Introduction

It is surprising to observe that Schumpeter (1954) does not mention the word of incentives in his monumental history of economic thought. How is it possible when, today, for many economists, economics is to a large extent a matter of incentives: Incentives to work hard, incentives to produce good quality products, incentives to study, incentives to invest, incentives to save,... How to design institutions in order to provide good incentives for economic agents has now become a central question of economics.

Maybe, Schumpeter's omission arose because economics mostly concentrated on understanding the theory of value in large economies. For that purpose, neoclassical economics postulates rational individual behavior in the market. In a perfectly competitive market, this assumption translates for firms' owners into profit maximization which, itself, implies cost minimization. In order words, the pressure of competitive markets solves the problem of incentives for cost minimization. Similarly, consumers faced with exogenous prices have the proper incentives for maximizing their utility levels. The major project of understanding how prices are formed in competitive markets can proceed without worrying about incentives.

However, by treating the firm as a black box, the theory remains silent on how the owners of firms succeed in aligning the objectives of its various members like workers, supervisors, managers with profit maximization. When economists began to look more carefully at the firm, either in agricultural economics or in managerial economics, incentives became the central focus of their analysis. Indeed, for various reasons, the owner of the firm must delegate several tasks to the members of the firm. This raises first the problems of managing information flows within the firm. This was the first research topic for economists, once they mastered behavior under uncertainty, thanks to Von Neumann and Morgenstern (1944). This line of research culminated in the theory of teams (Marschak and Radner (1972)). This theory recognized the decentralized nature of information, but postulated identical objective functions for the members of the firm considered as a "team". How to coordinate actions among the members of the team by the proper management of information was the central focus of this research. Incentive questions were still outside the scope of the analysis.

However, as soon as one acknowledges that the members of a firm may have different objectives, delegation becomes more problematic as recognized early by Marschak (1955) and also by Arrow (1963a) when he observes:

> "by definition the agent has been selected for his specialized knowledge and the principal can never hope to completely check the agent's performance."

Delegation of a task to an agent who has different objectives than the principal who delegates this task is problematic when information about the agent is imperfect. This is the essence of incentive questions. If the agent had a different objective function but no private information, the principal could propose a contract which perfectly controls the agent and induces the latter's actions to be what he would like to do himself in a world without delegation. Again, incentives issues would disappear.

Conflicting objectives and decentralized information are thus the two basic ingredients of incentive theory. That economic agents pursue at least to some extent their private interests is the essential paradigm for the analysis of market behavior by economists. What is proposed by incentive theory is to maintain this major assumption in the analysis of organizations, small number markets and any other kinds of collective decision. This paradigm has its own limits. Social behavior, in particular in small groups, is more complex, and norms of behavior culturally inculcated or developed over time play a large role in shaping societies. However, it would be foolish not to recognize the role of private incentives in motivating behavior in addition to these cultural phenomena. The purpose of this book is to synthesize what we have learned from the incentives paradigm.[1] We hope that the step by step approach taken here, as well as our attempt to present many different results in a unified framework, will help the readers not only to know more about incentive theory, but to appropriate this indispensable tool for thinking about society.

The starting point of incentive theory corresponds therefore to the problem of delegation of a task to an agent with private information. This private information can be of two types : either the agent can take an action unobserved by the principal, the case of *moral hazard* or *hidden action*; or the agent has some private knowledge about his cost or valuation that is ignored by the principal, the case of *adverse selection* or *hidden knowledge*. The theory studies when this private information is a problem for the principal, and what is the optimal way for the principal to cope with it. Another type of information problem has also been raised in the literature, the case of *non-verifiability* where the principal and the agent share ex post the same information but no third party and, in particular, no Court of Law can observe this information. One can study to which

[1] How do private incentives interact with cultural norms of behavior might be the next important step of research needed to be able to offer sensible advice on the design of institutions. It is our conviction nevertheless that for such a goal the mastering of incentive theory is a must.

extent the non-verifiability of information is also problematic for contractual design.

We will discover that, in general, these informational problems prevent society from achieving the first-best allocation of resources which could be possible in a world where all information would be common knowledge. The additional costs that must be incurred because of the strategic behavior of privately informed economic agents can be viewed as one category of the transaction costs emphasized by Williamson (1975). They do not exhaust all possible transaction costs, but economists have been rather successful during the last thirty years in modeling and analyzing this type of transaction costs, providing a good understanding of the limits put by these new costs on the allocation of resources. This work shows that the design of proper institutions for successful economic activities is more complex than one could have thought a priori. This line of research provides also a whole set of insights on how to proceed to take into account agents' responses to the incentives provided by institutions.

As the next chapter will illustrate, a brief look at the history of economy thought shows that incentive theory was pervasive in many areas of economics, even though it was not central in economic thinking. Before describing how we will proceed to present this theory, it may be worth mentioning how the major achievement of economics, namely the general equilibrium theory (GE), met incentives.

General equilibrium theory proved apt to powerful generalizations and able to deal with uncertainty, time, externalities, extending the validity of the *invisible-hand* as long as the appropriate competitive markets could be set up.[2] However, at the beginning of the seventies, works by Akerlof (1970), Spence (1974), and Rothschild and Stiglitz (1976) showed in various ways that asymmetric information was posing a much greater challenge and could not be satisfactorily imbedded in a proper generalization of the Arrow-Debreu theory. The problems encountered were so serious that a whole generation of general equilibrium theorists gave up momentarily the grandiose framework of GE to reconsider the problem of exchange under asymmetric information in its simplest form, i.e., between two traders, and in a sense went back to basics. They joined another group trained in game theory and in the theory of organizations to build the theory of incentives, that we take as encompassing contract theory and mechanism design.

We will present incentive theory in three progressive steps. Volume I is the first step, in which we consider the principal-agent model where the principal delegates an action to a single agent through the take-it-or-leave-it offer of a contract. Two implicit assumptions are made here. First, by postulating that it is the principal who makes a-take-it-or-leave-it contract offer to the agent, we put aside the bargaining issues which are a topic for game

[2]See Mas-Colell, Whinston and Green (1995) for a recent textbook exposition.

theory.[3] Second, we assume also the availability of a benevolent Court of Law which is able to enforce the contract and to impose penalties if one of the contractual partners adopts a behavior which deviates from the one specified in the contract.[4]

Three types of information problems will be considered, adverse selection, moral hazard and non-verifiability. Each of those informational problems leads to a different paradigm and, possibly, to a different kind of agency costs. On top of the usual technological constraints of neoclassical economics, these agency costs incorporate the informational constraints faced by the principal at the time of designing the contract.

In this volume, we will assume that there are no restrictions on the contracts that the principal can offer. As a consequence, the design of the principal's optimal contract reduces to a simple optimization problem.[5] This simple focus will turn out to be already enough to highlight the various trade-offs between allocative efficiency and the distribution of information rents arising under incomplete information. The mere existence of informational constraints may generally prevent the principal from achieving allocative efficiency. The main objective of the analysis undertaken in this volume is therefore the characterization of the allocative distortions that the principal finds desirable to implement in order to mitigate the impact of informational constraints.

Volume II will be the second step of our analysis. We will consider there situations with one principal and several agents, still without any restriction on the principal's contracts. Asymmetric information may not only affect the relationship between the principal and each of his agents, but it may also plague the relationships between agents. Moreover, maintaining the hypothesis that agents adopt an individualistic behavior, those organizational contexts require a new equilibrium concept, the Bayesian-Nash equilibrium, which describes the strategic interaction between agents under incomplete information. Three main themes arise in this context. First, the organization may have been built to facilitate a joint decision between the agents. In such a context, the principal must overcome the free-rider problems which may exist among agents when they must undertake a collective decision. Second, the principal may attempt to benefit from the competition between the agents to relax the informational constraints and better reduce the agents' information rents. Auctions, tournaments, yardstick competition and supervision of an

[3]See for example Osborne and Rubinstein (1994) and Mutho (1999).

[4]Let us stress here the importance of this assumption which is apparently innocuous because, in equilibrium, no penalty is ever paid and the role of the court looks minimal in what follows. However, judges may have to be given proper incentives to enforce contracts. We rely here on the idea that in repeated relationships the desire to maintain their reputation will provide the appropriate incentives. This implicit assumption is a little bit problematic since once could also appeal to the same reputation argument to justify that the principal-agent relationship may achieve allocative efficiency in repeated relationship even in the absence of any contract, with the appropriate cooperative behavior being self-enforcing. This assumption will be relaxed in Volume III.

[5]Hence, solving for the optimal contract requires only the simple tools of optimization theory.

agent by another one are all mechanisms designed by the principal with this purpose in mind. Third, the mere attempt by the principal to use competition between agents may also trigger their collusion against the principal. The principal must now worry not only about individual incentives, but also about group incentives in a multi-agent organization.

Volume III will be the third step of the analysis and will study the implications of various imperfections in the design of contracts: Informed principal, limited commitment, renegotiation, implicit incentives, imperfect coordination among various principals, incomplete contracting due to the non-verifiability of a parameter relevant for assessing the value of trade. The dynamics of some of these imperfect contractual relationships call for the extensive use of another equilibrium concept: the perfect Bayesian equilibrium. Equipped with this tool, we will be better able to describe the allocation of resources resulting from such imperfect contractual relationships.

In Volume I, we proceed as follows. Chapter 1 gives a brief account of the history of thought concerning incentive theory. It will show that incentives questions have been present in many areas of economics over the last two centuries even though it is only recently that their importance has been recognized and that economists have undertaken their systematic treatment. Chapter 2 presents the basic rent extraction-efficiency trade-off which arises in principal-agent models with adverse selection. Extensions of this framework to more complex environments are discussed in Chapter 3. Chapter 4 presents the two types of agency conflicts under moral hazard: the trade-offs between the extraction of a limited liability rent and efficiency or between insurance and efficiency. Again, extensions of this basic framework are discussed in Chapter 5. Chapter 6 considers the non-verifiability paradigm which in general does not call for economic distortions. Mixed models with adverse selection, moral hazard and non-verifiability are the subject of Chapter 7. The extension of principal-agent models with adverse selection and moral hazard to dynamic contexts with full commitment is given in Chapter 8. Finally, Chapter 9 discusses a number of simple extensions of the basic framework used all over the book.