A paraître dans *The New Institutional Economics*, E. Brousseau et JM. Glachant (Eds.) Cambridge University Press, 2005.

# Game Theory and Institutions

**Thierry Pénard (CREM, Université de Rennes 1)**[1]

**October 2005**

Since the seminal book by Von Neumann and Morgenstern entitled "Game Theory and Economic Behavior" published in 1944, game theory has progressively permeated all fields of economics (industrial organization, labor, financial and international economics) and extended its influence on the other social sciences (politics, sociology, law). It has become an essential tool for studying interpersonal relationships and provides a rigorous and useful methodology for modeling and analyzing strategic decisions. Game theory methodology has also incited profound and far-reaching changes in the way markets, organizations and institutions are viewed; it has contributed to better understanding the rationale of many private and public institutions, such as contracts, franchising, insurance, certification agencies and standardization committees.

This chapter is aimed at revealing the contribution of game theory to the analysis of institutions. The goal herein is not to propose an exhaustive course in game theory (good introductions to game theory are available in Gibbons (1992) or Osborne and Rubinstein (1995)), but rather to focus on the main results and lessons that can be drawn from game theory with respect to institutions[2].

Institutions can basically be characterized in two ways. They can first be defined as the set of fundamental political, social and legal ground rules that establish the bases for production, exchange and distribution (Davis and North, 1971). From this perspective, institutions appear as the rules of the game imposed on all economic actors. Yet, an institution may also be

[1] CREM, Université de Rennes 1, 7 place Hoche, 35065 Rennes Cedex, France, **thierry.penard@univ-rennes1.fr**

[2] This chapter will just consider non-cooperative games and leave aside cooperative games and evolutionary games; the latter will be developed in Chapter 15 (J.-P. Platteau) within this book.

considered in a more endogenous manner, in being defined as a player that can interact with the game's other strategic players, albeit with a specific status since it can influence or modify the rules of the games and directly affect the outcome, e.g. by helping players to coordinate their strategies or select an equilibrium. This chapter will adopt the latter approach in order to examine how institutions can be modeled within strategic games and what can be learned about the role of institutions.

For this purpose, it is helpful to define what exactly a strategic game is: a strategic game is characterized by its players (number of players, their features), the set of strategies assigned to each player, players' information set and payoff function (or utility function), and lastly by the sequence of moves (and scope of the game). This chapter is intended to show how institutions interact with these various dimensions. Institutions can actually extend or restrict the set of strategies; they can also modify the quality and quantity of information available to players or change the sequence of moves or number of periods during which the game is played. In basic terms, institutions have three main purposes: 1) facilitate coordination and commitment; 2) improve information; and 3) promote cooperation. In all three cases, the rationale behind institutional intervention is often aimed at improving efficiency in interpersonal relationships. From the viewpoint of game theory therefore, institutions appear as multi-purpose and multifaceted efficiency-enhancing devices.

The chapter has been organized into four sections. The next section will present coordination and trust games that emphasize the role of institutions as commitment devices. Section 2 will then consider signaling games in the presence of incomplete information and show that institutions may be helpful in improving information. Section 3 is devoted to repeated games, within which institutions are used as cooperation-enabling devices. Finally, Section 4 illustrates the multiple roles institutions play through two case studies: first, the case of the medieval merchant guilds inspired by the paper from Greif, Milgrom and Weingast (1994); and second, the case of eBay, the well-known online auction Website that has recently been receiving widespread attention from academics.

## Section 1: Coordination, commitment and institutions

Subsection 1.1 will examine the role of institutions in a classical coordination game, while Subsection 1.2 presents a more general framework based on a principal-agent relation in order

to introduce the problem of credibility in player announcements. Subsection 1.3 will illustrate how an institution can become a commitment device.

### 1.1 Coordination games

For game theorists, a strategic decision always constitutes a conditional decision, based not only on the intrinsic characteristics and preferences of the decision-maker, but also on his beliefs and expectations regarding the decisions of other decision-makers with whom he interacts. The predictable outcome of a strategic game is displayed as a Nash equilibrium; this corresponds to a stable situation in which no decision-maker has an incentive to change strategy given the strategies chosen by the others. A Nash equilibrium thus appears as a self-enforcing or self-organized state: i.e. a spontaneous order without any apparent link to institutions and organizations. What game theory does not mention however is how decision-makers reach this stable state or how they coordinate themselves around this outcome. The first lesson drawn from game theory is essentially that the rationale of many institutions and organizations is to help decision-makers coordinate their strategies on an equilibrium, especially when multiple equilibria exist. Institutions can facilitate the convergence of beliefs and expectations towards the same equilibrium. This role of coordination has been well illustrated in the following strategic game.

Consider a market where two firms compete to impose their own standard. Standardization (achieving a common standard) is collectively optimal because it stimulates demand through a network externality (as consumers highly appreciate compatibility). Yet each firm has elected to sponsor a different technology: Firm 1 promotes Standard 1, while Firm 2 promotes Standard 2. If Standard 1 were implemented, then Firm 1 would benefit to a greater extent thanks to its technological advantage in this standard and ownership of some patents that could be licensed. Table 1 shows the payoff associated with standardization on Technology 1: Firm 1 obtains a profit of 4, whereas Firm 2 obtains just 2. If Technology 2 were adopted, then the payoffs would be reversed: Firm 2 would receive more profit than Firm 1. In the absence of standardization (a *standards war*), both firms would obtain less profit than what they could obtain with standardization: in particular, they would receive no gain if each firm chose the rival standard (i.e. when Firm 1 adopts Standard 2 and Firm 2 adopts Standard 1)

and a profit equal to 1 when each firm opts for its own standard. This strategic game is called a "*coordination game*"[3].

In such a game, encounter two Nash equilibria in pure strategies[4]: a *de facto standardization* on Technology 1 and a *de facto standardization* on Technology 2. Which of these two Nash equilibria will prevail? In the absence of institutions, it is likely that neither will emerge: the firms may fail to coordinate on a common standard and decide to launch a standards war, which could lead to an inefficient and unstable situation.

This risk of coordination failure explains the emergence of private institutions (in the form of a Working group, or a National or International Standards Body) that help firms discuss and converge on either one of the sponsored technologies or a hybrid technology that mixes some elements of the two competing technologies. For example, the telecommunications industry is familiar with such standardization committees and forums (European Telecommunications Standards Institute, GSM Forum and 3G Forum for wireless telecommunications). Similar private institutions exist in other industries (electronics, computer science, aviation, etc.).

Another solution to the coordination problem is a mandated standardization, whereby a public authority (i.e. a public institution) imposes an equilibrium upon all actors. Such was the case for the third generation of wireless telephony in Europe, as the European Commission (with the agreement of every member country) decided to mandate a harmonized standard called W-CDMA; this standard was competing with another standard, CDMA-2000 (see Gandal, Salant and Waverman, 2003).

**Table 1: A standardization game**

| | | Firm 2 | |
|---|---|---|---|
| | | Standard 1 | Standard 2 |
| Firm 1 | Standard 1 | (4,2) | (1,1) |
| | Standard 2 | (0,0) | (2,4) |

In a strategic game, coordination issues are frequently linked with commitment problems, since players often seek to coordinate themselves by means of announcements, threats or

[3] A well-known version of such a game is the "battle of sexes", in which a man and a woman are trying to decide on evening entertainment. They prefer spending the evening together, but each has an entertainment preference he or she would like to impose on the other.

[4] An equilibrium in mixed strategies also exists, where Firm 1 chooses Standard 1 with a probability of 4/5 and Firm 2 adopts Standard 2 with a probability of 4/5.

promises. Another rationale for institutions is to make these commitments credible and efficient.

### 1.2 The commitment problem

A decision-maker may seek to manipulate a strategic situation by unilaterally changing some rules of the game, particularly the sequence of moves. For example, instead of moving simultaneously, one player can obtain a strategic advantage by moving either before or after the other players (depending on the nature of the game). In the previous coordination game, moving first provides a strategic advantage (it enables one firm to launch its technology ahead of time and to market its product earlier). The other firm will then have no other choice but to adopt the rival standard (*de facto standardization*).

A decision-maker may obtain a similar advantage by committing himself to take a specific action in the future (or not to take such action); for example, Firm 1 can commit itself to playing Technology 1 regardless of the decision of Firm 2 (or, on the other hand, commit to never playing Technology 2). Dixit and Nalebuff (1993) defined a commitment as a response rule that prescribes a response to the decisions of the other players. This rule is generally communicated at the beginning of the game. There are two broad categories of response rules, according to Dixit and Nalebuff: threats and promises. A threat is a response rule that punishes others who fail to cooperate: *"I will punish you if you don't play the specified action."* A promise is an offer to reward someone who cooperates: *"I will reward you if you play the specified action."* The goal of these announcements is to influence the actions of the others. To return to our standardization game example, Firm 1 can either threaten to trigger a price war if Firm 2 does not choose Technology 1, or promise a reward if Firm 2 adopts Technology 1. Through these commitments, a player can expect coordination on his most preferred equilibrium.

Such a commitment however can only influence the other player if the latter is convinced that threats or promises will indeed be carried out. In the standardization game, suppose that Firm 1 announces that it is committing to Technology 1. If Firm 2 considers that this announcement is of no value, then it has no reason to abandon its own technology. It can instead decide to launch its standard in order to force Firm 1 to adopt Standard 2, which implies that a pre-announcement has no strategic effect if it cannot be trusted or if the other

players doubt the credibility of such commitments. Game theory has highlighted these issues and incites us to rethink many institutions and organizations as credible commitment devices.

Technically speaking, Selten (1975) underscored the fact that some Nash equilibria may be inconsistent or fallacious if they rely on non-credible commitments. He proposed a more restrictive concept of equilibrium: the concept of subgame-perfect equilibrium. A Nash equilibrium satisfies the criterion of subgame-perfection if each strategy yields a best response to the strategies of other players in all of the subgames[5]. Players should not normally reach subgames that are out of equilibrium; but if such a subgame should were to be reached, each player would implement the prescribed strategy. If a player shows no interest in carrying out the prescribed strategy, then his strategy relies on non-credible commitments (reward or threat), and the other players will not take such a strategy into consideration when deciding on their own strategies to implement.

This idea can be illustrated by the following game: Consider two players, where one is the principal who delegates decisions to a second player, called the agent. This describes the classic principal-agent relation. The agent (Player 1) must undertake a project or action that has an impact on the utility or payoff of the principal (Player 2).

Now assume that the principal and the agent have divergent interests on the optimal actions to be undertaken. The agent should act in the interest of the principal, but he has the opportunity to cheat by choosing actions that yield higher utility for himself. The alternative for Player 1 is thus to behave either honestly (respecting the interest of Player 2) or opportunistically (cheating). Consider the decision tree and payoffs of this game, as displayed in Figure 1.

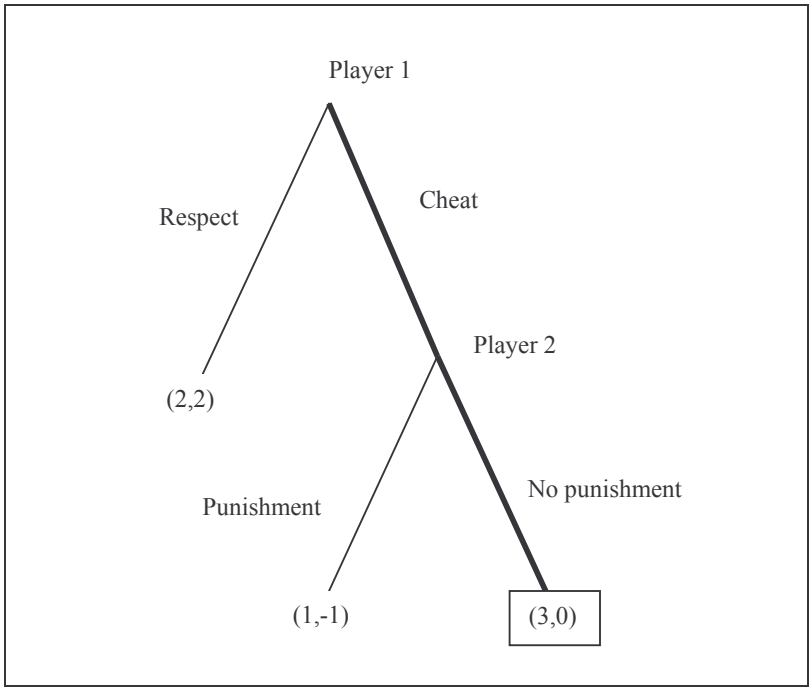
If Player 1 respects the interest of the principal, the game ends and each player receives a gain of 2. If Player 1 decides to behave opportunistically, then Player 2 can either punish him or play *laissez-faire* (i.e. tolerate the cheating). Punishing Player 1 is costly and leads to a payoff of -1 for Player 2. But Player 1 also receives a smaller gain in comparison with the situation in which he does not cheat. Lastly, if Player 2 does not punish Player 1, then Player 2 obtains a zero payoff and Player 1 reaps all of the benefits (a payoff of 3).

Let's consider the threat by Player 2 to punish Player 1 if the latter betrays his trust. Can this threat deter Player 1 from cheating? Player 1 is apparently better off being honest than being punished. Yet being honest cannot be part of a subgame-perfect equilibrium, since the threat of punishment by Player 2 is not credible. When Player 2 is in a situation to carry out his threat, he would not be inclined to do so. *Laissez-faire* proves to be a better strategy than

[5] A strategic game contains as many subgames as decision nodes. For each node, the subgame corresponds to the sequence of all actions or decisions that follow this node.

punishing Player 1 because punishments are costly to implement. Since the threat of reprisal is not credible, Player 1 will ignore this threat and cheat. The predictable outcome of this strategic game is thus (cheat, *laissez-faire*). In anticipating this outcome, the principal is likely to give up playing with the agent. The absence of credible commitment can therefore undermine a valuable relationship, in which the potential value created equals 4 (2 for Player 1 plus 2 for Player 2 when both cooperate).

**Figure 1: A principal-agent game**




How then can credibility be restored and the principal be convinced to continue interacting with the agent? Institutions might provide a response to the lack of credibility encountered in many strategic situations, such as the principal-agent relationship.

### 1.3 Institutions as commitment devices

One key function of many institutions and organizations is to help decision-makers credibly commit themselves. For Dixit and Nalebuff (1993), various ways exist for a decision-maker to establish credibility:

- First, he can *burn bridges behind him*: This means that he would eliminate all other options or actions, except those he wishes to exercise. This strategy has long been used in military tactics. "*In 1066, William the Conqueror's invading army burned its own ships, thus making an unconditional commitment to fight rather than retreat. Cortes followed the same strategy in his conquest of Mexico. Upon his arrival in Cempoalla, Mexico, he gave orders that led to all but one of his ships being burnt or disabled. Although his soldiers were vastly outnumbered, they had no other choice but to fight and win"* (Dixit and Nalebuff, 1993, p. 153). A similar strategy consists of *cutting-off communications* with the other players in order to deny oneself any opportunity to back down (i.e. to change the initial move).
- Second, he can *establish and utilize a reputation*: This strategy is only viable if the game is repeated, in which case the initial cost of building a reputation is more than offset by the benefit the player can reap from his strengthened credibility.
- Third, he can *write contracts* to reduce the cost of punishment and amplify the loss incurred by the cheater. A contract represents a means for partners to make their relationship enforceable. If one partner seeks to breach the terms of the contract, then the other party can sue him and force respect of the contract. Signing a contract eliminates the possibility of cheating even though contracting is a costly process (writing and negotiating the terms of a contract)[6].

Institutions definitely provide the means for a decision-maker to *establish a reputation, cut off communications with others, burn bridges* or *write contracts*. In a principal-agent relation for example, the principal can hire a supervisor in charge of punishing the agent in case of opportunistic behavior. The supervisor would then receive a fixed payment whether he punishes the agent or not. The threat of punishment thus becomes credible; in other words, the creation of hierarchy (i.e. an organization) overcomes the problem of commitment (*the principal cuts off direct communications with his agent*). It is also easier for a principal to build a reputation and then exploit it through setting up a hierarchical organization, thanks to the longevity of an organization.

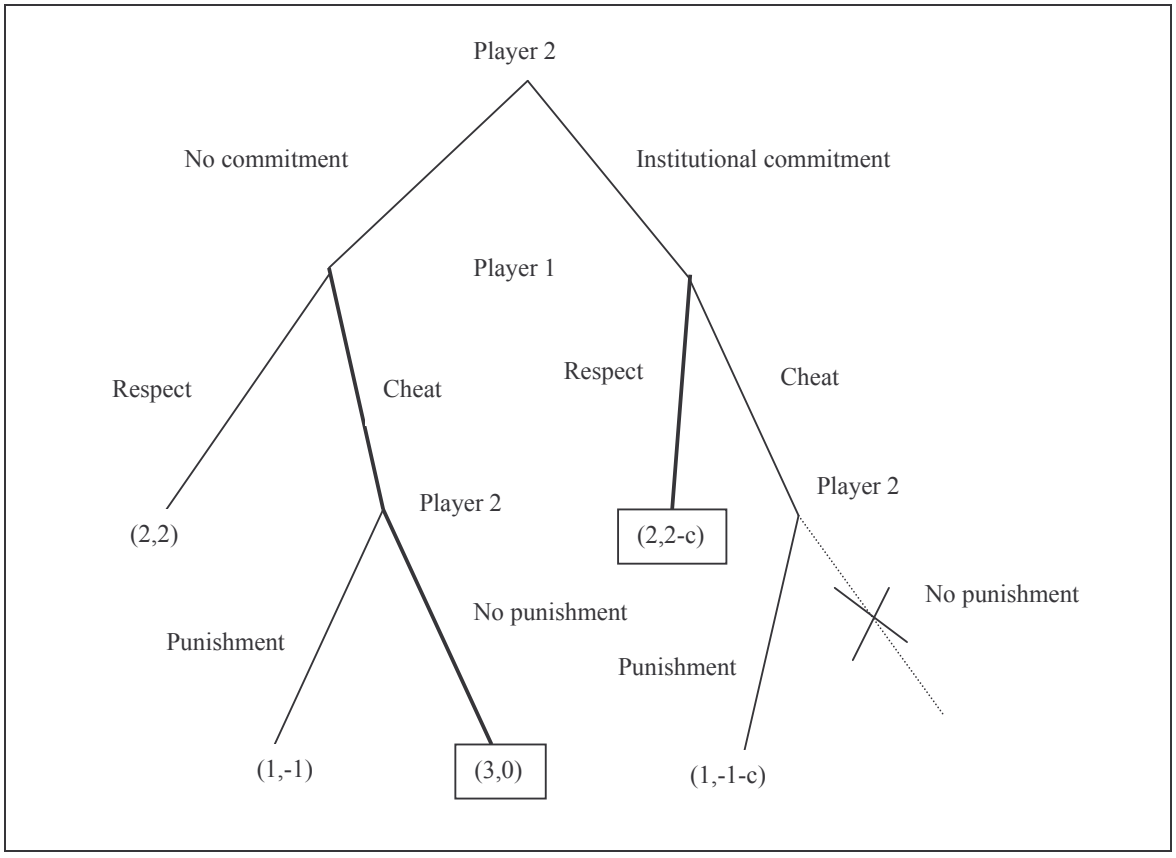
Insurance companies provide another example of institutions that enable *burning bridges*. When the principal uses the service of an insurance agent, he obtains a payoff guarantee (i.e. a fixed gain). Regardless of the agent's behavior, the principal is certain to earn a payoff of 2-P,

[6] However, one important issue is to prevent against renegotiations. The possibility of renegotiating the contract could weaken the credibility of commitment. To avoid this risk, some additional devices may be required, such as employing a neutral party with independent interests or building a strong reputation of no-renegotiation.

where P is the premium paid to the insurance agent. If Player 2 receives less than 2-P from Player 1, then the insurer compensates for the differential and is responsible for punishing Player 1 (obtaining reimbursement for the damages caused by Player 1).

In order to more formally represent the impact of an institutional device in the previous game, let's suppose that the principal can initially use an institutional device (contract, insurance, hierarchy, etc.) at a cost of $c$ (see Figure 2). This device enables him to credibly commit himself to punish his partner in case of cheating (by eliminating the "no punishment" option), so that Player 1 will behave honestly to avoid reprisals. When institutions are introduced into the principal-agent relation, the subgame-perfect equilibrium becomes (respect, punishment if cheating). Player 2 then obtains a gain of 2-c and Player 1 a gain of 2. Finally, Player 2 will invest in the institutional device if the cost $c$ is not excessive: by implementing the institutional solution, Player 2 anticipates a payoff of 2-c instead of 0 (without any device). If $c<2$, then commitment constitutes the best strategy for Player 2.

**Figure 2: A principal-agent relation with institutional commitment**



Let's now consider the second rationale for institutions: the reduction of imperfect information in interpersonal relationships.

## Section 2: Imperfect information and institutions

Game theory has strongly highlighted the importance of information in interpersonal relationships and has largely contributed to the current focus of microeconomics on information issues. Subsection 2.1 will consider a principal-agent relation submitted to imperfect information and will underscore the source of inefficiency in this context. Subsection 2.2 will show how institutions can overcome imperfect information.

### 2.1 A principal-agent game under asymmetric information

Consider the principal-agent game presented in the previous section and assume now that the agent (Player 1) can be highly-competent (type H) or incompetent (type L). These two types will influence the payoff of the principal (Player 2) since the principal can expect to obtain a higher gain with a type-H agent than with a type L.

Assume that the agent is type H with probability $\alpha$ and type L with probability $1-\alpha$. The true type is of course private information held by Player 1 (the choice of the nature is only observed by the agent). Given his type, Player 1 can decide to either cheat or be honest. Player 2 does not observe the action chosen by Player 1 (respect or cheat) but receives a signal imperfectly-correlated with Player 1's behavior. Player 2 can in fact receive three kinds of signals: positive, neutral and negative. When a negative signal is received, Player 2 can perfectly infer opportunistic behavior by an incompetent agent. Likewise, a positive signal indicates honest behavior by a highly-competent agent. However, a neutral signal is compatible with either honest behavior by a type L or opportunistic behavior by a type H. In this situation, the deterrence of cheating is more complex because even if Player 2 holds credible threats of reprisals, he must apply them cautiously. When observing the neutral signal, Player 2 can make two kinds of errors: punishing an honest agent, or playing *laissez-faire* with an opportunistic agent.

In such a game, one can restrict the set of predictable outcomes to the perfect Bayesian equilibria (PBE). Two potential classes of PBE emerge:

- Pooling equilibria, in which Player 1 sends the same signal regardless of his type (here, a neutral signal). The goal of this strategy is to maintain Player 2's initial beliefs (i.e. to maintain his initial ignorance);

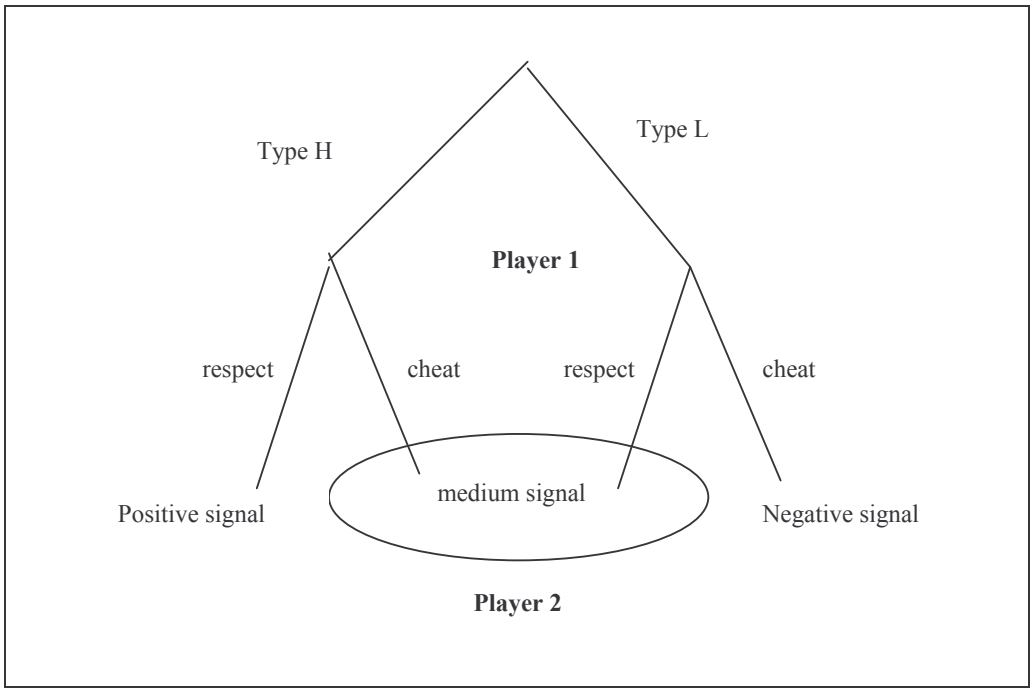
- Separating equilibria, in which Player 1 sends a different signal depending on his type: a positive signal when he is highly competent, and a neutral signal when incompetent. The goal of this strategy is to perfectly inform Player 2[7].

The first type of equilibrium is observed when Player 2 holds *a priori* beliefs that lead him to adopt a *laissez-faire* strategy (beliefs that $\alpha$ is low). Player 2 thus considers that the probability of a neutral signal stemming from an opportunistic agent is low and finds it optimal to react with a *laissez-faire* attitude. Consequently, it is in the interest of the agent to maintain the principal's *a priori* beliefs, which would allow a type-H agent to act opportunistically without being punished. In a pooling equilibrium, the rent captured by the agent (called "informational rent") represents a source of inefficiency. If this inefficiency becomes too great, it is then possible that the principal will put an end to the relationship in order to protect himself against opportunistic behavior by the agent.

The second type of equilibria (separating equilibria) appears when Player 2 holds *a priori* beliefs that $\alpha$ is high. In this case, Player 2 will punish Player 1 in response to a neutral signal. To avoid being systematically punished (in particular when he is of type L and behaves honestly), it is in the interest of Player 1 to reveal his private information. The signal received then enables Player 2 to revise his *a priori* beliefs. In a separating equilibrium however, the credibility of signals proves important: Player 1 must have no incentive to send a neutral signal when his type is H. The source of inefficiency may stem from either the absence of credibility or the cost incurred to make the signal credible.

[7] Hybrid or semi-separating equilibria can also exist, whereby the type-H agent randomizes between the two strategies - respect and cheat.

**Figure 3: Principal-agent game under imperfect information**




## 2.2 Institutions as information-enhancing devices

The role of institutions is to restore efficiency in relations characterized by imperfect information. Institutions can facilitate the transmission and dissemination of information among players; they can also incite players to reveal their private information or deter them from retaining information. Moreover, they can make signals more credible. The following examples serve to illustrate these various functions.

In the used car market, buyers do not perfectly observe the quality of cars; they have a positive probability of acquiring a poor-quality car. This problem was highlighted by Akerlof (1970) in his seminal paper entitled "The market for lemons" and leads to a phenomenon of adverse selection. Buyers react to imperfect information on quality by lowering their willingness to pay; however, this eliminates all high-quality cars for which sellers have a high reservation price. In the end, all quality cars are withdrawn from the market and only poor-quality cars can be found. Many institutions have been set up to overcome adverse selection problems; for example, the obligation of a technical verification before selling a car is an institutional device that partially handles this problem. The presence of car dealers (intermediaries) in the used car market can also be explained by the quality guarantee they provide for buyers.

A certification agency is another kind of institution created to reduce information asymmetry. Certification agencies have expanded their activity in recent decades, especially in the food processing industries. On the buyer's side, the role of these institutions is to facilitate screening between good and bad products. On the seller's side, they allow for credible quality signals of the seller's products.

Another institution that tackles the problem of information asymmetry is franchising. A franchise contract is an agreement in which a franchisee obtains the right to operate the business concept of a franchiser by agreeing to pay royalties, which are generally based on revenues. Franchising constitutes an organizational innovation in retailing and has tremendously grown in all developed countries since 1970. Its success can be attributed to its ability to reduce moral hazard in the agency relations that tie the owner of a business concept and the managers in charge of operations. The owner cannot perfectly monitor all his managers; he cannot be present at all times to ensure that they are furnishing the appropriate effort. Franchise contracts that provide the manager with the status of residual claimant deter shirking by the manager because they align the interests of the franchisee with those of the franchiser. A considerable body of literature has focused on the rationale of franchising as an institution that mitigates moral hazard issues and enhances business performance (Lafontaine, 1993; Mathewson and Winter, 1985).

The role of institutions is not only to overcome problems of adverse selection or moral hazard, but also to facilitate mutual cooperation in repeated relationships.

**Section 3: Repeated cooperation and institutions**

Many relationships in the real world are repeated over a finite time horizon (if people know exactly when the relationship will end) or, more generally, over an infinite horizon (should there be some uncertainty about the ending time). The theory of repeated games provides the simplest approach for considering the effects of long-term interactions. Subsection 3.1 will present the general framework of repeated games and Subsection 3.2 will highlight the role of institutions as a facilitating device for enforcing cooperation within a repeated relationship.

### 3.1 Prisoners' dilemma and repeated games

The prisoners' dilemma is the most famous and certainly the most useful game for analyzing social relationships. Let's consider a symmetric prisoners' dilemma game. Two players have the opportunity to cooperate or cheat. If both cooperate, they will earn more than if both decide to cheat ($r_c > r_{nc}$). If one cooperates and the other cheats, then the cheater earns more than the cheated ($r_d > r_p$). In this game, cheating (no cooperation) is a dominant strategy for both players and constitutes the predictable outcome. However this Nash equilibrium is sub-optimal since both players will be better off if they cooperate.

**Table 2: Prisoners' dilemma**

| | | Player 2: Cooperation | Player 2: No cooperation |
|---|---|---|---|
| Player 1 | Cooperation | $r_c, r_c$ | $r_p, r_d$ |
| | No cooperation | $r_d, r_p$ | $r_{nc}, r_{nc}$ |

How can trust be restored? In fact, cooperation can emerge from such a relationship if the game is repeated. Cooperation may be a Nash equilibrium of an infinitely-repeated game because actions beyond the short-run self-interest of a player may be consistent with his long-run self-interest. Two underlying effects may intervene:

- A reputation effect: "*One cooperates so as to incite the other to cooperate in the future"*
- A deterrent effect: "*One does not cheat because of fear of reprisal*".

To better understand these two effects, let's assume that the two players agree to follow decision rules that consist of cooperating as long as the other player has chosen cooperation during past periods. In case of cheating, they agree to revert indefinitely to the one-shot Nash equilibrium (no cooperation). The deterrent effect here lies in credibly threatening to punish the other player forever. This strategy is called a "trigger strategy" (Friedman, 1971)[8]. Under which conditions can a cooperative outcome be enforced by trigger strategies?

The expected value of respecting cooperation, given the other player is behaving honestly, is given by:

[8] Players can also adopt other strategies in order to sustain cooperation. For example, Abreu (1986) showed that *stick and carrot* strategies, where punishments consist of an extremely severe one-period sanction (stick)

$V^c = \sum_{t=0}^{\infty} \delta^t r_c = \left(\frac{1}{1-\delta}\right) r_c$, where $\delta$ is the discount factor,

and the expected value of cheating is given by:

$$V^d = r_d + \sum_{t=1}^{\infty} \delta^t r_{nc} = r_d + \left(\frac{\delta}{1-\delta}\right) r_{nc}$$

where the cheater obtains an instantaneous payoff of $r_d$ and then receives $r_{nc}$ indefinitely.

Each player will elect to cooperate if cooperating yields more than cheating over the long run, i.e. if:

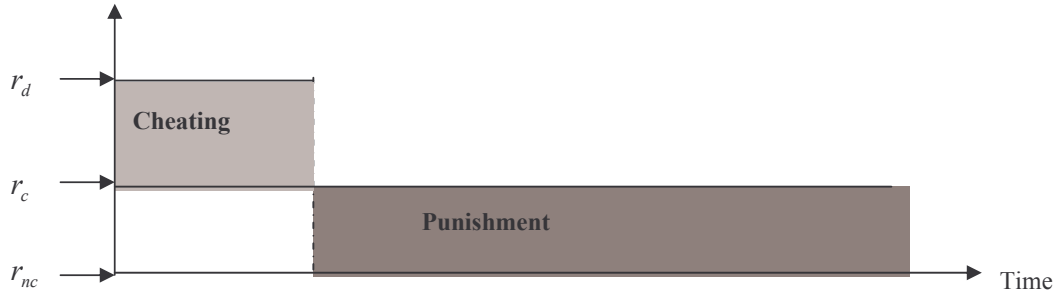
$$V^c > V^d$$

This incentive condition can be rewritten as follows:

$$r_d - r_c \leq \left(\frac{\delta}{1-\delta}\right)(r_c - r_{nc}) \qquad \text{(C1)}$$

The net benefit of cheating (the left-hand term) must be more than offset by the opportunity cost of being punished (the right-hand term). This tradeoff is well illustrated in Figure 4, where each player must balance the expected gain of betrayal with the cost of reprisal.

**Figure 4: The tradeoff between short-run temptation to cheat and long-run costs under trigger strategies**




This theoretical framework suggests that many institutions may be cooperation-enhancing devices capable of relaxing incentive conditions. Subsection 3.2 will examine how institutions can practically facilitate cooperation and will illustrate this role through the example of cartels (price-fixing agreements).

followed by an indefinite reversion to the cooperative outcome (carrot), are at least as efficient in sustaining cooperation as any more complex strategy.

### 3.2 Institutions as cooperation-enhancing institutions

Let's now consider the incentive condition (C1); it can be reformulated in terms of a threshold discount factor, i.e.:

$$\delta \geq \frac{(r_d - r_c)}{(r_d - r_{nc})} \qquad \text{(C2)}$$

The right-hand term corresponds to the minimum value of the discount factor below which players cannot sustain a cooperative outcome. If the discount factor lies above this threshold (which means that players' preference for the future is sufficiently high), then cooperation is likely to emerge as a Nash equilibrium of the repeated game.

The value of the threshold discount factor appears as a measure of the likelihood of a long-run cooperative relationship. The lower the threshold discount factor, the higher the propensity to cooperate. Note that this threshold value increases in both $r_d$ (the cheating payoff) and $r_{nc}$ (the punishment payoff), while it decreases in $r_c$ (the cooperation payoff).

From this perspective, an institution is likely to facilitate cooperation if it acts to decrease the threshold discount factor. To be more precise, an institution can facilitate cooperation by increasing the severity of punishments (by reducing the punishment payoff $r_{nc}$) or by reducing the cheating benefits (increasing the cooperative payoff $r_c$ and/or reducing the cheating payoff $r_d$).

Moreover, if players face imperfect monitoring (e.g. if they do not perfectly observe the past decisions of their partners), then cooperation will be more difficult to sustain, since a cheater is not systematically detected and may escape reprisals. The role of institutions might then be to improve the detection of cheating by facilitating the monitoring of partners' behavior and increasing the speed of reprisals.

In sum, repeated games capture an important aspect of interpersonal relations: most economic relationships are embedded within *relational contracts*, in which cooperation or agreement is enforced by reputation and trust rather than by the courts (Gibbons, 2000, 2001). These *relational contracts* are omnipresent in markets and inside firms (in labor relations, supplier relations, pricing agreements, etc.). They appear as a set of informal rules and codes of conduct backed by formal institutions or formal contracts to enhance performance in long-run relations (Kreps, 1990). This notion can be clearly illustrated by cartels and price-fixing agreements.

A cartel is a kind of institution designed to facilitate cooperation between competing firms. An oligopoly faces the same dilemma as players in the prisoners' dilemma: collectively, they should relax price competition, yet individually each firm has an incentive to undercut the rival's prices in order to increase market share and profit.

The role of cartels is to facilitate enforcement of a cooperative outcome. Considerable research has been devoted to understanding how firms organize collusion (Salop, 1986; Jacquemin and Slade, 1989). Much of the recent research is aimed at responding to the criticism of Fisher (1989) that too little effort has been dedicated to this issue: *"I think game theoretic oligopoly theorists are studying the wrong thing. They are accumulating a wealth of anecdotal material about one-shot oligopoly games when what one wants to know concerns the factors that lead to the collusive equilibrium to be chosen in repeated games"* (Fisher, 1989).

The facilitating factors (factors leading to the collusive equilibrium) emphasized by Fisher refer in part to institutions and organizations. In all cartels detected and sued in Europe over the recent decades, antitrust authorities actually had obtained evidence of committees, agencies or associations set up in which firms discussed their intentions (price, capacity, production, etc.) and exchanged information. These facilitating devices not only improve market transparency, but also enable firms to more efficiently negotiate collusive agreements and decision rules (especially the rules of punishment).

As an example, concerning the cement decisions (1994), the European Commission suspected the main European cement firms of having organized a cartel in order to restrain competition in Europe. The objective of the agreement was to regulate capacities and sales, as well as the "non-transshipment to the home markets of the Member States". After several investigations and audits, the Commission found that cement firms had created various institutions to sustain this cooperative agreement; they had set up an export committee and task force in charge of examining and preparing different dissuasive and persuasive measures to stabilize the market (*stick and carrot policy*). The main objective of this policy was to eliminate imports from Central Europe and Greece. These measures included the creation of the Joint Trading Company, an institution whose function was "*to capture the orders of the principal export markets supplied by countries threatening the stability of the Member Countries' markets; to purchase cement and clinker from countries threatening the stability of the Member Countries' markets; to market the quantities purchased through market intervention; to export cement and clinker to countries threatening the stability of the Member Countries; a committee was to designate the markets for purchases, marketing and exports and to set the*

*purchase and selling prices*[9]". All these institutions were intended to enhance the collusive agreement.

Along the same lines, the "UK Agricultural Tractor Registration Exchange" antitrust is most insightful. On January 4, 1988, the United Kingdom trade association of manufacturers and importers of agricultural machinery notified the existence of an information exchange agreement; this agreement concerned the registration of new tractors and permitted members to perfectly identify the volume of retail sales and market shares of the eight biggest manufacturers and importers of agricultural tractors on the United Kingdom market (sales by model, geographical area, weeks, etc.). The collection and dissemination of sales data were conducted by a specific institution, called the Agricultural Engineers Association Ltd (AEA). Even if firms argued that this information exchange was pro-competitive, the European Commission forbade the agreement because it suspected that this institutional arrangement was designed to relax price competition and promote a self-enforcing cartel:

*"The high market transparency between suppliers on the United Kingdom tractor market which is created by the Exchange takes the surprise effect out of a competitor's action, thus resulting in a shorter space of time for reactions with the effect that temporary advantages are greatly reduced."*[10]

Any temptation to cheat could thus be deterred thanks to this information agreement because each cheater was sure to be detected and punished. The rationale of this institution (AEA) was actually to promote price cooperation between competitors.

Evidence on the self-enforcing role of institutions in collusive agreements has also been provided by Dick (1996), who conducted an empirical study on the determinants of cartel longevity using a database on US exportation cartels. These cartels are legal under the Webb-Pomerene Act, but firms cannot enforce them via the courts. Dick found that these cartels were more stable when firms had created a common sales agency (an institution) that enabled discussing and monitoring their sales.

The next section illustrates the various roles of institutions (coordination, commitment, information and cooperation-enhancing devices) through two detailed case studies.

[9] Decision issued on November 30, 1994, *Ciment*, No. L343, December 30, 1994.

[10] Decision issued on February 17, 1992, *UK Agricultural Tractor Registration Exchange*, No. L68, March 13, 1992.

## Section 4: Rethinking institutions - Two case studies

The first case study is owed to Greif, Milgrom and Weingast (1994), who studied in detail the medieval merchant guilds in Europe (Subsection 4.1). The second one is more recent and concerns eBay, the well-known online auction Website (Subsection 4.2). These two examples will further our understanding of how institutions improve efficiency in interpersonal relationships and emphasize the multi-faceted roles of institutions.

### 4.1 Merchant guilds as a contract-enforcing device in medieval trade

In a survey on game theory applied to economic history, Greif (2000) noted that explicit models of repeated games with imperfect information have been successfully used to understand historical institutions[11]. In particular, Greif (1989, 1993) examined the agency relationships in Mediterranean and European long-distance trade between merchants and their overseas agents. Greif analyzed the institutional rules applied by Jewish Maghribi traders, who operated during the 11$^{th}$ century in the Muslim Mediterranean to overcome moral hazard by agents tempted to act opportunistically and steal a part of the merchants' goods.

Greif, Milgrom and Weingast (1994) examined the role of merchant guilds in medieval European trade. A guild was *"an administrative body that supervised the overseas operations of merchant residents of a specific territorial area and held certain regulatory powers within that territorial area (a town or city)"* (Greif *et al.*, 1994). These guilds were organized at the city level (in Italy) or at a regional level (German Hansa). According to Greif *et al.*, these guilds should not be seen as a monopoly that restrained trade, but rather as an institution that secured trade and solved contractual problems. During medieval times, rulers (local authorities) were indeed tempted to abuse the property of alien merchants; such abuse was detrimental to the development of trade. The lack of security was costly and inefficient for both traders and rulers.

Merchant guilds credibly threatened to boycott the city or territory of any cheating ruler. Guilds had the ability to coordinate an embargo and ensure traders' compliance with boycott decisions. They facilitated communications between traders and collected information about conflicts and abuse. Without this institution, merchants failed to enforce any embargo. Many

[11] *"Apart from indicating the empirical relevance of repeated games (with and without imperfect monitoring), these studies demonstrate the extent to which game theoretical analysis can highlight diverse aspects of a society, such as the interrelations between economic institutions and social structures"*, Greif (2000).

merchants were not informed of what happened to other merchants or had no incentive to respect an embargo; consequently, rulers never committed themselves to securing alien merchants' rights.

To better understand the impact of merchant guilds, let's consider the following model, inspired by Greif *et al.*, in which M individual merchants have the opportunity to trade with a city. The value of trade is $V(x)$, where $x$ is the number of active merchants. Let $c$ be the variable costs incurred by the ruler of the city for the purpose of securing traders. $V(x)(1-c)$ is the net value of trade. Let $x^*$ be the efficient level of trade ($x^* = \text{argmax } V(x)$). Assume that the ruler finances his service by charging a toll $\iota$ on the volume of transactions.

The ruler faces two possible decisions: to respect his obligation of guaranteeing security or to cheat. An opportunistic behavior for the ruler would thus consist of saving on protection costs. If the ruler upholds his promises, his payoff is $V(x)(\iota - c)$; if he decides to undermine a fraction $e$ of traders, his payoff becomes $V(x)(\iota - (1-e)c)$. For those traders not cheated, their payoff is $(1-\iota)V(x)/x$, while cheated traders incur a loss equal to $\iota V(x)/x$.

Now consider the infinitely-repeated games, in which at each period the ruler has to choose the level of security provided to the merchants, and the merchants decide whether to continue trading with the city or boycott it. How can merchants force the ruler to play honestly? Each merchant can unilaterally threaten to boycott the city forever if he is cheated (trigger strategy according to *Friedman*). This bilateral reputation mechanism however cannot sustain an efficient trade level since the ruler always maintains an incentive to violate the rights of a few traders. To prove this result, consider the incentive constraints of the ruler when the trade level is efficient ($x^*$ merchants). Mistreating $e$ merchants yields: $r_d = [\tau - c(1-e)]V(x^*)$, whereas cooperating yields a payoff of: $r_c = [\tau - c]V(x^*)$ (the cooperative payoff). An opportunistic ruler can obtain a net benefit of $r_d - r_c = ceV(x^*)$. But the reprisals of cheated merchants over future periods would lead to a payoff of $r_{nc} = [\tau - c]V(x^*(1-e))$ instead of $r_c = [\tau - c]V(x^*)$. The cost of reprisals per period is therefore equal to $r_{nc} - r_c$.

Given the unilateral threat of a boycott, according to condition C2 (see Subsection 3.2), a ruler will be deterred from cheating if for any $e \in [0, x^*]$:

$$\delta \geq \frac{ceV(x^*)}{[\tau - c(1-e)]V(x^*) - [\tau - c]V(x^*(1-e))}$$

The right-hand term represents the threshold discount factor, where $r_{nc}$, $r_c$ and $r_d$ in the general expression of the threshold discount factor $\frac{(r_d - r_c)}{(r_d - r_{nc})}$ have been replaced by their specific values. This expression can then be reformulated as follows:

$$\delta \geq \frac{cV(x^*)}{[\tau - c]\left(\frac{V(x^*) - V(x^*(1-e))}{e}\right) + cV(x^*)}$$

As *e* tends to 0, the threshold discount factor tends to 1[12], which means that the ruler still has an incentive to cheat *de minimis* (i.e. to despoil a few merchants) regardless of its preference for the future. The efficient trade level is therefore not sustainable. A bilateral reputation mechanism (based on individual boycotts) is not severe enough to deter the ruler from cheating given that the loss of one merchant is marginal for the city.

What then is the value of a guild? The guild can enhance trading by increasing the severity of reprisals. This institution can in fact impose a complete boycott as soon as one merchant is violated. The credibility of this collective embargo can be ensured by the ability of the guild to sanction those merchants not respecting the embargo. If the ruler wants to cheat, he would know that all merchants would have to be mistreated (*e*=1) since regardless of the number of cheated merchants, he will be boycotted (no activity and no revenue during the punishment phase). In such a case, the efficient trade level is sustainable if $\delta \geq \frac{c}{\tau}$[13]. Thanks to the guild, trading can be stimulated and reach the efficient volume provided the ruler is sufficiently patient. According to Greif *et al.*, the existence of merchant guilds lends an explanation for the expansion of medieval trading[14].

---

[12] Because $\lim_{e \to 0}\left(\frac{V(x^*) - V(x^*(1-e))}{e}\right) = V'(x^*) = 0$ since x* maximizes the net value of trade.

[13] Since $r_d = \tau V(x^*)$, $r_c = [\tau - c]V(x^*)$ and $r_{nc} = 0$.

[14] A similar example of institutional device is provided in Milgrom, North and Weingast (1990). They showed that the *Law Merchant* institution in Champagne Fairs (during medieval times) enhanced the multilateral reputation mechanism, which by itself was unable to enforce contracts between merchants.

### 4.2 eBay as a reputation-building device for online traders

The worldwide Internet network can create and convey an unlimited number of remote trading relationships. Yet the anonymity of Internet users and the ease of Internet entry and exit may also incite opportunistic behavior. A seller, for example, can agree to trade with a buyer, promising to deliver the good upon receipt of payment. But once payment has been received, the seller can decide to exit the Internet market without delivering the product or after delivering a poor-quality substitute product (not conforming to the original description provided). In such cases, the buyer has no ability to punish the seller since the latter can easily change his identity and return as a new seller without incurring any reprisal (a situation similar to the principal-agent game in which the principal is the buyer and the agent the seller). The ability to avoid punishment (due to the relative anonymity and distance separating the trading partners) can hinder the development of online commerce, and especially C-to-C (consumer-to-consumer) commerce, since C-to-C is highly subject to fraud and opportunism.

In order to resolve this problem, private institutions have emerged. eBay is the most famous of these institutions. Its role is to provide traders with credible means or tools to retaliate against opportunistic behavior. Its function is not to directly sanction an opportunistic seller or buyer, but rather to offer the means to make their trading self-enforceable. For this purpose, eBay has established a reputation mechanism that enables anyone to know whether or not a buyer or seller can be trusted. Following a transaction, each side can evaluate his partner by giving him a positive (+1), neutral (0) or negative (-1) score. Opportunistic behavior can, for example, be sanctioned by a negative evaluation. The evaluations are then summed to generate the agent's overall score: the higher this score, the more trustworthy the agent. Each evaluation can also be accompanied by comments, especially in the case of a negative assessment.

Resnick and Zeckhauser (2002) examined all transactions conducted on eBay between February and June 2001. They noted that 60% of all buyers evaluated their trading partners and that 99% of these evaluations were positive. They described this evaluation system as a public good that helps the community of eBay users in their future transactions[15]. This mechanism serves as a substitute for the classic mechanism of bilateral reputation, which is not adapted to online transaction since repeated transactions between two partners are rare. How does this reputation mechanism actually discipline eBay users?

[15] Dellarocas, Ming and Wood (2004) observed rare coin auctions on eBay and found that 77% of the sellers and 67% of the buyers left feedback.

Since sellers are more likely to behave opportunistically (given that buyers have to send payment before the good is delivered, their ability to cheat is very limited), we will only consider herein the incentives of the seller. His incentive to be honest stems from the expected long-term benefits of enjoying a good reputation. These benefits are twofold: first, a good reputation can increase the probability of attracting many bidders for his products; moreover, a good reputation can increase the bid amounts that buyers are posting for his products.

If the seller is sufficiently patient and if reputation matters for his future expected auctions, then he should be deterred from acting opportunistically. Even though he can obtain instantaneous benefit by cheating (i.e. keeping the product), he will receive a negative evaluation and buyers will not trust him in subsequent transactions[16].

Many empirical studies have sought to measure the impact of reputation on future transactions. Houser and Wooders (2005) found that a positive reputation on eBay enables a seller to obtain a better price. Buyers tend to accept paying a premium when the product is sold by a highly-reputed seller. This result is based on a data set of Pentium III chips sold in online auctions during fall 1999 (a total of 95 auctions were observed). The authors found that a seller who increases the number of positive evaluations by 10% can expect a price 0.17% higher for his chips, whereas a 10% increase in the number of neutral or negative evaluations lowers the price by 0.24%.

Resnick and Zeckhauser (2002) found a more ambiguous result. They observed that reputation has no impact on the bids for MP3 players. However, reputation influences the probability that the transaction takes place. For example, a seller with no reputation (score of 0) has a 72% probability of selling his MP3 player, whereas the same seller with a score of 70 has a 96% probability.

Another experimental study conducted by Resnick, Zeckhauser, Swanson and Lockwood (2003) found results similar to those derived by Houser and Wooders. The authors attempted to determine the willingness-to-pay for old postcards on eBay. They compared the bids for postcards offered by a seller with a good reputation to those for the same postcards offered by a seller without any reputation; the difference in bid prices amounted on average to 8%[17].

[16] Or he can change his identity, but in this case he would lose his entire reputation and would reappear as a no-reputation seller.

[17] See also the study by Lucking-Reiley *et al.* (2000) on coin auctions. For a survey of empirical studies on eBay, see Resnick, Zeckhauser, Swanson and Lockwood (2003).

## 5. Conclusion

This chapter has highlighted that game theory is highly useful in examining the rationale of institutions. Game theory is a rigorous framework for questioning the nature of interpersonal relationships (Who are the decision-makers? What sets of actions do they implement? What information is available?), and for capturing the essential feature of institutions along with their impact.

Integrating game theory and institutions offers a promising and fruitful avenue for the field of new institutional economics. In return, the body of institutional literature can also enrich game theory by giving insight into how players actually behave during strategic situations (how they choose their decision rules, form their expectations and select an equilibrium).

## References :


Abreu, Dilip. 1986. "Extremal Equilibria of Oligopolistic Supergames", Journal of Economic Theory, Vol. 39, pp.191-225.

Akerlof, Georges A. 1970. "The market for 'lemons' Quality Uncertainty and the Market Mechanism," Quarterly Journal of Economics 84(3), pp.488-500.

Davis, Lance and Douglass North. 1971. *Institutional Change and American Economic Growth*, NY: Cambridge University Press.

Dellarocas, Chris; Ming Fan, and Charles A. Wood. 2004. "Self-Interest, Reciprocity, and Participation in Online Reputation Systems", MIT Sloan Working Paper4500-04, http://ssrn.com/abstract=585402

Dick, Andrew. 1996. "When are Cartels Stable Contracts? ", *Journal of Law and Economics*, 39, pp.241-283.

Dixit, Avinash and Barry Nalebuff. 1993. *Thinking Strategically: The Competitive Edge in Business, Politics, and Every Day Life*, WW. Norton & Company: New York.

Fisher, Franklin M. 1989. "Games economists play: a non-cooperative view" *Rand Journal of Economics 20,* pp.113–124.

Friedman, James W. 1971. "A Non-Cooperative Equilibrium for Supergames", *Review of Economic Studies*, 28, pp.1-12.

Gandal, Neil; David Salant, and Leonard Waverman. 2003. "Standards in wireless Telephone Networks", *Telecommunications Policy* 27, pp.325-332.

Gibbons, Robert. 1992. *Game Theory for Applied economists*, Princeton, NJ, Princeton University Press.

Gibbons, Robert. 2000. "Why Organizations Are Such a Mess (and What an Economist Might do About it)?", MIT Sloan working paper, http://web.mit.edu/rgibbons/www/Org_mess.pdf.

Gibbons, Robert. 2001. "Trust in Social Structures: Hobbes and Coase Meet Repeated Games", in *Trust in Society*, K. Cook, editor, New York: Russell Sage Foundation, pp. 332-353.

Greif, Avner. 1993. "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition." *American Economic Review*, 83:3, June, pp. 525-548.

Greif, Avner. 1989. "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders", *Journal of Economic History* Vol XLIX, N°4, pp.857-882.

Greif, Avner. 2000. "Economic History and Game Theory: a Survey", in *Handbook of Game Theory*, Vol. III, Robert J. Aumann, and Sergiu Hart, eds. North Holland: Amsterdam, chapter 52.

Greif, Avner; Paul Milgrom and Barry Weingast. 1994. “Coordination, Commitment and Enforcement: The Case of the Merchant Guild”, *Journal of Political Economy*, vol. 102, No. 4, pp.745-76.

Houser, Daniel and John Wooders (2005) “Reputation in Auctions: Theory and Evidence from eBay”, forthcoming in *Journal of Economics and Management Strategy*.

Jacquemin, Alexis and Margaret E. Slade. 1989 “Cartels, Collusion, and Horizontal Merger”, in *Handbook of Industrial Organization*, Richard Schmalensee, and Robert D. Willig, eds. Amsterdam: North-Holland, pp. 415-473.

Kreps, David. 1990. “Corporate Culture and Economic Theory”, in *Perspectives on Positive Political Economy*, J. Alt, and K. Shepsle, eds. Cambridge: Cambridge University Press.

Lafontaine, Francine. 1993. “Contractual Arrangement as Signaling Devices: Evidence from Franchising ”, *Journal of Law, Economics & Organization*, 9, pp. 256-289.

Lucking-Reiley, David; Doug Bryan; Naghi Prasad, and Daniel Reeves. 2000. “Pennies from eBay: The Determinants of Price in Online Auctions”, Working Paper, University of Arizona (forthcoming in *Journal of Economics and Management Strategy*).

Mathewson G. Frank and Ralph Winter. 1985. “The Economics of Franchise Contracts”, *Journal of Law &* Economics, 28, pp.503-526.

Milgrom, Paul; Douglass North and Barry Weingast. 1990. “The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs”, Economics-and-Politics 2(1), pp.1-23.

Osborne, Martin J. and Ariel Rubinstein. 1995. A course in Game Theory, Cambridge: M.I.T. Press.

Resnick, Paul and Richard Zeckhauser. 2002. “Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System”, in *Volume 11 of Advances in Applied Microeconomics*, Michael Baye, editor. Amsterdam: Elsevier Science, pp.127-157.

Resnick, Paul; Richard Zeckhauser; John Swanson, and Kate Lockwood. 2003. “The Value of Reputation on eBay: A Controlled Experiment,” KSG Working Paper Series No. RWP03-007.

Salop, Steven C. 1986. “Practices that (credibly) Facilitate Oligopoly Co-ordination”. In New Developments in the Analysis of Market *Structure*, Joseph Stiglitz and G. Frank Mathewson, eds. Cambridge: MIT Press, pp. 265-290.

Selten, Richard. 1975. “Reexamination of the perfectness concept for equilibrium points in extensive games”, *International Journal of Game Theory* 4, pp.25-55.

Von Neumann, John and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*, Princenton University Press.