition of  $a^N$  and let  $U_1(a^{N^1})$  denote a random variable that assumes value

$$U_1(a^{N^1}) = \frac{1}{N^1} \sum_{t=1}^{N^1} u_1(a_{1t}^{N^1}; y_t)$$

with probability  $\frac{1}{2} \sum_{t=1}^{N^1} (y = y_j a_{2t}^{N^1})$ . The distribution of  $U_1(a^{N^1})$  gives the distribution of possible values of the average discounted payoff to player 1 when  $a^{N^1}$  is being played in the  $N^1$ -fold repeated game.

**Lemma 6** For every  $N, a^N \in A^N, \hat{\epsilon}_2 > 0$  there exists a  $\hat{\epsilon}_{12} < 1$  and a  $\hat{\epsilon}_1$  such that for all  $\epsilon_1 > \hat{\epsilon}_{12}$

$$\Pr(U_1(a^{N^1}) < u_1(a^N) + \hat{\epsilon}_2) < \hat{\epsilon}_2:$$

**Proof:** Let  $\mathcal{U}_1(a^K)$  denote a random variable that assumes value

$$\mathcal{U}_1(a^K) = \frac{1}{K} \sum_{t=1}^K u_1(a_{1t}^K; y_t)$$

with probability  $\prod_{t=1}^K \frac{1}{2} (y_t = y_t^j a_{2t}^K)$ . Then from the weak law of large numbers, for all  $\epsilon_2 = 3 > 0$  there exists a  $\epsilon_1$  such that

$$\Pr(U_1(a^{N^1}) < u_1(a^N) - \frac{\epsilon_2}{2}) < \frac{\epsilon_2}{3}.$$

Since for all  $N^1 \lim_{\epsilon_1 \rightarrow 1} U_1(a^{N^1}) = U_1(a^{N^1})$ , for all  $\epsilon_2 = 3 > 0$  there exists a  $\epsilon_{12} < 1$  such that for all  $\epsilon_1 > \epsilon_{12}$

$$|U_1(a^{N^1}) - U_1(a^{N^1})| < \frac{\epsilon_2}{3}$$

and the proof is complete.  $\square$

**Corollary 1** For every  $\epsilon_1; \epsilon_2 > 0$  there exist  $N, a^N \in A^N, \epsilon_1^*, \epsilon_2^*$  such that

- (i)  $\Pr(U_1(a^{N^1}) < U_1^* - \epsilon_1 - \epsilon_2) < \epsilon_2$  for all  $\epsilon_1 > \epsilon_1^*$ ;
- (ii)  $u_2(a^{N^1}; \epsilon_2) > u_2$  for all  $\epsilon_2 > \epsilon_2^*$

**Proof:** Immediate from Lemmas 5 and 6.

**Lemma 7** For all  $\epsilon_1; \epsilon_2; \epsilon_3 > 0$  there exist  $N_3$ , a strategy for player 1  $s_1^{N_3} \in S_1^{N_3}$ , a  $\epsilon_{23} < 1$ , and an  $\epsilon^2 > 0$  such that for all  $\epsilon_{13} < \epsilon_2 < 1$  there exists a  $\epsilon_{13} < 1$  such that for all  $a_{\pm 1} > \epsilon_{13}$  if  $\frac{3}{4} N_3 \in B_2(s_1^{N_3}; \epsilon_2)$ , the average discounted payoff to type  $\epsilon_0$  player 1 in the  $N_3$ -fold repeated game is at least  $U_1^* - \epsilon_1 - \epsilon_2 - \epsilon_3$ .

**Proof:** We want to show that for all  $\epsilon_1; \epsilon_2; \epsilon_3 > 0$  there exists a pure repeating strategy for player 1 in the  $N_3$ -fold repeated game,  $s_1^{N_3}$ , and a discount factor for player 2  $\epsilon_{23}$  such that if  $\epsilon_{23} < \epsilon_2 < 1$  and player 2 plays a  $\frac{3}{4} N_3 \in B_2(s_1^{N_3}; \epsilon_2)$ , then the loss in the  $N_3$ -fold repeated game to a sufficiently patient type  $\epsilon_0$  player 1 with respect to  $U_1^*$  is no more than  $\epsilon_1 + \epsilon_2 + \epsilon_3$ .

Let  $v_1^K(h_t)$  denote the average discounted payoff to player 1 in the last  $K$  periods under history  $h_t$ :

$$v_1^K(h_t) = \frac{1}{K} \sum_{k=t-K+1}^t \delta^{k-t} u_1(h_k)$$

For fixed  $\epsilon_1, \epsilon_2 > 0$ ,  $K \geq 1$ ,  $J \geq 1$  define the random variable  $s_t(h_{t-1})$  as following:

$\geq$  For  $t = 1; \dots; K$ ,  $s_t(h_{t-1}) = t$ , for all  $h_{t-1} \in H_{t-1}$ ;

$\geq$  For  $t = K + 1; K + 2; \dots$ :

$$s_t(h_{t-1}) = \begin{cases} s_{t-1}(h_{t-2}) + 1 & \text{if } s_{t-1}(h_{t-2}) = 0; \text{ for all } h_{t-1} \in H_{t-1} \\ & \text{or if } 0 < s_{t-1}(h_{t-2}) < K \text{ and } v_1^K(h_{t-1}) \geq U_1^{\epsilon_1}(\epsilon_1, \epsilon_2) \\ 1 & \text{if } s_{t-1}(h_{t-2}) = K \text{ and } v_1^K(h_{t-1}) \geq U_1^{\epsilon_1}(\epsilon_1, \epsilon_2) \\ j + 1 & \text{if } s_{t-1}(h_{t-2}) > 0 \text{ and } v_1^K(h_{t-1}) < U_1^{\epsilon_1}(\epsilon_1, \epsilon_2) \end{cases}$$

From Corollary 1 we know that for every  $\epsilon_1, \epsilon_2 > 0$  there exist  $N$ ,  $a^{N-1} \geq \underline{a}_1 < 1$   $\underline{a}_2 < 1$  such that for every  $a_1 > \underline{a}_1$

$$\Pr(U_1(a^{N-1}) < U_1^{\epsilon_1}(\epsilon_1, \epsilon_2)) < \epsilon_2$$

and for every  $a_2 > \underline{a}_2$ ,  $u_2(a^{N-1}; a_2) > \underline{u}_2$ . Let  $K = N-1$  and let  $a^K$  be an action profile that satisfies Corollary 1 for fixed  $\epsilon_1, \epsilon_2$ , and let  $a_1^{N-1} = (a_{11}^{N-1}; \dots; a_{1N-1}^{N-1})$  be player 1's component of this action profile.

Recall that  $\underline{a}_1$  was defined as an  $a_1 \in A_1$  such that

$$\max_{a_2 \in A_2} u_2(\underline{a}_1; a_2) = \underline{u}_2;$$

and consider the following strategy for player 1:

$$s_{1t}^{\epsilon_1}(h_{t-1}; K; J) = \begin{cases} a_{1, s_t(h_{t-1})}^K & \text{if } s_t(h_{t-1}) \geq 1 \\ \underline{a}_1 & \text{if } s_t(h_{t-1}) < 1 \end{cases}$$

Claim 1 Suppose that there exists a  $\underline{u}_2 < 1$  such that  $u_2(a_1^N; \underline{u}_2) > \underline{u}_2$  for  $\underline{u}_2 < \underline{u}_2 < 1$ , then for all  $\epsilon > 0$  there exist  $\delta, J, N_3, \epsilon > 0$  such that if  $\frac{1}{2}N_3$  is such that

$$\Pr(v^{N_1}(h_t) < \frac{1}{2} \delta) > \epsilon$$

for some history  $h_t$  that can be reached with strictly positive probability, for some  $t = N_1 + 1; \dots; N_3$ , then  $\frac{1}{2}N_3 \in B_2(s_1^{N_3}(N_1; J))$ .

Pf: Let  $H_t^s(N_1; J)$  be the set of time  $t$  histories consistent with player 1 playing  $s_1^s(N_1; J)$  for some  $s, J$ . Suppose strategy  $\frac{1}{2}N_3$  is such that for some history  $h_t \in H_t^s(N_1; J)$  that can be reached with strictly positive probability and such that  $s_t(h_{t-1}) > 1$

$$\Pr(v_1^{N_1}(h_{t-1}) < \frac{1}{2} \delta) > \epsilon:$$

Consider the following strategy for player 2,  $s_2^s(N_1; J) = (s_{21}^s(h_0; N_1; J); s_{22}^s(h_1; N_1; J); \dots)$  where

$$s_{2t}^s(h_{t-1}; K; J) = \begin{cases} a_{2, s_t(h_{t-1})}^N & \text{if } s_t(h_{t-1}) > 1 \\ \underline{a}_2 & \text{if } s_t(h_{t-1}) < 1 \end{cases}$$

Since player 2 can always play this strategy, for a strategy to be an  $\epsilon$  best response to  $s_1^s(N_1; J)$  it has to be the case that it gives more than the payoff of  $s_2^s(N_1; J)$  minus  $\epsilon$ .

For a finite action profile  $a^K \in A^K$  let

$$u_2(a_1^T; \underline{u}_2) = \sum_{t=1}^T \delta^{t-1} (\max_{a_2 \in A_2} u_2(a_{1t}^N; a_2)) \quad (3.9)$$

denote the highest discounted payoff player 2 can get if player 1 plays according to  $a^K$ . Consider the following inequality

$$u_2(a_1^N; \underline{u}_2) \frac{1 - (\delta \underline{u}_2)^{N_1} (1 - \delta)^J}{1 - \delta \underline{u}_2^{N_1} (1 - \delta)} + \frac{\underline{u}_2 - \delta \underline{u}_2^{N_1} \epsilon}{1 - \delta \underline{u}_2^{N_1} \epsilon} \frac{1 - (\delta \underline{u}_2)^{N_1} (1 - \delta)^J}{1 - \delta \underline{u}_2^{N_1} (1 - \delta)} \delta^{N_1(J+1)} \frac{1 - (\delta \underline{u}_2)^J}{\delta} + \frac{\delta^{N_1 J + 1}}{1 - \delta \underline{u}_2} \underline{u}_2 >$$



$$\begin{aligned}
&> \mathbf{b}_2(a_1^N; \pm_2) \frac{1 - (\pm_2^{N^1} (1 - \hat{\gamma}_2))^J}{1 - \pm_2^{N^1} (1 - \hat{\gamma}_2)} + \\
&+ \frac{\underline{u}_2}{1 - \pm_2} \pm_2^{N^1} \hat{\gamma}_2 \frac{1 - (\pm_2^{N^1} (1 - \hat{\gamma}_2))^J}{1 - \pm_2^{N^1} (1 - \hat{\gamma}_2)} i^{\pm_2^{N^1(J-1)}} \frac{1 - (1 - \hat{\gamma}_2)^J}{\hat{\gamma}_2} i^{\hat{\gamma}_2}
\end{aligned} \tag{3.10}$$

Some uninteresting algebra shows that the left hand side of inequality (3.10) gives an upper bound on the infinitely repeated game payoff to player 2 if player 1 plays a strategy whose  $JN^1$  truncation is  $s_1^{\pm_2^{JN^1}}(N^1; J)$  while player 2 plays a strategy such that there is a history  $h_t \in H_t^{\pm_2}(N^1; J)$  for which  $\Pr(v_1^{N^1}(h_t) < \hat{U}_1^{\pm_2} i^{\hat{\gamma}_2} (1 - \hat{\gamma}_2)) > \hat{\gamma}_4$ . Similarly the right hand side of (3.10) gives a lower bound on the infinitely repeated game payoff to player 2 when players 1 and 2 play respectively  $s_1^{\pm_2^{JN^1}}(N^1; J)$  and  $s_2^{\pm_2^{JN^1}}(N^1; J)$ . Inequality (3.10) is therefore a necessary condition for a strategy for player 2 such that for some history  $\Pr(v_1^{N^1}(h_t) < \hat{U}_1^{\pm_2} i^{\hat{\gamma}_2} (1 - \hat{\gamma}_2)) > \hat{\gamma}_4$  to be an  $\epsilon$  best response to  $s_1^{\pm_2^{JN^1}}$ .

Since by Lemma 6 for all  $N$  and  $a^N \in A^N$  there exist  $\epsilon_1, \epsilon_{12} < 1, \epsilon_{22} < 1$  such that for all  $\epsilon_1 > \epsilon_{12}, \hat{\gamma}_2 > \Pr(v_1^{N^1}(h_t) < \hat{U}_1^{\pm_2} i^{\hat{\gamma}_2} (1 - \hat{\gamma}_2))$  is arbitrarily small and for all  $\epsilon_2 > \epsilon_{22}$   $\mathbf{b}_2(a^{N^1}; \pm_2) > \epsilon_2$ , we can choose a  $\epsilon_1$  so as to make the right hand side arbitrarily close to  $\mathbf{b}_2(a^{N^1}; \pm_2)(1 - \pm_2^{JN^1}) = (1 - \pm_2) i^{\pm_2}$ , which is in turn larger than  $\underline{u}_2 i^{\hat{\gamma}_2}$ . This implies that there exists a  $\epsilon_{23} < 1$  such that there exists a  $J, \epsilon_2 > 0$  for which the previous inequality is violated. Letting  $N_3 > JN^1$  the Claim follows.  $\square$

From the previous Claim we conclude that for all  $\hat{\gamma}_4 > 0$  there exists a  $\epsilon_{23} < 1$  such that if player 2's discount factor is larger than  $\epsilon_{23}$  and player 2 plays an  $\epsilon$  best response to  $s_1^{\pm_2^{N_3}}$ ,  $\Pr(v_1^{N^1}(h_t) < \hat{U}_1^{\pm_2} i^{\hat{\gamma}_2} (1 - \hat{\gamma}_2)) < \hat{\gamma}_4$ , for all  $h_t \in H_t^{\pm_2}(N^1; J)$ . This implies that there exists a  $\epsilon_{13} < 1$  such that for all  $\epsilon_1 > \epsilon_{13}$  the discounted payoff to player 1 will be at least  $(\hat{U}_1^{\pm_2} i^{\hat{\gamma}_2} (1 - \hat{\gamma}_4))$  in all but the last  $J$  repetitions of the  $N^1$ -fold repeated game. Defining  $\hat{\gamma}_3 = \hat{U}_1^{\pm_2} \hat{\gamma}_4$  and choosing  $N_3$  sufficiently large the Lemma follows.  $\square$

We are now in the position to prove Theorem 3.

Theorem 3 Suppose Assumptions 4-6 are satisfied. Then

$$\lim_{\delta \rightarrow 1} \lim_{N_3 \rightarrow \infty} N_1(\delta_1; \delta_2) \geq U_1^*$$

Proof: From Lemma 7 for all  $\epsilon_1; \epsilon_2; \epsilon_3 > 0$  there exists an  $N_3$   $\delta_{23} < 1$  and a pure repeating strategy for player 1  $s_1^{N_3}$  such that for  $\delta_{23} < \delta_2 < 1$ , if player 2 plays an  $\delta_2$  best response to  $s_1^{N_3}$ , there exists a  $\delta_{13} < 1$  such that for all  $\delta_1 > \delta_{13}$  the average discounted payoff to player 1 in the  $N_3$ -fold repeated game is at least  $U_1^* - \epsilon_1 - \epsilon_2 - \epsilon_3$ . Since the  $\delta_2$  best response correspondence of player 2  $B_2(\delta_1)$  is upper hemi-continuous, for a fixed  $\delta_{23} < \delta_2 < 1$  there exists a  $\eta < 1$  such that if  $\delta_1^{N_3}$  is such that  $\delta_1^{N_3} > \eta$ , then  $B_2(\delta_1^{N_3}; \delta_2) \subset B_2(s_1^{N_3}; \delta_2)$ .

For all  $\epsilon > 0$  let  $N_3$  be such that there exists an  $\delta^2 > 0$  such that if player 2 plays a  $\delta^2$  best response to  $s_1^{N_3}$ , then there exists a  $\delta_{14}$  such that for all  $\delta_1 > \delta_{14}$  player 1 gets at least  $U_1^* - \epsilon$ . Let  $\delta$  in Lemma 4 equal  $N_3$ . Then we know that for all  $\delta^2 > 0$  and for all  $\eta < 1$  the probability that the number of periods in which  $\delta_1^{N_3} < \eta$  is larger than  $K_1 K_2(\delta^2) \delta$  is less than  $\delta^2$ .

Since  $K_1 K_2(\delta^2) \delta$  is finite for all  $\delta^2 > 0$ , for all  $\epsilon > 0$  there exists a  $\delta_1$  such that for all  $\delta_1 > \delta_1$  if player 1 always plays  $s_1^*$  then his infinitely repeated game average discounted payoff is at least  $U_1^* - \epsilon$ . Since strategy  $s_1^*$  is always feasible in any Bayesian Nash equilibrium, player 1 has to get at least what he would get by playing  $s_1^*$  and the theorem follows.  $\square$

Theorem 3 says that if player 2 is sufficiently patient, then an arbitrarily patient player 1 will get an equilibrium average payoff which is at least what he could get from any correlated strategy that gives player 2 strictly more than his pure strategy minmax payoff, since for an arbitrarily patient player a sequence of pure strategy profiles is equivalent to a correlated strategy.

This implies that allowing player 1 to establish a reputation for mixed strategies would give a higher bound on player 1's equilibrium payoff only if player 2's minmax payoff is strictly lower than his pure strategy minmax payoff. Therefore, in the cases in which minmax and pure strategy minmax payoff for player 2 coincide, the bound in Theorem 3 provides a tight characterization of equilibrium payoffs.

## 3.5 Special Cases

The purpose of this section is to discuss the relationship of the general result of Section 3.3 with the existing literature. Our goal will be to show that the results of Chapter 2 and of Schmidt (1993) can be derived as special cases of this more general model.

### 3.5.1 Reputation with a Short Run Opponent

As was discussed in the Introduction, Chapter 2 considers the case of a long run player who can establish a reputation for a particular behavior against a sequence of short run opponents who all observe the previous play of the game. The main goal of Chapter 2 was to show that introducing a perturbation on the types of the short run opponents that made sure that all informational nodes of a stage game would occur with strictly positive probability, the long run player could establish a reputation for playing the (short run) Stackelberg strategy, which in turn allows to use reputational arguments to characterize the set of Nash equilibrium payoffs by imposing a lower bound on the long run player's Nash equilibrium payoff.

Even though Chapter 2 was phrased in terms of perturbations on the types of the short run opponents, the result is actually driven by the fact that all nodes of the stage game are visited with strictly positive probability. In this sense,

the result can be rephrased as giving a lower bound on the long run player's payoff when Assumption 5 holds, i.e. when the play of the short run player is not observed, and when the support of the probability distributions over outcomes do not depend on the action chosen by the short run player.

When player 2's discount factor is equal to zero, the  $\delta_2$  best response correspondence  $B_2(\delta_1; \pi_2)$  coincides with the short run  $\delta_2$  best response correspondence, which implies that  $U_1^*$  is equal to the static Stackelberg payoff and Lemma 3 becomes trivial since it states that if player 1 plays the Stackelberg strategy and player 2 plays a short run best response to it, then player 1's payoff cannot be lower than the Stackelberg payoff.

If player 2 has zero discount factor, he will not care about the future, and therefore all that we need to show is that the probability that the number of periods in which player 2 expects player 1 to play at the current stage only the Stackelberg action with probability less than an arbitrary  $\epsilon < 1$  exceeds a given number, becomes arbitrarily small as the given number increases.

In other words, given that the number of periods each short run player is interested in is only 1, we need to state Lemma 4 for  $\delta = 1$ . It is immediate to see that the statement of Lemma 4 in Section 3.3 is equivalent to Lemma 2 in Chapter 2, since when  $\delta = 1$  the definitions of  $K_1$  and  $K_2^{(2)}$  are the same and therefore so are  $K_1 K_2^{(2)} \delta$  in Section 3.3 and  $K_1 K_2^{(2)}$  in Chapter 2.

### 3.5.2 Games of Conflicting Interest

Schmidt (1991) studied reputational arguments in the characterization of equilibrium when player 2 is a patient player and the stage game is a game of conflicting interest. A game of conflicting interest with respect to player 1, was defined as a game in which the (short run) Stackelberg action of player 1 is also an action

that minmaxes player 2.

Schmidt's (1993) argument is that if player 2 is a patient player, then the fact that he becomes convinced that player 1 will play the Stackelberg action at the current stage does not imply that he will play a short run best response to it, since he might believe that in the continuation of the game he will be minmaxed if he does play a short run best response to it. In other words, player 2 might never play a best response to the Stackelberg strategy and his estimate of the probability of being minmaxed if he does play a best response can stay bounded away from zero.

This argument was then shown to fail in games of conflicting interest with respect to player 1, since never playing a best response to the Stackelberg strategy gives player 2 a payoff<sup>®</sup> which is lower than the minmax payoff<sup>®</sup>.

As in the previous subsection, assuming that the stage game has conflicting interest with respect to player 1, implies that  $\hat{U}_1^\alpha(\pm_2)$  is equal to the (short run) Stackelberg payoff<sup>®</sup>, so that the limit result of Section 3.3 coincides with the limit of Schmidt's (1993) result as  $\alpha \rightarrow 1$ .

Finally, if player 2 is sufficiently patient, Theorem 3 provides a stronger result than the one in Schmidt (1993) since the result holds in a wider class of games than the games of conflicting interest with respect to player 1<sup>6</sup>, and, in particular, given that  $\hat{U}_1^\alpha$  is greater than or equal to the static Stackelberg payoff<sup>®</sup>.

## 3.6 Conclusion

In a repeated game between two patient players a player whose type is not common knowledge has been shown to be able to exploit the possibility of establishing a reputation for a (possibly) history dependent pure strategy if a type that is

---

<sup>6</sup>Apart from the perturbation on the outcome, games of conflicting interest are a strict subset of games satisfying Assumption 6.

committed to such strategy exists with strictly positive probability, and if there is even a small amount of noise in the observation of the other player that is such that all finite length histories can occur with strictly positive probability.

The results presented in this Chapter show that the implications of reputational arguments change dramatically when the actions of the player whose type is common knowledge are only imperfectly observed. In fact the introduction of a small amount of imperfect observability has been shown to be sufficient to extend previous results for games with two long run players (Schmidt, 1993) to a wider class of games, as well as to explicitly provide an even tighter characterization of Bayesian Nash equilibrium for an arbitrarily patient player playing against a sufficiently patient opponent.

Interesting applications of the results presented include the case in which actions are observed but players tremble, as well as a repeated principal agent problem in which the principal (player 1) cannot observe the action (effort) chosen by the agent (player 2). In the latter case, the possibility for the principal of establishing a reputation for making payments which are made contingent not only on the observed value of output, but also on the past history of the game, leads to the conclusion that if both the principal and the agent are sufficiently patient, then the principal will be able to get an average payoff which is equal to the net value of the output under the efficient effort level.

The framework presented here can be straightforwardly generalized to deal with the case in which also player 1's action is imperfectly observed, using the result on statistical inference introduced for this case by Fudenberg and Levine (1991). More importantly, the framework of Fudenberg and Levine (1991) can be used to introduce the possibility of establishing a reputation for a mixed strategy since this can increase player 1's Bayesian Nash equilibrium payoff lower bound

if player 1 can more successfully punish player 2 by using a mixed strategy than by using a pure strategy. In such cases, if the two players are sufficiently patient, introducing the possibility of establishing a reputation for a mixed strategy would actually provide a tight bound on player 1's equilibrium payoffs.

## Chapter 4

# Reputation in Dynamic Games



## 4.1 Introduction

Many economic problems have the feature that a state variable such as capital, debt or money, provides a link between present actions and future payoff opportunities. As an example, games that describe the strategic interaction between a government and households usually involve state variables. It is in this context that the problem of time consistency of optimal government policy arises: since ex-ante and ex-post optimal policies differ, even a benevolent government may not be able to achieve the optimal commitment outcome.

Recent work has turned attention to this kind of games. Dutta (1991) provides a Folk theorem for stochastic games. Chari and Kehoe (1990) and Stokey (1992) study the time inconsistency problem introduced by Kydland and Prescott (1977, 1980) and Fischer (1980) and characterize the set of equilibria in problems of optimal policy design when the government cannot commit. Both Chari and Kehoe (1990) and Stokey (1992) show that, if there is sufficiently little discounting, a desirable outcome (the Ramsey outcome in a capital taxation model, Ramsey (1927)), can arise in equilibrium. However, in their model the Ramsey outcome is only one of many equilibria.<sup>1</sup>

We consider a general class of dynamic games with one large player and a large number of small players. A deterministic transition law describes the evolution of the state variable. The large player has some private information about his type, i.e. the small players are uncertain about the type of large player they are facing. This uncertainty may be very small in the sense that the large player is of one particular type with a probability close to one. The goal of this Chapter is to find conditions under which a patient large player can exploit the uncertainty

---

<sup>1</sup>It is sometimes argued that in this case the government may be able to select its preferred equilibrium. However, Dekel and Farrell (1990) show that these selection arguments are inconsistent.

of his opponents and enforce an outcome that is essentially equivalent to publicly committing to an optimal strategy.

The introduction of uncertainty relative to the type of a player and the consequent possibility of acquiring a reputation for an appropriate behavior has received considerable attention in the literature. Starting with the work of Kreps and Wilson (1982) and Milgrom and Roberts (1982) the studies of reputation effects have focused exclusively on repeated games.

Fudenberg and Levine (1989) study a class of repeated games in which a long lived player faces a sequence of short lived opponents, each of whom plays only once but observes the entire history of the game. If there is a positive probability that the long lived player is a type who always plays the strategy to which the normal player would like to commit, then reputation effects lead to a sharp prediction for all Nash equilibria of the game: the large player will receive a payoff that is at least as large as what he would receive if he could publicly commit to his preferred strategy.

This result is robust in the sense that it does not rely on a refinement of Nash equilibrium and that it is unaffected by further perturbations of the information structure of the game, i.e. by the introduction of additional commitment types<sup>2</sup>.

The present Chapter uses reputational arguments and provides conditions under which results analogous to the ones obtained by Fudenberg and Levine (1989) apply to dynamic games and also provides conditions under which reputational arguments may fail.

Since we consider games with a large number (continuum) of small players, we will assume that the individual play of the small players is not observed. In a purely repeated game this assumption would imply that each small player behaves

---

<sup>2</sup>For extensions of Fudenberg and Levine (1989) see Fudenberg and Levine (1992), Schmidt (1993) and Cripps and Thomas (1992).

like a short-lived player, since his actions will affect neither his future payoffs nor the public history of the game. In a dynamic game the presence of state variables creates an intertemporal link and introduces a new strategic dimension to the problem. Even though a small player cannot influence his opponent's future play, he can change the value of his individual state variable, thereby affecting his own future payoff opportunities. Therefore a small player's behavior will depend on the (expected) future actions of the large player <sup>3</sup>.

For example, in a capital taxation model in order to choose a high investment level today, the households in the economy need to become convinced that the government will set low capital tax rates not only today but also in the future.

As is clear, the presence of state variables makes it more difficult for the large player to establish a reputation: small players have to become convinced that the large player will follow a particular strategy not only in the current period but also in the future. The more the small players' behavior is affected by play in the distant future the harder it will be for the large player to gain from establishing a reputation.

Our first result (Theorem 4) applies to the case where the small players have a fixed discount factor while the large player is arbitrarily patient. If there is a commitment type that plays the strategy to which the large player would want to commit then in any Nash equilibrium the large player is guaranteed at least the optimal commitment payoff<sup>4</sup>.

To obtain a result that holds for a wide range of interesting economic applications we allow the payoffs to the small players to depend on the aggregate play of the small players and the aggregate state variable, as well as on their own play

---

<sup>3</sup>See also Schmidt (1993), for a similar effect in games with two long run players.

<sup>4</sup>By the optimal commitment payoff, we mean the maximal time average that the large player could guarantee himself by publicly precommitting if the game started in the worst possible state from the large player's point of view.

and the play of the large player. In the terminology to be introduced, we allow for strategic externalities among small players. This has the surprising implication that, for a fixed discount factor, arbitrarily distant play of the large player may affect current behavior of the small players (see Example 2). If the optimal commitment strategy can be approximated by an eventually periodic sequence, i.e. a sequence that converges to some cycle of bounded length in finitely many periods, then also in this case reputation effects allow a precise characterization of equilibria. Assuming that the discount factor of the small players stays fixed while the large player gets arbitrarily patient, the large player will receive at least the optimal commitment payoff in all Nash equilibria (Theorem 5).

Finally, we consider the case where both the large and the small players are arbitrarily patient. This case is particularly relevant for policy games, in which, for example, the payoff function of the government is equal to the payoff function of the median voter. Then, if players are very patient, the small players' action may be affected by very distant future outcomes.

In this case it is shown (Theorem 6) that the large player will only be able to exploit his reputation if the following reversibility condition on the transition function is satisfied. A transition function is reversible if players can move from one state to another only if they can also return. This condition is satisfied in capital accumulation games, but is not satisfied, for example, in the standard durable goods monopoly<sup>5</sup>. Once a customer has purchased the durable good, he has reached an irreversible state. Example 3 shows how in the durable goods monopoly reputational arguments fail to guarantee the large player his optimal commitment payoff.

---

<sup>5</sup>See for example Coase (1972), Ausubel and Deneckere (1989), Stokey (1981), Fudenberg, Levine and Tirole (1985), Gul, Sonnenschein and Wilson (1986).

An interesting application of the case in which large and small players have the same discount factor is the classical time inconsistency problem in an intertemporal capital taxation model (Kydland and Prescott, 1977). Fischer (1980) describes a situation in which a benevolent government has to finance a public good by levying taxes on capital and labor. If the government could commit to a certain strategy, it could achieve the Ramsey outcome (Ramsey, 1927), i.e. the sequence of combinations of capital and labor tax rates that minimize distortions. Once capital has been accumulated, however, it is optimal for the government to raise as much revenue as possible from capital taxation that is ex-post non distortionary. If private investors expect the government to renege on its promise to set low capital tax rates, they will accumulate a suboptimal level of capital.

The result of the present Chapter is that if the prior probability of a particular commitment type is strictly positive, then in any Nash equilibrium the government will achieve a payoff close to the payoff corresponding to credible commitment to an optimal tax rate.

The structure of the Chapter is as follows. In section 4.2 we describe the complete information game. Section 4.3 introduces the perturbed game, i.e. the possibility of the large player to be one of many "types". Section 4.4 provides the first result for the case where the discount factor of the small players stays fixed while the large player is very patient. Section 4.5 gives the result for the case in which there are strategic externalities. Section 4.6 deals with the case where large and small players share a common discount factor and Section 4.7 provides conclusions. Proofs are presented in Section 4.8.

## 4.2 Description of the Game

The class of games we consider has two types of players: one large player denoted by  $b$ , and a continuum of identical small players  $i \in [0; 1] = I$ . The finite sets  $Y$  and  $X$  denote the actions of the large player and the small players respectively;  $y \in Y, x \in X$ . Furthermore we let  $\mathcal{S}$  denote the set of mixed actions of the large player.

Each small player has an individual state variable,  $z \in Z$ , where  $Z$  is the state space that is assumed to be finite and identical for all small players. Let  $\mathcal{P}$  denote the set of probability measure on  $Z, \mu \in \mathcal{P}$ , and  $M$  denote the set of probability measures on  $Z \times X, \nu \in M; \nu(z; x)$  is to be interpreted as the measure of small players with initial value of the state variable equal to  $z$  that choose action  $x$ . Finally let  $\nu^Z \in \mathcal{P}$  denote the marginal of  $\nu$  on  $Z$  and  $\nu^X$  the marginal of  $\nu$  on  $X$  ( $\nu^X$  belongs to the set of probability measures on  $X$ ).

The game is played in the following way: At the beginning of each period  $t = 1; 2; \dots$  the public history (to be described below) is observed by all players and each small player observes his own private history. Conditional on these observations, every small player takes an action  $x^i \in X$  and the large player simultaneously takes a (possibly mixed) action  $y_t \in \mathcal{S}$ , where  $\mathcal{S}$  denotes the set of probability distributions on  $Y$ .

After these actions have been selected, payoffs occur and all players observe the realization of the action of the large player  $y_t$  and the distribution  $\nu_t$  of actions of the small players. Note that this is a joint distribution over actions and states, i.e. after each period every player knows which proportion of players in state  $z$  played action  $x$ , for every  $z \in Z$ . Clearly this joint distribution has to be consistent with the state in the beginning of the period  $\mu_t \in \mathcal{P}$ , i.e.  $\nu_t^Z = \mu_t$ .

The law of motion for the individual state is described by the following func-

tion:

$$f : Y \times X \times Z \rightarrow Z$$

i.e.  $z_{t+1}^i = f(y_t; x_t^i; z_t^i)$ . In other words we assume that the value of the individual state variable at date  $t + 1$  does not depend on the aggregate distribution of the state variable or on the aggregate action played by the small players.

The aggregate law of motion is described by:

$$F : Y \times M \rightarrow \mathfrak{A}$$

where

$$F(y_t; \mu_t)(z) = \int_{f(x; z^0) = z} \mu_t(x; z^0) dx$$

Note that  $F$  is continuous.

Let the distance between  $\mu_t$  and  $\mu_t^0$  be defined as

$$j(\mu_t; \mu_t^0) = \int_{Z \times X} |j(\mu_t(x; z); \mu_t^0(x; z))| dx$$

and let

$$j_s(\mu_s; \mu_s^0) = \int_Z |j_s(\mu_s(z); \mu_s^0(z))| dz$$

A public history of the game at time  $t$  is the sequence of realizations of  $y_{t^0}; t^0 = 1; \dots; t-1, \mu_{t^0}; t^0 = 1; \dots; t-1$  and the aggregate state in period  $t, \mu_t$ . Since we will want to use a recursive definition of histories we also include  $\mu_{s^0}; s^0 = 1; \dots; t-1$  in the history at time  $t$ <sup>6</sup>. The set of histories in period  $t$  is denoted by  $H_t = (Y \times M \times \mathfrak{A})^{t-1} \times \mathfrak{A}$ , with  $h_t \in H_t$ ;  $h_1 = \mu_{s^0}$  and  $h_t = (h_{t-1}; y_{t-1}; \mu_{t-1}; \mu_s)$  for  $t > 1$ ;  $H = H_1$ . For the history from  $t^0$  to  $t, t^0 \leq t$  we write  $h_t; h_{t^0} \in H_{t^0}$ .

For a given sequence of play  $(y; \mu) = ((y_1; y_2; \dots); (\mu_1; \mu_2; \dots))$  the payoff<sup>®</sup> to the large player is:

$$V^b(\cdot; y; \mu) = \prod_{t=1}^{\infty} \beta^{-t} v^b(y_t; \mu_t)$$

<sup>6</sup>Notice that given the transition law  $\mu_s$  is determined by  $\mu_1; y_1; \dots; \mu_{t-1}; y_{t-1}$ .

Similarly for a given sequence  $(y; \pi; x; z) = ((y_t; \pi_t; x_t; z_t)_{t=1}^T)$  the payoff to a small player is:

$$V(\pi; y; \pi; x; z) = (1 - \delta)^{T-1} \sum_{t=1}^T \delta^{t-1} v(y_t; \pi_t; x_t; z_t)$$

Since the small players' payoffs depend on the individual state variable this formulation includes the case in which there is a finite number of different types of small players.

**Assumption 7**  $v^b$  and  $v$  are continuous on  $M$ . Moreover  $0 < v; v^b < \bar{v}$ .

A pure strategy for the large player is a mapping  $y_t : H_t \rightarrow Y$ ; a mixed (behavioral) strategy is a mapping  $\pi_t : H_t \rightarrow \Pi$ . Similarly a strategy for a small player is a mapping<sup>7</sup>  $x_t : H_t \times Z \rightarrow X$ . An aggregate strategy for the small players is a mapping  $\pi_t : H_t \rightarrow M$  that satisfies the following consistency requirement: For  $h_t = (h_{t-1}; \pi_{t-1}; y_{t-1}; z_t)$ , we have  $[\pi_t(h_t)]^Z = z_t$ . In other words, for every history the marginal distribution of  $\pi_t(h_t)$  over states has to coincide with the current state  $z_t$ . Finally,  $\pi = (\pi_1; \dots; \pi_T; \dots)$ ,  $x = (x_1; \dots; x_T; \dots)$ , and  $\pi = (\pi_1; \dots; \pi_T; \dots)$ .

In an abuse of notation we will often write  $V(\pi; \pi; \pi; x; h_t; z_t)$  as the expected payoff to a small player from playing  $x$ , starting at state  $z_t$  after history  $h_t$ . Similarly  $V^b(\pi; \pi; \pi; h_t)$  is the expected payoff to the large player after history  $h_t$ .

### 4.2.1 Best Response and Aggregate Best Response

For a given strategy of the large player and a given aggregate strategy for the small players (which no individual small player can influence) we define an <sup>2</sup> best response as follows:

---

<sup>7</sup>Given that private histories are unobservable, we assume that small players do not condition their play on their private history.



**Definition 3** ( $\epsilon^2$  Best Responses) The strategy  $x = (x_t)_{t=1}^1$ , is an  $\epsilon^2$  best response for player  $i$  to  $(\beta; 1)$  if for all  $h_t \in H_t; t = 1; \dots$ , such that  $h_t$  is a public history that is reached with strictly positive probability and for all  $z \in Z$ ,  $V(\pm; \beta; 1; x; h_t; z) \geq V(\pm; \beta; 1; x^0; h_t; z) - \epsilon^2$ , for all  $x^0$ . Let  $B^2(\beta; 1; \epsilon; z)$  denote the set of best responses given  $\beta, 1$ , and initial state  $\epsilon, z$ .

Let  $B_t^2(\beta; 1; \epsilon; h_t; z)$  be the  $\epsilon^2$  best response in period  $t$  only, i.e. :

$$B_t^2(\beta; 1; \epsilon; h_t; z) = \{x \in X \mid x(h_t; z) = x; \text{ for some } x \in B^2(\beta; 1; \epsilon; z)\}$$

Note that for  $\epsilon = 0$  we have the conventional best response.

**Definition 4** (Aggregate  $\epsilon^2$  Best Response) The aggregate strategy  $1 = (1_t)_{t=1}^1$ , is an aggregate  $\epsilon^2$  best response to  $\beta$  for initial state  $\epsilon$ , if for all  $h_t$  that are reached with strictly positive probability there is a  $1$  with  $\sum_j 1_j(h_t) < \epsilon^2$  such that  $x \in B_t^2(1; \beta; h_t; z)$ , for all  $(x; z) \in \text{supp}^1$ . Let  $E^2(\beta; \epsilon)$  denote the set of aggregate  $\epsilon^2$  best responses to  $\beta$  for initial state  $\epsilon$ .

Therefore an aggregate  $\epsilon^2$  best response to a strategy  $\beta$  of the large player is an aggregate strategy  $1$  such that almost all individual strategies in its support are an  $\epsilon^2$  best response to  $1$  and  $\beta$  for all reached histories.

Finally let  $E_t^2(\beta; h_t)$  be defined as the aggregate  $\epsilon^2$  best response in period  $t$  only, given a history  $h_t$ :

$$E_t^2(\beta; h_t) = \{1 \in M \mid \sum_j 1_j(h_t) = 1 \text{ for some } 1 \in E^2(\beta)\}$$

When  $t = 1, h_1 = \epsilon$ , therefore we will write  $E_1^2(\beta; \epsilon)$  instead of  $E_1^2(\beta; h_1)$ .

For  $\epsilon = 0$  we get the usual best response and aggregate best response. Let  $B; B_t; E; E_t$  denote the best response and aggregate best response for  $\epsilon = 0$ .

Note that according to Definition 2, all small players may be able to gain  $\epsilon^2$  every period if an aggregate  $\epsilon^2$  best response is played. Thus for an aggregate  $\epsilon^2$

best response to be close to an aggregate best response  $\epsilon^2$  has to be small relative to the discount factor  $\delta$  since  $\epsilon^2 = (1 - \delta)^2$  measures the maximum utility loss for a typical small player over the course of the game. While for a fixed discount factor an  $\epsilon^2$  can be chosen such that  $\epsilon^2 = (1 - \delta)^2$  is very small, when we will consider the case where the discount factor of the small players is arbitrarily close to one (Section 4.6), we will need to use a stronger notion of aggregate  $\epsilon^2$  best response.

### 4.3 The Perturbed Game

Now we consider a slight variation of the game defined above. Suppose that the small players are not completely sure about the large player's payoff function and in particular that they believe that with positive probability the large player's payoff function is different from the one described in the previous section. Let  $\Theta$  be the space of potential types with generic element  $\theta$ . Then the large player's payoff function will also depend on his type,  $V^b(\cdot; y; \theta; h_t)$ . Let  $\theta_0$  denote the event that the type of the large player is such that his payoff function is like in the unperturbed game, i.e.  $V^b(\cdot; y; \theta_0; h_t) = V^b(\cdot; y; \theta; h_t)$ . In the following we will call type  $\theta_0$  the rational or normal player.

Types other than  $\theta_0$  may have a possibly history dependent payoff function that makes a given pure strategy dominant in the infinite game. Such players will be called commitment players and for the sake of simplicity will be identified by the strategy they play rather than by their payoff function.

The existence of these commitment types captures uncertainty of the small players about the type of large player they are facing. The idea is that although the small players are almost certain they face the rational type, they cannot exclude the possibility that the large player perceives the game in a different way and hence will behave "irrationally". To account for the possibility that the large

player can be of different types we have to modify the definition of a strategy for the large player. A mixed (behavioral) strategy for the large player is now a mapping  $\sigma_t : H_t \rightarrow \Sigma$ .

Since the small players cannot observe the type of the large players, the definition of a strategy for the small players remains unchanged. We assume that the prior distribution of types is common knowledge. By  $\sigma^0(\theta)$  we denote the strategy that is obtained by substituting  $\sigma^0(\theta)$  for  $\sigma(\theta)$  in  $\sigma$ .

**Definition 5** A Nash equilibrium for initial state  $s_0$  is a  $(\sigma; \pi)$  with  $\pi_1^Z = s_0$ , such that  $\pi \in E(\sigma; s_0)$ , and for all  $\theta \in \Theta$ ,  $V^b(\cdot; \sigma; \pi; \theta; s_0) \geq V^b(\cdot; \sigma^0(\theta); \pi; \theta; s_0)$  for all  $\sigma^0(\theta)$ .

First we want to investigate the consequences of imitating a particular commitment type on the beliefs of the small players. In Lemma 8 we show that if the large player chooses to imitate a pure strategy of a particular commitment type, then in all but finitely many periods the small players will actually believe that with high probability the aggregate play will be consistent with this strategy being played in the next  $\ell$  periods. Both the formulation and the proof of Lemma 8 are an extension of a result in Fudenberg and Levine (1989).

Let  $y^\pi$  denote the pure strategy played by a particular commitment type  $\theta^\pi$ . Let  $h^\pi$  be the event that  $y_t = y_t^\pi(h_t)$  for all  $h_t$  that are reached following  $(y^\pi; \pi)$  starting from a given  $h_0 = s_0$ . Furthermore let  $p(\theta^\pi) = p^\pi$  denote the prior probability that  $\theta = \theta^\pi$ . Let  $\mu_t^{\pi, \ell}$  be the probability that in the next  $\ell$  periods the actions of the large player are consistent with  $y^\pi$ , i.e. the probability that in periods  $t; t+1; \dots; t+\ell-1$  aggregate play is consistent with  $y^\pi$  being played given the aggregate strategy  $\pi$ , i.e.  $\mu_t^{\pi, \ell} = \Pr[y_t = y_t^\pi(h_t); \dots; y_{t+\ell-1} = y_{t+\ell-1}^\pi(h_{t+\ell-1}) | h_{t-1}; \pi]$ . Finally, let  $n(\mu_t^{\pi, \ell} \cdot \mathbb{N})$  be the random variable denoting the number of periods in which  $\mu_t^{\pi, \ell} \cdot \mathbb{N}$ .

Lemma 8 Let  $0 < \eta < 1$  and suppose that  $p^a > 0$ , and that  $(\frac{3}{4}; 1)$  are such that  $\Pr(h^a | j^a) = 1$ . Then

$$\Pr \left( \sum_{t=1}^{\bar{A}} \eta^t \cdot \eta > \epsilon \frac{\log p^a}{\log \eta} j^a \right) = 0:$$

Remark: Note that since certain states may not be reached along a given history  $h^a$ , the small players will not get convinced that the large player actually uses the same strategy as the commitment type. However, since no individual small player can affect the aggregate state, the play in public histories that are not reached is irrelevant for any small player's decision problem.

## 4.4 The Case with No Strategic Externality

In this section we consider the simpler case in which the payoff of every small player is independent of the aggregate play of the small players.

Assumption 8 (No Strategic Externality)  $v$  is independent of  $1$ .

The following equation defines  $\bar{V}^b$  to be the limit of time averages of the large player's payoff. Since time averages need not converge we will take the limit inimum. Let

$$\bar{V}^b(y; 1; \eta) = \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T v^b(y_t; 1_t)$$

where  $(y_t; 1_t)$  is the history induced by  $y$  and  $1$  and  $\eta$ .

Define a strategy  $y$  such that  $y_t(h_t) = y_t$  for all  $h_t$  a simple strategy. A simple strategy is a strategy that does not depend on history but only on calendar time. With an abuse of notation in the following we will sometime identify a simple strategy with the infinite sequence of actions it prescribes,  $y = y$ .

Let  $\hat{V}^b(2; s)$  be the best time average the large player could guarantee to himself by committing to a given simple strategy subject to the small players playing an aggregate  $2$  best response. Then

$$\hat{V}^b(2; s) = \sup_{y \in Y^1} \inf_{1 \in 2E^2(y; s)} \hat{V}^b(y; 1; s)$$

Let  $\hat{V}^b(2) = \inf_s \hat{V}^b(2; s)$  and let  $\hat{V}^b = \lim_{\delta \rightarrow 0} \hat{V}^b(2)$ .

Now we define a collection of types (the Stackelberg types) which can be used by the rational large player to establish a reputation. Let  $y(2; \hat{s}; s) = (y_1(2; \hat{s}; s); \dots; y_t(2; \hat{s}; s); \dots)$  be a simple strategy that satisfies

$$\inf_{1 \in 2E^2(y(2; \hat{s}; s); s)} \hat{V}^b(y(2; \hat{s}; s); 1; s) \geq \hat{V}^b(2; s) - \delta$$

Hence  $y(2; \hat{s}; s)$  is an "almost" optimal sequence if the criterion is the limit of time averages.

The type  $! (2; \hat{s}; T)$  is defined by the following strategy:

$2$  In the first  $T$  periods this type follows  $y(2; \hat{s}; s_1)$ .

$2$  In case the small players reacted with an  $2$  best response in period  $1$ , this type continues with  $y(2; \hat{s}; s_1)$  in period  $T + 1$ . If the small players did not choose an action close to a best response in period  $1$ , this commitment type switches to  $y(2; \hat{s}; s_{T+1})$ .

$2$  The same pattern is repeated for all periods: The commitment type will continue following the sequence  $y(2; \hat{s}; s)$  if either it has been played for fewer than  $T$  periods or if  $T$  periods ago "almost" a best response was played. Otherwise a new sequence  $y(2; \hat{s}; s^0)$  will be started, where  $s^0$  is the current value of the state variable.

More precisely, let  $\mu_t \in \mathbb{N} \ t = 1, 2, \dots$  be defined as follows:

$$\mu_t = \begin{cases} 1 & \text{if } t = 1 \\ \mu_{t-1} + 1 & \text{if } \mu_{t-1} < T \text{ or if } t_{i-T} \in E_{t_i-T}^2(y(\beta; \gamma; s_{t_i-\mu_t}); h_{t_i-T}) \\ 1 & \text{otherwise} \end{cases}$$

Now let the type  $(\beta; \gamma; T)$  be defined by following strategy:

$$y_t(h_t) = y_{\mu_t}(\beta; \gamma; s_{t_i-\mu_t})$$

Note that the  $T$ -period lag in adjusting the optimal policy in the definition of the commitment types is crucial to avoid the time-inconsistency problem. A commitment type who chooses the optimal policy for the current state in every period is of little use to the large player since he wants to commit to ex-ante rather than ex-post optimal policies.

**Assumption 9** For all  $(\beta; \gamma)$  there is an  $\beta^0 < \beta; \gamma^0 < \gamma$  such that  $(\beta^0; \gamma^0; T) \in \Sigma$  has strictly positive prior probability for all finite  $T$ .

Assumption 9 says that there is a large variety of the described Stackelberg types. In particular, for arbitrarily small  $(\beta; \gamma)$  we can find a commitment type with positive prior for any finite lag parameter  $T$ .

In Theorem 4 we make two important assumptions that will be relaxed later. First we assume that  $v$  is independent of  $\beta$ , i.e. there is no "strategic externality" in the play of the small players. Second we assume that the small players' discount factor stays fixed while the large player's discount factor approaches 1. Theorem 4 states that if the type space includes a particular collection of commitment types then as the discount factor of the large player goes to 1 in any Nash equilibrium he gets at least  $V^b$ .

**Theorem 4** Suppose that Assumptions 7, 8 and 9 hold. Then in any Nash equilibrium  $(\beta; \gamma; 1)$  for initial state  $s$ ,  $\lim_{\beta \rightarrow 1} V(\beta; \gamma; 1; s) \geq V^b$ .

The intuition behind Theorem 4 is the following: Suppose that the large player imitates a commitment type  $(\beta; \hat{\gamma}; T)$ . Then Lemma 8 implies that after a finite number of periods the small players will actually believe that the large player will continue to play like the commitment type in the next  $\hat{T}$  periods with very high probability. Since the small players discount future payoffs there is a  $\hat{T}$  large enough such that if the commitment strategy is followed with large probability in the next  $\hat{T}$  periods then the small players will actually play a best response to the commitment strategy. Since we can find commitment types with positive prior probability for arbitrarily large lag parameter  $T$ , we can choose  $T$  in such a way that  $T > \hat{T}$ . In this case an aggregate best response to the commitment strategy implies an aggregate best response to an optimal sequence  $y(\beta; \hat{\gamma}; \delta)$ . The Theorem then follows from the fact that this argument can be repeated for arbitrarily small  $(\beta; \hat{\gamma})$ .

## 4.5 Including Strategic Externalities

In many economic problems the utility of individuals depends on their individual choice as well as on the aggregate behavior of the other individuals, for example through prices in a market game or through the aggregate level of capital in a capital accumulation problem with externalities. In this section we discuss reputational arguments in the general case in which the payoffs to individual small players may also depend on the aggregate play of the small players.

Allowing for this possibility complicates the analysis for the following reason: Even though small players discount future payoffs at a fixed rate  $\delta < 1$ , it is not true that the aggregate best response today does not change when the large player's action in the very distant future changes.

The large player can only exploit his reputation successfully if the small players

choose an aggregate  $\beta$  best response to the commitment strategy whenever the large player imitates this strategy long enough. This implies that we need to find a (uniform) bound  $T$  such that if the small players believe that the commitment strategy is played for the next  $T$  periods, then they will actually play an aggregate  $\beta$  best response to it. When  $v$  is independent of  $\beta$  discounting implies that we can find such a  $T$  uniformly over all strategies. If  $v$  depends on  $\beta$ , this property fails. The following example illustrates this point.

**Example 2** Consider an economy in which there is a continuum of private agents (small players) and a government (large player). Suppose that private agents can choose between becoming specialized or staying autarkic. Then the strategy space for each private agent is  $X = \{0, 1\}$  where  $x = 0$  symbolizes autarky and  $x = 1$  specialization. Any agent can either be in an experienced state,  $z = 1$ , or in an inexperienced state,  $z = 0$ . Experience is obtained after having specialized for one period. If an experienced agent fails to specialize, then he loses his experience. Hence the individual state variable transition can be summarized by:  $f(x; z) = x$ ,  $x = 0; 1$ .

Only experienced agents who specialize are productive. However, their payoff from specialization depends on how many other agents decide to specialize in the current period (irrespective of whether these agents are experienced or not). Let  $\beta^x(1)$  be the fraction of agents who specialize in the current period, then the value of the output produced by an agent who plays  $x$  and is in state  $z$  is:  $\beta^x(1)z x - cx$  where  $c$  is the cost of specialization.

The government has two policies: It can either do nothing ( $y = 0$ ) or it can reward all the experienced specializing agents by giving them a subsidy of 1 for each unit they produce ( $y = 1$ ). With this set-up the payoff function of a private



agent will be

$$v(1; y; x; z) = (1^x(1) + y)z - c$$

The government is benevolent but giving a subsidy is costly. Let  $1^x(1; 1)$  denote the proportion of experienced private agents (private agents in state  $z = 1$ ) who decide to specialize ( $x = 1$ ). Then the government's payoff function can be written as

$$v^b(y; 1) = 1^x(1; 1)(1 + (1 - k)y) - c$$

where  $k > 1$  is the unit cost of raising funds to pay the subsidies.

The following table summarizes the payoffs to the private agents. The column entries are combinations of actions and individual values of the state variable of the small player ( $z; x$ ); the row entries denote actions of the large player.

	(0; 0)	(0; 1)	(1; 0)	(1; 1)
$y = 0$	0	$-c$	0	$1^x_t(1) - c$
$y = 1$	0	$-c$	0	$1^x_t(1) + 1 - c$

Let  $c < 1 < k$ . Under policy  $y = 0$  the private agents will specialize ( $x = 1$ ) only if enough other small players specialize. Under policy  $y = 1$  there is a reward for experienced agents who specialize.

The government would like to play policy  $y = 0$  in every period and would like all small players to specialize (choose action 1). However this is not the only aggregate best response to  $y = (0; 0; \dots; 0; \dots)$ . Any sequence of the form:

$$1^x_t(1) = \begin{cases} 1 & \text{if } t \leq \bar{t} \\ 0 & \text{if } t > \bar{t} \end{cases}$$

for  $\bar{t} \geq 0$  is an aggregate best response. In particular  $1^x_t(1) = 0$  for all  $t$  is the worst aggregate best response.

Now suppose the government plays  $y = (0; \dots; 0; 1; 0; \dots; 0; 1; \dots)$ , where the sequence of consecutive 0's is arbitrarily large. Whenever the government gives

a subsidy ( $y = 1$  is played) in period  $i$  every small player wants to specialize ( $x_i = 1$ ) and be experienced ( $z_i = 1$ ). But this implies that in period  $i - 1$  every small player has to specialize ( $x_{i-1} = 1$ ), otherwise he would not be experienced in the following period. This in turn implies that also in period  $i - 1$  every small player can benefit from specialization ( $x_{i-1} = 1$ ) as long as he is experienced ( $z_{i-1} = 1$ ). But to be experienced in period  $i - 1$  ( $z_{i-1} = 1$ ) he has to specialize in  $i - 2$  ( $x_{i-2} = 1$ ) and so on.

Thus every small player will choose  $x_t = 1$  for  $t \geq i$ , which implies that the unique equilibrium is  $x_t^* = 1$ , for all  $t$ . In order to guarantee that the private agents will actually specialize ( $x = 1$ ) the large player has to give a subsidy (switch to policy  $y = 1$ ) every once in a while. <sup>2</sup>

Theorem 4 relied on the fact that we could find a uniform bound  $T$  such that the large player's actions more than  $T$  periods from now did not affect the small players' current behavior. The previous example shows that in the case with strategic externalities such a uniform bound does not exist<sup>8</sup>. This creates a problem for a large player who tries to exploit his reputations: the small players may have to be convinced that the large player follows a given strategy for very many future periods.

As an illustration, consider again Example 2. Suppose that the large player wants to establish a reputation for playing the sequence  $y = (A; B; A; A; B; A; A; A; B; \dots)$ . Clearly  $x_t^*(y) = 1$  is the unique aggregate best response to  $y$ . However, to ensure that this best response is played in period  $t$ ,  $\frac{1}{4} y_t^{T_t}$ , with  $T_t \rightarrow \infty$  has to be sufficiently large. If  $T_t$  goes to infinity very fast then the large player may actually never be able to establish a sufficiently "far-reaching" reputation so that the small

---

<sup>8</sup>In other words: the aggregate best response fails to be lower hemi continuous in the product topology.

players will play  $x_t^1(1) = 1$ .

To circumvent this problem we will assume that by committing to an "eventually periodic" sequence the large player can do almost as well as by committing to an arbitrary sequence. This allows us to restrict the Stackelberg type to a set of strategies for which we can find a uniform bound on the number of future periods that matter for the current behavior of the small players.

Recall that a pure strategy  $y$  for  $b$  is called simple, if  $y_t = y_{t+k}$ , for some  $y \in Y^1$ ; i.e. no matter what history is reached in period  $t$ , player  $b$  chooses  $y_t$  in period  $t+k$ . Define a simple strategy  $L$  periodic if for some  $L; k \in \mathbb{N}$ ,  $L < \infty$ , we have  $y_{t+L} = y_t$  for all  $t \geq k$ . Let  $Y(L)$  denote the set of  $L$  periodic simple strategies. The following assumption says that by committing to an  $L$  periodic sequence, the large player can guarantee himself almost the same payoff as by committing to an arbitrary sequence.

**Assumption 10** For all  $\epsilon > 0; \delta > 0$ , there is an  $L$  such that for all  $y; s$  there is a  $y^0 \in Y(L)$  such that  $\inf_{y^0 \in Y(L)} V^b(y^0; 1; s) \geq \inf_{y \in Y} V^b(y; 1; s) - \epsilon$ .

Note that Assumption 10 is satisfied in Example 2.

Commitment types (Stackelberg types) are constructed analogous to the ones in Section 4.4. The only difference is that we restrict the Stackelberg type to the use of  $L$  periodic sequences. Let  $y(2; \epsilon; s) \in Y(L)$  satisfy

$$\inf_{y(2; \epsilon; s)} V^b(y(2; \epsilon; s); 1; s) \geq V^b(2; s) - \epsilon$$

Assumption 10 guarantees the existence of such a sequence. As before we define  $\mu_t \in \mathbb{N}$   $t = 1; 2; \dots$  as follows:

$$\mu_t = \begin{cases} 1 & \text{if } t = 1 \\ \mu_{t-1} + 1 & \text{if } \mu_{t-1} < T \text{ or if } 1_{t \geq T} \in E_{t-1}^2(y(2; \epsilon; s_{t-1} \mu_{t-1}); h_{t-1}) \\ 1 & \text{otherwise} \end{cases}$$

Type  $(\beta; \mu; T)$  is committed to the strategy:

$$y_t(h_t) = y_{\mu_t}(\beta; \mu; s_t; \mu_t)$$

The interpretation of this strategy is the same as the one that was provided for the case with no strategic externality.

**Theorem 5** Suppose that Assumptions 7, 9, and 10 hold. In any Nash equilibrium  $(\beta; \mu)$  for initial state  $s$ ,  $\lim_{\beta \rightarrow 1} V^b(\beta; \mu; s) = \hat{V}^b$ .

Theorem 5 generalizes Theorem 4 to include the possibility of the small player's payoff to depend on the aggregate play of the small players. If Assumption 10 holds, then Theorem 5 says that as the discount factor of the large player goes to 1 in any equilibrium he gets at least the time average of payoffs corresponding to an optimal commitment.

## 4.6 Patient Small Players

In Theorem 4 we assumed that the discount factor of the small players stays fixed while the large player becomes arbitrarily patient. In applications like policy games the utility function of the large player frequently reflects the utility function of the small players (e.g. the large player's preferences are identical to the utility function of the "median voter"). Thus it is important to identify classes of games where reputation allows the large player to achieve essentially his commitment payoff when both the large and the small players become arbitrarily patient simultaneously.

The difficulty in establishing a reputation with patient small players lies in the fact that small players may become increasingly reluctant to take an action that leads to an irreversible state as they get more patient. Thus to convince a very

patient small player to take this action the large player may have to establish a reputation for following the commitment strategy for very many periods and hence it may take "too long" to establish a reputation that induces the small players to enter an irreversible state.

#### 4.6.1 The Failure of Reputation in the Durable Goods Monopoly

The following simple example of a monopolist selling a durable good to a population of buyers illustrates the failure of reputational arguments with irreversible states.

**Example 3** Suppose there are two types of buyers H and L. The reservation price of type H,  $r^H$ , for the durable good is 5, the reservation price of type L,  $r^L$ , is 2. There is mass 1/2 of both types of buyers. Each period the buyer takes either action 0 (he does not buy) or action 1 (he buys). Similarly the state of a buyer is either 0 (no purchase has occurred in the past) or 1 (a purchase has occurred in some previous period). Thus the transition function is defined as:

$$f(x_t; z_t) = \begin{cases} 0 & \text{if } x_t = 0 \text{ and } z_t = 0 \\ 1 & \text{otherwise} \end{cases}$$

The monopolist sets a price  $p_t$  every period, where  $p_t \in [0; 1] \times \dots \times [5; 1] \times \dots$ . If buyer  $i \in [H; L]$  purchases the durable good in period  $t$  then his payoff<sup>8</sup> is  $\delta^t(r^i - p_t)$ . More precisely, buyer  $i$ 's payoff<sup>8</sup> function is

$$v^i(p_t; x_t; z_t) = \begin{cases} \delta^t(r^i - p_t) & \text{if } z_t = 0 \text{ and } x_t = 1 \\ \delta^t p_t & \text{if } z_t = 1 \text{ and } x_t = 1 \\ 0 & \text{otherwise} \end{cases}$$

For any sequence of prices  $p$  and aggregate actions  $1$  the payoff<sup>9</sup> to the large player is<sup>9</sup>

$$V^b(p; 1) = \sum_{t=1}^{\infty} \delta^{t-1} p_t \phi_t(1; 0)$$

<sup>9</sup>We do not include  $\phi_t(1; 1)$  (the proportion of buyers that have already bought the durable good in the past that do so again) because no buyer will purchase twice in equilibrium.

Suppose there are three types of monopolists: one normal type characterized by the payoff function above, type  $H$  who sets  $p_t = (5n - 1)/n$  for all  $t$  and type  $L$  who follows the strategy:

$$p_t = \begin{cases} (5n - 1)/n & \text{if } t \leq T \\ (2n - 1)/n & \text{otherwise} \end{cases}$$

where  $\log(1/2) = \log(\delta) < T < \log(2/(3n - 1)) = \log \delta^{-1}$ . Both commitment types have prior probability  $\delta > 0$ .

The strategy of playing  $p_t = (2n - 1)/n$  (for the normal type) constitutes a sequential equilibrium for large  $\delta$ . To see this first note that  $p_t = (2n - 1)/n$  constitutes a subgame perfect Nash equilibrium in the game where there is only the normal type if  $\delta$  is sufficiently large. Thus it remains to show that the normal type does not have an incentive to imitate type  $H$ . Suppose  $b$  deviates and offers  $p_t = (5n - 1)/n$ . Since

$$\frac{\delta^T}{2\delta^2} \delta^{\pm T} \delta (5n - (2n - 1)/n) > 1/n$$

type  $H$  will not buy until period  $T + 1$ <sup>10</sup>. However

$$\delta^{\pm T} \frac{1}{2} \delta \frac{5n - 1}{n} + \delta^{\pm T+1} \frac{1}{2} \delta \frac{2n - 1}{n} < \frac{7n - 1}{4n} < \frac{2n - 1}{n}$$

for  $n \geq 6$ , where the first element of the chain of inequalities is an upper bound on the payoff to  $b$  from deviating and the last is the payoff from setting  $p_1 = (2n - 1)/n$ . This implies that deviation from  $p_1 = (2n - 1)/n$  does not pay. Thus in this game the large player is unable to exploit reputational effects to achieve the simple monopoly payoff  $(5n - 1)/2n$ .<sup>2</sup>

---

<sup>10</sup>The left hand side of the inequality is a lower bound on the expected payoff from waiting until  $T + 1$  and then buying at  $p_{T+1} = (2n - 1)/n$ , and the right hand side is the payoff from buying at  $(5n - 1)/n$  at  $t = 1$ .

## 4.6.2 No Irreversible Actions

The following Assumption says that no action that the small players can take has irreversible consequences.

**Assumption 11 (Reversibility of Accumulation Paths)** Suppose there is a sequence

$(y_1; \dots; y_N), (x_1; \dots; x_N)$  such that for  $z_1 = z$  and  $z_n = f(y_n; x_n; z_{n-1})$  we have  $z_N = z^0$ . Then for any other sequence  $(\hat{y}_1; \hat{y}_2; \dots)$  there is a sequence  $(\hat{x}_1; \dots; \hat{x}_{N^0})$  such that for  $z_1 = z^0$  and  $z_n = f(\hat{y}_n; \hat{x}_n; z_{n-1})$  we have  $z_{N^0} = z$ .

Using Assumption 11 we can partition the states  $Z$  into subsets  $Z^j$  such that the small players can only move between states in the same subset  $Z^j$  and furthermore there is an  $N$  such that for any pair  $(z; z^0)$  belonging to the same  $Z^j$  a small player can move from  $z$  to  $z^0$  in fewer than  $N$  periods (independent of  $y$ ).

Note that the definition of aggregate  $\epsilon$  best response (Definition 4) contains strategies in which every small player "loses"  $\epsilon$  units of utility as compared to a best response every period. Since  $\epsilon$  is fixed in Theorem 4, we can make  $\epsilon = (1 - \epsilon)$  arbitrarily small. (Note that  $\epsilon = (1 - \epsilon)$  denotes the overall "loss" of utility of a typical small player as compared to a best 1). Here we want to let  $\epsilon \rightarrow 1$  and therefore we need a stronger notion of aggregate  $\epsilon$  best response.

Denote by  $y^T = (y_1; \dots; y_T)$  a  $T$  period sequence of actions for the large player and similarly  $1^T = (1_1; \dots; 1_T); x^T = (x_1; \dots; x_T); z^T = (z_1; \dots; z_T)$ . Finally let  $G^T(y^T) = \{(x^T; z^T) : z_{t+1}^T = f(y_t^T; x_t^T; z_t^T)\}$  denote the set of sequences  $(x^T; z^T)$  of length  $T$  that are feasible under  $y^T$ . Now define a truncated aggregate  $\epsilon$  best response in the following way:

**Definition 6 (Truncated Aggregate  $\epsilon$  Best Responses)**  $1^T$  is a  $T$ -truncated aggregate  $\epsilon$  best response to  $y^T$  for initial state  $z$ , if  $z_{t+1}^T = F(y_t^T; 1_t^T); t = 1; \dots; T - 1$

and for all  $(x^T; z^T) \in \text{supp}^{1T} \setminus G^T(y^T)$

$$\frac{1}{T} \sum_{t=1}^T v(y_t^T; x_t^T; z_t^T) \geq \frac{1}{T} \sum_{t=1}^T v(y_t^T; x_t^{0T}; z_t^{0T})$$

for all  $(x^{0T}; z^{0T}) \in G(y^T)$  with  $z_1^{0T} = z_1^T$ . Let  $E^{T;2}(y^T; \epsilon)$  denote the set of  $T$ -truncated aggregate  $\epsilon$  best responses to  $y^T$  for initial state  $\epsilon$ .

This definition of a truncated aggregate  $\epsilon$  best response says that over the course of  $T$  periods the average payoff could not be increased by more than  $\epsilon$  by any other sequence of actions. Note that this definition requires  $\epsilon$  optimality (in a time average sense) over  $T$  periods and irrespective of the continuation of play and hence is a much stronger notion of aggregate  $\epsilon$  best response than the one used in Theorems 4 and 5.

Next we define the limit of the commitment payoffs of a sequence of truncated games. Consider a truncated game in which the large player commits to an optimal sequence and the small players choose a truncated aggregate  $\epsilon$  best response.  $\hat{V}^b$  denotes the limit of payoffs for the large player when the game is truncated farther and farther away in the future. Again, since time averages need not converge, we take the limit in  $\epsilon$ . Let

$$\hat{V}^b(\epsilon; \epsilon) = \lim_{T \rightarrow \infty} \inf_{y^T} \max_{y^T} \min_{z^T \in E^{T;2}(y^T; \epsilon)} \frac{1}{T} \sum_{t=1}^T v^b(y_t^T; z_t^T)$$

and

$$\hat{V}^b(\epsilon) = \inf_{\epsilon} \hat{V}^b(\epsilon; \epsilon)$$

$$\hat{V}^b = \lim_{\epsilon \rightarrow 0} \hat{V}^b(\epsilon)$$

Again we define a collection of commitment types who will allow the large player to establish a reputation. Let  $y^T(\epsilon; \epsilon)$  be a  $T$  period sequence that solves

$$\max_{y^T} \min_{z^T \in E^{T;2}(y^T; \epsilon)} \frac{1}{T} \sum_{t=1}^T v^b(y_t^T; z_t^T)$$



We define the type  $\sigma^T(\delta; T)$ , to play the strategy

$$y^T(\delta; T) = (y_1^T(\delta; s_1); \dots; y_T^T(\delta; s_1); y_1^T(\delta; s_{T+1}); \dots; y_T^T(\delta; s_{T+1}); \dots)$$

The commitment type  $\sigma^T(\delta; T)$  plays the optimal sequence in the  $T$ -period truncated game given the initial state at the beginning of the truncated game.

**Assumption 12** For all  $\delta > 0$  there is an  $\delta^0 < \delta$  such that  $\sigma^T(\delta^0; T)$  has strictly positive prior probability for every finite  $T$ .

Theorem 6 says that if both players are very patient and if the transition function is reversible then in any Nash equilibrium the large player will receive at least a payoff that is close to the maximal time average in the  $T$ -period truncated game for arbitrarily large  $T$ .

**Theorem 6** Suppose Assumptions 7, 8, 11, and 12 hold and all players have a common discount factor  $\delta$ . Then in any Nash equilibrium  $(\sigma; \sigma^1)$  for initial state  $s$ ,  $\lim_{\delta \rightarrow 1} V^b(\delta; \sigma; \sigma^1; s) = \hat{V}^b$ .

The idea behind the proof of Theorem 6 is that we split up the infinite game into finite "superstage games" of length  $T$ . Note that the effect of a current decision of a small player on the payoffs in future "superstage" games can be "undone" in  $N$  periods or less by the reversibility assumption. If  $N$  is small as compared to  $T$  then the small players will behave almost like short lived players in every superstage game, i.e. they will behave essentially as if they were alive only for one superstage game. Therefore the large player can exploit his reputation if he convinces the small players that he will follow the commitment strategy in the current superstage game. Thus it is sufficient for the large player to establish a reputation for a bounded number of future periods and Lemma 1 shows that this can be accomplished in finitely many periods.

## 4.7 Conclusions

Whenever current play affects future payoff opportunities, agents' current decisions depend not only on present but also on future expected behavior of their opponents.

To describe this situation an infinite dynamic game between a large player and a continuum of small players has been studied and it has been shown that the use of reputational arguments allows to characterize the set of equilibria by providing a lower bound on the equilibrium payoffs to the large player. This has been accomplished by noticing that, if there is uncertainty relative to the type of the large player, the large player can actually establish a reputation for behaving in a certain way in a finite horizon.

An example has been presented to show that when individual small players' payoffs also depend on the aggregate play of the small players, it is possible that arbitrarily distant play of the large player affects current aggregate behavior of the small players. Interestingly it turns out that even in cases like this reputational arguments do have a bite: it is argued that the large player can establish a reputation for playing repeatedly an appropriate finite sequence of actions which in turn allows him to get at least his commitment payoff.

Provided that the small players' actions do not have irreversible consequences, reputational arguments have been shown to work independently of the rate of patience of the small players. Even when the large player and the small players have the same discount factor (like in the case of a benevolent government), the fact that the large player can establish a reputation for playing a strategy that depends on the aggregate state variable only after a sufficiently long adjustment lag provides a lower bound on the large player's equilibrium payoffs.

A simple example of a durable goods monopoly problem has been presented

to illustrate the role of the assumption that the small players' actions do not have irreversible consequences: when some action profile leads to an absorbing state (purchase of the durable good in this case), then, if the small players are arbitrarily patient, the possibility of establishing a reputation may fail to improve the large player's payoff since no finite adjustment lag in the strategy of the large player would convince an arbitrarily patient small player to play a best response to the optimal commitment strategy. Only in a case like this is it possible that payoffs that are not close to the large player's optimal commitment payoff be equilibrium payoffs of the perturbed game.

## 4.8 Proofs:

### 4.8.1 Proof of Lemma 8

**Lemma 8** Let  $0 < \eta < 1$  and suppose that  $p^\pi > 0$ , and that  $(\delta; \beta)$  are such that  $\Pr(h^\pi | \pi) = 1$ . Then

$$\Pr(\eta^{\delta \ell} \cdot \eta) > \delta \frac{\log p^\pi}{\log \eta} \eta^\pi = 0$$

**Proof:** Let  $\pi^\pi$  denote the event that  $\pi \in \pi^\pi$ . Then by Bayes's law we have

$$\begin{aligned} \Pr(\pi^\pi | h_{t+1}) &= \Pr(\pi^\pi | y_t^\pi(h_t); h_t) \\ &= \frac{\Pr(\pi^\pi | h_t) \Pr(y_t^\pi(h_t) | \pi^\pi)}{\Pr(\pi^\pi | h_t) \Pr(y_t^\pi(h_t) | \pi^\pi) + (1 - \Pr(\pi^\pi | h_t)) \Pr(y_t^\pi(h_t) | \pi)} \end{aligned} \quad (4.1)$$

Notice that  $\Pr(y_t^\pi(h_t) | \pi^\pi) = 1$  and that the denominator of (4.1) is equal to  $\Pr(y_t^\pi(h_t))$ . Therefore (4.1) can be rewritten as

$$\Pr(\pi^\pi | h_{t+1}) = \frac{\Pr(\pi^\pi | h_t)}{\Pr(y_t^\pi(h_t))} \quad (4.2)$$

Notice that for any  $\ell$ ,  $\Pr(y_t = y_t^\pi(h_t)) = \Pr(y_{t^0} = y_{t^0}^\pi(h_{t^0}); t^0 = t; \dots; t + \ell - 1) + \Pr(y_t^\pi(h_t); y_{t^0} \in y_{t^0}^\pi(h_{t^0}); \text{for some } t^0 = t + 1; \dots; t + \ell - 1)$ . Recall

$\frac{1}{4}^{\beta \ell} = \Pr(y_{t^0} = y_{t^0}^{\beta}(h_{t^0}); t^0 = t; \dots; t + \ell - 1)$ , and let  $\frac{1}{4}^{\beta \ell} = \Pr(y_t = y_t^{\beta}(h_t); y_{t^0} \in y_{t^0}^{\beta}(h_{t^0}); \text{ for some } t^0 = t + 1; \dots; t + \ell - 1)$ , i.e.  $\frac{1}{4}^{\beta \ell}$  is the probability that the large player's play is in accordance with  $y^{\beta}$  at time  $t$ , but differ at some point in the next  $\ell - 1$  periods. Then, for any fixed  $\ell$  (4.2) can be rewritten as

$$\Pr(!^{\beta} j h_{t+1}) = \frac{\Pr(!^{\beta} j h_t)}{\frac{1}{4}^{\beta \ell} + \frac{1}{4}^{\beta \ell}}$$

Suppose that  $\frac{1}{4}^{\beta \ell} < \frac{1}{4}$  for all  $t^0 = t; \dots; t + \ell - 1$ . Then if the large player plays like the commitment type for  $t^0 = t; \dots; t + \ell - 1$  (i.e.  $y_{t^0} = y_{t^0}^{\beta}(h_{t^0})$ ), then the probability that he is type  $!^{\beta}$  has to go up by a factor of at least  $1 - \frac{1}{4}$  (because if  $y_{t^0} = y_{t^0}^{\beta}(h_{t^0})$  all  $t^0 = t; \dots; t + \ell - 1$ , then at some  $t^0 = t; \dots; t + \ell - 1$   $\frac{1}{4}^{\beta \ell}$  will be updated to zero. Given that  $\Pr(!^{\beta} j h_1) = p^{\beta}$ , after  $\ell$  periods

$$\Pr(!^{\beta} j h_{\ell+1}) \geq p^{\beta} \frac{1}{4}$$

If  $\frac{1}{4}^{\beta \ell} < \frac{1}{4}$  for  $\ell < K$  periods during which  $y_t = y^{\beta}(h_t)$ , all  $t$ , then

$$\Pr(!^{\beta} j h_{K\ell+1}) \geq p^{\beta} \frac{1}{4}^K$$

However, since

$$\Pr(!^{\beta} j h_t) < 1 \tag{4.3}$$

if

$$p^{\beta} \frac{1}{4}^K > 1 \tag{4.4}$$

inequality (4.3) is violated and a contradiction to the hypothesis that  $\frac{1}{4}^{\beta \ell} < \frac{1}{4}$  for all  $t^0 = t; \dots; t + K \ell - 1$  is obtained.

Taking the log of (4.4) the condition becomes

$$K > \log p^{\beta} = \log \frac{1}{4}$$

and the proof is complete.  $\square$

## 4.8.2 Proof of Theorem 4

**Theorem 4** Suppose that Assumptions 7, 8 and 9 hold. Then in any Nash equilibrium  $(\sigma; \pi)$  for initial state  $s$ ,  $\lim_{\delta \rightarrow 1} \pi(\sigma; \pi; s) = \pi^b$ .

The strategy of the proof will be to show that if the large player imitates the Stackelberg type  $(\sigma; \pi; T)$  for appropriately chosen  $T$ , then eventually the small players will play a best response to the Stackelberg strategy.

In the following we present a Lemma that shows that if the small players believe that the large player follows a given sequence of actions for a sufficient number of periods with a sufficiently large probability, then the small players will play an aggregate best response to this sequence of actions.

For a pure strategy  $y$ , let  $\mu_t^{yT}$  be the probability that  $y$  is played in each of the next  $T$  periods, i.e. in the periods  $t; t+1; t+2; \dots; t+T-1$ .

**Lemma 9** Suppose Assumptions 7 and 8 hold. For every  $\epsilon > 0$  and  $T > (\log \frac{2}{\epsilon} i + \log v) / \log \delta$ , there is an  $\epsilon'$  such that for every simple strategy  $y$ , if  $\mu_t^{yT} > 1 - \epsilon'$ , then in equilibrium  $\pi_t \in E_t^2(y; s)$  for all  $s$  and all  $t$ .

**Proof:** In the case with no strategic externality to prove that  $\pi_t \in E_t^2(y; s)$  it suffices to show that for all  $(x; z) \in \text{supp} \pi_t$ ,  $x \in B_t^2(y; z)$  (the aggregate action  $\pi^1$  has been dropped as an argument of  $B(\cdot)$  since by Assumption 8  $v$  is independent of  $\pi^1$ ).

Choose a  $T$  such that

$$\delta^T v < \epsilon'$$

or, taking logs,

$$T > \frac{\log \epsilon' + \log v}{\log \delta}$$

Let  $x_t \in B_t^2(y; z)$  and  $x_t^0 \notin B_t^2(y; z)$ . Let  $V_t$  be the expected payoff along the equilibrium path if  $x_t$  is chosen in period  $t$  and the player behaves optimally

otherwise and let  $V_t^0$  be the expected payoff along the equilibrium path if  $x_t^0$  is chosen and the player behaves optimally otherwise. Then

$$V_t \geq V_t^0 - (1 - \beta)^2 \beta^2 \psi \beta^{\tau - t} \quad (4.5)$$

Note that this inequality holds independent of the particular choice of  $x_t$  and  $x_t^0$ . To show that an  $\beta^2$  best response to  $y$  is played in equilibrium we need to show that there is an  $\beta$  such that  $V_t \geq V_t^0 > 0$ . From (4.5) a sufficient condition for that to happen is:

$$(1 - \beta)^2 \beta^2 \psi \beta^{\tau - t} > 0 \quad (4.6)$$

For  $\beta < \frac{1}{2}$  there is an  $\beta$  such that (4.6) is satisfied and the Lemma follows.  $\square$

Lemma 9 shows that if the small players believe that the large player will play a given sequence of actions for a sufficiently long period of time with a high probability, then they will play an aggregate  $\beta^2$  best response to it. Lemma 8 on the other hand showed that if the large player played a certain strategy long enough then the small players would become convinced that he will continue to play that strategy for the following  $T$  periods with an arbitrarily high probability.

The following Lemma applies Lemma 8 and Lemma 9 to the Stackelberg strategy described above to show that in all but a finite number of periods the small players will play an aggregate  $\beta^2$  best response to an optimal sequence if the large player imitates a Stackelberg type.

Let  $T^\beta > (\log \frac{2}{1 - \beta} / \log \beta) = \log_{\beta} 2$  and let  $y^\beta$  denote the strategy played by commitment type  $b$  ( $2; \beta; T^\beta$ ). Let  $H^\beta$  be the set of histories consistent with  $y^\beta$  being played by  $b$ . Further let

$$H_t^\beta(2; \beta; \tau) = \{h \in H^\beta \mid y = y(2; \beta; \tau); 1 \leq t \leq E(y(2; \beta; \tau)); \tau_1 = \tau\}$$

be the histories for which the sequence  $y(2; \beta; \tau)$  and an aggregate  $\beta^2$  best response to this sequence have been played.

Lemma 10 Suppose  $h \in H^a$ . Then there is a number  $N$ , independent of  $h$ , such that the number of periods for which  $1_t \notin E_k^2(y(2; \hat{\cdot}; s); h_k)$ , for all  $h_k \in H_k^a(2; \hat{\cdot}; s)$ , for all  $k \in t$  and for all  $s$  is bounded by  $N$  with probability 1.

Proof: For the proof of this Lemma we keep  $(2; \hat{\cdot})$  fixed and therefore we will drop  $(2; \hat{\cdot})$  as arguments in  $y(\cdot)$  and  $H^a(\cdot)$ . Suppose that for all  $k \in t$ ,  $1_t \notin E_k^2(y(s); h_k)$ ,  $h_k \in H_k^a(s)$ . Then there is a  $t^0 \in (t - T + 1; \dots; t)$  such that  $\frac{1}{4} \epsilon^{aT} < (1 - \epsilon)^{aT}$  since otherwise  $h_t n_{t^0} \in H_{t^0}^a(s, t^0)$  for some  $0 < t^0 < t$  and hence the large player will continue to play  $y(s, t^0)$  for the next  $T$  periods with probability greater than  $(1 - \epsilon)^{aT}$  and therefore Lemma 9 implies that  $1_t \in E_k^2(y(s); h_k)$ .

But  $\frac{1}{4} \epsilon^{aT} < (1 - \epsilon)^{aT}$  at most  $T \frac{\log p^a}{\log(1 - \epsilon)^{aT}}$  times with probability 1 (Lemma 8). Thus  $N \leq T \frac{\log p^a}{\log(1 - \epsilon)^{aT}}$  with probability 1.  $\square$

The following Lemma says that by imitating the commitment type constructed in the section 4.4 the large player can get a payoff at least  $V^b(2) - \epsilon$ . Since  $\epsilon$  and  $\hat{\cdot}$  are arbitrary Lemma 11 proves Theorem 4.

Lemma 11 Suppose Assumptions 7 and 8 hold. Further suppose that  $\epsilon^a(2; \hat{\cdot})$  has prior probability  $p^a > 0$  then in any Nash equilibrium  $(\sigma; \pi)$  for initial state  $s$ ,  $\lim_{\epsilon \rightarrow 1} V^b(\sigma; \sigma; \pi; s) \geq V^b(2) - \epsilon$ .

Proof: Consider the strategy for  $b$  of always following  $y^a$  (corresponding to  $\epsilon^a = \epsilon^a(2; \hat{\cdot}; T)$ ). Then  $1_t \notin E_k^2(y(2; \hat{\cdot}; s); h_k)$  for fewer than  $N$  periods by Lemma 3. For a given  $s$  let

$$v^t(s) = \inf_{1 \in E^2(y(s))} (1 - \epsilon)^{-k} \sum_{k=1}^{\infty} \epsilon^k V(y_k(2; \hat{\cdot}; s); 1_k);$$

for  $1_1^Z = s$  and let  $v^t = \inf_s v^t(s)$ . Then a lower bound for  $b$ 's Nash equilibrium payoff can be described as:

$$v^{t_1} + 0 + \epsilon^{-t_1+1} v^{t_2} + \dots + \epsilon^{-t_1+\dots+t_{N-1}+N} v^{t_N} + 0 + \epsilon^{-t_1+\dots+t_N+N} v^1$$

$$v^t + 0 + \dots + v^{t+1} + \dots + v^{t+N} + 0 + \dots + v^{t+N}$$

for some  $t$  where  $0 < t < 1$ . Then

$$\begin{aligned} V^b(\dots) &\leq v^t(1 + \dots + v^{-(t+1)} + \dots + v^{-(N+1)t}) + v^{-(t+1)} \\ &\leq v^{t+1}(1 + \dots + v^{-(t+1)} + \dots + v^{-(N+1)(t+1)}) + v^{-(t+1)} \\ &= v^{t+1} \frac{1 - v^{-(t+1)(N+1)}}{1 - v^{-(t+1)}} + v^{-(t+1)} \end{aligned} \quad (4.7)$$

Let

$$\hat{v}^t = \frac{v^t}{(1 - v^{-t})} = \frac{\prod_{k=1}^t v^{-k}}{\prod_{k=1}^t v^{-k}}$$

Then (4.7) becomes

$$V^b(\dots) \leq \hat{v}^{t+1}(1 - v^{-(t+1)(N+1)}) + v^{-(t+1)} \frac{1 - v^{-(t+1)(N+1)}}{1 - v^{-(t+1)}} \quad (4.8)$$

Now we want to let  $\epsilon \rightarrow 1$ . Notice that the  $t$  that appears in (4.8) is a function of  $\epsilon$ . If  $t(\epsilon)$  stays bounded by some  $T < 1$  as  $\epsilon \rightarrow 1$ , then the first and the last term (4.8) tend to zero and the result follows since

$$\lim_{\epsilon \rightarrow 1} v^{-(t+1)(N+1)} \leq V^b(2) \epsilon^{-1}$$

If  $t(\epsilon)$  does not stay bounded as  $\epsilon \rightarrow 1$ , i.e.  $t(\epsilon) \rightarrow 1$ , then we have for some  $0 < \mu < 1$ :

$$\lim_{\epsilon \rightarrow 1} V^b(\dots) \leq (1 - \mu) \liminf_{\epsilon \rightarrow 1} \hat{v}^{t(\epsilon)+1} + \mu \lim_{\epsilon \rightarrow 1} v^1 \leq V^b(2) \epsilon^{-1}$$

since both  $\liminf_{\epsilon \rightarrow 1} \hat{v}^{t(\epsilon)+1}$  and  $\lim_{\epsilon \rightarrow 1} v^1$  are greater than or equal to  $V^b(2) \epsilon^{-1}$ .

2

### 4.8.3 Proof of Theorem 5

Theorem 5 Suppose that Assumptions 7,9, and 10 hold. In any Nash equilibrium  $(\dots)$  for initial state  $\dots$ ,  $\lim_{\epsilon \rightarrow 1} V^b(\dots) \leq V^b$ .



First we will need a preliminary Lemma.

**Lemma 10** For all  $\epsilon > 0$ , there is a pair  $(\delta; \epsilon^3)$  such that if  $\mu_1 \in E_1^3(\delta; \epsilon)$  and  $\|j^{\mu_1} - j^{\mu_1^0}\| < \delta$ , then  $\mu_1^0 \in E_1^2(\delta; \epsilon)$ .

**Proof:**  $\mu_1 \in E_1^3(\delta; \epsilon)$  means that there exists a  $\mu^1 \in E^3(\delta)$  :  $\mu^1(h_1) = \mu_1$ , where  $h_1 = \epsilon$ . Moreover for each realization of  $y \in Y^1$ ,  $\mu^1$  implies a  $\mu^1 \in M^1$ . Clearly we can construct a  $\mu^0$  such that for each realization of  $y \in Y^1$  we get a  $\mu^0 \in M^1$  with  $\|j^{\mu^0} - j^{\mu^1}\| < \delta$ . By the definition of aggregate  $\mu^2$  best response there exists a  $\mu_1^0$  with  $\|j^{\mu_1^0} - j^{\mu^1}(h_1)\| < \epsilon^3$  such that:

$$x \in B_1^3(\mu_1^0; \delta; h_1; z); \mu(x; z) \in \text{supp} \mu_1^0.$$

However, since  $\|j^{\mu_1^0} - j^{\mu^1}\| = \|j^{\mu_1^0} - j^{\mu^1}(h_1)\| < \delta$  we have  $\|j^{\mu_1^0} - j^{\mu_1^0}\| < \delta + \epsilon^3$ . By continuity of  $v$ , for all  $\epsilon > 0$  there exist  $\delta^3$  and  $\delta^4$  such that

$$x \in B_1^2(\mu_1^0; \delta^4; h_t; z); \mu(x; z) \in \text{supp} \mu_1^0$$

which means that  $\mu_1^0 \in E_1^2(\delta^4; \epsilon)$ .

The next Lemma is a weaker version of Lemma 9 for the case where  $v$  depends on  $\mu$ . Recall that for a pure strategy  $y$ ,  $\mu_t^{yT}$  denotes the probability that  $y$  is played in each of the next  $T$  periods, i.e. in the periods  $t; t+1; t+2; \dots; t+T$ .

**Lemma 12** Given  $\epsilon > 0$ , for every  $L$  there is a  $(T; \epsilon)$  such that for  $\mu_t^{yT} > \epsilon$  then in equilibrium for all  $y \in Y(L)$ ,  $\mu_t \in E_t^2(y; h_t)$  where  $(T; \epsilon)$  is independent of  $t; h_t$ .

**Proof:** Note that since  $Y(L)$  contains a finite number of elements and since  $(y_t; y_{t+1}; \dots) \in Y(L)$  if  $y \in Y(L)$  it is sufficient to show that for every pure strategy  $y$  we find a  $(T; \epsilon)$  such that if  $\mu_t^{yT} > \epsilon$ , then  $\mu_1 \in E_1^2(y; \epsilon)$ , for all  $\epsilon$ . Let  $S^T(y; \epsilon) = \{y_1; \dots; y_T\}$  is played with probability  $\epsilon$ . If  $\mu_t^{yT} > \epsilon$  then in equilibrium  $\mu \in E(\delta; \epsilon)$  for some  $\delta \in S^T(y; \epsilon)$ .

Let  $(\mathcal{Y}^T)$  be a sequence such that  $\mathcal{Y}^T \in \mathcal{S}^T(y; \theta^T); \theta^T \rightarrow 0$  as  $T \rightarrow \infty$ . Let  $(\mathcal{Y}^T; \mathcal{Y}^T)$  be a sequence such that  $\mathcal{Y}^T \in E(\mathcal{Y}^T; \mathcal{Y}^T)$  and let  $\mathcal{Y}^T = \mathcal{Y}^T(h^T)$  where  $h^T$  is the history where  $\mathcal{Y}_t = y_t$  for all  $t$ .

Claim: If  $\mathcal{Y}^T$  is a limit point of  $(\mathcal{Y}^T; \mathcal{Y}^T)$ , then  $\mathcal{Y}^T \in E(y; \mathcal{Y})$ .

Pf: Suppose  $\mathcal{Y}^T \notin E(y; \mathcal{Y})$ . Then there is a  $\delta$  and a set  $D \subset Z \times X$  with  $\mathbb{P}_{(z;x) \in D} \mathcal{Y}^T(z; x) > \delta > 0$  and for all  $(z; x) \in D$

$$x \notin B_\delta(y; \mathcal{Y}^T; h_\delta; z)$$

Thus if  $(z; x) \in D$  then for all  $x$  with  $x(h_\delta) = x$  there exists an  $x^0$  such that:

$$V_\delta(y; \mathcal{Y}^T; x; z; h_\delta) - V_\delta(y; \mathcal{Y}^T; x^0; z; h_\delta) \geq \delta$$

Choose  $T$  such that

$$\delta^T \geq \delta$$

Since  $\mathcal{Y}_t^T \rightarrow \mathcal{Y}_t$  uniformly for  $t \leq \delta + T$  and since  $\theta^T \rightarrow 0$  as  $\mathcal{Y}^T \rightarrow y$ , by continuity of  $v$  in  $\mathcal{Y}$  and in  $\theta$  at  $\theta = 0$ , it follows that for large  $T$  and for  $(z; x) \in D$  there is an  $x^0$  such that

$$V_\delta(\mathcal{Y}^T; \mathcal{Y}^T; x; z; h_\delta) - V_\delta(\mathcal{Y}^T; \mathcal{Y}^T; x^0; z; h_\delta) \geq \delta \tag{4.9}$$

However, for large  $T$ ,  $\mathbb{P}_{(z;x) \in D} \mathcal{Y}^T(z; x) < \delta$  and hence (4.9) contradicts the fact that  $\mathcal{Y}^T \in E(\mathcal{Y}^T; \mathcal{Y}^T)$ .

Next we want to show that as  $T \rightarrow \infty$  the distance between the first element of an aggregate best response to  $\mathcal{Y}^T$ ,  $\mathcal{Y}_1^T \in E_1(\mathcal{Y}^T; \mathcal{Y})$ , and the set of first elements of the aggregate best response to  $y$ ,  $E_1^3(y; \mathcal{Y})$ , tends to zero. More precisely, we want to show that for all  $(\delta; \epsilon)$  there exists a  $T^\delta$  such that for all  $T > T^\delta$  for  $M^T(\mathcal{Y}) = \mathcal{Y}_1^T \in E_1(\mathcal{Y}^T; \mathcal{Y})$  for some  $\mathcal{Y}^T \in \mathcal{S}(y; \theta^T)$

$$\sup_{\mathcal{Y}^T \in \mathcal{S}(y; \theta^T)} \sup_{\mathcal{Y}_1^T \in E_1(\mathcal{Y}^T; \mathcal{Y})} \inf_{\mathcal{Y}_1^3 \in E_1^3(y; \mathcal{Y})} \|\mathcal{Y}_1^T - \mathcal{Y}_1^3\| \leq \delta$$

This is satisfied if

$$\limsup_{T \rightarrow \infty} \sup_{\substack{y \in Y \\ \|y - y^T\| \leq \frac{1}{2M^T(\cdot)} \|\sum_{t=1}^T E_1^3(y; \cdot)\|}} \inf_{j \in \{1, \dots, n\}} \|j^T - y^T\| = 0:$$

From above we know that  $y^T$  is a limit point of a sequence  $y^T$  with  $y^T \in E(\mathcal{Y}^T; \cdot)$  has to belong to  $E(y; \cdot) \cap E^3(y; \cdot)$  which implies that the limit above is zero.

This implies (by Lemma 0) that by choosing  $\epsilon$  and  $\delta$  appropriately,  $y^T \in E_1^2(y; \cdot)$  for  $T > T^\epsilon$ , for all  $\delta > 0$ .

**Proof of Theorem 5:** Let  $\sigma(\epsilon; \delta) = \sigma(\delta; \epsilon; T^\epsilon)$  denote the commitment type that plays the Stackelberg strategy described above, where  $T^\epsilon$  satisfies Lemma 5 uniformly for all  $y \in Y(L)$  and  $L$  is chosen sufficiently large so that there is a  $y(\delta; \epsilon; \delta) \in Y(L)$  for all  $\delta$ . Now we can apply Lemma 10. Given that Lemma 10 holds so does Lemma 11. Note that  $\epsilon; \delta$  can be chosen arbitrarily by Assumption 9. Thus Lemma 11 proves Theorem 2.  $\square$

#### 4.8.4 Proof of Theorem 6

**Theorem 6** Suppose Assumptions 7, 8, 11, and 12 hold and all players have a common discount factor  $\delta$ . Then in any Nash equilibrium  $(\mathcal{Y}; \pi)$  for initial state  $y_0$ ,  $\lim_{\delta \rightarrow 1} \pi V^b(\delta; \mathcal{Y}; \pi; y_0) = \hat{V}^b$ .

**Proof:** Step 1: Let  $T$  be such that for all  $\delta$

$$\min_{y^T \in E^T(\mathcal{Y}^T; y_0)} \frac{1}{T} \sum_{t=1}^T v^b(y_t^T(\delta; y_0); \pi_t) \geq \hat{V}^b(\delta; y_0) - \epsilon$$

Note that for all  $\delta > 0$  there is a  $T < \infty$  such that  $y^T(\delta; y_0)$  satisfies the above inequality for all  $\delta$ . This is the case since

$$\| \hat{V}^b(\delta; y_0) - \hat{V}^b(\delta; y_0^0) \| < \psi \delta \| y_0 - y_0^0 \|$$

since  $v$  is independent of  $\delta$ <sup>11</sup>.

<sup>11</sup>The constant  $\psi$  is the upper bound on the payoffs of the small and the large players (Assumption 7).

Step 2: **Claim** Let  $y$  be a given pure strategy. Independent of  $y$ , for any  $\delta > 0$  there are  $\epsilon > 0$  and  $\beta < 1$  and a  $T < \infty$ , such that if the probability that  $y$  is followed in the first  $T$  periods is greater than  $1 - \epsilon$ , then for all  $\beta \leq \beta$  in any Nash equilibrium  $(x_1; \dots; x_T) \in E^{\delta, \beta}(y; z_1)$ .

**Pf:** Let

$$B^{\delta, \beta}(y; z) = \{x_t; z_t\}_{t=1}^T \in G^T(y^T) | z_1 = z \text{ and for all } (x_t^0; z_t^0) \in G^T(y^T) \text{ with } z_1^0 = z \\ \frac{1}{T} \sum_{t=1}^T v(y_t^T; x_t; z_t) \geq \frac{1}{T} \sum_{t=1}^T v(y_t^T; x_t^0; z_t^0) - \delta\}$$

Let

$$v^{\text{BT}}(z) = \frac{1}{T} \max_{x_t} \sum_{t=1}^T v(y_t; x_t; z_t)$$

with  $z_1 = z$ .

There are 3 reasons why a small player may not want to play an element in  $B^{\delta, \beta}(y; z)$ . First,  $y$  will only be followed with probability  $1 - \epsilon$ ; second, playing a best response may cause the player to reach a state in period  $T$  which is not the optimal state for the play thereafter and third, the player discounts future payoffs, instead of using the time-average criterion.

Let  $\sum_{t=1}^T \beta^{t-1} v_t$  be the expected payoff of the small player along the equilibrium path in the next  $T$  periods and let  $z_{T+1}$  be the state in which player  $i$  is in period  $T + 1$  along the equilibrium path. For  $(x^T; z^T) \in B^{\delta, \beta}(y; z)$  we have:

$$\sum_{t=1}^T \beta^{t-1} v_t \geq (1 - \epsilon) \sum_{t=1}^T \beta^{t-1} v^{\text{BT}}(z) + \epsilon \sum_{t=1}^T \beta^{t-1} v_t + \beta^T \sum_{t=1}^T \beta^{t-1} \frac{1 - \beta}{1 - \beta^T} v_t$$

On the other hand, the player can use the following sequence: for the first  $T - N$  periods play a sequence that maximizes the average payoff in the first  $T$  periods against  $y$ , in the last  $N$  periods, adjust the state so that in period  $T + 1$  the state  $z_{T+1}$  is reached. This gives a lower bound on the payoff:

$$\sum_{t=1}^T \beta^{t-1} v_t \geq (1 - \epsilon) \sum_{t=1}^T \beta^{t-1} v^{\text{BT}}(z) + \epsilon \sum_{t=1}^T \beta^{t-1} \frac{1 - \beta}{1 - \beta^T} v_t + \beta^T \frac{N}{T} v_{T+1}$$

Now note that if  $T$  is large and  $\epsilon$  is close to zero and  $\delta$  is close to one then the prescribed strategy is an element in  $B^{2;T}(y; z)$ . Furthermore it gives a larger payoff  $\epsilon$  than any strategy that is not an element of  $B^{2;T}(y^T)$  since

$$(1 - \epsilon)^2 > 2\psi \left( \prod_{t=1}^T \frac{1}{T} \right) \delta^{t-1} \frac{1 - \delta}{1 - \delta^T} + \psi \frac{N}{T} + \epsilon \psi$$

But this implies that for all  $(z_t; x_t)_{t=1}^T \in \text{supp}(1_1; \dots; 1_T)$  such that  $z_{t+1} = f(y_t; z_t; x_t)$  we have  $(z_t; x_t)_{t=1}^T \in B^{2;T}(y; z_1)$ , which proves the claim.  $\square$

Step 3: Let  $\sigma^1$  denote the sequence of  $(1_t)$  induced by the history when player  $b$  imitates the type  $1$  ( $2; T$ ) and let  $\mu_t^{2;T}$  be the probability that  $y(2; T)$  is played in the periods  $t; t + 1; \dots; t + T - 1$ . For every  $\epsilon > 0$ ,  $\mu_{kT+1}^{2;T} < 1 - \epsilon$  for fewer than  $N(\epsilon; T)$  different  $k$  (Lemma 8). Thus for all but  $N(\epsilon; T)$  different  $k$  we have  $(1_{kT+1}; \dots; 1_{kT+T}) \in E^{2;T}(y^T(2; \cdot_{kT+1}); \cdot_{kT+1})$ . But this implies that for all but  $N$  periods of length  $T$  the undiscounted payoff  $\epsilon$  of  $b$  is larger than  $\hat{V}^b(2) - \epsilon$ . Since  $\epsilon$  can be chosen arbitrarily small (Assumption 12), the Theorem follows.  $\square$

# Chapter 5

## References

1. Ausubel L. M. and R. Deneckere, (1989), "Reputation in Bargaining and Durable Goods Monopoly", *Econometrica* 57, 511-531.
2. Chari, V.V. and P. Kehoe (1990), "Sustainable Plans", *Journal of Political Economy*, 98, 783-802.
3. Coase, R. (1972), "Durability and Monopoly", *Journal of Law and Economics*, 15, 143-149.
4. Cripps, M., and J. Thomas (1992), "Reputation and Commitment in Two-Person Repeated Games", mimeo, University of Warwick.
5. Dekel, E. and J. Farrell (1990), "One-Sided Patience with One-Sided Communication Does Not Justify Stackelberg Equilibrium", *Games and Economic Behavior*, 2, 299-303.
6. Dutta, P. (1991), "A Folk Theorem for Stochastic Games", mimeo, Columbia University.
7. Fischer, S. (1980), "Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government", *Journal of Economic Dynamics and Control*, 2, 93-107.
8. Fudenberg, D., D. Kreps, and D. Levine, "On the Robustness of Equilibrium Refinements", *Journal of Economic Theory* 44 (1988), 354-80.
9. Fudenberg, D., and D. Levine (1989), "Reputation and Equilibrium Selection in Games with a Single Patient Player", *Econometrica* 57, 759-78.
10. Fudenberg, D., and D. Levine (1992), "Maintaining a Reputation when Strategies are Unobserved", *Review of Economic Studies*, 59(3), 561-580.

11. Fudenberg, D., D. Levine, and J. Tirole (1985), "Infinite Horizon Models of Bargaining with One-Sided Incomplete Information", in A. Roth (ed.) *Game Theoretic Models of Bargaining*, Cambridge University Press.
12. Gul, F., H. Sonnenschein, and R. Wilson, (1986), "Foundations of Dynamic Monopoly and the Coase Conjecture", *Journal of Economic Theory*, 39, 155-190.
13. Kreps, D., and R. Wilson, (1982) "Reputation and Imperfect Information", *Journal of Economic Theory* 27, 253-79.
14. Kydland F., and E. Prescott (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans", *Journal of Political Economy*, 85, 473-91.
15. Kydland F., and E. Prescott (1980), "Dynamic Optimal Taxation, Rational Expectations and Optimal Control", *Journal of Economic Dynamics and Control*, 2, 79-91.
16. Milgrom, P., and J. Roberts (1982), "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory* 27, 280-312.
17. Ramsey, F. P. (1927), "A Contribution to the Theory of Taxation", *Economic Journal* 37, 47-61.
18. Schmidt, K. (1993), "Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interest", *Econometrica*, 61, pp. 325-352.
19. Stokey N. (1981) "Rational Expectations and Durable Goods Pricing", *Bell Journal of Economics* 12, 112-128.
20. Stokey, N. (1992) "Credible Public Policy", *Journal of Economic Dynamics and Control*, 15, 627-656.