

# Introduction

## 1.1 The Model of a Mechanism

A fundamental problem in economic theory is to explain how acceptable choices can be made by a group despite the fact that only a small portion of the information that may a priori seem relevant can be taken into account. This problem arises in many settings, ranging from the largest scale problems of macroeconomic systems to the smallest problems of coordination among individuals in an organization. A market economy, for instance, coordinates production by firms and purchases by consumers through prices and quantities. The enormous amount of information held by each firm concerning its production processes and the knowledge of each consumer concerning his own tastes are not communicated among the participants in a market. General equilibrium theory, however, explains a sense in which the production plan selected by each firm and the purchases of each consumer in a market equilibrium are optimal despite the fact that a vast amount of the information known by firms and consumers remains private. A similar phenomenon arises within organizations. Employees cannot communicate all that they know to their manager, and if they could, then the manager could not possibly absorb all of this information. Communication is instead typically limited to conversations and memos. Determining exactly what information should be transmitted to a manager in order to allow him to make good decisions is a fundamental problem in the design of organizations and in the theory of accounting. Firms successfully function, however, despite this limited communication among its layers of management.

The root of the problem is *informational decentralization*, i.e., the information relevant to a group's collective decision may be dispersed among the group members and communication among members is costly or limited. In the two examples cited above, for instance, a consumer cannot possibly

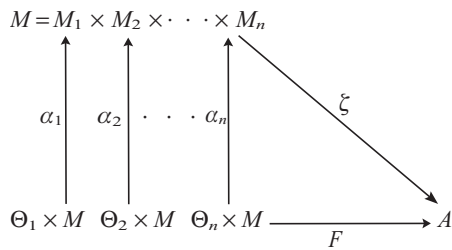


Figure 1.1. The Mount–Reiter diagram.

communicate his comparative ranking of all the different bundles of goods that he can afford, and an employee who communicates more and more about his job to his manager takes an increasing amount of his and his manager’s time away from other tasks. Communication by necessity is limited to a reduced set of signals, or *messages*: consumers and firms communicate in a market by means of prices and proposed trades, and an employee may communicate particular statistics (e.g., the profit of a division, costs, or return on equity) to his manager.

An essential issue in the selection of the *mechanism* (or protocol) for group decision making is thus the choice of the set of messages with which group members communicate with each other. In his seminal 1960 paper, Hurwicz initiated a model of communication among agents that encompasses both macro- and microeconomic problems.<sup>1</sup> The problem that motivated Hurwicz was to precisely define the term “mechanism”; prior to Hurwicz’s work, it was common to discuss the comparative properties of different mechanisms in various economic situations, and even to claim that a particular mechanism was “optimal,” without specifying the alternative mechanisms with which the given mechanism was being compared. The accomplishment of Hurwicz’s paper was thus not only the development of a model that has since been widely used to investigate the comparative properties of mechanisms, but more fundamentally to push economists toward a higher level of rigor in evaluating mechanisms.

The model of mechanisms that is considered in this text is depicted in the Mount–Reiter diagram (Mount and Reiter, 1974) of Figure 1.1. Consider first the left side of this diagram. There are  $n$  agents. Associated with each agent is a set of parameters  $\Theta_i$  and a set of messages  $M_i$ . An element  $\theta_i \in \Theta_i$  represents the *information* known by agent  $i$  but not by the other agents.

<sup>1</sup> See Reiter (1977) for an introduction to the economic purposes of this model and Hurwicz (1986) for a survey of the literature on informational decentralization in economic mechanisms.

## 1.1 The Model of a Mechanism

3

The elements of  $M_i$  represent the range of *messages* that the  $i$ th agent may select in order to communicate his information. Let

$$\Theta \equiv \prod_{i=1}^n \Theta_i$$

and

$$M \equiv \prod_{i=1}^n M_i. \quad (1.1)$$

An element  $\theta \in \Theta$  is the *state of the world* or *state*. An element  $m \in M$  is a vector of publicly observable messages

$$m = (m_1, m_2, \dots, m_n).$$

The model considered in this text originated as the equilibrium state of a myopic, dynamic adjustment process. Let  $m(t) = (m_i(t))_{1 \leq i \leq n}$  denote the messages at time  $t$ . Agent  $i$ 's message  $m_i(t+1)$  at time  $t+1$  is determined by a *message adjustment rule*

$$m_i(t+1) - m_i(t) = \alpha_i(\theta_i, m(t)), \quad (1.2)$$

which is a function of his private information  $\theta_i$  and the messages  $m(t)$  that he observes from all the agents at time  $t$ . The  $i$ th agent's *message correspondence*  $\mu_i$  specifies for each  $\theta_i$  the set of messages  $m$  at which the agent would not alter his message  $m_i$ :

$$\mu_i(\theta_i) \equiv \{m \mid 0 = \alpha_i(\theta_i, m)\}. \quad (1.3)$$

Agent  $i$ 's message correspondence  $\mu_i$  is *privacy preserving* in the sense that it depends only on his information  $\theta_i$  and not on the information  $\theta_{-i}$  of the other agents.<sup>2</sup> Equilibrium is given by the *equilibrium correspondence*

$$\mu(\theta) \equiv \bigcap_{i=1}^n \mu_i(\theta_i). \quad (1.4)$$

A correspondence  $\mu : \Theta \rightarrow M$  is *privacy preserving* if it can be expressed as an intersection of form (1.4).<sup>3</sup>

<sup>2</sup> The use of "privacy" in this sense originates in Hurwicz (1972).

<sup>3</sup> A correspondence  $\mu$  of the form (1.4) is also referred to as a *coordinate correspondence*. Mount and Reiter (1974) identified this property of correspondences and its central role in modeling informational decentralization.

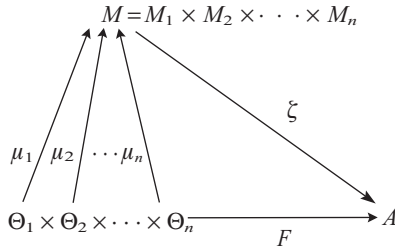


Figure 1.2. A mechanism  $\mathcal{M} = (\Theta, M, (\mu_i)_{1 \leq i \leq n}, \zeta)$  that realizes the objective  $F$ .

The equilibrium state is depicted in Figure 1.2. Notice that in equilibrium, it is not necessary to assume that  $M$  is a Cartesian product of the form (1.1). In the general formulation considered in this text, an arbitrary set  $M$  can serve as the set of equilibrium messages and equilibrium is addressed without reference to the process by which it may be achieved.<sup>4</sup>

The triple

$$\mathcal{MP} = (\Theta, M, (\mu_i)_{1 \leq i \leq n}) \tag{1.5}$$

is a *message process*. On the right side of Figure 1.2 is the set  $A$ , the *set of alternatives* for the group. Selecting an element of this set represents a collective decision. The mapping  $\zeta : M \rightarrow A$  is the *outcome mapping*, which represents the selection of an alternative based on the messages of the agents. Together,

$$\mathcal{M} = (\Theta, M, (\mu_i)_{1 \leq i \leq n}, \zeta) \tag{1.6}$$

is a *mechanism*. The mechanism is *informationally decentralized* in the sense that each agent  $i$ 's message correspondence  $\mu_i$  depends only on his parameter vector  $\theta_i$  and not on those of the other agents. It thus respects the constraint of information decentralization on the group's decision problem that is the focus of this text.

The correspondence  $F : \Theta \rightarrow A$  is the *objective*. The objective may be specified by a welfare criterion such as Pareto optimality; in communication among divisions of a firm, the objective may be maximizing the firm's profit. In the central problem of mechanism design, the objective  $F$  is given

<sup>4</sup> Resource allocation mechanisms are formulated in Hurwicz (1960) as dynamic processes in the sense described above. Mount and Reiter (1974) focus on the equilibrium state, which generalizes mechanisms by allowing  $M$  to be an arbitrary set. This text builds most directly on the equilibrium approach of Mount and Reiter.

## 1.1 The Model of a Mechanism

5

and the task is to compare mechanisms  $\mathcal{M}$  that *realize*  $F$  in the sense that<sup>5</sup>

$$\zeta \circ \mu(\theta) = F(\theta) \quad (1.7)$$

for all  $\theta \in \Theta$ . This means that the mechanism collects a sufficient amount of information about  $\theta$  from the agents in its equilibrium state to calculate the alternative  $F(\theta)$ . Equality (1.7) is sometimes weakened to the requirement that

$$\zeta \circ \mu(\theta) \subset F(\theta) \quad (1.8)$$

for all  $\theta \in \Theta$ . In this case, the mechanism *weakly realizes*  $F$ . Alternatively, particular message spaces  $M$  and outcome mappings  $\zeta$  are sometimes of interest because they model aspects of real mechanisms. In this case, the problem is to determine the equilibrium correspondence  $\mu$  that is appropriate for the model. The goal is then to evaluate the welfare properties of the objective  $F$  that is defined by the mechanism through formula (1.7).

**The Revelation and Parameter Transfer Mechanisms**

There are two types of mechanisms that realize any objective  $F$ . In the *revelation mechanism*, each agent  $i$ 's message  $m_i$  is his information  $\theta_i$  and the outcome mapping  $\zeta$  is the objective  $F$ . Formally, the revelation mechanism is defined as follows:<sup>6</sup>

$$\text{for all } i, M_i = \Theta_i \text{ and } \mu_i(\theta_i) = \{m \mid m_i = \theta_i\};$$

$$\zeta(m) = F(m).$$

The second mechanism is actually a family of  $n$  mechanisms, indexed by the selection of one of the  $n$  agents. Select agent  $j$ . In the *parameter transfer mechanism*, all agents except agent  $j$  transfer their parameters to the message space  $M$ , which then permits agent  $j$  to compute the objective. The mechanism is defined formally as follows:

$$\text{for } i \neq j, M_i = \Theta_i \text{ and } \mu_i(\theta_i) = \{m \mid m_i = \theta_i\};$$

$$M_j = A \text{ and } \mu_j(\theta_j) = \{m \mid m_j = F(\theta_j, m_{-j})\};$$

$$\zeta(m) = m_j.$$

<sup>5</sup> The use of the term “realize” in this sense originated in Mount and Reiter (1974).

<sup>6</sup> Despite its simplicity, the revelation mechanism is fundamental in game theory because of the revelation principle. See Myerson (1991), for instance, for a discussion of this topic.

Agent  $j$  has special status in this mechanism as a “head” who receives information from the other agents and then realizes the objective. Parameter transfer is in this sense an elementary model of hierarchical decision making.<sup>7</sup>

The revelation and the parameter transfer mechanisms demonstrate that mechanisms exist that realize any given objective. The realization problem is thus never vacuous. These mechanisms share the flaw, however, of requiring all but at most one of the agents to communicate all his information. They are therefore implausible in many real settings, and they fail to address the fundamental problem posed at the beginning of this chapter.

### 1.1.1 Example: The Competitive Mechanism

Modeling the mechanism of perfect competition is one of the problems that inspired the development of mechanism design. The competitive mechanism is a special case of the equilibrium model depicted in Figure 1.2; notably, this model also provides a framework for designing and evaluating alternatives to the competitive mechanism for achieving gains from trade among agents. This subsection presents a model of the competitive mechanism in the case of an exchange economy. The goal is to familiarize the reader with the general model through examining this central mechanism of microeconomic theory. The competitive mechanism is addressed further in this text in Subsections 3.1.2, 4.3.1, and 4.3.3 and in Sections 4.5 and 4.6.

There are  $n$  agents and  $l$  goods that may be traded. Each agent  $i$ 's information consists of a pair  $\theta_i = (U_i, w_i)$ , where  $U_i : \mathbb{R}_+^l \rightarrow \mathbb{R}$  is his utility function over his consumption space  $\mathbb{R}_+^l$  and  $w_i \in \mathbb{R}_+^l$  is his initial endowment of the goods. The set of possible utility–endowment pairs is

$$\Theta_i = \mathcal{U} \times \mathbb{R}_+^l,$$

where  $\mathcal{U}$  is the set of all increasing, continuous, and concave functions on  $\mathbb{R}_+^l$ . A state

$$\theta = ((U_i, w_i))_{1 \leq i \leq n}$$

in this context is commonly referred to as an *economy*.

The objective of trading is the assignment of a net trade vector  $\Delta x_i \in \mathbb{R}^l$  to each of the agents in each economy  $\theta$ . Trader  $i$  then receives  $w_i + \Delta x_i$  as his final allocation. Let  $\Delta x = (\Delta x_i)_{1 \leq i \leq n}$  denote a vector of net trades for

<sup>7</sup> This interpretation is pursued in Section 4.9.

## 1.1 The Model of a Mechanism

7

the agents. The set of alternatives  $A$  consists of all balanced net trades for the  $n$  agents,

$$A = \left\{ \Delta x = (\Delta x_i)_{1 \leq i \leq n} \in \mathbb{R}^{nl} \left| \sum_{i=1}^n \Delta x_i = 0 \right. \right\}.$$

It is desirable that the allocation should be feasible, Pareto efficient, and individually rational. A net trade vector  $\Delta x$  is *feasible* for the economy  $((U_i, w_i))_{1 \leq i \leq n}$  if

$$w_i + \Delta x_i \in \mathbb{R}_+^l,$$

i.e., no trader is assigned a negative amount of some good in his final allocation. *Pareto efficiency of  $\Delta x$  for the economy  $((U_i, w_i))_{1 \leq i \leq n}$*  is the requirement that no feasible net trade  $\Delta x'$  exists for this economy such that

$$U_i(w_i + \Delta x'_i) \geq U_i(w_i + \Delta x_i), \quad (1.9)$$

with strict inequality in (1.9) for at least one agent  $i$ . It is thus not possible to select a feasible net trade  $\Delta x'$  that is as good for every agent as  $\Delta x$  and strictly better for at least one agent  $i$ . *Individual rationality* is the requirement that no trader is made worse off by trading, i.e.,

$$U_i(w_i + \Delta x_i) \geq U_i(w_i)$$

for each agent  $i$ . This insures that each agent will voluntarily participate in trading. Let  $F^*: \Theta \rightarrow A$  denote the correspondence that assigns to each economy  $((U_i, w_i))_{1 \leq i \leq n}$  the set of all possible, feasible, balanced, individually rational, and Pareto efficient net trade vectors  $\Delta x$  for that economy.

Let  $p \in \mathbb{R}_+^l$  denote a vector of prices for the  $l$  goods.<sup>8</sup> The message space for the competitive mechanism is  $A \times \mathbb{R}_+^l$  with a message denoted as  $(\Delta x, p)$ . Agent  $i$ 's message correspondence  $\mu_i(U_i, w_i)$  specifies those net trade vectors  $\Delta x^*$  and price vectors  $p^*$  such that his final allocation  $w_i + \Delta x_i^*$  maximizes his utility  $U_i$  subject to his budget constraint:

$$\mu_i(U_i, w_i) = \left\{ (\Delta x^*, p^*) \in \left| \Delta x_i^* \in \arg \max_{\Delta x_i \in \mathbb{R}_+^l} U_i(w_i + \Delta x_i) \text{ s.t. } p \cdot \Delta x_i = 0 \right. \right\}.$$

<sup>8</sup> While prices can be normalized in a number of different ways, normalization is not needed for this introductory discussion of the competitive mechanism.

The equilibrium correspondence  $\mu$  specifies for the economy  $((U_i, w_i))_{1 \leq i \leq n}$  all pairs  $(\Delta x^*, p^*)$  such that  $w_i + \Delta x_i^*$  maximizes each trader  $i$ 's utility  $U_i$  subject to his budget constraint. Because  $\Delta x^*$  satisfies the balance condition, such a pair  $(\Delta x^*, p^*)$  is a *Walrasian equilibrium*<sup>9</sup> for the economy  $((U_i, w_i))_{1 \leq i \leq n}$ .

The outcome mapping of the competitive mechanism is the projection mapping  $\zeta(\Delta x, p) = \Delta x$ . Classic results in general equilibrium theory imply that (i) an equilibrium message  $(\Delta x^*, p^*)$  exists for each economy  $((U_i, w_i))_{1 \leq i \leq n}$ ; (ii) if  $(\Delta x^*, p^*)$  is an equilibrium message for  $((U_i, w_i))_{1 \leq i \leq n}$ , then  $\Delta x^*$  is Pareto optimal and feasible for  $((U_i, w_i))_{1 \leq i \leq n}$ . The competitive mechanism thus weakly realizes the objective  $F^*$  in the sense of (1.8).

### 1.1.2 Example: Mechanisms and Noncooperative Solution Concepts

The equilibrium model depicted in Figure 1.2 is sufficiently general to include most noncooperative games as special cases. In this interpretation,  $M_i$  is agent  $i$ 's strategy set and  $\zeta$  is the outcome mapping of the game. Agent  $i$ 's message correspondence  $\mu_i : \Theta_i \rightarrow M$  reflects his strategic choice. The definition of a game is completed by specifying a function  $U_i(a, \theta)$  for each of the agents that computes his payoff based on the alternative  $a \in A$  and the state  $\theta$ . A *solution concept* is a theory that explains each agent  $i$ 's strategic choice of a message in terms of his information and his self-interest. Examples that fit this model include the dominant, ex-post Nash, and Bayesian Nash solution concepts.<sup>10</sup>

These three solution concepts are now reviewed in order to illustrate the relationship between this text and noncooperative game theory. The Nash correspondence is then discussed at the end of this subsection. It fits the model of this text if the agents' preferences are *private* in a sense that is discussed below.

In all four of these cases, existence of equilibrium (i.e., nonemptiness of the equilibrium correspondence) is a significant issue that depends on

<sup>9</sup> This is also referred to as a *competitive equilibrium* in the literature. In order to distinguish clearly between an objective and a particular mechanism that may be used to realize it, "Walrasian" in this text refers to the allocation and prices in the standard equilibrium of an exchange economy (i.e., an objective of trading), while "competitive" refers to a particular mechanism that realizes a Walrasian objective.

<sup>10</sup> In game theory, the problem of devising a mechanism  $\mathcal{M}$  that realizes a given objective  $F$  is modified by the addition of the constraint that the message correspondences must be consistent with a particular solution concept. *Implementation* is the special case of the realization problem defined in this way by adding an incentive constraint.



the problem under consideration. Each of these four solution concepts has been used to model incentives in situations in which agents have private information. Noncooperative game theory typically focuses on the rationale for the selection of particular message correspondences (or strategies) by the agents. The emphasis in this text is instead on the different ways in which each agent  $i$  can encode his private information through the choice of the message set  $M$  and his message correspondence  $\mu_i$ . The point of this section is that this can be a more general topic that includes noncooperative game theory as a special case in the sense that a solution of a game may define such a communication structure. A particular issue of interest in this text is that different solution concepts may present different opportunities for encoding of information. This issue is addressed in Sections 1.2 and 4.8.

### *Dominant, Ex-Post Nash, and Bayesian Nash Equilibrium*

These three solution concepts can be motivated by interpreting information as revealed over time: agent  $i$  chooses his message  $m_i$  after learning  $\theta_i$  but without knowing  $\theta_{-i}$ . An equilibrium in each case therefore posits a strategy  $\sigma_i : \Theta_i \rightarrow M_i$  for agent  $i$ 's selection of a message  $m_i$  based on his information. The strategy  $\sigma_i$  is *privacy preserving* in the case of these three solution concepts in the sense that it depends only on agent  $i$ 's information  $\theta_i$ . In each of these three cases, the message correspondence  $\mu_i : \Theta_i \rightarrow M$  is defined from the strategy  $\sigma_i$  by the formula

$$\mu_i(\theta_i) \equiv \{\sigma_i(\theta_i)\} \times M_{-i}, \quad (1.10)$$

and the equilibrium correspondence is the function

$$\mu(\theta) \equiv (\sigma_i(\theta_i))_{1 \leq i \leq n}. \quad (1.11)$$

Privacy preserving strategies thus define a privacy preserving equilibrium correspondence.

Dominance is the strongest notion of incentive compatibility in the sense that an agent's choice is optimal given his information regardless of the information and messages of the other agents. A strategy  $\sigma_i^d$  for agent  $i$  is *dominant* if the message  $\sigma_i^d(\theta_i)$  that it specifies for each  $\theta_i \in \Theta_i$  maximizes his payoff regardless of the messages  $m_{-i}$  chosen by the other agents or their information  $\theta_{-i} \in \Theta_{-i}$ : for each  $\theta_i \in \Theta_i$ ,

$$\sigma_i^d(\theta_i) \in \arg \max_{m_i \in M_i} U_i(\zeta(m_i, m_{-i}), (\theta_i, \theta_{-i})) \quad (1.12)$$

for all  $\theta_{-i} \in \Theta_{-i}$  and  $m_{-i} \in M_{-i}$ . An  $n$ -tuple of strategies  $(\sigma_i^d)_{1 \leq i \leq n}$  is a *dominant equilibrium* if  $\sigma_i^d$  is a dominant strategy for each agent  $i$ .

Optimality of an agent's choice in each of the next two solution concepts is weaker than in dominance in that it depends on the strategies chosen by the other agents. An  $n$ -tuple of strategies  $(\sigma_i^x)_{1 \leq i \leq n}$  is an *ex-post Nash equilibrium* if, for each agent  $i$  and each  $\theta_i \in \Theta_i$ , the message  $\sigma_i^x(\theta_i)$  maximizes agent  $i$ 's payoff given the choice of  $\sigma_{-i}^b(\theta_{-i})$  by the other agents for every value  $\theta_{-i} \in \Theta_{-i}$  of their information: for each agent  $i$  and each  $\theta_i \in \Theta_i$ ,

$$\sigma_i^x(\theta_i) \in \arg \max_{m_i \in M_i} U_i(\zeta(m_i, \sigma_{-i}^b(\theta_{-i})), (\theta_i, \theta_{-i})) \quad (1.13)$$

for all  $\theta_{-i} \in \Theta_{-i}$ .

An  $n$ -tuple of strategies  $(\sigma_i^b)_{1 \leq i \leq n}$  defines a *Bayesian Nash equilibrium* if, for each agent  $i$  and each  $\theta_i \in \Theta_i$ , the message  $\sigma_i^b(\theta_i) = m_i$  maximizes agent  $i$ 's conditional expected payoff assuming that each other agent  $j$  uses the strategy  $\sigma_j$ : for each agent  $i$  and each  $\theta_i \in \Theta_i$ ,

$$\sigma_i^x(\theta_i) \in \arg \max_{m_i \in M_i} E_{\theta_{-i}} [U_i(\zeta(m_i, \sigma_{-i}^b(\theta_{-i})), (\theta_i, \theta_{-i})) | \theta_i, \sigma_{-i}^b]. \quad (1.14)$$

The expected payoff in (1.14) is calculated with respect to a probability distribution on  $\Theta$  that is postulated as part of the Bayesian approach. It is typically assumed with both the ex-post Nash and Bayesian Nash solution concepts that agent  $i$  knows the strategies of the other agents so that he may verify the optimality of his message.<sup>11</sup>

### Nash Correspondence

The *Nash correspondence*  $\mu^{ne} : \Theta \rightarrow M$  specifies all Nash equilibria in each state  $\theta$ .<sup>12</sup> An  $n$ -tuple  $m' = (m'_i)_{1 \leq i \leq n}$  lies in  $\mu^{ne}(\theta)$  if each  $m'_i$  maximizes agent  $i$ 's utility given  $\theta$  and the choice of  $m'_{-i}$  by the other agents:

$$\mu^{ne}(\theta) \equiv \left\{ m' \mid m'_i \in \arg \max_{m_i \in M_i} U_i(\zeta(m_i, m'_{-i}), \theta), 1 \leq i \leq n \right\}.$$

<sup>11</sup> There is now a large literature in game theory concerning how players learn to play equilibria. Much of this literature studies the dynamic stability of message adjustment rules of the form (1.2) in a variety of special cases. A special case is typically defined by assumptions concerning the incentives of the players, their information, and how they learn. A more abstract approach that focuses on the informational requirements for the local stability of message adjustment rules was initiated by Reiter in 1979, with subsequent contributions by Jordan (1987), Mount and Reiter (1987), Saari and Williams (1986), and Williams (1985).

<sup>12</sup> A similar analysis can also be made for the *dominant correspondence*, which specifies all dominant strategy equilibria in each state  $\theta$ . Ex-post Nash and Bayesian Nash equilibria, however, are not defined by a single state  $\theta$ ; the condition for the optimality of agent  $i$ 's choice in each case depends on the other agents' strategies over for all  $\theta_{-i} \in \Theta_{-i}$ . Equilibrium correspondences are therefore not defined for these solution concepts in the same sense as they are for the dominant and Nash solution concepts.