1 Introduction

The theory of general equilibrium is one of the most impressive achievements in the history of economic thought. In the 1950s and 1960s the proof of the existence of equilibrium and of the close correspondences among equilibria, Pareto optima, and the core seemed to open the way for a reconstruction of the whole of economic theory around these concepts. However, it quickly appeared that the general equilibrium model was not a fully satisfactory descriptive tool. Strategic interactions between agents are heavily constrained in that model. This is because agents only interact through the price system, which the pure competition assumption says they cannot influence. In the logical limit one gets the models of the Aumann-Hildenbrand school in which there is a continuum of nonatomic agents, none of which can influence equilibrium prices and allocations. Similarly the organization of the many institutions that govern economic relationships is entirely absent from these models. This is particularly striking in the case of firms, which are modeled as a production set. This makes the very existence of firms difficult to justify in the context of general equilibrium models, since all interactions are expected to take place through the price system in these models. As Coase said long ago in one of his most influential papers (Coase 1937), "The distinguishing mark of the firm is the supersession of the price mechanism."

Creating general equilibrium models that could account for informational asymmetries presented another challenge. Arrow and Debreu had shown that it is fairly straightforward to extend the

general equilibrium model to cover uncertainty as long as information is kept symmetric. Unfortunately, asymmetries of information are pervasive in economic relationships. That is to say, customers know more about their tastes than firms, firms know more about their costs than the government, and all agents take actions that are at least partly unobservable. So rational expectations equilibria were conceived, at least in part, to encompass asymmetric information. However, while they offered interesting insights on the revelation of information by prices, their treatment of asymmetric information did not prove satisfactory. A *homo œconomicus* who possesses private information can be expected to try to manipulate that information, since he has in effect a monopoly over his own piece of private information. If we want to take this into account, we must forsake general equilibrium models. We then need to resort to other tools and, in particular, to game-theoretic tools.

The theory of contracts thus evolved from the failures of general equilibrium theory. In the 1970s several economists settled on a new way to study economic relationships. The idea was to turn away temporarily from general equilibrium models, whose description of the economy is consistent but not realistic enough, and to focus on necessarily partial models that take into account the full complexity of strategic interactions between privately informed agents in well-defined institutional settings. It was hoped then that lessons drawn from these studies could later be integrated inside a better theory of general equilibrium.

The theory of contracts, and more generally what was called the "economics of information," were the tools used to explore this new domain. Because they are just that—tools—it is somewhat difficult to define their goals other than by contrasting their shared characteristics with previous approaches:

- For the most part, the models are partial equilibrium models. They isolate the markets for one good (sometimes two goods) from the rest of the economy.

- The models describe the interactions of a small number of agents (often just two, one of whom possesses some private information and is call the "informed party").

- The models sum up the constraints imposed by the prevailing institutional setting through a *contract*. The contract may be explicit and in the form of a written agreement, or may be implicit and depend on a system of behavioral norms. An explicit contract will be guaranteed by a "third party" (e.g., a court or a mediator) or by the desire agents to maintain a reputation for fair trading. An implicit contract is sustained by an equilibrium tacitly observed in the interactions between the agreeing parties.

- The models make an intensive use of noncooperative game theory with asymmetric information, although their description of the bargaining process generally calls for a simplistic device known as the Principal-Agent model (on which more is provided later in this introduction). They are embedded in a Bayesian universe in which parties have an a priori belief on the information they do not possess, and they revise this belief as the interaction unfolds. The equilibrium concept they use in fact belongs to the family of perfect Bayesian equilibria.

The theory of contracts obviously covers a lot of ground and many varied situations. As a consequence early empirical studies were mostly case studies. Only recently has a body of literature emerged that tries to test the main conclusions of the theory of contracts using standard econometric techniques, as is discussed in chapter 8.

1.1 The Great Families of Models

The models of the theory of contracts can be distinguished along several axes, depending on whether they are static or dynamic, whether they involve complete or incomplete contracts, whether they

describe a bilateral or multilateral relationship, and so on. A large class of models, which can easily be divided into three families, is that where an informed party meets an uninformed party. I have chosen, somewhat arbitrarily of course, to classify these models according to two criteria. First is to distiguish whether the private information bears on

- what the agent *does*, the decisions he takes ("hidden action"),
- who the agent *is*, what his characteristics are ("hidden information").

Second, as in the form of the strategic game, is to distinguish the models in which the initiative belongs to the uninformed party from those in which it belongs to the informed party.

This classification yields three important families[1]:

- *Adverse selection* models. The uninformed party is imperfectly informed of the characteristics of the informed party; the uninformed party moves first.
- *Signaling* models. The informational situation is the same but the informed party moves first.
- *Moral hazard* models. The uninformed party moves first and is imperfectly informed of the actions of the informed party.

In chapters 2 to 5, I will study the basic structure of each of the three families. I should mention here, however, that one important class of models does not fit this system: models of incomplete contracting. This is because these models have so far only been developed in situations of symmetric information. They are studied in chapter 7.

1. The fourth case is that where the uninformed party does not observe the actions of the informed party. The informed party then takes the initiative of the contract. It is difficult to imagine a real-world application of such a model, and I do not know of any paper that uses it.

1.2 The Principal–Agent Model

Most of this book will use the Principal–Agent paradigm. There are two economic agents in this model: the informed party, whose information is relevant for the common welfare, and the uninformed party. Since this is a bilateral monopoly situation, we cannot go very far unless we specify how the parties are going to bargain over the terms of exchange. Unfortunately, the study of bargaining under asymmetric information is very complex.[2] The Principal–Agent model is a simplifying device that avoids these difficulties by allocating all bargaining power to one of the parties. This party will propose a "take it or leave it" contract and therefore request a "yes or no" answer; the other party is not free to propose another contract.

The Principal–Agent game is therefore a Stackelberg game in which the *leader* (who proposes the contract) is called the Principal and the *follower* (the party who just has to accept or reject the contract) is called the Agent.[3] While this modeling choice makes things much simpler, the reader should keep in mind that actual bargaining procedures are likely to be much more complex. For instance, if the Agent rejects the contract, the interaction would stop in the Principal–Agent model, whereas in the real world it would be expected to continue.

Because much of the book's discussion is informed by the Principal–Agent model, let us explore it a bit. One way to justify the Principal–Agent paradigm is to observe that the set of (constrained) Pareto optima can always be obtained by maximizing the utility of one

2. The main difficulty is that the natural equilibrium concept, perfect Bayesian equilibrium, leads to a large multiplicity of equilibria. See Ausubel-Cramton-Deneckere (2002) for a recent survey of bargaining models with asymmetric information.

3. I have tried to use consistent notation throughout the book: thus the "Agent" will always be the follower in a Principal–Agent game, while an "agent" is simply an economic agent, so that the Principal is also an agent. I hope this will create no confusion.

agent while the other is held to a given utility level. This is precisely what the Principal–Agent model does; so if we are only interested in common properties of the optima and not in one particular optimum, this approach brings no loss of generality. On the other hand, it may be that reasons outside the model should make us fix the Agent's reservation utility at some given level; if, for instance, the Principal is an employer and the Agent a prospective employee, the level of unemployment benefits and/or the market wage determine his reservation utility. In that case the peculiar properties of the Principal–Agent bargaining solution—it gives all surplus to the Principal—may make it less attractive as it picks a single point on the utility possibility frontier.

Finally, the choice of the words "Principal" and "Agent" should not be taken to imply that one of the parties works for the other or that the modeler is more interested in one than in the other. Each model has its own logic and should be interpreted accordingly. I should also point out that this terminology is taken by several authors, starting with the pioneering paper by Ross (1973), to refer to what they call the problem of agency, which is a moral hazard problem. My use of the Principal–Agent paradigm is both wider and more abstract; to me, it basically means that a Stackelberg game is being played.

1.3 Overview of the Book

An exhaustive look at the theory of contracts and its applications would make a very thick book. Such is not my ambition here. I merely want to present the main models of the theory of contracts, and particularly the basic models of the three great families described in section 1.1. It is not always easy to determine what belongs to the theory of contracts and what belongs to the wider field of the economics of information. I have chosen to include a brief description of auction models because their study relies on the same tools as the theory of contracts. On the other hand, I have preferred not to give a central role to models of insurance markets, even though their historical importance in shaping the field is well-established. As I will argue in section 3.1.3, these models have some peculiar features, and they deserve a fuller treatment than I can give them in a short book.

I have deliberately chosen to emphasize the methods used to analyze the models of the theory of contracts rather than the many applications that it has generated in various fields of economics. I have included brief introductions to these applications, but without any claim to completeness; most of the applications are not elaborated in the text. The reader interested in a particular application is urged to peruse the lists of references and to read the original papers. My goal in writing this book was to give the basic tools that allow the reader to understand the basic models and to come up with his own. I have tried to include recent developments, except where this could have led to overtechnical analyses. In most cases the lists of references will be sufficiently rich to allow the reader to find his way through this burgeoning literature.

Chapter 2 presents the general theory of adverse selection models. It starts with a brief summary of mechanism design, and proceeds to solve a basic model of second-degree price discrimination of two types. It then presents the solution in a more general continuous-type model. Several examples of applications and some more recent extensions are studied in chapter 3.

Chapter 4 turns to signaling models, and considers both signals that are costly and that are free. The basic moral hazard model, and some of its extensions and its application to insurance and wage contracts, are studied in chapter 5.

Chapter 6 is dedicated to the dynamic aspects of the theory of complete contracts. It introduces concepts like commitment and renegotiation that have been at the forefront of recent research. Because this field is very technical, I have not tried to provide complete proofs of