# Reputation From a Game Theoretic Point of View

Mario Gilli*
University of Milano-Bicocca - Department of Economics
Piazza Ateneo Nuovo 1 - 20126 Milano - Italy
tel: ++39 02 461916
fax: ++39 02 700439225
email: mario.gilli@unimib.it

March 27, 2009

`very preliminary and incomplete - comments welcome`

**Abstract**

This paper provides a general game theoretic framework to analyze reputation. The aim of this paper is to show the value added by a game theoretic approach to informal analysis of reputation. This object is pursued first through an exemplification of the contribution of game theory to the understanding of reputation, then explaining the formal mathematical machinery used to understand the way of building reputation. In particular I will develop an extensive analysis of three simple cases to show how the game theoretic approach to reputation works and its possible application to academic institutions. I conclude comparing the answers provided by game theory to open question in the analysis of reputation and showing how game theory stress subtle aspects of the way reputation is built and works.

*JEL*: C72, D83, D82
*Keywords*: reputation, game theory.

*This paper has been written following a suggestion by Emma Zavarrone, so all responsibility for possible mistakes are hers.

*"Until you have lost your reputation, you never realize what a burden it was or what a freedom really is"*
Margaret Mitchell, Gone with the Wind.

*" It takes 20 years to build a reputation and five minutes to ruin it. If you think about that, you'll do things differently"*
Warren Buffet, The Essays of Warren Buffett : Lessons for Corporate America.

# 1 Introduction

"Reputation is the general opinion about the character, qualities, etc of somebody or something"[1]. Therefore the crucial aspects of **reputation** are:

1. it is an opinion;
2. it is shared by a group of agent;
3. it regards hidden characteristics
4. of a person or a group of people or an organization.

Consequently these aspects can not be analyzed outside a context of social interaction. Since the role of game theory is to provide an abstract analysis of the implications of people interacting behaviour when agents' personal welfare depends on everyone behaviour, game theory is the natural mathematical language to analyze reputation. Indeed, game theory provides a formal model for each of the defining characteristics of reputation:

1. opinion are modelled as players' beliefs, i.e. probability measure on opponents' characteristics and/or behaviour;
2. in any game theoretic equilibrium players' beliefs should be shared;

[1] Hornby 1987.

3. players' hidden characteristics are defined as "types", i.e. players' private information on the defining aspects of the game: payoffs, players and strategies;

4. persons or organizations are modelled as "players" with their well defined objective function.

Therefore it is not surprising that reputation is one of the most developed and interesting field of application of game theory[2].

I will not provide a comprehensive survey of the game theoretic approach to reputation since it would be too long and anyway there are optimal surveys[3] so that it is not necessary to add a further similar work to the existing stock. The aim of this paper instead is to show the value added by a game theoretic approach to an informal analysis of reputation. This object is pursued first through an exemplification of the contribution of game theory to the understanding of reputation, then explaining the formal mathematical machinery used to understand the way of building reputation. In particular I will develop an extensive analysis of three simple cases to show how the game theoretic approach to reputation works and its possible application to academic institutions. First I will analyze these simple games informally, then after the presentation of the necessary mathematical tools I will fully develop their game theoretic analysis to enlighten the value added by game theory to informal analysis. I conclude comparing the answers provided by game theory to open question in the analysis of reputation and showing how game theory stress subtle aspects of the way reputation is built and works.

The paper consequently is organized as follows: section 2 is devoted to a quick review of the informal analysis of reputation, section 3 provides the exemplification of the game theoretic approach to reputation then providing notations and results, while section 4 applies these tools to the formal analysis of previous examples. Section

[2] To mention just few papers: Kreps-Wilson 1982; Kreps *et al.* 1982; Milgrom-Roberts 1982; Fudenberg-Levine 1989; Fudenberg-Levine 1992; Ely-Valimaki 2003, Ely *et al.* 2004.

[3] See in particular Fudenberg 1992 and Mailath-Samuelson 2006.

5 concludes emphasizing the importance of game theory to understand subtle aspects of the way reputation works in specific social context. Finally the appendix contains the mathematical details of the stochastic processes involved in this analysis.

## 2 The Informal Analysis of Reputation[4]

As *Informal Analysis of Reputation* I mean the approach elaborated by different authors, as exemplified by the previous references, which stress the social relevance of *Reputation, Identity, Trust*, all connected concepts that are particular relevant for the dynamic of social interaction. This approach is developed through empirical analysis, intuition and general argumentations, but a formal model is never presented.

According to this approach, reputation refers to an organizing principle of a socially recognized agent by which actions are linked into a common assessment. In particular reputation is a collective representation enacted in social relations. As a consequence reputation is connected to forms of communication and it is tied to a community.

Reputation operates in several different domains: personal, mass-mediated, organizational, historical. Usually reputation begin within circles of personal intimates, then they spread it outward.

Therefore, several aspects of reputation should be stressed.

First, agents build and share reputation of those who are within their social circle. Personal reputation has immediate consequences because of the options opened and closed changing possible interaction outcomes: those identities that we are given channel those identities that we can select. Therefore agents engage in form of self-presentation and impression management to modify their images in the eyes of others.

[4] See for example Dasgupta 1988, Good 1988, Kramer 1999, Ritzer 2007, Shenkar-Yuchtman-Yaar 1997, Strathdee 2009, Zucker 1986.

Second, the media help to determine who people should know, how people should care about and the social opinion people should confront with. The reputation of public agents are consequently used in formal and informal transactions among strangers.

Obviously, as social agents we mean not only individuals, but organizations too. In particular organizations strategically develop reputations that influence (positively or negatively) their effectiveness: the growth of public relations and of rating agencies are an important part of this process[5]. History, constituting narratives of personal and institutional biographies, may serve a similar role in a more sedimented way, as an institutionally sanctioned process: in fact from this point of view history represents settled cultural discourse about the past, determined by culturally literate experts; moreover this knowledge too is acquired through social institutions such as school and media.

From this point of view, reputation attempt to teach how agents should think about likely behaviours and thus likely social outcomes. People share memory and opinions because of what they have socially learned, in this way reputational knowledge reduces the complexity of the social world helping agents to focus on specific likely social outcomes. Therefore, reputation is particularly valuable in an uncertain world where it helps to focus agents' individual expectations on focal points, working as a potential selection device altering opportunities and constraints. In other words, because of this role as collector and summary of different opinions, reputation is the most valuable the highest the social uncertainty on the hidden characteristics of an agent. For example, this might explain why the rise of fashionable dot.com companies in the nineties has been accompanied by the expansion of rating agencies. In this setting clearly how to build a reputation is as much important as knowing how to protect and salvage a lost reputation. Indeed many surveys

[5] From 1990 to 2008, major U.S. media's usage of the term *reputation* has tripled, as shown in Leslie Gaines-Ross 2008.

enlighten the fact that agents regard the loss of reputation as the most dangerous risk for a company, exceeding all others, including market risk, natural hazards and physical or political security. Actually, an important aspect of reputation is that it can fallen suddenly and precipitously, but this does not mean that reputation recovery is not possible, as many real life examples show. To be more precise, consider the following facts:

1. only three companies from *Fortune*'s America's Top 10 Most Admired Companies in 2000 were among the Top 10 Most Admired in 2006;

2. IBM was once ranked first in *Fortune*'s America's Top 10 Most Admired Companies, fell to rank 354 in 1993 and it was again in the first Top 10 in 2003.

As this quick illustration shows the analysis of reputation is thus closely linked to the examination of collective memory and social mnemonic, and therefore it builds on cognitive sociology, on social movement research, and on sociology of knowledge. What we want to show in next section, is that **all** these aspects can be formally and rigorously studied within the language of mathematical game theory enlightening subtle aspects in the process of reputation building, loss and recovery.

From this discussion clearly emerges that the analysis of reputation building and maintenance, of loss and recovery of reputation must consider the following issues:

1. How broad is the context where the agent is trying to build a reputation for? Are we considering single or multitasks situations, specific or generic actions?

2. What have the agents at stake?

3. How informative is the observations of past outcomes in order to predict future behaviour?

4. Can we tie specific actors to observed and future outcomes?
5. Is it possible to measure reputation and, if the answer is positive, how can we do it?
6. What algorithms can let us make accurate predictions based on these reputation measures?

After the illustration of the game theoretic approach in sections 3 and 4, the conclusion will listen the precise answers provided by game theory to each of these open questions.

# 3 The game Theoretic Approach to Reputation

As the previous section argues, reputation is a tool to predict behaviour based on past actions and characteristics and it is based on linkability, in the sense that reputation should allow to link past actions to a specific set of possible identities so that future actions by the same set of possible identities are linked to future behaviour. These links are what help to make predictions about the agents' future choices. The previous section also shows that reputation involves behaviour that one might not expect in an isolated interaction. Finally, previous observations show that reputation is especially valuable in environments in which there is asymmetric information.

These considerations imply that the branch of game theory that can help us to build mathematical models of reputation is **repeated game theory with incomplete information**, strategic situations where a bunch of players is called to repeatedly act observing, may be imperfectly, players' past behaviour but where these players ignore some hidden characteristics of the interacting agents.

## 3.1 Example 1

Consider the game shown in figure 1:

| | $h$ | $l$ |
|---|---|---|
| $H$ | **2, 3** | **0, 2** |
| $L$ | **3, 0** | **1, 1** |

**Figure 1**

A possible story behind this trivial game is the following: player raw (1) is a university who can exert either high effort (H) or low effort (L) in the production of its output. Player column (2) is a student who can buy either a high-priced education (h) or a low priced one (l). Payoffs reflect likely players' preferences. In a simultaneous moves game as the one of figure 1, the University can not observably choose H before student's choice, the same holds for player 2. Therefore this game has a unique (rationalizable) solution: (L,l), which is Pareto inefficient.

To see reputation at work, first suppose that there is a succession of short-lived player 2, each of whom plays the game only once, but such that each of them can observe the players' previous strategic choices. Clearly **short-lived players** are concerned only with their payoffs in the current period. One interpretation is that in each period, a new short-lived player enters the game, is active for only that period and then leaves the game. An alternative interpretation is that each short-lived player represents a continuum of small anonymous long-lived agents such that each agent's payoff depends on its own action, the action of the large player and the average of the small players' actions. Moreover observable histories of play are assumed to include the actions of the large players and only the distribution of play produced by the small players. Since small players are negligible, change in the action of a single agent does not affect the distribution of play, and so does not influence future behavior of any small or large players, generating myopic rational behaviour[6].

Then as long as player 1 is sufficiently patient, (H, h) is an equilibrium of this repeated game where deviation to $L$ in order to boost payoff is deterred by future punishment trough the playing of the inefficient equilibrium (L,l). Unfortunately this repeated game has multiple equilibria, including of course the inefficient one, but the efficient equilibrium $(H, h)$ can be interpreted as being focal because of player 1 positive reputation, which in turn is based on the threat of loosing reputation after bad behaviour. In other words is the existence of player 1 reputation that justifies the focus of the theorists on the Pareto efficient equilibrium as the focal one. But then, this means that this approach can not justify the building, the loss and the restoration of reputation, it simply shows the consequence of its existence and of the fear of loosing it. Therefore this approach based on simple repeated games fails on the analysis of what we indicated as the most relevant aspects of any theory of reputation, it can simply be used to interpret a repeated game equilibrium strategy profile, but otherwise adding nothing to the formal or informal analysis.

The second approach, called **adverse selection approach** and pioneered by Kreps-Wilson 1982 and Kreps *et al. 1982*, is based on the idea that a player might be uncertain about key aspects of its opponent, in particular its payoffs and consequently its rational behaviour. Going back to previous informal analysis, we stressed that a crucial element of any reputational model should be the focus on hidden characteristics of the agents, and this is exactly the starting point of this approach. The crucial role of **incomplete information** is that it **introduces an intrinsic connection between past observations and beliefs on unknown payoff characteristics, which in turn affect expectation on future outcomes**.

[6] Se Gilli 2002 for a formal treatment of the case of negligible and anonymous long-lived players in repeated games leading to myopic rational behaviour.

As we will see incomplete information places constraints on the set of possible equilibria since small departures from complete information can have large effects on the set of equilibrium payoffs, even coarsening such set.

To show the effectiveness of this approach, suppose that in game 1 students are not entirely certain of the characteristics of the university: they attach high probability to the university being "normal" meaning that it has the payoffs of figure 1, but they entertain some very small probability that they face a university that has intrinsic motivation to play H. This kind of player 2 is called a **commitment type**.

The introduction of this small uncertainty on players' characteristics is enough to generate new interesting results.

First, two periods are enough to destroy the inefficient equilibrium; second, as long as the university is sufficiently patient, in any Nash equilibrium of the infinitely repeated game the university's payoff must be arbitrary close to 2, no matter how unlikely is the commitment type.

The reasons for these results are simply. Consider when the game is played twice, the payoffs are added over the two periods and suppose that the normal type of player 1 plays L in both periods. Then player 2 plays l in the first period since the probability of the commitment type is small, but its behaviour in the second period will depend on its observations: after observing L it will conclude that it is facing a normal type and will play l in the second period too, but if it observes H it will conclude that it is playing with a commitment type and it will best response with h; but then playing L by normal 1 is not an equilibrium since deviating and masquerading as the commitment type would increase the payoff in the next period. Therefore this game repeated twice has not a pure strategy equilibrium but a mixed one. But if the game is repeated infinitely many times, then there exist a pure strategy pooling equilibrium where both types play H. In fact in any equilibrium with player 1 payoff smaller than 2-$\epsilon$, the two types of player 1 should behave differently otherwise the students would choose h yielding a payoff of 2 for player 1. Therefore the normal type mimicking the commitment type behaviour over a sufficiently long period of time is able to convince the students that it is the commitment type inducing their best response h. Once this happens, the normal university thereafter earns 2, and the initially zero payoffs will not matter since the university is patient enough and the game is repeated many times.

### 3.2 Example 2

Now consider the following sequential game with a manipulable outcome function, i.e. a situation where the outcome actually observed by the players depends on player 2 behaviour[7].

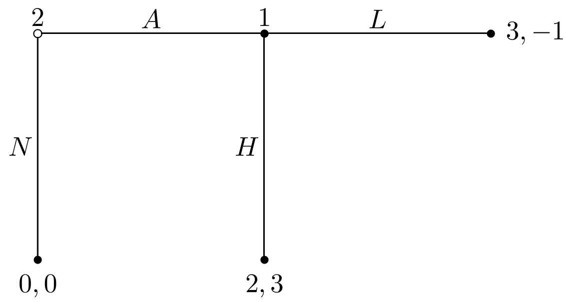


**Figure 2**

A possible story behind this simple dynamic game is a slightly change in the previous one: player 2 is a student who should decide whether to apply (A) or not (N) to a university (player 1), which in turn **after the application** can decide whether to exert high effort (H) or low effort (L) in the production of its output. Payoffs reflect likely players' preferences. The main effect of the non trivial sequential structure of this game is that the choice of N by player

[7] See Gilli 1999 for a complete treatment of signal functions and the role of manipulable signal functions foe players' equilibrium behaviour.

2 implies that the university choice is not publicly observable but remains 1's private information. In other words, this is a game with **imperfect public monitoring**. Does this characteristic change the way of working of reputational effects?

Note that this game has a unique (subgame perfect) equilibrium: (L,N), which is Pareto inefficient.

As before, to see reputation at work suppose that there is a succession of short-lived player 2, each of whom plays the game only once, but such that each of them can observe the players' previous choices.

Then we can apply both the approaches we saw for game 1.

If game 2 is repeated infinitely many times, then $(H, A)$ is an equilibrium of this repeated game where deviation to $L$ in order to boost payoff is deterred by future punishment trough the playing of N. Of course, as before this repeated game has multiple equilibria, including the inefficient one.

To apply the **adverse selection approach**, as before suppose that the students are not entirely certain of the characteristics of the university: they attach high probability to the university being "normal" meaning that it has the payoffs of figure 2, but they entertain some very small probability that they face a commitment type that always plays H. In this situation, the reasoning used on game 1 can not show anymore that repeating the game with short-lived player 2 is enough to destroy the inefficient equilibrium, independently from players' patience as long as the commitment type is unlikely enough. The reason for this dramatic change in the possibility of reputation building is simply the change in the informative structure: player 2 will observe player 1's behaviour only playing A, but if prior probabilities are such that the students' myopic best responses are always to play N then they will never have any new information on university's type and thus priors will not change. Moreover, the fact that students are short-lived means that they don't have any reason to experiment playing A since they will not have the opportunity to exploit such new information. Of course a mistake or a crazy type of player 1 or a long-lived player 1 would support the efficient equilibrium (H, A): after observing L, all future students will conclude that they are facing a normal type and will play N, but after observing H they will conclude that they are playing with a commitment type and will best response with A; but then **if given the possibility** playing L by normal 1 is not an equilibrium since deviating and masquerading as the commitment type would increase the payoff in the next period. Therefore everything would work as in game 1: this means that to re-establish previous results on reputation building and effectiveness in setting with imperfect public monitoring we need to introduce a crazy short-lived player. In other words, the existence of "noise" players ensures that the long-lived player can build a track record and through it a reputation.

### 3.3 Example 3

Example 2 is particularly important because it might be applied to any general strategic setting where there is imperfect public or private monitoring unless the information structure allow to avoid non informative behavior. This idea of inducing public revelation of long-lived player action as a consequence of the short-lived player choice, suggests a possible way out from this impasse: the introduction of a new observable action for the long-lived player that can costly signal its type inducing the participation of some short-lived player whose behavior would generate a positive informative externality allowing that kind of mimicking behavior that generate positive reputation.

This possibility can be represented in the following picture:

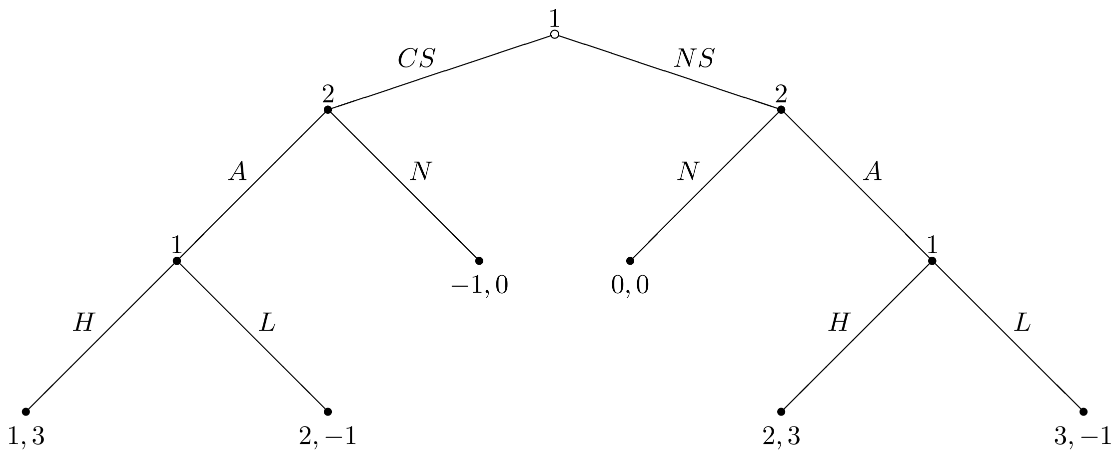


**Figure 3**

Finally, it should be stressed that the realistic assumption of noisy signals significantly complicates the analysis, as it is intuitive.

### 3.4 Notation and Theory

The previous examples have motivated the importance of modeling different monitoring possibilities to study reputation. Therefore we will use as building block of our formal language the notion of **Imperfect Monitoring Game**.

Consider the following generalization of the standard notion of repeated game with discount and almost perfect information. The players play a fixed stage game finitely or infinitely many times. The **stage game** is described by a finite **Imperfect Monitoring Game with Incomplete Information** (IMG)[8] defined as follows:

$$G(\eta, \Xi) := (N, S_i, \Xi_i, u_i, \eta_i)$$

where:

- $N = \{1, \cdots N\}$ is the set of players,

[8] See Gilli 1995 for a comprehensive analysis of IMGs.

- $S_i$ is the finite set of player $i$'s pure strategies; moreover $S := \times_{i \in N} S_i$, $\Lambda_i := \Delta(S_i)$ and $\Lambda := \otimes_{i \in N} \Lambda_i$, where $\Delta(\cdot)$ is the set of all probability measure on a set $\cdot$ and $\otimes_{i \in N} \Delta(\cdot_i)$ is the set of all independent probability measure on $\times_{i \in N} \cdot_i$;

- $\Xi_i$ is the set of possible player $i$'s types: formally a player's set of type is a random variable with probability distribution $\mu_{-i} \in \Delta(\Xi_i)$[9], its realization $\xi_i$ is a player's type representing the player's private information, while $\mu_{-i}$ is the common prior beliefs of players $j \neq i$ on $i$'s private information. In other words a type $\xi$ is a full description of

  - Player's beliefs on states of nature, e.g. players' payoffs;
  - Beliefs on other players' beliefs on states of mature and its own beliefs
  - Etc.

  Clearly there is a circular element in the definition of type, which is unavoidable in interactive situations

- $u_i : \Delta(S) \times \Xi_i \to \mathbf{R}$ is player $i$'s utility function, which we suppose represents preferences satisfying von Neumann and Morgenstern axioms and therefore such that: $u_i(\alpha; \xi_i) = E_\alpha[u_i(s; \xi_i)], \quad \alpha \in \Delta(S)$;

- $\eta_i : S \to M_i$ is player $i$'s signal function: $\eta_i(\bar{s}) = m_i \in M_i$ is the signal privately received by $i$ as a consequence of the strategy profile $\bar{s}$ played. Trivially how $m_i$ translates in information on opponents' behavior depends on the particular specification of $\eta_i$ and on the strategy played by $i$ herself (active learning).

Note that imperfect public monitoring and perfect monitoring are obtained for specific case of the signal function:

[9] We assume that $\Xi_i$ is measurable for all $i \in N$, so that probability measures are well defined.

- an IMG $G(\eta)$ has **imperfect public monitoring** if and only if $\forall i \in N \;\; M_i = M$

- an IMG $G(\eta)$ has **perfect monitoring** if and only if $\forall i \in N \;\; \eta_i$ is bijective.

As usual, when we will omit the index $i$ it means that we consider a profile belonging to the Cartesian product of the sets considered, the subscript $-i$ denotes the $j$s different from $i$ and $(-i, i)$ indicates a complete profile, stressing the $i$ component.

Define as follows the probability distribution $\rho_i(\alpha) \in \Delta(M_i)$ induced on $M_i$ by a probability measure $\alpha \in \Delta(S)$:

$$\forall i \in N, \;\; \forall m_i \in M_i \quad \rho_i(m_i)[\alpha] := \int_{\{s|\eta_i(s)=m_i\}} \alpha(s).$$

**Assumption 1** *The signal is defined to contain all of the information player $i$ receives about opponents' choices. Therefore*

$$\rho_i[s_i, \alpha_{-i}] = \rho_i[s_i, \alpha'_{-i}] \quad \Rightarrow \quad u_i(s_i, \alpha_{-i}) = u_i(s_i, \alpha'_{-i}).$$

This assumption means that each player receives her payoff after the stage game and that each player knows her own move.

Now suppose that the IMG is played many times, possibly infinite, and at the end of each period each player $i$ observes a stochastic outcome $m_i$, which is drawn from a finite set $M_i$ according to a probability distribution $\rho_i[\alpha]$, for some $\alpha \in \Delta(S)$. Therefore consider the following generalization of the usual notions defined for repeated games.

The **set of (finite) histories** for player $i$, $H_i$, is defined as follows:

$$H_i^0 := \Xi_i, \quad \forall t \geq 1 \;\; h_i^t := (\xi_i, m_i^1, \cdots, m_i^t) \in H_i^t := H_i^0 \times M_i^{(t)} = \Xi_i \times M_i^{(t)}$$

$$H_i := \bigcup_{t=0}^{\infty} H_i^t$$

where $m_i^t$ and $H_i^t$ indicate respectively $i$'s message and set of histories at time $t$, and the superscript $(t)$ denotes the $t$-fold Cartesian product of the set. Therefore a history at time $t$ for player $i$, $h_i^t$, is the private information received by player $i$ in the periods before $t$.

The pure superstrategies are defined in the usual way: the only difference from the traditional case is that the strategic choice are contingent to the players' private information: the **set of pure superstrategies** for player $i$, $F_i$, is defined as follows:

$$F_i := \{f_i \,| f_i = \{f_i^t(h_i^{t-1})\}_{t=1}^T \quad \text{with} \;\; f_i^t : H_i^{t-1} \to S_i\}$$

Where $T$ is either finite or infinite.

Let $F_i^t$ be the set of player $i$ times $t$ superstrategies: $F_i^t := \{f_i^t | f_i^t : H_i^{t-1} \to S_i\}$. Thus $F_i = \times_{t\in\mathbf{N}} F_i^t$.

As was first noticed by Aumann 1964, the definition of mixed and behavior strategies in this context requires some care when $T = \infty$, since in this case $F_i$ has the cardinality of the continuum (see Kolmogorov-Fomin 1975). Hence defining mixed superstrategies as probability distributions on pure superstrategies would not be straightforward. Following Aumann 1964, we think of a mixed strategy as a random device for choosing a pure strategy, i.e. as a random variable. Therefore consider an abstract probability space $(\Omega_i, \mathcal{A}_i, \alpha_i)$ and define a **mixed superstrategy** for player $i$ as a sequence $\phi_i \in \Phi_i$ of functions $\phi_i^t$ $\mathcal{A}_i$- measurable with

$$\phi_i^t : \Omega_i \times H_i^{t-1} \to S_i.$$

In general denote by $x \in \Delta(F)$ a probability distribution on $F$ constructed in this way and by $(F, \mathcal{F}, x)$ a probability space where $\mathcal{F}$ is the Borel $\sigma$-algebra on $F$. Similarly denote by $(F_i, \mathcal{F}_i, x_i)$ and $(F_{-i}, \mathcal{F}_{-i}, x_{-i})$ the probability spaces obtained through the opportune marginalizations.

The **set of behavior superstrategies** for a player $i$, $B_i$, is similarly defined, asking however for an additional restriction: $b_i \in B_i$ if and only if $b_i = (b_i^t)_{t=1}^{\infty}$ with $b_i^t : \Omega \times H_i^{t-1} \to S_i$ where $b_i^t(\cdot, h_i^{t-1}) : \Omega \to S_i$ is measurable and $b_i^t(\cdot, h_i^{t-1}), \quad b_i^\tau(\cdot, h_i^{\tau-1})$ are mutually independent random variables $\forall t \neq \tau, \ \ \forall h_i^t, \ h_i^\tau$. Therefore $B_i \subset \Phi_i$.

To simplify we might denote a mixed superstrategy as $\phi_i \in \Delta(F_i)$ and a behaviour superstrategy as $b_i : H_i \to \Delta(S_i)$.

The **outcome** at time $t$ for player $i$, $O_i^t(f)$, is defined inductively as a function of the superstrategies chosen:

$$O_i^0(f) := \xi_i$$

$$\forall t \geq 1 \quad O_i^t(f) := \eta_i[f^t(O^0(f), \cdots, O^{t-1}(f))] =$$
$$= \eta_i[f_1^t(O_1^0(f), \cdots, O_1^{t-1}(f)), \cdots, f_N^t(O_N^0(f), \cdots, O_N^{t-1}(f))] \ \in \ M_i.$$

Then the **outcome path** of the imperfect monitoring game **at time** $t$ for player $i$, $P_i^t(f)$, is defined as follows:

$$\forall t \geq 0 \quad P_i^t(f) := \{O_i^\tau(f)\}_{\tau=0}^t \ \in \ \Xi_i \times M_i^{(t)}$$

and the **outcome path** of the imperfect monitoring game for player $i$ is

$$P_i(f) := \{O_i^t(f)\}_{t=0}^T \ \in \ M_i^{(T)} := \Xi_i \times M_i^0 \times M_i^1 \times \cdots$$

where $T$ is either finite or infinite.

Therefore an outcome path is the sequence of outcomes induced by the players' behavior $f \in F$. As usual, when the dependence from the superstrategy profile $f$ is omitted this should be interpreted as a “realisation" of the mapping considered, for example $O_i^t$ is the outcome for player $i$ realised at time $t$ and $P_i \in M_i^{(\infty)}$ is a generic realisation of the outcome path $P_i(f)$.

Previous definition of outcome path implicitly assumes that players have perfect recall, therefore KuhnŠs theorem[10] holds and therefore *w.l.g.* we can restrict players behavior to behavior superstrategy.

[10] See for example Osborne-Rubinstein 1994.

Player $i$'s **intertemporal payoff function** $U_i : \Delta(F) \times \Xi_i \to \mathbf{R}$ is so defined:

$$U_i(b;\xi_i) := E_b \sum_{t=1}^{T} \delta_i^t u_i(f^t(P^t(f));\xi_i)$$

where: $b \in B,\ \xi_i \in \Xi_i,\ \delta_i \in [0,1)$ and $T$ is either finite or infinite.

Summing up, the repeated strategic situation is modeled by means of a **Repeated Imperfect Monitoring Game with Incomplete Information** (RIMG) denoted by

$$G^T(\delta,\eta,\Xi) = (N, F_i, \Xi_i, U_i, \eta_i).$$

In the appendix I provide the formal details of the objective and subjective strategic environment where the players make their choices. In particular the discussion on players' beliefs in the RIMG can be summmarized in the following assumption:

**Assumption 2** *In the RIMG every player $i \in N$ updates her beliefs $\beta_i$ according to the following expression:*

$$\forall f_i \in F_i, \quad \forall A \in \mathcal{F}_{-i}, \quad \forall t \in \mathbf{N} \qquad \beta_i^t[f_i](A) = E[\chi_A(f_{-i})|\mathcal{F}_{-i}^t(f_i)].$$

**Remark**: note that these beliefs on opponents' behaviour depend on players' beliefs on opponents' types.

Before applying these tools to the analysis of reputation, I sum up the notation introduced so far.

| NOTATION | |
|---|---|
| **Expression** | **Meaning** |
| $N$ | set of players |
| $S_i$ | set of player $i$'s pure strategies |
| $\Delta(\cdot)$ | set of probability measures on $\cdot$ |
| $\Sigma_i = \Delta(S_i)$ | player $i$ set of mixed strategies |
| $\Xi_i$ | set of types of player $i$ |
| $\mu_i(\cdot\|\xi_i) \in \Delta(\Xi_{-i})$ | player $i$ beliefs given its type on opponents' private information |
| $u_i : \Delta(S) \to \mathbf{R}$ | player $i$' utility function |
| $\eta_i : S \to M_i$ | player $i$' signal function |
| $\rho_i[\alpha] \in \Delta(M_i)$ | distribution of signals induced by a $\alpha \in \Delta(S)$ |
| $H_i = \cup_{t=0}^{\infty} H_i^t$ | set of possible histories for player $i$ |
| $F_i = \times_{t \in \mathbf{N}} F_i^t$ | set of possible superstrategies for player $i$ |
| $\Phi_i, \quad B_i$ | set of $i$'s mixed, behavioral superstrategies |
| $\beta_i \in \Delta(F_{-i})$ | probability evaluations on opponents' behaviour |
| $O_i^t(f), \quad P_i^t(f)$ | outcome, outcome path at time $t$ for player $i$ |
| $U_i : \Delta(F) \to \mathbf{R}$ | $i$'s utility function for the repeated game. |

# 4 Building a Reputation: Results and Applications

To simplify and according to our previous example, I consider two players games only, where

1. **player 1** is the long-lived player, and
2. **player 2** is the short-lived player, representing either a succession of players living one period or a continuum of small and anonymous infinitely living players.

Moreover we assume that the type of player 2 is common knowledge, while the type of player 1 is unknown to player 2, therefore *w.l.g.* $\Xi_1 = \Xi$, which is assumed to be finite or countable to simplify technical details. Player 1 set of types is partitioned into

1. **payoff types** $\Xi_P$ and
2. **commitment type** $\Xi_C = \Xi \setminus \Xi_P$.

**Payoff types** are payoffs characteristics such that the player maximizes $U_1$:

$$\forall \hat{b}_1 \in B_1, \quad \xi(\hat{b}_1) \in \Xi_P \text{ if and only if}$$

$$\exists b_2 \in B_2 \;\; s.t. \;\; \hat{b}_1 \in \arg\max_{b_1 \in B_1} U_1(b_1, b_2; \xi(\hat{b}_1)).$$

A specific payoff type is the **normal type**: $\xi_N \in \Xi_P$ if and only if

$$u_1(s, \xi_N, t) = u_1(s) \quad \forall s \in S, \; \forall t \in \mathbf{N},$$

i.e. the normal type has the standard payoff of the complete information stage game and thus $U_1(b, \xi_N) = U_1(b) \quad \forall b \in B$.

**Commitment types** have payoffs such that a specified superstrategy is strictly dominant and thus is certainly played by any rational player:

$$\forall \hat{b}_1 \in B_1, \quad \xi(\hat{b}_1) \in Xi_C \text{ if and only if}$$

$$\forall b_2 \in B_2 \quad \hat{b}_1 \in \arg\max_{b_1 \in B_1} U_1(b_1, b_2; \xi(\hat{b}_1)).$$

A **Reputation Game** (RG) is a Repeated Imperfect Monitoring Game with Incomplete Information $G^T(\delta, \eta, \Xi) = (N, F_i, \Xi_i, U_i, \eta_i)$ satisfying the following restrictions:

1. $N = \{1, 2\}$
2. $\Xi_1 = \Xi_P \cup \Xi_C$ and $\Xi_2 = \emptyset$
3. $U_1(b; \xi) = E_b \; \sum_{t=1}^{T} \delta^t u_1(f^t(P^t(f)); \xi)$ with $\xi \in \Xi_P \cup \Xi_C$
4. $U_2(b; \xi) = E_b u_2(f^t(P^t(f)))$.

The equilibrium concept we are going to use to present the results of the effect of reputation in strategic situations is the standard Nash equilibrium.

A strategy profile $(b_1^*, b_2^*)$ is a **Nash equilibrium** of a Reputation Game $G^T(\delta, \eta, \Xi) = (N, F_i, \Xi_i, U_i, \eta_i)$ if and only if

1. $\forall \xi \in \Xi_P, \quad b_1^* \in \arg\max_{b_1 \in B_1} U_1(b_1, b_2^*; \xi)$

2. $\forall t, \ \forall h^t \in H \quad s.t. \quad \mathbf{P}_{b_1^*, b_2^*, \mu}(h^t) > 0$

$$b_2^*(h_2^t) \in \arg \max_{b_2(h_2^t) \in \Delta(S_2)} E_{\beta_2}[u_2(b_1^*(h_1^t), b_2(h_2^t))].$$

Let denote the set of Nash equilibria of a Reputation Game $G^T(\delta, \eta, \Xi)$ by $NE(G^T(\delta, \eta, \Xi))$.

### 4.1 Building a Reputation When There is Perfect Monitoring

To present the important results of this approach applied to games with perfect monitoring, we need four further definitions:

1. Player 1's **pure strategy Stackelberg payoff** is defined as follows:
$$v_1^* := \sup_{s_1 \in S_1} \min_{\sigma_2 \in Br^u(s_1)} u_1(s_1, \sigma_2)$$
where $Br^u(s_1) := \arg\max_{\sigma_2' \in \Sigma_2} u_2(s_1, \sigma_2')$ i.e. it the set of player 2 myopic best reply to $s_1$;

   **Remark**: this is the best payoff that player 1 could get through a precommitment assuming that the opponent would reply rationally.

2. Player 1's **Stackelberg pure strategy** if it exists is defined as follows:
$$s_1^* := \arg\max_{s_1 \in S_1} \min_{\sigma_2 \in Br^u(s_1)} u_1(s_1, \sigma_2).$$

   **Remark**: this is a pure strategy to which player 1 would commit, if it had the chance to do so, given that such a commitment induces a best reply by 2.

3. Player 1's **commitment Stackelberg type** is defined as follows: $\xi^* := \xi(s_1^*)$, i.e. as the type that would always play the Stackelberg pure strategy.

4. Normal type of player 1's **lower bound payoff** is defined as follows: $\underline{v}_1(\xi_N, \mu, \delta) := (1-\delta) \inf_{(b_1,b_2)\in NE(G^T(\delta,\eta,\Xi))} U_1(b_1, b_2)$.

**REMARK:** it is possible to consider Stackelberg mixed strategy, which would reinforce reputational effects, but the analysis would be significantly more complex.

The main result for the case of perfect monitoring is the following.

**Theorem 1** *Suppose* $\mu \in \Delta(\Xi)$ *assign positive probability to some sequence of simple types* $\{\xi(s_1^k)\}$ *such that* $\lim_{k\to\infty} v_1^*(s_1^k) = v_1^*$, *then*

$$\forall \epsilon > 0,\ \exists \delta' \in (0,1) \quad s.t.\ \forall \delta \in (\delta', 1) \quad \underline{v}_1(\xi_N, \mu, \delta) \geq v_1^* - \epsilon.$$

**REMARKS:**

1. I will not provide a detailed proof[11], but I will show the implications of this theorem and I will illustrate the behavior of players beliefs in connection with previous examples.

2. If there a Stackelberg strategy and the associate Stackelberg type has positive probability under $\mu$, then the hypotheses of theorem 1 are trivially satisfied. In this case, the normal type of player 1 builds a reputation for **playing** like the Stackelberg type. Note that it builds this reputation despite the fact that there are many other possible commitments types.

3. Theorem 1 does not tell much about equilibrium strategies, in particular it does not imply that it is optimal for the normal type of player 1 to choose the Stackeberg strategy in each period, which in general is not the case.

4. The discount factor plays two roles in this results:

   (a) it makes future payoffs relatively more important, as it is standard in folk theorem arguments in repeated game theory

[11] A complete proof is in Mailath-Samuelson 2006.

(b) it discounts into insignificance the initial sequence of periods during which it may be costly for player 1 to mimic the commitment type, which is a new aspect relevant for reputation models only.

5. The proof of theorem 1 relies on the behavior of the posterior probability on player 1's type and on the probability of next period choice given today history. The key of the proof is to show that the observation of a strategy of player 1 increases the probability of the type committed to this action and consequently the probability of observing this strategy next period. This does not mean that the posterior probability of facing this commitment type is going to 1, since it is possible that player 1 is normal but plays like the commitment type, as we saw in the previous example.

The logic behind this result can be appreciated referring to the **example of figure 1**, that is reported here to simplify reading:

| | $h$ | $l$ |
|---|---|---|
| $H$ | **2, 3** | **0, 2** |
| $L$ | **3, 0** | **1, 1** |

**Figure 1**

Clearly in this game the pure strategy Stackelberg payoff is $v_1^* = 2$ since

$$Br^u(s_1) = \begin{cases} h & s_1 = H \\ l & s_1 = L. \end{cases}$$

and thus $\min_{\sigma_2 \in Br^u(s_1)} u_1(s_1, \sigma_2) = \{2, 1\}$ that implies $\sup_{s_1 \in S_1}\{2, 1\} = 2$. Consequently player 1's **Stackelberg pure strategy** is $s_1^* = H$.

To construct a **Reputation game** starting from game 1, suppose that there is perfect monitoring, i.e. $\eta$ is bijective and therefore the players observe the strategy profile played, and that there is incomplete information on the type of player 1. In particular suppose that the set of types is $\Xi_1 = \{\xi_N, \xi(H), \xi(L)\}$. Then the following strategy profile $(b_1^{NE}, b_2^{NE})$ is a Nash Equilibrium of the Reputation Game:

$$b_1^{NE}(\xi, h^t) = \begin{cases} H & \text{if } \xi = \xi(H) \text{ and } h^t \in H_1 \\ H & \text{if } \xi = \xi_N \text{ and } m^\tau = (Hh)\ \forall \tau < t \text{ or if } t = 1 \\ L & \text{otherwise.} \end{cases}$$

$$b_2^{NE}(h^t) = \begin{cases} h & \text{if } m^\tau = (Hh)\ \forall \tau < t \text{ or if } t = 1 \\ l & \text{otherwise.} \end{cases}$$

For $\delta \geq 1/2$ and $\mu(\xi(L)) < 1/2$ it is easy to show that this is a Nash equilibrium of the Reputation Game.

First, consider player 2: at $t = 1$, it will choose $h$ if and only if $Eu_2(s_1, h) = 3(1 - \mu(\xi(L))) \geq Eu_2(s_1, l) = 2(1 - \mu(\xi(L)) + 1\mu(\xi(L))$ which is satisfied when $\mu(\xi(L)) < 1/2$. Now consider its beliefs in subsequent periods; these are easily derived by Bayes rules and satisfy the following conditions:

$$\mu^t(\xi|h^t) = \begin{cases} 1 & \text{if } \xi = \xi(L) \text{ and } h^t = h^1 = (L\cdot) \\ 1 & \text{if } \xi = \xi_N \text{ and } m^t = (L\cdot) \text{ with } t > 1 \\ \frac{\mu(\xi(H))}{\mu(\xi(H)) + \mu(\xi_N)} & \text{if } \xi = \xi(H) \text{ and } m^t = (H\cdot) \text{ with } t \geq 1. \end{cases}$$

Note that to show that $(b_1^{NE}, b_2^{NE})$ is a Nash Equilibrium we need to consider beliefs only on the equilibrium path, therefore the previous calculation of $mu(\xi|h^t)$ is enough. To work on refinements such as Sequential equilibria we would need to specify also out-of-equilibrium beliefs where Bayes rules does not apply.

Given these beliefs and player 1's equilibrium strategy $b_1^{NE}$, through Bayes rule it is possible to derive player 2 beliefs on player 1's choice of $H$ next period:

$$\beta_2^t(\{f_1(\xi,h^t)=H\}|h^t)=\mu^t(\xi(H)|h^t)\times b_1^{NE}(H|\xi(H),h^t)+\mu^t(\xi_N|h^t)\times b_1^{NE}(H|\xi_N,h^t)$$

that implies

$$\beta_2^t(\{f_1(\xi,h^t)=H\}|h^t)=\begin{cases}\frac{\mu(\xi(H))}{\mu(\xi(H))+\mu(\xi_N)}\times 1+\frac{\mu(\xi_N)}{\mu(\xi(H))+\mu(\xi_N)}\times 1=1 & \text{if } h^t=(Hh)^{(t)}\\ 0 & \text{otherwise.}\end{cases}$$

Therefore the Nash equilibrium condition for player 2 is satisfied since $\forall t$, $\forall h^t\in H_2$ such that $\mathbf{P}_{b_1^{NE},b_2^{NE},\mu}(h^t)>0$, if $h^t=(Hh)^{(t)}$ then $E_{\beta_2}[u_2(b_1^{NE}(h^t),h)]=3>E_{\beta_2}[u_2(b_1^{NE}(h^t),l)]=2$ while if $h^t\neq(Hh)^{(t)}$ then $E_{\beta_2}[u_2(b_1^{NE}(h^t),l)]=1>E_{\beta_2}[u_2(b_1^{NE}(h^t),h)]=0$.

Finally, consider the normal type of player 1, since the behaviour of committed types $\xi(H)$ and $\xi(L)$ is trivial. Player 1 has complete information and it is long-lived, therefore it easy to see that in equilibrium it gets $U_1(b_1^{NE},b_2^{NE})=(1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau\times 2=2$, while deviating at most it gets $U_1(b_1,b_2^{NE})=(1-\delta)\times 3+\delta(1-delta)\sum_{\tau=0}^{\infty}\delta^\tau\times 1=3(1-\delta)+\delta$. Then $U_1(b_1^{NE},b_2^{NE})\geq U_1(b_1,b_2^{NE})$ if $\delta\geq\frac{1}{2}$. This shows that the restriction on the discount factor is needed to ensures that the normal type of player 1 has sufficient incentives to make $H$ optimal mimicking $\xi(H)$.

The two most important aspects of this example regard player 1's beliefs on opponent's type and on opponent's behaviour. The posterior probability that 2 assigns to the Stackelberg type **does not** converge to 1, it is actually bounded away from 1 since it converges to $\frac{\mu(\xi(H))}{\mu(\xi(H))+\mu(\xi_N)}$ if $h^1=(Hh)$ after which is constant unless $(L\cdot)$ is observed. But this does not forbid the fact that the posterior probability that 1 assigns to observing $H$ next period is going to 1 as long as $h^t=(Hh)^{(t)}$: as I showed before $\beta_2^t(\{f_1(\xi,h^t)=H\}|h^t=(Hh)^{(t)})=1$ and this is what drives the result.

A particularly interesting property of the Reputation Game associated to figure 1 is that playing $Ll$ forever is NOT a Nash equilibrium, even if it is a(subgame perfect) equilibrium of the repeated game with complete information. This is exactly the content of theorem 1, and I will show how it works in this example. Formally suppose that the set of types is $\Xi_1 = \{\xi_N, \xi(H), \xi(L)\}$ such that that $\mu(\xi(H)) < 1/3$ and $\mu(\xi(L)) < 1/3$.

Now consider the first period: clearly player 2 best reply is either l or h depending on the values of beliefs on opponent's behaviour:

$$BR^u(\beta_2^1(\{f_1(\xi) = H\}|h^0) = \begin{cases} h & \text{if } \beta_2^1(\{f_1(\xi) = H\}|h^0) \geq 1/2 \\ l & \text{if } \beta_2^1(\{f_1(\xi) = H\}|h^0) \leq 1/2. \end{cases}$$

Player 2 beliefs can easily be derived:

$$\beta_2^1(\{f_1(\xi) = H\}|h^0) = \mu(\xi(H)) \times 1 + \mu(\xi_N) \times b_1(H|\xi_N).$$

Therefore player 2 in the first period best responds to $b_1(H|\xi_N)$, i.e. to the normal type of player 1.

Then consider period 2: applying Bayes rule it easy to derive that

$$\mu^2(\xi|h^1) = \begin{cases} \frac{\mu(\xi(H))}{\mu(\xi(H))+\mu(\xi_N)} < 1/2 & \text{if } \xi = \xi(H) \text{ and } m^1 = (Hh) \\ \frac{\mu(\xi(L))}{\mu(\xi(L))+\mu(\xi_N)} < 1/2 & \text{if } \xi = \xi(L) \text{ and } m^1 = (Ll). \end{cases}$$

Therefore in the second period the strategic situation of player 2 is very similar to the situation of previous period: whether $(Hh)$ or $(Ll)$ being observed, player 2 in the second period best responds to the normal type of player 1. Therefore if player 1 chooses H in period 1, then player 2 will choose $h$ in all subsequent period, having concluded that 1 is not type $\xi(L)$ and this implies that for the long-lived player 1 is worth to pay at most a cost of $u_1(H, l)$ in one period in order to get $u(H, h)$ in all subsequent periods. Consequently in the Reputation game there is no equilibrium where $(L, L)$ is played forever: incomplete information reduces the set of possible equilibria outcomes.

## 4.2 Building a Reputation When There is Imperfect Monitoring

To present the important results of this approach applied to games with imperfect monitoring, I need two further definitions:

1. Player 2's $\epsilon$**-confirmed best reply correspondence** is defined as follows:

$$Br_\epsilon : \Sigma_1 \Rightarrow \Sigma_2 \;\; s.t. \;\; \forall \sigma_1 \in \Sigma_1$$

$$Br_\epsilon(\sigma_1) := \{\sigma_2 \in \Sigma_2 \mid \exists \sigma_1' \;\; \text{such that (a) and (b) are satisfied}\}$$

$$(a) \qquad SUPP(\sigma_2) \subseteq \arg\max_{s_2'} u_2(\sigma_1', s_2')$$

$$(b) \qquad |\rho_2[\sigma_1, \sigma_2] - \rho_2[\sigma_1', \sigma_2]| \leq \epsilon$$

where $SUPP(\cdot)$ denote the support of the probability measure $\cdot$. **REMARKS:**

(a) Note that it is possible to have

$$\sigma_2 \in Br_\epsilon(\sigma_1) \quad and \quad SUPP[\sigma_2] \nsubseteq Br_\epsilon(\sigma_1).$$

To show this possibility consider for example the game of figure 1 with imperfect public monitoring such that $M = \{m_L, m_H\}$ and with the following probability distribution $\rho(m_H|Hh) = \rho(m_H|Ll) = 1$ and zero otherwise. Then

$$\frac{1}{2}[h] \oplus \frac{1}{2}[l] \in Br_0(H) \quad but \quad l \notin Br_0(H);$$

(b) If there exist two different strategies $\sigma_1$ and $\sigma_1'$ such that $\rho[\sigma_1, \sigma_2] = \rho[\sigma_1', \sigma_2]$, then it is possible that $Br^u(\sigma_1) \subset Br_0(\sigma_1)$. To show this possibility, consider the game of figure 2: from the extensive form $\rho[L, N] = \rho[H, N]$; moreover $Br^u(H) = \{A\} \subset Br_0(H) = \{N, A\}$ since $N \in \arg\max u_2(L, \cdot)$ and $L$ imply the same signal distribution of $H$.

Define $Br_\epsilon^*(\sigma_1) := \{\sigma_2 | SUPP(\sigma_2) \subset Br_\epsilon(\sigma_1)\}$.

2. Player 1's **maximum rational payoff** is defined as follows:

$$v_1^{**} := \sup_{\sigma_1 \in \Sigma_1} \min_{\sigma_2 \in Br_0^*(\sigma_1)} u_1(\sigma_1, \sigma_2).$$

**REMARK:** a game with perfect monitoring is a special case of an imperfect monitoring game when $M = s$ and $\rho(m|s) = 1$ if and only if $m = s$. In this case $Br_0^* = Br_0 = Br^u$, therefore $v^{**} \geq v^*$. Thus this section extends the reputation results not only to imperfect monitoring but also to the case of mixed commitment types with perfect monitoring, obtaining a stronger bound on payoffs for the perfect monitoring case. More generally $v^{**}$ can be greater or smaller than $v^*$ depending on the information structure, as we will see.

The main result for the case of imperfect monitoring is the following.

**Theorem 2** *Suppose* $\mu \in \Delta(\Xi)$ *assign positive probability to some sequence of simple types* $\{\xi(\sigma_1^k)\}$ *such that* $\lim_{k\to\infty} \min_{\sigma_2 \in BR_0^*(\sigma_1^k} u_1^*(\sigma_1^k, \sigma_2) = v_1^{**}$*, then*

$$\forall \epsilon > 0, \ \exists \delta' \in (0, 1) \quad s.t. \ \forall \delta \in (\delta', 1) \quad \underline{v}_1(\xi_N, \mu, \delta) \geq v_1^{**} - \epsilon.$$

**REMARKS:**

1. I will not provide a detailed proof[12], but I will show the implications of this theorem and I will illustrate the behavior of players' beliefs in connection with previous examples.

2. Similarly to theorem 1, the normal player 1 effectively builds a reputation for playing like a commitment type, and this occurs despite the presence of many other possible commitment types.

3. The remarks previously discussed for theorem 1 again can be applied to this result.

As for theorem 1, even the proof of theorem 2 relies on the behavior of the posterior probability on player 1's type and on the probability of next period choice given today history. But here we have a new crucial problem in proving the theorem since we have imperfect

[12] A complete proof is in Mailath-Samuelson 2006.

monitoring possible with manipulable signal functions. The root of the problem can easily be shown referring to the game of figure 2, which is reported here to simplify reading:

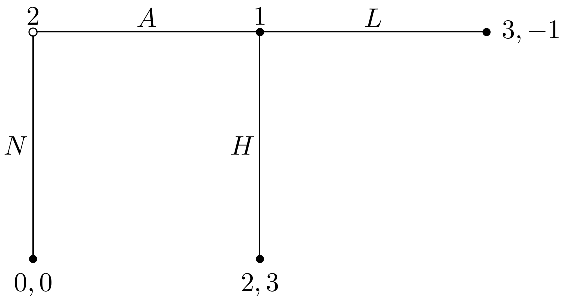


**Figure 2**

Clearly in this game the pure strategy Stackelberg payoff is $v_1^* = 2$ since

$$BR^u(s_1) = \begin{cases} \sigma_2(A) = 1 & s_1 = H \\ \sigma_2(A) = 0 & s_1 = L. \end{cases}$$

and thus $\min_{\sigma_2 \in Br^u(s_1)} u_1(s_1, \sigma_2) = \{2, 0\}$ that implies $\sup_{s_1 \in S_1}\{2, 0\} = 2$. Consequently player 1's **Stackelberg pure strategy** is $s_1^* = H$.

To construct a **Reputation game** starting from game 2, suppose that there is incomplete information on the type of player 1. In particular suppose that the set of types is $\Xi_1 = \{\xi_N, \xi(H)\}$.

Note that $Br_0(N) = \Sigma_1$ since $|\rho[\sigma_1, N] - \rho[\sigma_1', N]| = 0$ for all $\sigma_1, \sigma_1' \in \Sigma_1$. Clearly in this game the maximum rational payoff $v_1^{**} = 0$ since

$$\forall \sigma_1 \in \Sigma_1 \quad BR_0^*(\sigma_1) = \Sigma_2$$

and thus $\min_{\sigma_2 \in Br_0^*(\sigma_1) = \Sigma_2} u_1(\sigma_1, \sigma_2) = \{0\}$ that implies $\sup_{\sigma_1 \in \Sigma_1}\{0\} = 0$. This means that in this case theorem 2 is actually not relevant since $v_1^{**}$ is equal to the equilibrium payoff of the complete information game. Moreover differently from the case of game 1, it is easy to construct a Nash equilibrium of the reputation game corresponding to the equilibrium of the complete information game. Consider the following strategy profile $(b_1^{NE}, b_2^{NE})$:

$$b_1^{NE}(\xi, h^t) = \begin{cases} H & \text{if } \xi = \xi(H) \text{ and } h^t \in H \\ L & \text{otherwise.} \end{cases}$$

$$b_2^{NE}(h^t) = N \quad \forall h^t \in H.$$

For any $\delta$ and $\mu(\xi(H)) < 1/4$ it is easy to show that this is a Nash equilibrium of the Reputation Game.

First, consider player 2: at $t = 1$, it will choose $N$ if and only if $Eu_2(s_1, N) = 0 \geq Eu_2(s_1, A) = -1(1 - \mu(\xi(H)) + 3\mu(\xi(H))$ which is satisfied when $\mu(\xi(L)) < 1/4$. Now consider its beliefs in subsequent periods; these are easily derived by Bayes rules and equal to the prior since no new observation on possible player 1's types has been collected. Note that to show that $(f_1^{NE}, f_2^{NE})$ is a Nash Equilibrium we need to consider beliefs only on the equilibrium path, therefore the previous calculation of $mu(\xi|h^t)$ is enough. To work on refinements such as Sequential equilibria we would need to specify also out-of-equilibrium beliefs where Bayes rules does not apply.

Given these beliefs and player 1's equilibrium strategy $b_1^{NE}$, through Bayes rule it is possible to derive player 2 beliefs on player 1's choice of $H$ next period:

$$\beta_2^t(\{f_1(\xi, h^t) = H\}|h^t) = \mu^t(\xi(H)|h^t) \times b_1^{NE}(H|\xi(H), h^t) +$$

$$+\mu^t(\xi_N|h^t) \times b_1^{NE}(H|\xi_N, h^t) = \mu(\xi(H)).$$

Therefore the Nash equilibrium condition for player 2 is satisfied since $\forall t$, $\forall h^t \in H$ such that $\mathbf{P}_{b_1^{NE}, b_2^{NE}, \mu}(h^t) > 0$ $E_{\beta_2}[u_2(b_1^{NE}(h^t), N)] = 0 > E_{\beta_2}[u_2(b_1^{NE}(h^t), A)] = 3\mu(H) - 1(1 - \mu(H)) < 0$ if $\mu(H) < 1/4$. Finally, consider the normal type of player 1, since the behaviour of committed type $\xi(H)$ is trivial. Player 1 has complete information and it is long-lived, therefore it easy to see that in equilibrium it gets $U_1(b_1^{NE}, b_2^{NE}) = (1 - \delta) \sum_{\tau=0}^{\infty} \delta^\tau \times 0 = 0$, while deviating at most it gets $U_1(b_1, b_2^{NE}) = (1-\delta) \times 0 + \delta(1 - delta) \sum_{\tau=0}^{\infty} \delta^{\tau} \times 0 = 0$. Then $U_1(b_1^{NE}, b_2^{NE}) \geq U_1(b_1, b_2^{NE})\ \forall \delta$.

The most important aspects of this example is the fact that the transformation of the complete information game into a reputation game does not change the properties of the set of Nash Equilibria. Therefore it is particularly important to understand the root of this change in the properties of the strategic situation. Clearly the reason of this new results regard player 1's beliefs on opponent's type and on opponent's behaviour. The posterior probability that 2 assigns to the Stackelberg type **does not** change through time because there is no collection of new information. But this forbid the possibility for the normal type of player 1 of mimicking the commitment types, i.e. to build its reputation. Therefore the posterior probability that 1 assigns to observing $H$ next period is bounded away from 1 as long as $h^t = (N)^{(t)}$ and this is what drives the result.

This suggest that the crucial aspect is the information structure and its relation with players behavior. To show that this is actually the case, consider the strategic form game associated to game assuming that there is perfect monitoring, i.e. $\eta(s) = s$:

| | $N$ | $A$ |
|---|---|---|
| $H$ | **0, 0** | **2, 3** |
| $L$ | **0, 0** | **3, -1** |

**Figure 4**

In this game the pure strategy Stackelberg payoff is $v_1^* = 2$ since

$$Br^u(s_1) = \begin{cases} A & s_1 = H \\ N & s_1 = L. \end{cases}$$

and thus $\min_{\sigma_2 \in Br^u(s_1)} u_1(s_1, \sigma_2) = \{2, 0\}$ that implies $\sup_{s_1 \in S_1}\{2, 0\} = 2$. Consequently player 1's **Stackelberg pure strategy** is $s_1^* = H$.

To construct a **Reputation game** starting from game 3, suppose that there is perfect monitoring, i.e. $\eta$ is bijective and therefore the players observe the strategy profile played, and that there is incomplete information on the type of player 1. In particular suppose that the set of types is $\Xi_1 = \{\xi_N, \xi(H), \xi(L)\}$.

A particularly interesting property of the Reputation Game associated to figure 1 is that playing $LN$ forever is NOT a Nash equilibrium, even if it is a(subgame perfect) equilibrium of the repeated game with complete information. This is again the content of theorem 1, and it is interesting to show that it works or not depending on the information structure of the game. Formally suppose that the set of types is $\Xi_1 = \{\xi_N, \xi(H), \xi(L)\}$ such that that $\mu(\xi(H)) < 1/3$ and $\mu(\xi(L)) < 1/3$.

Now consider the first period: clearly player 2 best reply is either N or A depending on the values of beliefs on opponent's behaviour:

$$Br^u(\beta_2^1(\{f_1(\xi) = H\}|h^0) = \begin{cases} N & \text{if } \beta_2^1(\{f_1(\xi) = H\}|h^0) \leq 1/4 \\ A & \text{if } \beta_2^1(\{f_1(\xi) = H\}|h^0) \geq 1/4. \end{cases}$$

Player 2 beliefs can easily be derived:

$$\beta_2^1(\{f_1(\xi) = H\}|h^0) = \mu(\xi(H)) \times 1 + \mu(\xi_N) \times b_1(H|\xi_N).$$

Therefore player 2 in the first period best responds to $b_1(H|\xi_N)$, i.e. to the normal type of player 1.

Then consider period 2: applying Bayes rule it easy to derive that

$$\mu^2(\xi|h^1) = \begin{cases} \frac{\mu(\xi(H))}{\mu(\xi(H))+\mu(\xi_N)} < 1/2 & \text{if } \xi = \xi(H) \text{ and } m^1 = (HA) \\ \frac{\mu(\xi(L))}{\mu(\xi(L))+\mu(\xi_N)} < 1/2 & \text{if } \xi = \xi(L) \text{ and } m^1 = (LN). \end{cases}$$

Therefore in the second period the strategic situation of player 2 is very similar to the situation of previous period: whether $(HA)$ or $(LN)$ being observed, player 2 in the second period best responds to the normal type of player 1. Therefore if player 1 chooses H in period 1, then player 2 will choose $A$ in all subsequent period, having concluded that 1 is not type $\xi(L)$ and this implies that for the long-lived player 1 is worth to pay at most a cost of $u_1(H, N)$ in one period in order to get $u_1(H, A)$ in all subsequent periods. Consequently in the Reputation game there is no equilibrium where $(L, N)$ is played forever: incomplete information reduces the set of possible equilibria outcomes.

Note that this argumentation is the exact replica of the one used for game 1: this means that it is not the payoff structure that matters but the information structure. In particular the problem with imperfect monitoring is the fact that there may exists strategies by player 2 such that the signal, whether public or private, reveal no information about player 1's choice. Therefore if such strategy $s_2^{NI}$ is a best reply to anything, it belongs to $Br_0(\sigma_1)$ for all possible $\sigma_1$. To get $Br^u(\sigma_1) = Br_0(\sigma_1)$ it is necessary to rule out such non informative strategies, for example introducing some noise. A possible assumption is the following:

**Assumption 3** *For all $\sigma_2 \in \Sigma_2$, the collection of probability distributions*

$$\{\rho[s_1, \sigma_2] | s_1 \in S_1\}$$

*is linearly independent.*

Then an immediate result is the following:

**Theorem 3** *If assumption 3 holds, then*

$$Br^u(\sigma_1) = Br_0^*(\sigma_1) = BR_0(\sigma_1)$$

*and $v_1^{**}$ equals the mixed-strategy Stackelberg payoff.*

The proof is omitted since it is immediate.

## 5 Conclusion

The examples, results and calculations of this paper show the importance of the game theoretic approach to reputation building.

Considering the question posed at the beginning of the paper, now it is possible to consider the specific rigorous answers provided by game theory:

1. The context should be such that the choices of the agent that wish to build a reputation are publicly observable, may be noisily: as we have seen this might be a non trivial request;
2. Using reputation all agents can improve their payoffs, in particular the agent building its reputation has at stake the possibility of reaching the best payoff that it could get through a precommitment when facing a rational opponent;
3. the information players can infer from the observations of past outcomes to predict future behaviour crucially depend on two factors:
   (a) the structure of strategic interaction (see example 2 and example 3)
   (b) the possibility of facing "commitment" types, the possibility of existence of incomplete information on hidden characteristics of the long-lived player.
4. to tie actors to observed and future outcomes it is necessary to have an identity that last through time, i.e. a long lived agent;
5. reputation can be measured in term of likely future behaviour and players beliefs are exactly this measure;
6. Bayes rule and probability rules are the well defined algorithms that connect reputation measures and forecasts on future likely outcomes.

But the value added by game theory to the rigorous analysis of reputation is shown by some other new interesting and important perspectives that are opened by this point of view:

1. Reputation is possible only if in the population of possible types there exists at least one type committed to "desirable" behaviour: it is its simple existence that generates that informational positive externality that allow the possibility of building reputation;

2. paradoxically enough, reputation is not based on revealing and learning the players true characteristics, but exactly on the opposite: on the existence and persistence of pooling equilibria where the "normal" type can publicly mimic the "good" one. Therefore there should be enough public information to allow this mimicking and the consequent reputation building but also enough lack of information to avoid the possibility of perfectly learning the players' types;

3. A player's reputation can be modelled and measured as beliefs on its future behaviour not on its type: these beliefs in equilibrium are shared by all players and are correct, even if the players usually will never learn precisely opponents types. Actually, if the players type would be perfectly learned, then their reputation would disappear since it would be useless to behave as a commitment type;

4. Reputation works actually as a selection device: among all possible equilibria of a repeated game, it selects that equilibria that give rise to the Stackelberg payoff;

5. Reputation helps the player to reach a better payoff **reducing** its choice opportunities, it works as a precommitment device: in example 1 player 1 is able to convince player 2 that it will play H since otherwise it would loose its reputation in future

plays of the game and this "promise" is credible because of the possible loss of reputation;

6. To work as a precommitment device, the loss of reputation should depend on observable actions: transparency and the role of media is absolutely crucial to allow a virtuous work of this machinery, collusion and noise should not be enough to generate confusion of player 1 behaviour;

7. Reputation to work requires that the agent is afraid of loosing it in future interactions, therefore it regards long-lived agent: this is one further justification for the existence of institutions such as firms or universities besides economies of scale and scope, as a way of building and transmitting reputation from one period to the next[13]

[13] See Kreps 1990.

## References

[1] Aumann, R.: Mixed and Behavior Strategies in Infinite Extensive Games. In Dresher, M., Shapley, L., Tucker, A. (eds.) Advances in Game Theory, pp. 627-650. Princeton: Princeton University Press 1964.

[2] Billingsley, P.: Convergence of Probability Measure. New York: Wiley 1968.

[3] Billingsley, P.: Probability and Measure. New York: Wiley 1986.

[4] Dasgupta, P. 1988: Trust as a Commodity, in Gambetta, D. (ed.) *Trust: Making and Breaking Cooperative Relations*. Oxford: Oxford University Press, chapter 4, pp. 49-72.

[5] Ely J., Fudenberg D. and D. Levine 2004: When is Reputation Bad?, mimeo.

[6] Ely, J. and J. Valimaki 2003: "Bad Reputation", *Quarterly Journal of Economics*, 3: 785-814.

[7] Fudenberg, D. and D. Levine 1989: "Reputation and Equilibrium Selection in Games with a Patient Player", *Econometrica*, 50: 443-460.

[8] Fudenberg, D. and D. Levine 1989: "Mantaining a Reputation when Strategies are Imperfectly Observed", *Review of Economic Studies*, 59: 561-580.

[9] Fudenberg, D. 1992: Explaining Cooperation and Commitment in Repeated Games, in Laffont, J. (ed.) *Advances in Economic Theory. Sixth World Congress*. Cambridge, Cambridge University Press, chapter 3, pp. 89-130.

[10] Gilli, M. 1995: Modelling Strategic Interaction: a Case for Imperfect Monitoring Games. Working Paper 17/95, DAE Cambridge.

[11] Gilli, M., 1999, "On Non-Nash Equilibria", *Games and Economic Behavior,* 27: 184-203.

[12] Gilli, M. 2002: Rational Learning in Imperfect Monitoring Games. Working Paper 46/2002 Department of Economics, University of Milano-.

[13] Good, D. 1988: Individuals, Interpersonal Relations, and Trust, in Gambetta, D. (ed.) *Trust: Making and Breaking Cooperative Relations*. Oxford, Oxford University Press, chapter 3, pp. 31-48.

[14] Hornby, A. S.: Oxford Advanced Learner's Dictionary of Current English. Oxford: Oxford University Press, 1987.

[15] Kolmogorov, A.N., and S.V. Fomin: Introductory real analysis. New York: Dover Publications 1975.

[16] Kramer, R. 1999: "Trust and Distrust in Organizations: Emerging perspectives, Enduring Questions", *Annual Review of Psychology* 50: 569- 598.

[17] Kreps, D. 1990: Corporate Culture and Economic Theory, in Alt J. and K. A. Shepsle (eds.) *Perspectives on Positive Political Economy*. Cambridge, MA: Cambridge University Press, pp. 90 Ů 143.

[18] Kreps, D. and R. Wilson, 1982, "Reputation and Imperfect Information", *Journal of Economic Theory,* 27: 253-279.

[19] Kreps, D., P. Milgrom, J. Roberts and R. Wilson, 1982: "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma", *Journal of Economic Theory,* 27: 245-252.

[20] Kuratowski, K.: Topology. New York: Academic Press 1968.

[21] Lehrer, E., 1991, "Internal Correlation in Repeated Games", *International Journal of Game Theory* 19: 431-456.

[22] Mailath G.J. and L. Samuelson: Repeated Games and Repuations. Oxford: Oxford University Press 2006.

[23] Milgrom, P. and J. Roberts 1982: "Predation, Reputation, and Entry Deterrence", *Journmal of Economic Theory* 27: 280-312.

[24] Osborne, M.J. and A. Rubinstein: A Course in Game Theory. Cambridge Mass.: The MIT Press 1994.

[25] Parthasarathy, K. R.: Probability Measures on Metric Spaces. New York: Academic Press 1967.

[26] Ritzer, G. (ed): Blackwell Encyclopedia of Sociology, Oxford: Blackwell, 2007.

[27] Shenkar, O., and E. Yuchtman-Yaar. 1997: "Reputation, image, prestige, and goodwill: An interdisciplinary approach to organizational standing", *Human Relations* 50: 1361Ű1381.

[28] Strathdee, R. 2009: "Reputation in the sociology of education", *British Journal of Sociology of Education*,30: 83 - 96 .

[29] Zucker, L. 1986: "Production of trust: Institutional sources of economic structure, 1840Ű1920", *Research in Organisational Behaviour* 8: 53Ű111.

# 6 Beliefs and Stochastic Outcomes in the Repeated Imperfect Monitoring Game with Incomplete Information

Consider a Repeated Imperfect Monitoring Game with Incomplete Information (RIMG) denoted by

$$G^T(\delta, \eta, \Xi) = (N, F_i, \Xi_i, U_i, \eta_i).$$

Starting from this strategic setting, I would construct the stochastic environment that describes the objective and the subjective situation.

Let $(F, \mathcal{F}, x)$ be a probability space , where $F$ is the set of pure superstrategies, $\mathcal{F}$ the Borel $\sigma$-algebra of $F$ and $x \in \Delta(F)$ a generic probability measure on $F$. Note that in general $x$ will depend on $\mu \in Delta(\Xi)$, the probability of players' type, since the superstrategies actually played will depend on players' types. Tychonov product theorem (see e.g. Kuratowski 1968) imply that $F$ is a compact metric space in the product topology and thus $\Delta(F)$ is a compact metric space if endowed with the weak topology and with the metric being the Prohorov metric (see Billingsley 1968). Then consider the probability space $(M_i^{(\infty)}, \mathcal{H}_i, \mathbf{P}_x^i)$. The construction of this probability space involves some steps. Let $P_i \in M_i^{(\infty)}$ be a possible outcome path for player $i$ and define for each $t \in \mathbf{N}$ a mapping $Z_t : M_i^{(\infty)} \to M_i$ such that $Z_t(P_i) := O_i^t$, that is $Z_t$ is the projection of $P_i$ on its $t$ element. Consider the class $\mathcal{H}_i^C$ consisting of the cylinders, that is of the sets of the form $\{P_i \in M_i^{(\infty)} | (Z_{t_1}(P_i), \cdots, Z_{t_k}(P_i)) \in C\}$, where $k$ is an integer, $(t_1, \cdots, t_k)$ is a $k$-tuple in $\mathbf{N}$ and $C$ belongs to the Borel $\sigma$-algebra generated by $M_i^{(k)}$. Then it is possible to prove (see e.g. Billingsley 1986) that $\mathcal{H}_i^C$ is a field such that $\mathcal{H}_i$ is the $\sigma$-field generated by it. Therefore, since the $Z_t$ are measurable functions on $(M_i^{(\infty)}, \mathcal{H}_i)$, if $\mathbf{P}$ is a probability measure on $\mathcal{H}_i$, then $\{Z_t\}_{t \in \mathbf{N}}$ is a stochastic process on $(M_i^{(\infty)}, \mathcal{H}_i, \mathbf{P})$.

Now consider the probability distribution inductively defined according to the following rules: $\mathbf{P}_x^{i,(0)}(h_i^0) = \mathbf{P}_x^{i,(0)}(\xi_i) = 1$ and $\forall t \geq 1 \quad \mathbf{P}_x^{i,(t)}(h_i^{t-1}, m_i) = \mathbf{P}_x^{i,(t-1)}(h_i^{t-1}) \times [\int_{\{f|P_i^t(f_i,f_{-i})=h_i^{t-1},m_i\}} x(df)]$.

It is immediate to check that $\mathbf{P}_x^{i,(t_1)}, \cdots, \mathbf{P}_x^{i,(t_k)}$ are a system of probability distributions satisfying the Kolmogorov's consistency conditions. Therefore there exists a probability measure $\mathbf{P}_x^i$ on $\mathcal{H}_i$ such that the stochastic process $\{Z_t\}_{t\in\mathbf{N}}$ on $(M_i^{(\infty)}, \mathcal{H}_i, \mathbf{P}_x^i)$ has the $\mathbf{P}_x^{i,(t_1)}, \cdots, \mathbf{P}_x^{i,(t_k)}$ as its finite-dimensional distributions.

Now, consider the **subjective situation of player** $i$, given this stochastic environment. Player $i$ wish to maximize either $U_i$ or $u_i$ depending on being either long-lived or short-lived. In both situations player $i$ is uncertain about opponents' behavior $\phi_{-i} \in \Phi_{-i}$. To maximize utility the uncertainty relative to the opponents' random behavior $\phi_{-i}$ is equivalent to the uncertainty about opponents' pure superstrategies[14]. Therefore from $i$'s point of view the set of the states of the word is represented by $F_{-i}$ and thus a Bayesian player $i$ is endowed with a prior belief $\beta_i \in \Delta(F_{-i})$, where the probability space $(F_{-i}, \mathcal{F}_{-i}, \beta_i)$ is constructed deriving the marginal distributions from $(F, \mathcal{F}, x)$. This subjective assessment may exhibit correlation, but this does not contradict the fact that the actual strategy choices are independent. This correlation is due to $i$'s uncertainty: even if $i$ believes that the opponents choose their strategies independently, she may feel that they have common characteristics which partially resolve the strategic uncertainty. Moreover, the correlation may endogenously develop as the result of correlated observations and because of imperfect monitoring it can asymptotically persist even if the agents play independently[15].

Consider the information player $i$ collects by playing. Her beliefs are updated at time $t$ using this information, i.e.

$$P_i^t(f_{-i}, f_i). \tag{1}$$

[14] See Pearce 1984 lemma 2.

[15] See Lehrer 1991.

Note that this information depends on $f_i$, i.e. on $i$'s behavior. Moreover consider $f_{-i}$: even the opponents' strategic choices depend on player $i$ superstrategy since

$$f_{-i} \quad = \quad \{f_{-i}^t(h_{-i}^{t-1})\}_{t=1}^{\infty} \quad = \quad \{f_{-i}^t(P_{-i}^t(f_{-i}, f_i))\}_{t=1}^{\infty}. \qquad (2)$$

As a consequence of expressions (2) and (3), **player $i$'s beliefs depend on $f_i$** for two different reasons:

1. $f_i$ takes part in determining the information that $i$ receives at each stage, i.e. $P_i^t$ is a function of $f_i$, as shown by expression (2). This aspect regards the **"informative links between periods"**, that generate the possibility of experimentation, i.e. of active learning behavior;

2. $f_i$ takes part in determining the information that $i$'s opponents receive at each stage, i.e. $P_{-i}^t$ is a function of $f_i$, as shown by expression (3). I will refer to the second aspect using the label **"strategic links between periods"** since it is connected to players' behavior in repeated games.

During the play, $i$ is refining her information about opponents' behavior (passive learning), but the actual amount of information obtained depends on the superstrategy followed (active learning). For a fixed $f_i$ construct the natural filtration of the stochastic process given by the outcome path:

$$\mathcal{F}_{-i}^t(f_i) := \sigma(P_i^t(f_{-i}, f_i)),$$

where $\sigma(X(\omega))$ denotes the $\sigma$-algebra generated by the random variable $X(\omega)$. Intuitively $\mathcal{F}_{-i}^t(f_i)$ represents all the possible information about opponents' superstrategy that $i$ could collect at $t$ following the dynamic superstrategy $f_i$. In fact $\sigma(X(\omega))$ consists precisely of those events $A$ for which, for each and every $\omega$, player $i$ can decide whether or not $A$ has occurred, i.e. whether or not $\omega \in A$, on the basis of the observed value of the random variable

$X$. Formally a filtration $\{\mathcal{F}^t_{-i}\}$ is an increasing sequence of sub-$\sigma$-algebras of $\mathcal{F}_{-i}$, i.e. $\mathcal{F}^1_{-i}(f_i) \subseteq \mathcal{F}^2_{-i}(f_i) \subseteq \cdots \subseteq \mathcal{F}_{-i}$, and the natural filtration of a stochastic process $\{O^t_i(f_i)\}_t$ is the filtration generated by it in the sense that $\mathcal{F}^t_{-i}(f_i) := \sigma(O^0_i, \cdots, O^t_i)$. Finally, define $\mathcal{F}^\infty_{-i}(f_i) := \sigma(\bigcup_{t\in\mathbf{N}} \mathcal{F}^t_{-i}(f_i)) \subseteq \mathcal{F}_{-i}$. Then, for a fixed $f_i$, $P_i(f_{-i}, f_i) := \{O^t_i(f_{-i}, f_i)\}_{t\in\mathbf{N}}$ is a stochastic process adapted to the natural filtration $\{\mathcal{F}^t_{-i}(f_i)\}$, because by definition $O^t_i(f_{-i}, f_i)$ is $\mathcal{F}^t_{-i}(f_i)$-measurable. Therefore for every $t$ and for every $A \in \mathcal{F}_{-i}$ there exists a version of the conditional expectation $E[\chi_A(f_{-i})|\mathcal{F}^t_{-i}(f_i)]$, where $\chi_A$ is the indicator function for the set $A$. Indicate such a version with $\beta^t_i[f_i](A)$; then $\beta^t_i[f_i] \in \Delta(F_{-i})$ is a regular conditional probability distribution (Theorem 8.1 of Parthasarathy 1967). As the notation stress, *such a probability measure depends on* $f_i$. This probability measure represents the updated beliefs of player $i$ at time $t$, given that she is following the super-strategy $f_i$.

This discussion on players' beliefs in the RIMG can be summmarized in the following assumption:

**Assumption 4** *In the RIMG every player $i \in N$ updates her beliefs $\beta_i$ according to the following expression:*

$$\forall f_i \in F_i, \quad \forall A \in \mathcal{F}_{-i}, \quad \forall t \in \mathbf{N} \qquad \beta^t_i[f_i](A) = E[\chi_A(f_{-i})|\mathcal{F}^t_{-i}(f_i)].$$

**Remark**: this assumption is meaningful because of the existence of a regular conditional probability.