Economics and the Theory of Games: A Survey

By

ANDREW SCHOTTER
New York University

and

GERHARD SCHWÖDIAUER
University of Bielefeld

We are indebted to many people for their encouragement and advice in undertaking this project. Above all, however, we must thank the editor of this journal, whose practical and scientific judgment kept us on track through what turned out to be a long and at times difficult task. In addition, we must thank William Baumol, Gerald Faulhaber, Michael Intriligator, George Lee, Stephen Littlechild, Fritz Machlup, Thomas Marschak, Hervé Moulin, Andrew Postlewaite, Juan Eugenio de la Rama, Robert Rosenthal, David Schmeidler, Anne Schotter, and Martin Shubik, all of whom have read the manuscript either in part or in whole and have given us extremely useful scientific and editorial advice. Professor Schotter's participation in the project was made possible by a grant from the Office of Naval Research Contract No. N00014-78-C-0598, whose support is gratefully acknowledged. All errors in the work that follows are those of the authors alone.

I. Introduction

Game theory has been with us for 36 years, and in those years its popularity has risen and fallen with almost fixed periodicity. The first wave of interest came in 1944 when the first edition of The Theory of Games and Economic Behavior by John von Neumann and Oskar Morgenstern appeared and met with favora-
tions of strategic interdependence were paramount and so this new theory seemed natural. By the late 1950's the consensus in the profession was clear: game theory was to be a theory for the small numbers case in economics. As soon as the theory was pigeonholed, however, its popularity waned, for mathematical theorists turned their attention toward the axiomatic analysis of general equilibrium theory. And yet in this analysis, paradoxically, game theory found a new life. The vehicle for this renaissance was Martin Shubik's article "Edgeworth Market Games" (1959b), in which he demonstrated that Edgeworth's classic contract curve was identical to a game theoretical solution concept named "the core" that had been recently developed by Lloyd Shapley (1952) and Donald B. Gillies (1959). The impact of this equivalence was immediate. While the then current analysis of the general equilibrium problem being conducted by Kenneth Arrow, Gerard Debreu, Lionel McKenzie, and others was strictly Walrasian in orientation, treating all agents as perfect price-taking maximizers tied together by a price-making auctioneer, the new game theoretical analysis was Edgeworthian, taking inspiration from Edgeworth's *Mathematical Psychics* (1881), which viewed the price-formation process as the outcome of a large multilateral bargaining procedure. A long discussion ensued in the literature, in which it was proven that in the limit, as the economies studied get "large" in an appropriate manner, both the Walrasian and the game theoretical (Edgeworthian) analysis converge to the same solution. While this result was quite elegant, it spelled the end of the first renaissance in game theory: It seemed that the game theoretical analysis (which employed strictly cooperative game theoretical concepts) was too demanding informationally to be of any intuitive appeal. Since it yielded no new results, little could be gained through its use. Hence, by the late 1960's the theory was once again slipping from popularity (except among economists and game theorists interested in such topics as the bargaining problem, where the literature grew constantly throughout the period). This dip in popularity was short-lived, however, because during the 1960's and early 1970's economists began to turn their attention to a new set of problems for which game theory seemed to be the ultimate modelling tool. These new problems were concerned, in broad terms, with the design and operation of "satisfactory" economic and social institutions, and in narrow terms, with the design and implementation of allocating and voting mechanisms. There are three main sources of this new interest. One is again Martin Shubik, who, after initiating the cooperative game theoretical analysis of general equilibrium theory, became disillusioned with its entire apparatus, since it was totally lacking in any institutional detail. The new approach he, along with Lloyd Shapley and Pradeep Dubey, offered was to phrase the general equilibrium problem as one involving a set of isolated price-making agents who form prices by bidding for goods with money. In the construction of such models, a wide variety of institutions such as banks, money, market mechanisms, laws governing bankruptcy, etc., were described and rigorously introduced into the analysis, and their impact became the object of investigation. At about the same time, Leonid Hurwicz was dealing with a different type of question, which, ironically enough, also dealt with the question of institutional analysis, especially in informationally decentralized economies. The problem Hurwicz was dealing with involved the construction of allocating or planning mechanisms (or institutions) that determined what he considered to be "satisfactory" outcomes. In pursuing this goal, Hurwicz first defined a set of characteris-
tics that any good allocation mechanism should have and then discovered that for some sets of quite reasonable characteristics no mechanism could be found that satisfied them (1973). This startling impossibility result prepared the way for a great body of literature. The reason why game theory is so important here, as we will see in Section III, is that in an informationally decentralized economy, each allocating mechanism can be shown to define a different n-person non-cooperative game. Hence, by studying the properties of the equilibria of these games, we can study the properties of the allocation mechanism (or institution) that implicitly defines them. On a behavioral level, then, the study of economic and social institutions becomes equivalent to the analysis of the equilibrium properties of n-person games.

Finally, the social choice literature has furnished yet another source of renewed interest in game theory. Allan Gibbard (1973) and Mark Satterthwaite (1975) independently asked the question of what happens when the agents studied by Kenneth Arrow in Social Choice and Individual Values (1963) decide to vote strategically or in a manner that is not isomorphic to their true preferences. The answer, of course, is that the voting rule they are using may lead to social choices that are not Pareto optimal. As a result both Gibbard and Satterthwaite searched for a voting rule or mechanism (or what Gibbard called a ”game-form”) that created incentives for all agents to report their preferences truthfully when they voted. What they found was that there do not exist any such strategy-proof voting mechanisms that also satisfy a set of reasonable “democratic” criteria. This result set the stage for a systematic search for voting rules that yield satisfactory results or ”implement” social choice rules (see Eric Maskin [1978b] and Bezalel Peleg [1978]).

The end result of these independent institutional investigations is a remarkable third wave of interest in game theory that, as far as we can see, is likely to establish a secure place for it within economic theory. This is true because it is in the design of institutions and their stability properties that game theory offers a natural tool for analysis and offers a set of solutions or equilibrium notions that are quite useful for their analysis. It is ironic, however, that at least for the moment game theory has settled into investigating questions of institutional analysis, since if one were to reread the first chapter of The Theory of Games and Economic Behavior, one would immediately see that the authors had exactly such institutional questions in mind when they started their analysis. They saw game theory as, in sum, the theory of the emergence of stable institutional arrangements or ”standards of behavior” in a given physical situation or game. In other words, the theory tries to predict what stable institutional form will emerge from a given economic background and what the resulting value relationships will be. As a result, the theory does not assume that any particular institutional arrangement exists at the outset, as does the neoclassical theory, but starts out by describing the tastes and technologies of the agents in an institutional ”state of nature” from which it predicts what stable institutional arrangements or standards of behavior will evolve. Von Neumann and Morgenstern are very clear on this point. If, in the quotation that follows, one replaces the words ”order of society” and ”standards of behavior” with the words ”institutional arrangements,” one sees immediately the point we have been making:

The question whether several stable ”orders of society” or ”standards of behavior” based on the same physical background are possible or not, is highly controversial. There is little hope that it will be settled by the usual methods because of the enormous complexity of this
problem among other reasons. But we shall give specific examples of games of three or four persons, where one game possesses several solutions in the sense of 4.5.3 [the stable-set solution]. And some of these examples will be seen to be the models for certain simple economic problems. [Von Neumann and Morgenstern (1947), p. 43.]

The important point that von Neumann and Morgenstern are making here is that social institutions must be seen as the equilibrium outcome of games of strategy whose descriptions are given by the physical capabilities of the agents in the game—the "empirical background." They are an outcome of the theory rather than an input into it. If, however, we are presented with a fixed institutional setting in which the rules of conduct delimit the actions that agents can follow a priori, as is true of the literature described above, game theory can be used effectively to study the manner in which agents will behave in this institutional setting and hence allow us to study the comparative properties of alternative institutions.

In the survey below we will follow in reverse order the historical development just described, starting from the institutional models and working our way backwards. On the way we will survey the literature on incentive-compatible allocating mechanisms and the design of "optimal" games, strategy-proof voting mechanisms, externalities and public goods, public utility pricing and the allocation of joint costs, general equilibrium theory, and oligopoly theory. While our survey is quasi-historical in organization, it is not historical in presentation, since rather than attempting to trace each application through its long and complicated history, we present the main problems attacked in each field and the main results. In the process many topics are excluded, either because they have been surveyed thoroughly elsewhere or because they are so new that any survey would be obsolete within a matter of months. Hence, the bargaining problem, which has recently been surveyed with great care by Alvin E. Roth (1979), is not covered here, nor is the new and exciting literature on bidding and auctioning and the design of optimal auctions. In addition, we could not include any discussion of the new experimental literature that has developed in the past 20 years. Finally, our survey is limited to economics only, and we have not attempted to include any of the game theoretical work done in political science, philosophy, sociology, or social psychology, though they are many times very closely related. All of these gaps are unfortunate, but our space limitations make them necessary.

Finally, the articles we have chosen to describe or reference by no means represent an exhaustive list of all important contributions to game theory and economics. Rather they constitute a list of references that fit best into the themes of the survey we constructed. Hence, we apologize to those authors whose work we have omitted.

In this paper we will proceed as follows: In Section II we will present a quick and intuitive introduction to the basic game-theoretical concepts that will be used throughout the paper. This will, we hope, lower the most common barrier that scholars and students meet when reading game-theoretical material—the plethora of terms—and permit the reader free access to the material presented. Those already familiar with the theory should go straight to Section III.

Section III contains our survey of the application of game theory to economics. We begin with a discussion of the literature surrounding the question of public utility pricing, cross-subsidization, and the allocation of joint costs (Section 3.1) because it is an area that has traditionally
been studied by neoclassical analysis and hence provides us with a starting point familiar to many readers. While beginning with this topic does temporarily take us outside of our historical organization, we feel the detour is worthwhile, since the topic is of considerable real world significance and has not been surveyed elsewhere. In Section 3.2 we consider the question of the design of incentive-compatible allocating and voting mechanisms or equivalently the design of optimal games. We then follow with a discussion (Section 3.3) of the Gibbard-Staatterthwaite theorem and the questions surrounding the implementation of social choice rules. In Section 3.4 we turn our attention to the game theoretical analysis of the public goods and externality problem, where we discuss a major problem involved in the cooperative game— Theoretical treatment of these questions. Finally, in Section 3.5 we investigate the literature surrounding the problem of multilateral exchange and general equilibrium and follow this up (Section 3.6) with a discussion of the oligopoly problem and the new literature recently initiated by William Novshek and Hugo Sonnenschein (1978) and Shubik (1971/72; 1972) on general equilibrium analysis with what can be called strategically active agents (Section 3.7).

II. A Quick Introduction to Game Theory

This section is not intended to be a substitute for a text on the mathematical theory of games (see, e.g., R. Duncan Luce and Howard Raiffa, 1957; Ewald Burger, 1963; Guillermo Owen, 1968; T. Parthasarathy and T. E. S. Raghavan, 1971; Nikolai Vorob’ev, 1977). Rather, we hope to present a brief and informal account of the most important game theoretical concepts that will be referred to in the sections dealing with economic applications.

2.1 Describing Games

There are many levels on which one can describe a game, each of which involves a different degree of detail; the amount of detail an investigator wishes to use will depend on the type of analysis he wishes to undertake and on the results he envisages. For instance, when a detailed description of a situation of strategic interdependence is required, we may rely on what is called the extensive form of the game, which pays very close attention to the rules and details of the game and concentrates on the description of the game’s dynamic sequential movement. At other times, however, we may be content with a less stringent step-by-step description of the game’s rules and examine only the actions or strategies available to the players and the payoffs associated with such strategies. Such a description is called the normal or strategic form of the game. At yet other times, even less information about the game will be required, and we may merely want to know what payoff or set of payoffs a single player or coalition of players can guarantee themselves if they act in concert, no matter what the remaining players in the game do against them. This description is called the coalitional or characteristic function form of the game.

In addition to describing games differently, we can also categorize them as cooperative or non-cooperative, depending upon whether the players can communicate with each other and make binding agreements with respect to how they will correlate their actions (cooperative games) or whether such communication and contracting possibilities are ruled out (non-cooperative games). Also, if by a suitable order-preserving linear transformation of the players’ utility functions a representation can be found for which the sum of the players’ utility payoffs is constant for all strategy combinations, we can speak of an n-person constant-sum game or, since the constant may be zero, an n-person zero-sum game. Intuitively in such games the interests of any player or group of players are totally opposite to those of
the players remaining in the game. If no such utility transformation exists, the


form, then the resulting game is called a supergame (see James W. Friedman, 1977), and the strategy sets involve infinite sequences of moves.

2.1.3. Games in coalitional form

As we have said above, there may exist situations in which we will be content to study games from a distance and ignore the details of the game’s rules or even strategy spaces and concentrate merely on a description of what each coalition of players can guarantee itself if it formed and acted in concert. Such a description is contained in the coalitional or characteristic function form of the game. It is given by the so-called characteristic function \( V \), which associates with each coalition \( K \subset I, K \neq \emptyset \), the set of payoff vectors \( V(K) \subset \mathbb{R}^k \) (a \( k \)-dimensional Euclidian space) \( k = \# K \) (number of elements in \( K \)), that are attainable by \( K \). It is usually assumed that for two disjoint coalitions, \( K \) and \( L \):

(i) \( V(K) \) is non-empty, closed, and convex;
(ii) \( x \in V(K), y \in \mathbb{R}^k, y \leq x \rightarrow y \in V(K) \);
(iii) \( K \cap L = \emptyset \rightarrow V(K) + V(L) \subset V(K \cup L) \).

While the first two axioms are more of a technical character, the interpretation of the third axiom, the so-called superadditivity assumption, is of substantive significance: It means that if two coalitions of players which have no member in common join their forces, they can do at least as well as they could have if they were alone.

In investigating this game we may define \( V_a(K) \) as the set of all payoff \( k \)-vectors \( x_K \) such that there exists a correlated strategy for coalition \( K \) that guarantees at least \( x_K \), independently of the (correlated) strategy chosen by the complementary coalition \( K^c \). An alternative definition would be to denote by \( V_{\beta}(K) \), the set of all payoff

\( k \)-vectors \( x_K \) such that there does not exist a correlated strategy for \( K^c = I \setminus K \), the set of agents remaining after subtracting the players in coalition \( K \), which prevents \( K \) from receiving at least \( x_K \). Both characteristic functions \( V_a \) and \( V_{\beta} \) fulfill the axioms stated above, and it is easy to see that for any underlying normal-form game, \( V_a(K) \subset V_{\beta}(K) \) always holds for all non-empty \( K \subset I \). This definition will be important when we talk about the application of game theory to the externality problem in Section 3.4.1.

The description of games in coalitional form can be greatly simplified by the assumption of side payments and transferable utility. We speak of games with side payments if some means of payment, often called “money,” is available to the players (usually in unrestricted quantities), which can be freely transferred between them before or after the play, and for which the increment to the payoff of a player generated by a transfer of money is proportional to the amount of money transferred (by suitable linear transformations of the utility scales we can always make the proportionality constants equal for all players such that a redistribution of money among the players in coalition \( K \) will not change the sum of their payoffs associated with a certain outcome of the game; see Robert J. Aumann, 1960). The maximum attainable sum of utility for a coalition \( K \) in a game with transferable utility, \( u(K) \) is called the worth of coalition \( K \).

2.2. Solution Concepts

When it comes to defining what a solution to a game is, one must first specify the type and descriptive form of the game under investigation. For instance, if the game is a cooperative \( n \)-person side payment game in coalitional form, one may be looking for a utility or payoff vector \( x \) that disperses a sum of utility to the players and which is stable in some meaningful
sense. In this case the job of game theory would be to furnish such stability notions. On the other hand, if the game were a noncooperative n-person non-zero-sum game in normal form, then the game theorist would search for a set of strategy n-tuples which, when taken together, would form an equilibrium for the game. Again, the contributions of game theory would be in defining the proper equilibrium notions to be used. Let us look at these notions one at a time.

2.2.1. Noncooperative solutions

A game is called a noncooperative game if no player possesses the commitment power to enter into binding contracts with other players or to make credible promises and threats. In other words, whether the rules of the game allow for pre-play communication or not, a noncooperative game is characterized by the complete absence of trust among the players. In such a situation, a player i who expects the other players to choose strategies \( s_i, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n \) can himself be reckoned upon to select a strategy \( s_i \), maximizing his payoff given his anticipation of the other players' behavior. Thus, denoting by \( s_i | s_i \) a strategy n-tuple in which the \( i^{th} \) component has been replaced by \( s_i \), we define a best reply \( s_i^* \) for player \( i \) by \( P_i(s_i | s_i^*) \geq P_i(s) \) for all \( s_i \in S_i \). However, if he ascribes the same degree of rationality to the other players, player \( i \) will expect them to follow the same decision rule. This may lead into an expectational circle of the kind "If they know that I know that they know \( \ldots \)," which is typical of the decision problem any individual has to face when he recognizes the interdependence of actions in a game. Only if a strategy combination \( s^* \) existed for which

\[
P_i(s^*) \geq P_i(s_i^* | s_i)
\]

for all \( s_i \in S_i \) and all \( i \in I \),

is this expectational circle broken, since in \( s^* \) each player's strategy \( s_i^* \) is a best reply to the other players' strategies. Such a strategy n-tuple \( s^* \) is called a noncooperative equilibrium or Nash equilibrium point of the game (John F. Nash, 1950). It is the basic noncooperative solution concept and may be applied to games in strategic form or games in extensive form. For a continuous n-person game in normal form, an equilibrium point always exists (but need not be unique) if the players' strategy set consists of convex and compact (closed and bounded) subsets of finite-dimensional Euclidean spaces, \( S_i \subset R^{n_i} \) and their payoff functions are continuous functions with \( P_i \) being concave in \( s_i \) (theorem of Hukai Kaneko Nikaido and Isoda; see, e.g., Burger, 1963). If the game is finite, the above theorem implies Nash's theorem (1950), which demonstrates that every finite n-person game has an equilibrium point in mixed strategies (probability mixtures over pure strategies). A further specialization is von Neumann's famous minimax theorem, which guarantees the existence of a mixed-strategy equilibrium point for every matrix game (i.e., finite two-person zero-sum game).

The definition of a noncooperative equilibrium implies that no single player has an incentive to deviate from it. But what about coalitions? What if, say, by chance a group of players deviates and finds that each of them benefits from this deviation if the rest of the players retain their strategies? This motivates the definition of a strong equilibrium point as a strategy n-tuple, \( s^* \), for which

\[
[P_i(s^*)]_{i \in K} \geq [P_i(s_\varepsilon, s_{\varepsilon-K})]_{i \in K}
\]

holds for all \( K \subset I \), and where \( (s_\varepsilon, s_{\varepsilon-K}) \) denotes a strategy combination which is obtained from \( s^* \) by replacing the \( i^{th} \) components for all \( i \in K \) by \( s_i \). The set of strong equilibrium points is obviously a subset of the set of ordinary equilibrium points (and also of the set of self-policing Pareto-efficient points). Usually, of course, such strong equilibrium points will not exist. However, for properly defined super-
games the set of payoffs associated with the strong equilibrium point of the supergame coincides with the \( \beta \)-core of the original (or iterated) game (Aumann, 1959).

2.2.2. Cooperative solutions

As we have said before, a game is called cooperative if the players are free to communicate with each other and have full trust that their agreements with other participants will be binding. However, since the players can freely communicate with each other, we can expect that the outcome of the game under investigation will in essence be the outcome of the face-to-face bargaining that results from the situation. Hence, let us look at the solution concepts that have been devised to capture what a stable outcome to such a bargaining problem might look like. In order to keep things simple we present only the most important solution concepts for games with side payments and transferable utility; furthermore, we impose no restrictions on coalition formation; i.e., in view of the superadditivity of \( v \), we assume that the grand coalition \( I \) will form. Consequently, the only remaining problem is the distribution of the worth \( v(I) \) of the grand coalition among the players in a manner that is stable. To investigate this problem, let \( x \in \mathbb{R}^n \) denote a payoff vector, and \( x(K) = \sum_{i \in K} x_i \). We call an \( x \) satisfying:

(i) \( x_i \geq v(i) \) for all \( i \in I \),

(ii) \( x(I) = v(I) \),

an imputation for a given game \( v \). Thus, an imputation is a utility distribution exhausting the worth of the grand coalition and assigning to each player at least the amount that he can guarantee for himself without cooperation. We denote by \( X(v) \subset \mathbb{R}^n \) the set of all imputations for a given game \( v \). A solution is a subset of \( X(v) \) consisting of imputations that will not be eliminated in the bargaining process. Different hypotheses about the conditions for an imputation to be eliminated by some other imputation lead to different solution concepts.

Let us assume that in the course of bargaining an imputation \( x \) comes up that does not satisfy some player \( k \). We say that a player has an objection \( (y,K) \), \( k \in K \subseteq I \), against \( x \) if

(i) \( y_i > x_i \) for all \( i \in K \),

(ii) \( y(K) \leq v(K) \).

If (i) and (ii) hold, we also say that \( y \) dominates \( x \) (or, \( x \) is dominated by \( y \)) via coalition \( K \). We say that a player \( l \in K^c \) has a counterobjection \( (z,L) \), \( l \in L, k \in L, (y,K) \) if

(i) \( z_i \geq x_i \) for all \( i \in L \setminus K \),

(ii) \( z_i \geq y_i \) for all \( i \in L \cap K \),

(iii) \( z(L) \leq v(L) \),

i.e., if player \( l \) can point to a coalition \( L \), which would be willing to help him protect his payoff \( x_i \) in \( x \). An objection is said to be justified if there is no counterobjection to it.

The hypothesis that in the course of bargaining an imputation will be already eliminated if some objection to it can be found leads to the concept of core. The core \( Co(v) \subset X(v) \) of a game \( v \) is defined as a set of imputations to which no objections exist or, equivalently, are not dominated by any coalition. Let us call \( e(x,K) = v(K) - x(K) \) the excess or complaint of \( K \) in an imputation \( x \). It is easy to see that

\[
Co(v) = \{ x \in X(v) | \max_{K \subset I} e(x,K) \leq 0 \}.
\]

The assumption that an imputation will be dropped as soon as any objection to it arises is obviously rather strong. Consequently, the core of a game may be empty; e.g., all \( n \)-person constant-sum games with \( n \geq 3 \) have an empty core. If we assume instead that an imputation is eliminated only when a justified objection is raised, we are led to the concept of bargaining set (see Robert Aumann and Michael Mas-
Another solution concept, the nucleolus, was defined by David Schmeidler (1969). In essence, the nucleolus chooses that imputation which minimizes the maximum complaint that any coalition could have against any imputation. It has been shown that the nucleolus $Nu(v)$ is always non-empty, consists of a single point and, moreover, is a continuous function of $v$. If the core is non-empty, then the nucleolus is a member of the core.

Von Neumann and Morgenstern (1947) put forth a solution concept which, though using the domination relation, is more sophisticated and complicated than the other equilibrium set solutions so far described. A von Neumann-Morgenstern solution $NM(v) \subseteq X(v)$ to a game $v$ is defined as a set of imputations satisfying two conditions:

(i) any two imputations in $NM(v)$ must not dominate each other ("internal consistency");

(ii) any imputation not belonging to $NM(v)$ must be dominated by some imputation in $NM(v)$ ("external protection").

While the core, e.g., does fulfill the internal consistency requirement, it is in general not true that every non-core imputation is dominated by some imputation in the core. If a solution $NM(v)$ exists, the core $Co(v)$ is always a subset of $NM(v)$. However, though for most games there is a multiplicity of $NM$-solutions, each one characterizing a different stable standard of behavior (or institutional arrangement) for the players of the game, there are games, with possibly non-empty cores, which do not possess an $NM$-solution (William Lucas, 1968; Shapley and Shubik, 1969b).

2.2.3. Value solutions

The motivation of the value approach to the solution of cooperative games is the idea that a "rational" player should be able to evaluate the prospect of having to play a game in terms of a unique utility assignment, as he is able to do this with other prospects. Thus, intuitively, the value of a cooperative game is the (uniquely determined) payoff vector that the rules of the game hold in store for the players if they behave rationally.

Shapley was the first to define axiomatically a valuation function for $n$-person games in characteristic function form with side payments and transferable utility (1953). This function associates with every $v$ a payoff vector $\Phi(v)$ called the Shapley value of the game $v$. It fulfills a set of axioms that basically require that the value be an imputation and be an additive symmetric function. There is one and only one function fulfilling these axioms, and it assigns to each player $i$ in $v$ the payoff

$$\Phi(v)_i = \sum_{k \in I} \frac{(n - k)!(k - 1)!}{n!} \cdot [v(K) - v(K - \{i\})]$$

where $k = \# K$. This formula suggests a probabilistic interpretation: $v(K) - v(K - \{i\})$ is the marginal contribution of player $i$ to the worth of coalition $K$; $(n - k)!(k - 1)!/n!$ is the probability that in a random build-up of the grand coalition $I$, player $i$ will be the player joining a coalition consisting of the first $k - 1$ players in that random ordering. Thus, the Shapley value may be interpreted as giving each player his expected marginal contribution.

3.1. Public Utility Pricing, Cross Subsidization, and the Allocation of Joint Costs

$N$-person cooperative game theory is essentially a theory of $n$-person multilateral bargaining. Consequently, it is not surprising that problems in the theory of
public utility regulation and pricing should have been analyzed game theoretically, since the public utility pricing problem involves the question of how members of society (i.e., different classes of customers) will jointly split the costs of providing public utility services. While this problem is not a new one, game theorists have turned their attention to it only in the past twenty years, starting with the work of Karl Borch (1962), Shubik (1962), Jan Mossin (1965), Stephen Littlechild (1970b, 1975), Edna Loehman and Andrew Whinston (1971; 1974a; 1974b), Gerald Faulhaber (1975), Mikio Nakayama and Mitsuo Suzuki (1977), Louis Billera, David Heath, and Joseph Raanan (1978), and Faulhaber and Edward Zajac (1976), to mention just a few.

3.1.1. The problem

To get an intuitive view of the problem, let us follow an example given by Faulhaber and Zajac (1976). Assume that the cost of building a dam for irrigation purposes alone is $10 million, while the cost of building the dam for flood control alone is $8 million and that the total cost of the dam for both purposes is $15 million. Seven million dollars is then the extra or incremental cost of providing irrigation, given that the dam is already built for flood control, while $5 million is the incremental cost of building the dam for flood control, given that the dam is already built for irrigation. These costs are directly attributable to the projects separately. What about the joint $3 million? How should the costs be imputed so that neither feels that he is subsidizing the other? The argument put forward by economists is that as long as a group of customers contributes enough to cover its incremental cost, then the other group cannot have a complaint against the cost sharing scheme since, if the original group exactly covers its incremental costs, the other group is no worse off than it would be alone, and if the other covers more than incremental costs, its own contribution is diminished. More formally, if \( C_i \) \( i = 1,2 \) represents the incremental costs of user group \( i \) and \( R(i) \) is its revenue contribution,

\[
R(i) \geq C(i) \quad i = 1,2
\]

is a necessary condition for a set of fees to be subsidy-free. Put differently, if the total amount of revenue collected equals the total cost of the project and if \( C(i) \) is the cost of group \( i \) building the dam totally on its own (i.e., $10 million and $8 million), then subsidy-free prices require that each group contribute less in revenues than its "stand-alone" costs would be, yielding

\[
R(i) \leq C(i). \quad (2)
\]

The problem becomes more complex when there are more than two consumer groups. Assume that in addition to irrigation and flood control, the dam will also be used for electric power generation. Now we have three user groups indexed \( i = 1,2,3 \). Upon reflection, it quickly becomes apparent that the stand-alone test,

\[
R(i) \leq C(i) \quad i = 1,2,3 \quad (3)
\]

is no longer sufficient to guarantee the absence of cross-subsidization, since while (3) guarantees that no individual group is paying more than it would if it provided the service to itself alone, this is not true for coalitions of groups.

Consequently, the existence of subsidy-free prices under these circumstances would require the simultaneous satisfaction not only of (3) but of

\[
\sum_{i \in S} R(i) \leq C(S) \quad (4)
\]

or

\[
\sum_{i \in S} R(i) \geq C(S) \quad \text{for all } S \subseteq I, \quad (5)
\]
where \( S \) is any coalition of players taken from the set \( I \), the set of all players. Faulhaber and Zajac (1976) call (4) and (5) the “Generalized Stand Alone Test” and the “Generalized Incremental Cost Test,” respectively. They are equivalent if \( \sum_{i \in I} R(i) = C(I) \).

This setup of the cost-sharing problem is immediately adaptable to game theoretical modeling. The cost function determines the characteristic function of a “cost-sharing game” which, because of its opposite sign, is assumed to be subadditive instead of superadditive, while the vector of revenue contributions represents a payoff vector. If the vector satisfies (3) and the cost coverage condition,

\[
\sum_{i \in I} R(i) = C(I),
\]

the vector is called an imputation. Furthermore, the set of vectors that simultaneously satisfy (3), (6), and (4)—or (5)—represents the set of revenue contributions in the core of the game, since these imputations cannot be blocked by any coalition of players, in the sense that each coalition would find it less costly to contribute its \( R(S) \) than build the project itself at a cost of \( C(S) \), since \( R(S) \leq C(S) \).

Research in this area divides itself naturally into three branches, each one employing another solution concept to solve the joint cost-imputation problem. One line of research, followed by Shubik (1962), Mossin (1968), Littlechild (1970b), Loehman and Whinston (1971; 1974a; 1974b), and Billera, Heath, and Raanan (1978), proposes to allocate costs by employing a variant of the Shapley value to the cost-sharing game. Others (Littlechild and G. Thompson, 1977; Littlechild and K. Vaidya, 1976; Nayakama and Suzuki, 1977) use the nucleolus, while still others (Littlechild, 1975; Faulhaber, 1975; John Sorensen, Tschirhart, and Whinston, 1976; 1978) investigate the core. (Actually, Littlechild and Thompson employ all three concepts [1977].) Consequently, we will break our survey up into these three categories.

3.1.2. The Shapley value

One of the first attempts at applying game theory to a problem similar to the one described above was Shubik's article, “Incentives, Decentralized Control, the Assignment of Joint Costs and Internal Pricing” (1962). Here Shubik discusses the problems involved in allocating joint costs to the various departments of a complex organization. Phrased in a somewhat team theoretic manner in which the incentive problem is obvious, Shubik criticizes conventional cost accounting practices for not providing the proper set of decentralized incentives to departments to innovate. To provide such incentives, Shubik suggests that the profits of the organization be divided as follows:

\[
\Pi_i = \sum_{S: i \in S} \frac{(n - 1)!}{n!} \cdot [V(S) - \alpha(S - [i])],
\]

where \( \Pi_i \) is the profit for department \( i \), \( n \) is the number of departments in the organization, and \( V(S) \) is the joint profit of a coalition of \( s \) departments, assuming the other \( n - s \) do not operate. However, since \( \Pi_i = \Pi_i - C(i) \), where \( \Pi_i \) is the profit of department \( i \) before its joint cost charge, and \( C(i) \) is the charge to department \( i \), this Shapley value \( \Pi_i \) defines joint cost shares for the departments \( i = 1, \ldots, n \) in the organization; they are the Shapley values associated with the profit game defined by the organization. If the organization allocated profits this way, then: (1) each department is allocated at least as much profit as it would receive if it were to produce by itself (the stand-alone test); (2) any action (i.e., cost-saving technological innovation) considered to be undertaken by any department that
increases the organization’s profits will never cause the profit share of that department to fall.

In a problem closer to the one explained at the outset of this section, Loehman and Whinston offer a method of public utility pricing, which is based on the incremental cost principle but which provides a rigorous method for defining incremental cost (1971). They demonstrate that the prices determined by their method reduce to marginal cost pricing, average cost pricing, and multi-part pricing under various conditions. The Loehman–Whinston argument is simple. Since marginal cost pricing fails to be totally desirable when decreasing costs are present, a case can be made for incremental cost pricing as described by Arthur Hazelwood (1951), Royall Brandis (1953), and Ronald H. Coase (1946). However, a problem exists in rigorously defining incremental costs when capital costs are joint and consumers are heterogeneous, since the increment in cost created by a given consumer is a function of when one decides to include his demand in the calculation of total demand. To handle this problem, Loehman and Whinston employ a Shapley value-like procedure using the joint cost function of the problem and ask all users to pay their expected marginal or incremental contribution to joint costs. In the case where there are \( n \) users with \( k_1, \ldots, k_n \) demands \( (K = k_1 + k_2 + \ldots + k_n) \), the Loehman–Whinston cost scheme would have each user pay

\[
\alpha(i) = \sum_{\begin{subarray}{l} \sigma \subseteq G \\ i \notin \sigma \end{subarray}} \frac{(n-g)![(g-1)!]}{n!} [\alpha(k_\sigma) - \alpha(k_{\sigma-i})],
\]

where \( I \) is the group of all users, \( G \) is a subgroup of \( g \) members of \( N \), and \( K_\sigma = \sum_{k_i \in \sigma} k_i \).

Here again, in general \( \sum_{i \in G} \alpha(i) = \alpha(K) \). They demonstrate that under certain conditions this scheme reverts to either a marginal cost pricing scheme or an average cost pricing scheme.

Finally, Littlechild (1970b), in a more game theoretical context, uses the Shapley value to calculate “fair” and efficient prices for telephone service. The continuous Aumann–Shapley value is used for telephone pricing by Billera, Heath, and Raanan (1978).

Unfortunately, despite the many appealing properties of schemes employing Shapley value-like calculations and despite the fact that they all satisfy the stand-alone test for individuals, they may fail to satisfy the stand-alone test for coalitions or fail to satisfy the “Generalized Stand Alone Test” of Faulhaber and Zajac, since the Shapley value of a game may not be in the core of the game and only core imputations satisfy both the individual and generalized stand-alone test. As a result, there are two ways left for investigators to proceed. One way is to investigate the properties of various schemes that determine core capital charges. Here, however, the charges defined may not be unique. This approach is followed by Littlechild (1975), Littlechild and Thompson (1977), Faulhaber and Zajac (1976), Faulhaber (1975), and John Sorenson, John Tschirhart, and Andrew Whinston (1976). Another approach would be to search for a unique imputation that is “fair” and subsidy free while being in the core of the cost-sharing game if it is non-empty. This approach, followed by Littlechild and Thompson (1977), Nakayama and Suzuki (1977), and others, uses the nucleolus as a solution concept. Others, such as Dermot Gately (1974) and Littlechild and Vaidya (1976) have developed their own concepts. Let us look at these approaches.

3.1.3. The nucleolus

As was discussed in Section 2, the nucleolus is a solution concept first suggested by Schmiedler (1969), which presents another fairness criterion upon which to
share joint costs. In brief, it is that cost-sharing imputation that minimizes the maximum complaint that any coalition could have against it, where a complaint in this context measures the difference between the sum of the fees charged to a coalition and the total cost to that coalition of providing the service under consideration to itself. If the game under investigation has a non-empty core, then the imputation associated with the nucleolus would be in it. Consequently, if the core is non-empty, the cost-sharing scheme proposed by the nucleolus would satisfy both the stand-alone and the generalized stand-alone tests of Faulhaber and Zajac. Furthermore, in terms of equity, since the nucleolus aims to minimize the maximum complaint of any coalition, it has been suggested by Nakayama and Suzuki (1977) and Littlechild and Thompson (1977) that the nucleolus is comparable to a Rawlsian cost-sharing scheme whose purpose is to maximize lexicographically the payoff of the worst off player in the game (although it is not obvious that they are strictly equivalent).

Despite these drawbacks, Nakayama and Suzuki (1977) and Littlechild and Thompson (1977) have used the nucleolus to compute "fair" cost shares in situations involving joint costs. Nakayama and Suzuki employ this method to study the case of the Kanagawa prefecture in Japan (1977).

Littlechild and Thompson use the nucleolus to calculate fair and efficient landing and takeoff fees for the Birmingham, England, airport (1977). Treating each aircraft landing and takeoff as a separate player, they construct a game with 13,572 players in which the core, due to economies of scale in runway construction, is non-empty. They then calculate the nucleolus for this game, which prescribes the fees for all aircraft of differing sizes. Looking at the capital cost components of the fees actually charged during 1968–69, Littlechild and Thompson calculate that of the 11 aircraft types that land at Birmingham, the smallest 8 types are being charged £108,148 for runway facilities, which could have been provided for £104,849—meaning that for that coalition of aircraft types, the charges levied failed to satisfy the generalized stand-alone test. Also, the larger aircraft were seen to be slightly subsidized by the smaller aircraft.

3.1.4. The core

From all that has been said about public utility pricing in this section, it is not surprising that theorists have investigated the properties of the cores of various cost-sharing games. The basic models to follow in this analysis are offered by Littlechild (1975) and Faulhaber (1975). In the Littlechild model linear programming is used to calculate joint cost fees for individuals. Coalitions are viewed as voluntary clubs that can form at will and supply themselves with the good under consideration.

From the programs we can calculate the fixed capital charges for each player that satisfy the constraints, which are basically the generalized stand-alone tests for coalitions and what Littlechild calls the "ability-to-pay" constraints for individuals. These charges characterize Jack Wiseman's "competitive producer club" charges (1957). Littlechild then shows how these charges are related to the core of an associated game.

Faulhaber's paper concerns itself with a similar question, namely the question of finding core stable prices (1975). However, it is pointed out that subadditivity of costs in technology is not a sufficient condition for non-empty core in the cost-sharing game he analyzes. In addition, Faulhaber demonstrates that in general, there is no reason to think that second-best optimal prices are subsidy free (in the core). Consequently, if we are going to elect to charge efficient prices, which maximize social welfare, we may have to allow
a certain bit of cross subsidization in fees.

Sorensen, Tshirhart, and Whinston study the problem of designing optimal two-part tariffs for industries with decreasing average costs and peak-off-peak demands (1976). Setting up what they call the peak-load pricing game, they demonstrate that while the core of such games is always non-empty, the set of imputations in the core cannot always be achieved by a two-part tariff pricing system. William Sharkey has introduced a more sophisticated set of demand conditions into these models in order to investigate conditions for the non-emptiness of the core (1979).

3.2. Incentive-Compatible Allocating and Voting Mechanisms: The Design of Optimal Games

As was stated in the introduction, there is a close relationship between game theoretical analysis and questions involving economic institutions. Mostly inspired by the earlier work of Leonid Hurwicz (1960; 1972; 1973), economists have turned their attention to the question of designing allocating mechanisms or institutions that are "satisfactory" in a sense to be defined later. To fix ideas, let us define exactly what an allocating mechanism is.

To begin, assume that a planner exists and that there is a language $M$ with which agents communicate to him. (You may think of him as a Walrasian auctioneer if you like, with all agents sending quantity messages, although $M$ may be more general and include a set of demand functions for various goods, maximum bids, or whatever.) $m_i \in M$ is the message sent by agent $i$. For each agent $i$ in such an economy, we could specify a response rule $f^i$ instructing him how to respond (with what message) if the message vector $m = (m_1, \ldots, m_n) \in M^n$ was sent and the environment is $e \in E$ ($E$ defines the type of economies we are dealing with, i.e., "neoclassical economies" with all the proper convexity assumptions, etc., or economies with externalities). $f^i(m, \ldots, m_n; e)$ is then the message to be sent by $i$ if $(m_1, \ldots, m_n)$ is sent by all agents and $e$ is the environment. It describes how he will respond. $\bar{m}$ is an equilibrium $n$-tuple of messages if $\bar{m} = f(\bar{m}, e)$, i.e., $\bar{m}$ is a fixed point under the response mapping. $\bar{M}(e)$ is the set of equilibrium $n$-tuples for economy $e$.

The outcome of the economy (i.e., the set of final resource allocations) will be determined at the equilibrium by an outcome function $\Phi$: $\bar{M}(e) \rightarrow \mathbb{R}^n$ defined by the planner, which maps or translates equilibrium messages $\bar{m}$ into equilibrium allocations $x \in \mathbb{R}^n$, the commodity space. Since the outcome is defined only at an equilibrium, this mechanism is a type of tâtonnement mechanism in which actions are not taken out of equilibrium. The triple $\langle f, \Phi, M \rangle$ is what we shall call an allocating mechanism.

Now, depending on the outcome function, each agent can be assumed to behave in a manner that maximizes his utility. However, since his only available strategy is to send messages, every mechanism determines a different $n$-person game in which the strategy sets of the players consist of the set of response rules they can employ, which use language $M$, and the payoffs are determined by outcomes reached at the equilibrium of the game.

If we then attempt to compare various mechanisms, we will in essence be comparing various different $n$-person games, and for this purpose we will need some criteria upon which to do this. Fortunately, Hurwicz (1973) has provided us with some very appealing ones of which we will just mention a few. To begin, we may define a mechanism as "privacy-respecting" if the agents in the economy have direct information (as opposed to information they receive from messages) only about themselves (i.e., only about their own preferences and endowments).
In addition, a mechanism could be called "non-wasteful" or optimal if its outcomes are Pareto optimal, and unbiased if given any Pareto optimal position of the economy, there exists an income distribution that yields that Pareto-optimal distribution as an outcome of the mechanism. A mechanism may also be called individually rational if its outcomes are better for, or at least as good as, the nontrade or non-participation outcomes.

One interesting game theoretical problem arises here when we begin to ask whether the mechanisms analyzed are "incentive-compatible" in the sense that they contain equilibria in which the messages sent are the truthful messages of the agents in the economy, since it is only under these circumstances that we can guarantee ourselves that the outcomes of the mechanism (which are defined only over the equilibrium states) are truly Pareto optimal. More formally, a mechanism is "individually incentive compatible" if telling the truth for all agents is a Nash equilibrium.

Now, with these minimal criteria, Hurwicz has proven a rather disturbing impossibility theorem, which states that in pure exchange economies there exists no allocating mechanism that always yields individually rational Pareto-optimal allocations and that is both privacy-respecting and individually incentive-compatible (1972). This is particularly disturbing, since the result holds for exchange economies with private goods only and holds even without the existence of public goods.

It is probably natural to expect that this line of research would someday intersect with the research of Jacob Marschak and Roy Radner (1972), who were simultaneously developing the economic theory of teams. Teams are sets of economic agents who have the same common objective function, but must coordinate their activities in the presence of uncertainty and ignorance of each other's actions.

The intersection of these two lines of research first appeared in a seminal article by Theodore Groves entitled "Incentives in Teams" (1973) (see also Groves, 1970). Basically, Groves asked the question of what happens in a team when all or at least some of the agents on it have divergent interests and preferences or, more specifically, what happens when their utility is not a function of the total team profit but rather of their individual share of the team's profits. The question then is how can we design a profit-sharing scheme or incentive structure that will provide an incentive for the team members to transmit truthful information to the team's administrator. With these divergent interests among the team members, the team problem quickly degenerates into an $n + 1$-person game with $n$ team members and an administrator, and the choice of an optimal incentive structure by the administrator is simultaneously a choice of the optimal $n$-person game he wants his team members to play, where "optimal" means the game that will have the team members transmit truthful equilibrium messages to him so that he can maximize the team's real profits.

Since the team problem was phrased for environments that were not neoclassical exchange economies but rather environments with production and public goods, the full impact of Hurwicz's impossibility theorem did not apply, although a related version of it (Hurwicz, 1975; Jerry Green and Jean-Jacques Laffont, 1976) reappears. This version states that there do not exist allocating mechanisms for economies with public goods, which determine allocations that satisfy the Samuelson-Lindahl condition; are balanced, i.e., collect through taxes an amount exactly equal to what is paid in subsidies or transfers; and have telling the truth as a
dominant strategy. To get an intuitive feeling for one type of incentive-compatible mechanism and a general idea of why it works, we will take a look at a particular mechanism, which is one representative of a general class of mechanisms called “Groves mechanisms” by Green and Laffont (1977) and popularized as “Demand-Revealing Processes” by T. Nicholas Tidemann and Gordon Tullock (1976) (see also Edward Clarke, 1971, 1972). The mechanism we will investigate is one offered by Groves and Martin Loeb (1975) and is relevant to a partial equilibrium setting in which the income elasticity of demand for public goods is zero. Many more mechanisms exist, but this particular one is used for illustrative purposes.

The problem that Groves and Loeb investigate is as follows: Consider a set of firms which use as an input into their production technologies one that is non-exclusionary—a public input. If this is so, and if, as we know, the firms are heterogeneous, efficiency in the Lindahl-Samuelson sense dictates that each firm be charged a personalized price, which is proportional to their true willingness to pay for the use of this public input. Consequently, a strong incentive exists to disguise one’s willingness to pay and underreport. We are back to the “free rider” problem.

Groves and Loeb provide a mechanism in which a central agent elicits profit functions from firms, sets the optimal level for the public input, and charges the firms a cost share that has the surprising property that reporting the truthful profit function to the central agent (called the Center) is a dominant strategy for the game defined by the mechanism (1975). To describe the mechanism, let us consider \( n \) firms indexed \( i = 1, \ldots, n \). Let \( k \) be a public input whose level is to be determined and let \( r_i \) be the net revenue of firm \( i \) (gross revenue less input costs except for the cost of \( k \)), and let \( r_i \) be a function of \( k \) and some private decision \( L_i \), concerning the inputs of other (non-public) inputs. We can then write

\[
\eta = r_i(k, L_i) \quad i = 1, \ldots, n \tag{9}
\]

Assume that (a) the quantity of \( k \) is known when firms choose \( L_i \); (b) that for every quantity of \( k \) there exists a level of \( L_i \) that maximizes net revenues; (c) firms are profit maximizers; and (d) the difference between the firm’s actual profits and net revenues is independent of their choice of \( L_i \), so that all marginal decisions concern \( L_i \) given \( k \). We may then write

\[
\Pi_i(k) = \max_{L_i} r_i(k, L_i) \tag{10}
\]

Assume for all \( i \), \( \Pi_i(k) \) is strictly concave and everywhere differentiable for \( k \geq 0 \), with \( \lim_{k \to 0} \Pi_i(k) = 0 \). Let \( \pi \) denote all such functions.

Since the free rider problem is a real one in this context, we may think of trying to solve it by centralizing decision making in the hands of a central authority, or “center” for short. The center would then try to maximize joint profits by choosing \( k^* \geq 0 \) to maximize the sum of joint profits \( \Sigma \Pi_i(k) - p \cdot k \) (where \( p \) is the marginal cost of the public input). The problem is, of course, that the center does not know the true profits functions, \( \Pi_i(k) \), of the firms and must rely on their messages of reported profit functions to make its decision. Consequently, let \( m_i \) be a function from the set \( \pi \) (of all net revenue functions), which the center will interpret as its revenue functions. Thus, for a given \( n \)-tuple \( m = (m_1, \ldots, m_n) \), the center will maximize the “reported” joint profits by choosing \( k(m) \geq 0 \) to maximize \( \Sigma m_i(k) - pk \). If all firms send their true message, then \( k(m^*) = k^* \). The question is, can we devise a scheme that induces the firms to send truthful messages.

To demonstrate the Groves-Loeb
scheme, let \( c_i = c_i(m) \) be the cost share of the \( i^{th} \) firm, which is a function of the messages sent by the firms. After paying \( c_i \) the \( i^{th} \) firm’s final profits are

\[
\omega_i(m;c_i) = \Pi_i[k(m)] - c_i(m), \quad i = 1, \ldots, n.
\]

The incentive problem is then formulated by Groves and Loeb in the following terms: Do there exist cost sharing functions, \( c_i(m) \) \( i = 1, \ldots, n \), such that for every \( i \), \( m^*_i \) maximizes \( \omega_i(m,c_i) \) over all \( m_i \) in \( \pi \) for every \( m \setminus m_i = (m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_n) \), that is, \( \omega_i(m/m^*_i,c^*_i) \geq \omega_i(m,c_i) \) for all \( m_i \) in \( \pi \) where \( m/m^*_i = (m_1, \ldots, m_i^*, \ldots, m_n) \). Any \( n \)-tuple of cost share functions \( c^* = (c^*_1, \ldots, c^*_n) \) solving the incentive problem is an optimal incentive structure.

Notice that the incentive problem is solved by devising cost share functions that make telling the truth a dominant strategy. Clearly this is stronger than requiring that sending truthful messages be Nash equilibrium strategies. To solve the incentive problem, Groves and Loeb offer the following cost share function,

\[
c_i(m) = -\sum_{j \neq i} m_j[k(m)] + p \cdot k(m) + A(m \setminus m_i), \quad (11)
\]

where \( A(m \setminus m_i) \) is merely a constant that depends on the \((n-1)\)-tuple of messages sent by all of the other firms except the \( i^{th} \).

The final payoff to any firm \( i \) then becomes

\[
\omega_i(m,e) = \Pi_i[k(m)] + \sum_{j \neq i} m_j[k(m)] - pk(m) - A_i(m \setminus m_i), \quad (12)
\]

This simply says that firm \( i \) receives as its final profits, its true profits \( \Pi_i[k(m)] \) defined by the amount of public input produced by the center as determined by the \( n \)-tuple of messages \( m = (m_1, \ldots, m_n) \), plus the sum of all reported profits of all the firms except itself, \( \sum_{j \neq i} m_j[k(m)] \), minus the total cost of the public input, \( pk(m) \), minus a constant, \( A_i(m \setminus m_i) \), that is independent of the message of firm \( i \).

The cost-sharing functions solve the incentive problem, since given the messages of the other \( n - 1 \) firms in the game, a truthful message \( m_i^* \) by firm \( i \) always maximizes its net profits, \( \omega_i(m,e) = \Pi_i[k(m)] + \sum_{j \neq i} m_j[k(m)] - pk - A_i(m \setminus m_i) \). In other words, we could think of the scheme as one in which all firms \( i \) are faced with a fait accompli in the form of an \((n-1)\)-tuple of messages offered by the other firms and given that \((n-1)\)-tuple their message will alter the already determined quantity of the public input. Consequently, they must decide if they want to make a change in the already determined quantity. In making this decision, the marginal profitability to them of changing the amount of \( k \) to be built by offering a message is:

\[
\frac{\partial \omega_i}{\partial k} = \Pi_i(k) + \sum_{j \neq i} m_j(k) - p = 0, \quad (13)
\]

which states that the \( i^{th} \) firm should keep on changing \( k \) until (13) is equal to zero. However, (13) states that the firm should change \( k \) until its marginal private net profits are equal to its marginal social costs, i.e., the marginal cost of the public good construction \( p \), plus the externalities it places on all other firms \( \sum_{j \neq i} m_j(k) \). This, however, can only be satisfied by telling the truth. The incentive scheme properly internalizes all costs involved in sending a message.

Now, the Demand–Revealing Mechanism—or “Pivot Mechanism,” as Green and Laffont (1977) call it (for situations with discrete or lumpy choices) does have significant limitations. First, it only works when the utility or profit functions of the agents using it have zero-income elasticity for public goods or public inputs because, since the mechanisms involve transfers of income, only in those circumstances is the marginal willingness to pay of any agent
for the public good (public input) independent of the transfer he will receive. Second, while the mechanism does elicit truthful preferences, it does not necessarily determine Pareto optimal (i.e., balanced) outcomes, since all that is collected in taxes is not distributed as transfers. Third, the taxes collected by the center may be so great as to drive some agents into bankruptcy, and no contingency for this eventuality is provided for in the scheme. Finally, these schemes, while not manipulable by individuals, may indeed be manipulable by coalitions of individuals.

The reason why these Demand-Revealing Mechanisms run into such problems is clear in the light of the Green–Laffont (1976) and Hurwicz (1975) impossibility theorems, which state that no mechanism exists that simultaneously yields truthful revelation as a dominant strategy equilibrium, satisfies the Lindahl–Samuelson public goods conditions, and is balanced (i.e., does not collect more—or less—in taxes than it distributes as transfers). Consequently, the Demand-Revealing Mechanism, while it does yield truthful preference revelation as a dominant strategy equilibrium, needs a special zero-income elasticity assumption to do it, is not balanced, and runs the risk of bankrupting some agents.

To rectify these problems, it is clear that at least one of these properties has to be relaxed. Groves and John D. Ledyard in an effort to preserve the balancedness condition restricted the message space allowable for the agents to "quadratic approximations" of their true willingness to pay function and demonstrated that the resulting mechanism both insures balancedness and has truthful revelation of preferences as a Nash equilibrium (1977). Consequently, if one is willing to weaken the equilibrium requirements for preference revelation from dominance to Nash, one can use the Groves-Ledyard mechanism and achieve balanced outcomes at Nash equilibria.

Unfortunately, the Groves–Ledyard (1977) "optimal mechanism" cannot be considered the last word for three reasons: First, the Nash equilibrium condition is a weaker condition than the dominance condition. Even worse, there may not exist a dynamic procedure that will converge to such a Nash equilibrium because at each step of such a procedure the agents may misrepresent their preferences in an attempt to affect the future path of the process. In this connection, however, Jacques H. Drèze and D. de la Vallee Poussin have formulated a dynamic process for which it is a minimax strategy to tell the truth at each step and which converges to a Nash equilibrium (1971). A stronger result is derived by John Roberts (1979), who substitutes myopic Nash behavior at each step for minimax behavior. Although theoretical results in this area are not totally positive, Vernon Smith (1979) has offered experimental evidence demonstrating that the Groves-Ledyard (1977) mechanism does dynamically converge to a Nash equilibrium.

Second, the Groves–Ledyard mechanism may not yield individually rational outcomes so that some agents may wind up worse off from participating in the scheme than they would if they had simply consumed their initial endowment. Green has investigated this problem (1976).

Finally, even if all of these difficulties could be eliminated, the problem of coalitional manipulability still exists, since although the Groves–Ledyard mechanism does yield balanced outcomes at Nash equilibrium, a coalition of individuals may still be able, through coordination, to successfully manipulate the mechanism (the equilibrium may not be a strong Nash equilibrium).

To confront these difficulties, a closely related literature has arisen, which asks
a question for which the incentive problem is slightly submerged. The question asked here, by Hurwicz (1975; 1977a; 1977b); Schmeidler (1978); Ehud Kalai, Andrew Postlewaite, and John Roberts (1978), and others, is not whether there exist mechanisms that are balanced, satisfy the Lindahl–Samuelson or Pareto conditions, and are incentive-compatible, but whether there are individually rational, balanced mechanisms all of whose Nash equilibria yield either Walrasian (in the case of private goods), Lindahl (in the case of public goods), or core allocations. A survey of these questions, however, will be postponed and discussed in our section on Externalities and Public Goods.

In conclusion, game theory has played a major role in the design of allocating mechanisms that are incentive-compatible and consequently has gone a long way in solving the age-old “free rider” problem. The phrasing of the problem in game theoretical terms has allowed for its rigorous analysis and solution.

3.3. Strategy-Proof Voting Mechanisms, the Gibbard-Satterthwaite Theorems, Arrow's Problem, Implementation, and Game Theory

Closely connected with the question of whether allocating mechanisms exist that are incentive-compatible is the question of whether nondondictatorial voting mechanisms or Social Decision Functions (SDF) exist, which are strategy proof or have the property that telling the truth and submitting one's truthful preferences is a dominant strategy in the n-person game defined by the mechanism. Since the problems are somewhat similar, it is not surprising that the results derived have a negative flavor similar to the ones that were derived by Hurwicz (1975) and Green and Laffont (1976) concerning the existence of Pareto-optimal, balanced, incentive-compatible allocating mechanisms.

In short, the central theorems in this literature were offered independently by Gibbard (1973) and Satterthwaite (1975) and state that voting mechanisms do indeed exist that have the property that telling the truth is a dominant equilibrium in the games they define. The only problem is that all such mechanisms must be dictatorial, so that in a society where democratic values are cherished, these voting mechanisms are useless.

To investigate these results let us define a voting mechanism as strategy proof if in the voting game that is implicitly defined by that mechanism it is a dominant strategy for all voters to vote their truthful preferences. Let us call it dictatorial if there exists some voter that can determine the outcome of the voting no matter how the other voters vote. (For a fuller description of these terms see Groves, 1979.) A social decision function is called cheat-proof if for any configuration of truthful preferences, each individual prefers the outcome that is defined by the Social Decision Function when all players report truthfully to the outcome that would result if everyone else reported truthfully and he alone lied.

The central theorem in the literature, known as the Gibbard-Satterthwaite Theorem, can be stated as follows: (a) if a mechanism has at least three possible outcomes and is strategy proof, then it is dictatorial; (b) if a SDF has a range of at least three alternatives and is cheat-proof, then it is dictatorial.

The import of this theorem is that quite devastating, since it states that any voting mechanism, unless it is of the undesirable dictatorial variety, is manipulable by voters who disguise their preferences. In other words, the game theoretical problem involved in making social choices is inescapable.

One of the interesting aspects of this result is the connection that it has to Arrow's famous Impossibility Theorem.
The relationship is spelled out by Satterthwaite (1975) in a correspondence theorem that states that there is a one-to-one correspondence between every strategy-proof voting mechanism and every social welfare function satisfying Arrow's four conditions. This means that if a social welfare function violates Arrow's conditions, then the voting mechanism that that function naturally defines is not strategy proof. However, since no social welfare function exists that satisfies Arrow's four conditions, no nondictatorial strategy-proof voting mechanisms exist (see also Schmeidler and Hugo Sonnenschein, 1976). The two results are equivalent given the correspondence result.

Now although these impossibility theorems carry with them a pessimistic prognosis, there are several avenues of escape. The first is to follow the tradition existing in the Arrow literature and put restrictions on the domain of the SDF by restricting the type of preferences allowed as admissible. Richard Zeckhauser (1973) was one of the first to point out that if preferences of voters were single-peaked, then the manipulation problem would disappear, and since this is exactly the preference restriction used by Duncan Black (1958) to yield satisfactory social welfare functions, it is natural to expect scholars to have followed this approach and to have searched for preference restrictions that yield nondictatorial strategy-proof voting mechanisms when defined over this restricted domain. The most thorough treatment of this approach can be found in Prasanta K. Pattanaik (1978).

Another approach, however, and one that is more game theoretical in its thrust, is to change the type of equilibrium notion and require that the Nash equilibrium of the game defined by the mechanism under consideration both exist and be Pareto optimal. This is identical to the approach followed in the preference revelation literature on allocating mechanisms. When this equilibrium requirement is weakened, Hurwicz and Schmeidler (1976) have shown that the "king-maker" mechanism is a nondictatorial voting procedure that can be used and does have the property that for every profile of preferences there is Nash equilibrium and every Nash equilibrium is Pareto optimal (see also Maskin, 1979). Consequently, strategic dissembling when using this mechanism cannot yield sub-optimal results at Nash equilibria.

The mechanism is not totally satisfactory for several reasons. First, the concept of Nash equilibrium is not as strong as we might like. Second, although the mechanism is non-dictatorial, it does not treat each player in the game symmetrically, since the king-maker is clearly favored. Recently, Maskin has attacked this problem (1980).

3.3.1. Implementability

In recent years Maskin (1978a; 1978b; 1979; 1980); Hervé Moulin (1979; 1980); Peleg (1978); and others have turned their attention to a more general version of the incentive compatibility problem. Basically, the question that is asked here is as follows: Assume that a social planner were able to decide upon a social choice correspondence that would allow him to choose those social states he felt were best given the true preferences of the agents. In other words, assume that the preferential aggregation problem were solved. If this were true, the only question remaining for the planner would be how he was going to implement this social choice correspondence in practice, since in order to use his social choice correspondence the planner needs preference information and, as we already know, agents will misrepresent their preferences if it is in their interest to do so. Maskin's research program in this direction is to contrive what he (and Gibbard [1973]) call a game-form or a set of rules which, if played by the
players, would yield the predetermined set of "optimal" outcomes as equilibria to the game devised. Obviously, one cannot proceed in this line of research until he has specified what equilibrium concept is to be used, and Maskin considers the dominant strategy equilibrium, the minimax, the Nash, the strong Nash, and the Bayesian equilibria as candidates.

Since this research is still in its infancy, we will not spend a great deal of time surveying it. However, it has already yielded some results worth noting. First, one question that is investigated is what types of social choice correspondences can be implemented by game forms using the various equilibrium concepts stated above. One positive result offered by Maskin (1978b) is to demonstrate that if one uses the Nash equilibrium concept as the equilibrium concept, then any social choice correspondence defined on an arbitrary domain can be implemented by a game form if it satisfies two common restrictions: monotonicity and no veto power. If, on the other hand, one wanted to use the strong Nash equilibrium concept as the equilibrium concept, then Maskin shows that social choice correspondences satisfying no veto power restriction cannot, in general, be implemented if the number of alternatives is greater than three (1978a). Peleg has also achieved positive results here (1978). For further results along these lines, see Hurwicz, Postlewaite, and Maskin (1979).

Finally, it must be pointed out that this line of approach is quite general and need not be limited to questions of social choice, but rather can be applied to questions of implementing Walrasian and Lindahl allocations.

3.3.2. Norms, conventions, and incentive-compatibility

When the analysis of all of these mechanisms is over, a question still arises as to whether the problem attacked is of more theoretical than empirical interest. In short, how widespread is the incentive problem in situations where the same sets of agents recurrently face the same preference revelation problem? It is the contention of Schotter (1980) and Simeon Berman and Schotter (1979) that in such situations whether or not the preference revelation problem is a real one depends upon whether or not a convention of behavior (or social institution) is created in which truth-telling is stable. In other words, in many recurrent preference revelation problems, agents do not resort to manipulation because a social convention is created in which truth-telling is an equilibrium. This convention is built upon a system of norms, which are created by the common observation of each other's behavior by the agents. In Berman and Schotter a model is presented that can be applied to the preference revelation problem and which formulates the problem as a stochastic diffusion process in which the absorbing states represent the various (truth-telling or manipulating) social conventions (1979). Consequently, whether or not the preference revelation problem exists in each community is a stochastic event, and Berman and Schotter describe what the probabilities are that a manipulating social convention arises in any given recurrent preference revelation problem (1979). In another vein, Roberts and Postlewaite (1976) and Peter Hammond (1979) demonstrate that as the number of members of a society increases, the incentive problem diminishes so that, contrary to common belief, large anonymous societies are not necessarily more selfish than small ones and the preference revelation problem may not be important.

3.4. Externalities and Public Goods

It is probably natural to lump the problems of externalities and public goods together for analysis, since public goods, at least pure public goods, are prime examples of beneficial externalities. While this is of course true, the questions raised by
the two phenomena are quite different, since with public goods the problem is one of a collective decision on the provision of some non-excludable good, while with externalities the problem is one of preventing or altering a private decision made by some individual agent in the society. Despite this difference, however, one would think that if game theory were to be relevant at all in economics, it would be relevant to those situations in which externalities or public goods existed, since it is exactly those situations which exhibit the most fundamental strategic interdependence among economic agents. As we will see, however, it is in the treatment of externalities and public goods that game theory is most unsatisfactory, especially its treatment of games in characteristic function form.

3.4.1. Externalities

Even a cursory inspection of the literature on externalities reveals two divergent schools of thought on the subject. One school, the traditional or neoclassical Pigovian school, sees externalities as creating a divergence between private and social costs (or benefits), a divergence that must be corrected for by government taxes (subsidies). The basic proposition is one in which prices are failing to transmit the proper information about marginal rates of transformation and consequently their informational content must be altered.

The other school of thought can be called the “Coasian School,” since it grew out of the literature inspired by Ronald Coase’s article, “The Problem of Social Cost” (1960) (see also James M. Buchanan (1975), for a strong supporting view of the use of game theory in a Coasian-contractarian view of the economic problem). In the Coasian tradition, the problem is shifted away from the governmentally assisted invisible hand problem to a problem of bargaining in which the affected party or parties must negotiate compensation from or subsidies to a generator of externality. Consequently, the problem is merely one of determining imputations in the core of some properly defined n-person cooperative game with side payments, an analysis tailor-made for game theory.

By offering the Coasian school a tool of analysis, namely n-person side payment games in characteristic function form, game theory can make a major contribution to the modelling and analysis of the externality problem. In addition, game theory has been used as an analytical tool in pointing out a limitation of the Pigouvian analysis. The problem, raised by Otto A. Davis and Andrew Whinston (1962), is as follows: Assume that we have two firms whose cost functions are related in the sense that the cost function of firm 1 has as its arguments both its own output and the output of firm 2 and similarly the cost function of firm 2 has the output levels of both firms as its arguments. If this is the case, Davis and Whinston argue that the Pigouvian taxation/subsidy solution could only be useful in determining the Pareto optimal output levels of the two firms’ output when the cost functions of the firms are separable. The reason why separability holds out hope for the Pigouvian solution is simple. If the cost functions are separable, then the marginal costs for firm i are independent of the output of any other firm j. Consequently, in setting its output level, firm i produces that quantity at which price equals marginal cost, and since under perfect competition price is a parameter, and under separability marginal cost is independent of firm j’s output, firm i has an optimal dominant output strategy, which is to produce at marginal cost no matter what firm j does. Since the decisions are independent (at least with respect to marginal amounts), taxes and subsidies can be used to guide the firms to a Pareto optimal arrangement.

In the non-separable case the Pigouvian
solution breaks down because non-separability implies that the marginal cost functions of the firms are dependent on each other's output levels. As a result, the game that results does not have a dominant equilibrium in pure strategies and, even if the government placed taxes and subsidies on the firms, it could not be certain that they would choose Pareto optimal outputs, since there is no way to insure that these output levels are dominant equilibria. Consequently, the Pigou–Marshall–Meade solution is brought into some doubt.

Although from what has been said so far in this section it would appear that game theory has found its niche in the Coasian analysis of externalities, things are not quite so easy. To begin, there is a major conceptual problem that exists in defining a characteristic function for situations involving externalities. The problem arises from a distinction made by Shapley and Shubik between what they call “orthogonal” and “non-orthogonal” games (1973). Orthogonal games are games in which the set of utility vectors or the amount of utility (in the transferable case) that a coalition S can guarantee itself in a game is independent of the actions taken by players in S; the complement of S in I, the set of all players. Classical market games for economies with private goods are typical examples of such orthogonal games. It is for such games that the characteristic function is unambiguously defined and a meaningful construct. Non-orthogonal games are games in which the payoffs to a coalition S when they form are not independent of the actions of coalition S'. Clearly, social situations involving externalities and public goods are such cases, since the payoffs to a coalition in such a game are clearly affected by the amount of externality generated by S'. Consequently, the value of a coalition S in such a game is conditional on the actions of coalition S' and in order to meaningfully define the value of coalition S, some type of behavioral assumption must be made for S'. Such assumptions, however, involve a psychological judgment that is extragame theoretical.

To get around this problem, several alternatives have been offered. Aumann and Peleg (1960) and Aumann (1961) distinguish between the V_a and V_b definition of the value of a coalition V(S) for games with nontransferable utility (see Section 2.1.3). Robert W. Rosenthal (1971) has taken a different approach to the problem. He first demonstrates that the a-core concept is not satisfactory for non-orthogonal games, since it may involve an assumption that coalitions take actions that are extremely self-destructive in order simply to punish members of the complementary coalition. He then suggests that it might be meaningful to try to define several distinct types of behavior for complementary coalitions, S'; when a coalition S forms, which do not require that it engage in such spiteful behavior and delineates four such types. These concepts have been used successfully by Donald Richter (1974) in public goods economies and by Schotter (1978) in treating externality problems that exist in urban housing markets.

In summation, then, there does not as yet seem to exist a meaningful way to model economic and social situations exhibiting externalities, since the characteristic function form of the game is not unambiguously defined in these circumstances. The irony is, of course, that it is exactly in these circumstances that we would expect game theory to be of most use.

3.4.2. Public goods

In the theoretical analysis of economies with public goods the concept of a Lindahl equilibrium plays a role similar to the role played by the Walrasian equilibrium for economies with private goods. Conse-
quently, it is understandable that economists would try to prove an equivalence or limit theorem relating to the set of Lindahl equilibria and the set of core imputations. In other words, it is not surprising that some effort would be spent trying to prove a generalized version of Edgeworth’s conjecture in a public goods context. This effort is not senseless, however, since if the conjecture were true for public goods economies, the Lindahl equilibrium would play a role completely analogous to that of the Walrasian equilibrium, and in “large” economies no institutional artificiality would be introduced into the model if we assumed that all agents acted “competitively” or as if a central auctioneer existed to whom all agents reacted parametrically. If the conjecture is false, however, then even in the limit there is more room for bargaining and a need for the creation of institutions other than market institutions in the allocation of public and private goods.

The main result of the research, to date, has been negative. First, following the Aumann-type of model (1964), Thomas Muench has shown that for nonpathological economies with a continuum of agents, the core contains the Lindahl equilibria as a proper subset (1972). In other words, large public good economies are institutionally rich and cannot be treated as if agents behaved according to one and only one institutional arrangement—competitive markets. Second, following the Scarf-Debreu type replication arguments, Paul Champsaur, John Roberts, and Robert Rossenthal demonstrate that unlike the private goods case, where replication (or increasing the number of each type of trader proportionately) can only cause the core to shrink, replication in public goods economies may lead to larger cores, thereby making convergence between the core and the set of Lindahl equilibria less likely (1975). All of these results were anticipated though not proven by Duncan Foley in his article, which first presented a formal game theoretical model of a public goods economy (1970).

The reason for this lack of equivalence arises primarily from the problem of defining a proper notion of blocking for coalitions in public goods economies. In the pure public goods case, the problem is quite evident. If goods are pure public goods (in the Paul Samuelson [1954] sense), the consumption of the good by one more consumer does not at all deplete the amount of the good remaining to be consumed by the rest of society (i.e., the good is provided at zero marginal cost). In addition, no one can be excluded from its use once it is constructed. Consequently, if this definition is to be taken seriously, we are faced with the problem of defining the amount of utility or set of utility vectors that a coalition can guarantee itself when it forms, since we must make an assumption about the amount of the public good they expect to be built by their complementary coalition in I. This is true because of the non-excludability property of pure public goods. The reason, then, for the lack of equivalence between the core and the set of Lindahl equilibria is that, in the literature, it is assumed that when a coalition S forms it can rely on its complement, S^c, in I to produce a zero level of the public good (i.e., it can expect no free ride at all). Clearly, if such an assumption is made, it becomes extremely hard for coalitions (especially small coalitions) to block imputations, since in order to block an imputation a coalition would have to provide itself with an amount of the public good equivalent to what is being offered by society, without expecting any free ride from its complement. If blocking is so hard, it is no surprise that the core is large and that even in the limit it is larger than the set of Lindahl equilibria. The thrust, then, in the literature has been to redefine the blocking notion in order to make blocking
easier and hence to make the core smaller and hopefully equivalent to the Lindahl equilibria in the limit. Two approaches have been followed. The first, attributable to Rosenthal (1971), has already been discussed in our section on externalities and basically tries to impose some type of rationality on a coalition when their complement in I forms. Using this approach, Richter presents a simple example of a public goods economy (1974), which has an empty core under any three of Rosenthal’s four assumptions, and then presents necessary conditions for an empty core in such economies. Further work by Schotter (1979) has demonstrated that under two of Rosenthal’s assumptions it may be disadvantageous for syndicates to form in such economies. Such disadvantageousness results were already demonstrated before by Aumann (1973) and Postlewaite and Rosenthal (1974) for private goods economies.

Another approach to the problem investigates whether the equivalence between the Lindahl equilibria and core allocations exists when goods are only semi-public and when crowding exists. This emphasis on semi-public or local public goods is associated with the work of Barry Ellickson (1973), Donald John Roberts (1974), Mark Pauly (1967), Martin McGuire (1972; 1974), Charles Tiebout (1956), Myrna Wooders (1978), and Roger Guesnerie and Claude Oddou (1979). Basically, a good is semi-public if, given a coalition or neighborhood (if we are dealing with geographical models), all players not in the coalition can be excluded from using the goods at zero cost, but all members of the coalition cannot. By positing semi-publicness or crowding, these investigators felt that they could make blocking easier, since free riding would be harder in an environment where exclusion was possible. Ellickson (1973), however, has demonstrated that this approach may be too drastic, since in economies with crowding, the core may be empty or, if nonempty, it may be totally disjoint from the set of utility vectors associated with Lindahl equilibria. Facing this problem, Roberts (1974) demonstrated a modified form of Edgeworth conjecture, which states that any Lindahl equilibria treating all consumers equally belong to the core and that all the core allocations that treat equals equally can be supported by Lindahl prices. Consequently, for equal treatment of Lindahl equilibria we get an equivalence result, but since many Lindahl equilibria are not equal treatment equilibria, the result is not as strong as we might have hoped.

When costless exclusion enters the model, the public goods problem possesses many of the characteristics of a club-goods problem or local public goods problem analyzed non-game-theoretically by Buchanan (1965), McGuire (1972; 1974), and Tiebout (1956) and game theoretically by Pauly (1967) and Wooders (1978).

It is our feeling that research which tries to force an equivalence theorem onto the relationship between the core of a public good economy and its Lindahl equilibria is in essence trying to create an equivalence that really does not belong there in the first place. A more natural direction to proceed would be that taken by Hurwicz (1975, 1977a, 1977b, 1979) and Schmeidler (1978), in which allocation mechanisms are sought for which the Nash equilibria associated with the games they define are Lindahl (or Walrasian) equilibria and vice versa.

In other words, an equivalence is sought for the set of Nash and Lindahl equilibria.
This is, of course, a more logical type of equivalence to seek for large economies, since the information requirements necessary to establish a core imputation in such economies are so extreme that the core ceases to be an empirically significant solution concept. Just the opposite is true of the Nash equilibrium, however, since that concept is truly meaningful only when transactions costs are so large as to prevent the economically feasible formation of coalitions.

One of the first such mechanisms dealing with mechanisms whose Nash equilibria correspond to Walrasian equilibria was presented by Schmeidler (1976), who presented a balanced outcome function realizing Walrasian equilibria at Nash equilibria. The mechanism was not individually feasible and was discontinuous. In a later paper Schmeidler refined this outcome function and produced one that yielded Walrasian allocations which were strong Nash equilibria (1978). This function was balanced but again not individually feasible when out of equilibrium. Consequently, the possibility of individual bankruptcy out of equilibrium exists, and the fact of this possibility is not treated.

Hurwicz presents a mechanism with an outcome function yielding Walrasian and Lindahl allocations at the equilibrium (1979). His function, while balanced and smooth, is not individually feasible when out of equilibrium. The mechanism works by having an outcome function that creates penalties for all agents who name prices that diverge from the average of the prices named by all other agents, thereby forcing all price vectors to be identical and yielding Walrasian allocations. For the Lindahl allocation version, the penalties are levied for agents whose stated quantity of the public good to be consumed by all agents diverges from the average, thereby forcing the equilibrium quantity of the public good to be equal for all as is required.

In addition to these outcome functions, the Groves–Ledyard mechanism (1977), with its quadratic outcome function, is also a mechanism yielding Lindahl equilibria at Nash equilibria. As we have said before, however, while this function is smooth and balanced, it is not individually feasible.

The literature on this topic is now quite substantial. Among the models existing are ones by Shapley (1977), Shapley and Shubik (1977), and E. Fazner and Schmeidler (undated), who discuss allocating mechanisms using money, whose outcomes are not always in the Pareto-optimal set. In addition, because these outcome functions depend on the initial endowments of money of the traders, the mechanism is not informationally decentralized. Postlewaite and Schmeidler present an associated mechanism that does not use money, but which also only yields approximate efficiency (1978).

A paper by Kalai, Postlewaite, and Roberts presents a mechanism for a public goods economy that yields strong Nash equilibria whose allocations coincide with the core of the economy (1978). Also, because core allocations are individually rational, so are these outcomes.

Finally, let us end our discussion here by stating that all of these questions seem to be heading for a more general treatment where the questions asked above are rephrased by asking whether there exist allocating mechanisms (or game forms) that “implement” various predetermined outcomes (i.e., Lindahl equilibria). (See Hurwicz, Maskin, and Postlewaite [1979] and Postlewaite and Schmeidler [1979]).

3.5. Models of Multilateral Exchange; Games and Markets

3.5.1. Core and competitive equilibria

3.5.1.1. Finite markets. One of the most significant conceptual breakthroughs of cooperative n-person game theory has been the description of an n-person economy (with and without production) as a
game. This description is significant because it allows us to view the economic process without forcing all transactions through the limited institutional perspective of competitive markets. The cooperative game theoretical analysis allows us to investigate the general equilibrium problem in an institution-free context, limited only by an implicit property rights assumption that people have a right to their labor and/or initial goods endowment. The remarkable result proven in the literature is that, as we replicate the economies under investigation, the only institutional structures that remain viable are competitive markets, so that in the limit the game theoretical and the Walrasian analyses merge. Let us see how this occurs.

Consider a (mutilateral) market or (exchange) economy $E$ defined by four components: (1) a set of traders $I$, $i = 1, \ldots, n$; (2) a commodity space, which for the sake of simplicity is taken to be the nonnegative orthant $\mathbb{R}^m_{+}$ of the $m$-dimensional Euclidean space ($m$ being the number of different commodities); (3) the traders’ preferences with respect to commodity bundles, $(\varepsilon_i)_{i\in I}$, usually assumed to possess all the properties well known from general equilibrium theory (completeness, reflexivity, transitivity; continuity; convexity; insatiation or, stronger, monotonicity); and (4) the traders’ initial endowments, $(e_i)_{i\in I}$, $0 < e_i \in \mathbb{R}^m_{+}$, i.e., the commodity bundles owned by the traders. We will also consider so-called $r$-fold replicas $E^r$, which are defined as compound economies made up of $r$ identical subeconomies $E$. Thus, a replica economy $E^r$ is a market consisting of $n$ types of traders, characterized by preferences $\varepsilon_i$ and endowments $e_i$ with $r$ identical traders of each type. We assume that the traders are able to communicate freely and to make voluntary exchange agreements about a redistribution of their initial endowments, but we assume no other institutional constraints. A contract among a group or coalition $K \subset I$ is a provisional agreement on a feasible redistribution $(a^t)_{I\setminus K}$, $a^t \in \mathbb{R}^m_{+}$, $\sum_{i \in K} a^t_i = \sum_{i \in I} e^t_i$, of their initial endowments. For any partition $\mathcal{B}$ of the set of traders, the contracts of groups $K \in \mathcal{B}$ generate a feasible allocation $a = (a^t)_{I\setminus K}$, $\sum_{i \in I} a^t_i = \sum_{i \in I} e^t_i$, for the economy $E$. When a trader $j$, faced with some allocation $a$, finds out that there is a possible contract $(b^t)_{I\setminus L}$, $\sum_{i \in I} b^t_i = \sum_{i \in I} e^t_i$, with trading partners in $L \ni j$, which is strictly preferable to him, $a^t_j \neq b^t_j$, and not worse for the other traders in $L$, $a^t_i \leq b^t_i$, $i \in L$, then he will not be satisfied with $a$ but will recontract with his partners in $L$ to achieve the redistribution $(b^t)_{I\setminus L}$. This is often described as an allocation $a$ being “blocked” via coalition $L$. The process of contracting and recontracting will only stop when an allocation $c$ is reached that cannot be “blocked” via any coalition $K \subset I$. The set of all such allocations $c$ was termed the contract curve by Edgeworth (1881) and is today called the core $Co(E)$ of the economy $E$ (Debreu and Herbert Scarf, 1963; 1972; Hukukane Nikaido, 1968; Arrow and Frank H. Hahn, 1971; Schotter, 1973; Werner Hildenbrand and Alan Kirman, 1975; Leif Johansen, 1978). Obviously any core allocation $c \in Co(E)$ will be Pareto efficient for the whole economy $E$ and, moreover, $(c^t)_{I\setminus K}$ will be Pareto efficient from the respective subeconomy consisting only of traders $K \subset I$. The core $Co(E^r)$ of a replica economy $E^r$ displays an “equal treatment” property insofar as in each core allocation traders of the same type receive a commodity bundle yielding equal utilities. Now in relating this analysis to the neoclassical analysis, we have the interesting result that the core $Co(E)$ always contains the so-called competitive allocations $c^*$ of $E$ or those allocations which result when each trader maximizes his preferences under a budget constraint, $p^* c^* \leq p^* e^t$ (implied by a price system $p^* \in \mathbb{R}^m_{+}$, $p^* e^t \leq c^t \leq c^*$) and the market clearing condition $\sum_{i \in I} c^* = \sum_{i \in I} e^t$. 

Copyright © 2001. All Rights Reserved.
$e^*$ is called a (Walrasian) equilibrium price system, and $(p^*, e^*)$ a (Walrasian) competitive equilibrium (e.g., Debreu and Scarf, 1963). There is, however, a still more remarkable relationship between the core and the competitive equilibrium of an economy, which has enshrined the notion of core to the hearts of general equilibrium theorists: By virtue of the special equal treatment property mentioned above, the core $Co(E')$ of an $r$-replica economy can be described by allocations to representatives of types of traders only, i.e., can be conceived of as subsets of $R^mm$ instead of $R^m$. It is easy to show that, in this sense, $Co(E'^{r+1}) \subset Co(E^r)$. If we consider a sequence of replica economies $E = E^1, E^2, \ldots, E^r, \ldots$ the respective cores $Co(E^1) \supset Co(E^2) \supset \ldots \supset Co(E^r) \ldots$ form a nested sequence having their intersection as the limit $Co(E^\infty)$, when $r$ becomes indefinitely large. Debreu and Scarf have shown that every allocation of $E$ that is in $Co(E')$ for all $r$ is competitive (1963), i.e., $Co(E^\infty)$ equals the set of competitive allocations for $E$. In this sense, "the core converges to the competitive equilibria."

The core of a market is one of those concepts of mathematical economics that were well known in economic theory long before the theory of games, but have since undergone a thorough generalization. Edgeworth studied markets with two commodities and two types of traders and arrived (1881), for this special case, at all the results mentioned above, including the "shrinking" of the contract curve to the competitive equilibrium. Shubik (1959b) was the first to draw attention to the close conceptual relationship between Edgeworth's contract curve and the core of a game (with side payments and transferable utility). Shapley and Shubik introduced a class of cooperative $n$-person games in characteristic function form (1969b; 1975), called "market games," which are derived from transferable-utility markets $[I, R^m, (u_t)_{u_t}, (e^*)_{u_t}]$ in which the ordinal preferences $(e^*)_{u_t}$ of the traders are replaced by the stronger, "cardinal" utility functions, which are linear and separable in "money." Because of the form of the utility function, these Shapley–Shubik markets (sometimes also called markets with money) become immediately amenable to the application of all solution concepts defined for classical characteristic function games. Since externalities are assumed to be absent in these markets, a Shapley–Shubik market game $v$ is "orthogonal" (Shapley and Shubik, 1971), i.e., the worth of a trading coalition $K$ in no way depends on the actions of its complement $\Lambda K$, and the question of the plausibility (in terms of credible threats, etc.) of the characteristic function does not arise in this context. Shapley and Shubik have shown that their market games are always "balanced," i.e., always possess non-empty cores (1969b).

One of the major advantages of choosing a game representation of markets is, on the one hand, that conceptually appealing solution concepts beside the core (e.g., the Shapley value or the $NM$-solution) are applicable, and, on the other hand, that markets not fulfilling some of the neoclassical regularity assumptions (e.g., convexity of preferences, perfect divisibility of commodities, absence of externalities, which jointly ensure the existence of Walrasian competitive equilibria) remain amenable to analysis. Thus, cores and other solutions of games generated by markets with non-convex preferences (Shapley and Shubik, 1966), externalities (Shapley and Shubik, 1969a; Rosenthal, 1971), and indivisible commodities (von Neumann and Morgenstern, 1947; Shapley, 1959; Shapley and Shubik, 1972; Telsoner, 1972) have been studied. A famous example of a market with indivisible commodities is Eugen von Böhm-Bawerk's horse market (1888). It is interesting to note that the core of the corresponding
side-payment market game coincides with the range of prices bounded by the so-called “marginal pair,” Böhm-Bawerk’s solution (Shapley and Shubik, 1972). Telser (1972) demonstrated that in the case of multi-unit trade, when the individuals own or wish to own more than one unit of the commodity, the core of the market game reveals features not captured by an analysis in terms of supply and demand curves or Böhm-Bawerk’s marginal-pair theory. For a single-unit exchange problem with several different commodities, Shapley and Scarf (1974) studied the core in terms of ordinal preferences (see also Roth and Postlewaite, 1977).

It should also be mentioned that the concept of a market or economy \( E \) was extended to include productive capabilities of individuals and coalitions (see Arrow and Hahn, 1971; Volker Böhm, 1973; 1974a; Hildenbrand, 1974).

A new interesting variant of the core of an economy is the concept of sequential core (Douglas Gale, 1978). It refers to a sequential economy where commodities are explicitly distinguished with respect to the date of their delivery. In such an economy we have allocations of goods for spot and future delivery, which appear attainable in periods 1, 2, etc. Thus, the concept of sequential core formalizes the notions of trustworthiness and of an exchange economy without trust, thereby shedding some light on the fundamental role of money. An allocation that was untrustworthy in a nonmonetary economy (did not belong to its sequential core) becomes trustworthy in the corresponding monetary economy for an appropriate distribution of the fiat, government-backed money.

3.5.1.2. Large markets. The fundamental assumption of Walrasian competitive analysis—that all individuals act as price-takers—had always been justified by the intuitive appeal to a number of traders large enough to make each single one rela-

tively insignificant and unable to influence the relations of exchange established in equilibrium. The first person who gave a precise mathematical formulation to this aspect of the idea of “perfect competition” was Aumann (1964), who put forth the notion of an economy with a continuum of traders as a mathematical idealization of a perfectly competitive economy. Although it is of course of analytical interest to study the properties of game theoretic solutions in the limiting case of an arbitrary number of players without individual influence, we will touch on this direction of research only briefly. For, in our opinion, it can certainly not be the main contribution of game theory to economics to demonstrate under what conditions its solution concepts coincide with a non-game theoretic notion like the Walrasian competitive equilibrium with its limited institutional framework. The continuum economies have found their definitive expression in a measure-theoretic framework as markets with a non-atomic measure space of traders (Hildenbrand, 1974), for which the so-called “core equivalence theorem” states the identity of the core and the set of equilibrium (Walras) allocations. For this result to be of any relevance, one has to demonstrate that what is true for the continuum economy holds also for large but finite economies. The sequences of replica economies whose cores converge to the set of equilibrium allocations represent a very special method of enlarging the number of traders. Although an arbitrary procedure of enlarging a given market will not produce the desired result, more general convergence theorems not relying on replication have been proved (see, e.g., the discussion in Hildenbrand (1974) or Hildenbrand-Kirman (1975)). For an example demonstrating that increasing returns to scale may ruin the equivalence results derived, see Dieter Sondermann (1974).

The properties of the cores of markets
with a measure of space of traders consisting of a non-atomic ("oceanic") part and some "atoms," i.e., traders who are not insignificant but have positive individual weights in terms of their shares in the total resources of the economy, have gained considerable attention. The atoms in these "mixed markets" have been interpreted as monopolists and oligopolists, respectively, depending on whether the whole amount of one commodity is concentrated in their hands or not. Benjamin Shitovitz has shown that the equivalence of the core and the set of competitive allocations, which holds when the set of atoms is of measure zero, no longer necessarily holds when this set has positive measure (1973). However, if the set of atoms consists of identical traders, the equivalence is restored. Jean Jaskold-Gabszewicz (1977) demonstrated that these phenomena do not depend on the abstract measure space representation, but also occur when the mixture of small and large traders is obtained via sequences of finite economies (see also M. Ali Khan, 1976). This feature of the Shitovitz model is in keeping with the general approach of the theory of cores of large markets, which derives the competitive equilibrium nature of core allocations not from, say, high communication costs forcing the traders in the oceanic part of the market to behave as nonstrategical price takers, but quite to the contrary, from the very assumption—certainly not a "realistic" one—that also in large markets all "blocking" coalitions are permitted to form. For this reason, the interpretation of the "core equivalence theorem" as an explanation of the emergence of competitive prices is questionable; it just offers an argument for the "stability" of a competitive equilibrium (once miraculously established) against collusion among the traders, even if costs of communication and coalition formation were zero. Shitovitz studies the possibilities that the "oligopolists" have to exploit the other traders and gives conditions guaranteeing that in mixed markets every core allocation is competitive from the point of view of the small traders (belonging to the "ocean") (1974). For homogeneous markets with a monopolist (i.e., markets in which all traders have the same, homogeneous utility function), Shitovitz has shown that the unique competitive allocation is, from the point of view of the monopolist, the worst allocation in the generally quite large core (1973). Thus, as might have been suspected, it is an advantage to be a monopolist. However, Aumann offered quite unpathological examples of mixed markets involving two commodities and one atom who initially holds a "corner" on one of the two commodities for which the core contains allocations that are worse for the monopolist than any of the competitive ones (1973). Thus, it seems that monopolies can be disadvantageous. Aumann argued that these examples of disadvantageous monopolies revealed an analytical weakness in the concept of core (1973). He suspected that a concept like the Shapley value might be better suited to bringing out the intuitive idea of a monopolist's advantage in terms of power and threat possibilities, although this conjecture was later disproved.

3.5.2. Barriers to recontracting

Aumann's counterexamples motivated an alternative interpretation of atomic traders in mixed markets, viz., as so-called syndicates of small traders. A syndicate is a group of traders of the same type, which participates in the contracting and recontracting process as a whole only, so that it can be regarded as a single agent; in other words, only coalitions that contain the whole syndicate or do not contain any of its members are allowed as blocking coalitions (Jaskold-Gabszewicz and Drèze, 1971). Postlewaite and Rosenthal (1974) came to the rescue of the core by
presenting a simple example of a market game in which syndication is disadvantageous in Aumann’s sense, but nevertheless has economic meaning. However, they likewise regarded syndication as something exogenously imposed and not endogenously determined within the game. It is interesting to note, however, that from the point of view of another solution concept, the bargaining set \( M^0 \) (for the grand coalition), monopolistic syndication is not disadvantageous in the example given by Postlewaite and Rosenthal. Michael Maschler (1976), to whom we owe this observation, interprets it as an indication of a certain conceptual superiority of the bargaining set over the core. In our opinion, however, the Postlewaite–Rosenthal example does underline a profound distinction between a monopoly proper, which is rightly treated as a single player, and a so-called collective monopoly or cartel, consisting of several distinct decision-makers who do have the option of leaving the cartel whenever they see fit. Within a true syndicate or cartel, distributional problems will arise or, in other words, as Paul Champasaur and Guy Laroque have pointed out (1976), the assumption made by Shitovitz (1973; 1974) and also by Aumann (1973) that the atoms possess complete preference orderings may be considered unjustified for syndicates.

Terje Hansen and Jaskold-Gabsczewicz have analyzed a related problem in the context of an economy with production possibilities by a somewhat more sophisticated application of the concept of core (1972). Their aim was to make precise a conjecture by Samuelson (1967) according to which “Even under universal constant returns to scale, the competitive configuration is unstable with respect to (costless) collusion of factors of production.” Their findings support Samuelson’s conjecture: The large number of economic agents is not sufficient to immunize the competitive configuration against some forms of collusion of factor owners, since the core may contain allocations that are advantageous compared to the competitive allocations for the organized factor owners.

An approach to the analysis of collusion or cartelization, which radically differs from that of just comparing the cores of market games with and without exogenously introduced syndicates or combinations, is suggested by the concept of \( NM \)-solution. From the point of view of stability there is one serious flaw to the core: There exist, in general, imputations outside the core that are not dominated by any imputation in the core; i.e., the core (as well as the bargaining set) lacks the \( NM \)-solution’s property of “external stability.” For example, take the case of a monopolistic market game in which the core consists of the unique imputation in which the monopolist extracts all consumer surplus to himself, leaving the buyers with zero net gains or imputations. Let this imputation be defined by \( x = (1;0,0, \ldots ,0) \) where \( 1 \) is the monopolist’s payoff and zero the payoff of each buyer. Assume that the bargaining starts with an imputation \( x \neq (1,0, \ldots ,0) \) in which some buyers are necessarily better off than in the core and the buyers are aware of the circumstance that the procedure of free contracting and recontracting (i.e., competition among them) works by offering only temporary gains to some of them while making all worse off in the long run. In this case they can be expected to conclude that it is profitable to stop the process of recontracting at some imputation outside the core, which they are able to agree upon. Thus, collusion may be viewed as the practice of stabilizing dominated imputations (or, allocations that could be "blocked") by means of combinations (like cartels, trade unions, etc.), while competition in this sense may be defined as the absence of any combinations. Results have been obtained supporting the dynamic stability of the core as a set of equilibrium states with respect to the process of recontracting (Green, 1974; Allan M. Feldman,
1974). However, these stability investigations took for granted the viability of the dynamics of recontracting. The above example, however, suggests a different view: If the traders were rational, in the sense of always striving for higher payoffs, and if they knew that the core was the only stable outcome of the recontracting process, then the core would not be stable. Thus, competition can in general only be expected to prevail if the behavior of the traders is characterized by a peculiar mixture of rationality, complete information about the opportunities the market offers, and short-sightedness (Morgenstern and Schrödter, 1976). It has to be pointed out again that in contrast to the core approach to the analysis of cartels, the N*M-solution approach makes no a priori assumptions on how the economic community organizes itself. Which cartels will form is determined by the structure of the N*M-solutions of the original, unorganized market (Sergiu Hart, 1974). For finite markets in two divisible commodities, one of which serves also as a means of side payments and utility transfer, with one seller and one up to three buyers, Morgenstern and Schrödter have determined all symmetric solutions and interpreted them as consistent and defensible rules of division of joint profit to be observed in viable cartel agreements (1976). Gaming experiments based on this model seem to corroborate this interpretation (Yale Braunstein and Schotter, 1978).

3.5.3. Values of market games

The Shapley value of a market game solves the problem of imputation by appraising the marginal contributions of factors of production of traders or owners to the worth of the coalition comprising all agents, the value of the total product, so to speak. In this sense, it may be that game-theoretic solution concept that comes nearest to some early neoclassical, marginalist ideas.

Values of markets with a non-atomic continuum of traders were studied by Aumann and Shapley for the transferable utility case (1974); by Aumann for markets without side payments and for ordinal preferences only (1975); by Hart for both types of markets (1977; 1979); and by Andreu Mas-Colell (1977) to prove an asymptotic version of Aumann’s results (1975). The major conclusion is that for nonatomic economies every value allocation is competitive, and the set of value allocations coincides with the set of competitive allocations (“value equivalence theorem”) only if certain differentiability assumptions on the traders’ preferences (somewhat stronger than Debreu’s [1972] “smooth preferences”) are fulfilled. It has also been shown that in the case of nonatomic markets, the values obtained for concave cardinal utility functions and for purely ordinal preferences coincide.

As mentioned above, Aumann in his critique of the core was more optimistic in his belief that the value would reflect advantageousness of a monopolistic atom over its competitive nonatomic counterpart (1973). This proved indeed true for the bilateral markets studied by Shapley and Shubik (1967) and Champsaur (1975). However, Roy Gardner discovered a broad class of disadvantageous monopolistic situations where rewards are measured in terms of the Shapley values (1976). Guesnerie examined various concepts of disadvantageousness and gave examples, using the replication technique of approaching the limit, of monopolies that prove disadvantageous from the point of view of the Shapley value (1976).

3.6. Models of Oligopolistic Competition and Collusion

3.6.1. Single-stage oligopoly models

Since the crucial element of oligopoly is, by definition, the interdependence of the fortunes of the firms combined with their awareness of this interdependence,
oligopolistic decision making appears as a natural domain of game-theoretical analysis. From a game-theoretical point of view, oligopoly models can be categorized according to the solution concepts applied, the types of actions at the disposal of the players, and the role of time. In this section we review the development of noncooperative oligopoly models, i.e., models where the sellers are also assumed to be devoid of any commitment power, in which each oligopolist can make essentially only one choice during the course of the whole game. Here, however, game theory is not doing much more than generalizing the results obtained already by A. A. Cournot (1838), Joseph Bertrand (1883), Wilhelm Lauhnhardt (1885), Edgeworth (1897), Harold Hotelling (1929), Edward Chamberlin (1933), and Heinrich von Stackelberg (1934). For instance, the solution proposed by Lauhnhardt, Hotelling, and Chamberlin for the price setting problem in differentiated oligopolies is equivalent to a Nash noncooperative equilibrium in pure strategies of a game given in normal form (Shubik, 1959a; Wilhelm Krelle, 1961; 1976; Telser, 1972; Friedman, 1977). The theory of oligopolies with price-making multi-product firms was advanced along these lines mainly by Reinhard Selten (1970). Stackelberg’s so-called leadership solution is equivalent to the pure-strategy noncooperative equilibrium of an oligopoly game in extensive form in which the followers have perfect information about the leaders’ choices and do not have commitment power.

Edgeworth (1897) had already discovered that the introduction of capacity constraints in a price duopoly with homogeneous products may lead to the non-existence of a noncooperative equilibrium point in pure strategies à la Bertrand. In this case, however, game theory has been able to make a novel contribution by providing the notion of mixed strategies. Shubik (1959a) analyzed mixed-strategy equilibria for certain types of Bertrand–Edgeworth price duopoly games. The proof that there exists a unique equilibrium point in mixed strategies for a class of such games has been given by Shapley (1957). Martin Beckmann (1965) gave a mixed-strategy equilibrium as a solution to an integral equation for a model similar to that suggested by Shubik (1959a). These problems were further pursued mainly by Richard Levitan and Shubik (1972; 1978). They studied a simple price duopoly model with a special type of linear contingent demand function, which has the attractive property that, as one lets the capacity constraints vary, one gets the Cournot pure quantity-strategy equilibrium at one end of the interval and, at the other end, the Bertrand pure price-strategy equilibrium. Between the two extremes, an equilibrium in mixed strategies is shown to exist (Levitan and Shubik, 1972). In a later paper it was demonstrated that the introduction of inventory-holding costs in a Bertrand price duopoly has effects similar to those of capacity constraints (Levitan and Shubik, 1978). Shapley and Shubik examined the relationship between a Chamberlinian price strategy oligopoly with differentiated products and an Edgeworth-type oligopoly, demonstrating that product differentiation is not sufficient to guarantee the existence of a pure-strategy equilibrium in terms of price if capacity constraints are present (1969c). By studying the behavior of the model as the degree of product differentiation approaches zero and the number of competitors increases, the analysis of oligopoly is linked with that of pure competition. Yuval Shilony has also offered a game-theoretical hypothesis in terms of mixed-strategy equilibria to account for the phenomena of price dispersion in some markets (1977). It should also be interesting to note that under appropriate conditions the adding of a random component to an oligopoly’s demand

schedule may have a stabilizing effect in the sense of restoring the existence of a pure price-strategy equilibrium where, without the stochastic disturbance term, no such equilibrium existed (Levitan and Shubik, 1971). In addition to prices and output levels, other instruments of oligopolistic competition have been studied in game-theoretical models, such as advertising, product innovation (R. Reichardt, 1967; Ambar Rao and Melvin Shakun, 1972; Shubik, 1959a), and organizational structure (Shakun, 1968).

3.6.2. *Multistage oligopoly models*

The behavior of oligopolies during sequences of repeated plays of one and the same oligopoly game was studied for purposes of demonstrating the dynamic stability of the Cournot solution (Cournot, 1838; R. D. Theorcharis, 1960; Franklin Fisher, 1961; Maurice McManus and Richard E. Quandt, 1961; Hahn, 1962; Robert L. Bishop, 1962; Charles Frank and Quandt, 1963; K. Okuguchi, 1964). In these models, however, the players do not adopt overall supergame strategies, but are assumed to follow the extremely myopic policy of maximizing their profits on a period-by-period basis on the assumption that in each stage their rivals will stick to the production levels chosen in the previous period. For linear demand and cost functions, the Cournot duopoly solution is always dynamically stable; while for three sellers, oscillations about the equilibrium occur and for larger oligopolies instability always results. A host of generalizations, concerning adjustment speeds, demand and cost functions, and incomplete information about demand (Yasu Hosomatsu, 1969; Kirman, 1975; D. Gates, J. Rickard, and D. Wilson, 1977; 1978) have been achieved. Richard M. Cyert and M. H. DeGroot have given a dynamic model with alternating decisions of the duopolists (1970). For general payoff functions it was shown that even in the duopoly case the Cournot dynamics generated trajectories, which not only do not converge to an equilibrium but are so chaotic that any outside observer would be forced to choose a statistical approach to the description of its behavior (David Rand, 1978). In any case, Cournot’s dynamical hypothesis does not seem satisfactory from a rational-behavior point of view. This observation has led Friedman (1968; 1973; 1976; 1977) to study multistage oligopoly games in which the players maximize the sum of their discounted future profits and, instead of expecting their rivals in each period to leave their policies unchanged, assume that their competitors employed more general reaction functions in the following sense: In the strategy equilibrium reached via the dynamic process generated by the usage of the equilibrium reaction functions as predictors of the behavior of each player’s rivals, the reaction functions are confirmed by the oligopolists’ actual behavior. Outside of equilibrium, however, even the equilibrium reaction functions do not prove good predictors. Thus, the fundamental problem of rational versus nonrational, adaptive expectations (see also Telser, 1972) in dynamic oligopoly models has not been solved by Friedman’s concept of reaction function (also, Krelle, 1961; 1976) but has only been shifted to another level.

An alternative to analyzing an oligopoly supergame in terms of step-by-step decision making (which would be adequate if for some reason the oligopolists were not able to commit themselves to long-run plans) is to look for noncooperative equilibrium points involving supergame strategies. Supergames may be stationary (consisting of identical constituent games) or non-stationary. In the latter case they may be time-dependent in a systematic way (if the parameters of a constituent game depend systematically on past actions), and they may be stochastic (in case the
transition from one constituent game to another is governed by a random mechanism. Selten (1965) has studied perfect equilibrium points of a time-dependent price oligopoly supergame, in which the time-dependence is due to the inert response of demand to the competitors' pricing policies and the oligopolists are maximizing the sum of their discounted profits over time. An early model of a stochastic oligopoly supergame, termed "game of economic survival," was formulated by Shubik and Gerald Thompson (1959); it is a two-person zero-sum game involving production and financial operations of the duopolists. Shubik has outlined the structure of an n-person game of economic survival (1959a). More recent developments in the field of stochastic supergames are reviewed in Friedman (1977). (See also S. Deshmukh and W. Winston, 1978, for the model of stochastic duopoly games.) Telser studies equilibria of time-dependent oligopoly supergames whose dynamic demand structure is generated via price expectations depending on past prices and stocks of durable goods (1972). (For the time-dependence implied by investment in fixed capital, see Friedman, 1977.) A zero-sum differential game model of duopoly in which the duopolists aim at outselling each other over a given period, given some constraints on profits, is investigated by S. Clemhout, G. Leitman, and H. Y. Wan, Jr. (1971). Friedman examines noncooperative equilibria of infinite supergames and proves the existence of a particular noncooperative equilibrium for stationary supergames, which he calls "balanced temptation equilibrium" (1971; 1972; 1974; 1977). His investigation is very much in the spirit of Aumann's concept of strong equilibrium points of infinite supergames (1959), and his definition of a supergame differs from Aumann's only as far as a player's payoff function is concerned. While Aumann's supergame payoffs are obtained as suitably defined averages of the payoffs in the constituent games, Friedman takes as supergame payoffs the discounted sums of the constituent game payoffs. Friedman shows that the noncooperative balanced temptation equilibrium points yield Pareto-optimal supergame payoff vectors in stationary supergames. This is a remarkable result, since an essentially cooperative outcome emerges without any cooperative assumptions about trust, commitment power, etc.; the only thing needed to realize such a self-policing Pareto efficient outcome would be some form of pre-play communication. A noncooperative balanced temptation equilibrium may be regarded as a kind of tacit collusion (William Fellner, 1949). The enforcement mechanism, which does not require any commitment power, is provided simply by the circumstances that the original game is repeated an arbitrary number of times. A balanced temptation equilibrium strategy for a player prescribes a course of action corresponding to a Pareto-efficient payoff vector as long as the other players do the same, but as soon as some of the players fail to behave "cooperatively" in some time period, then, from the next period onward, all the other players choose actions corresponding to best replies in the original (constituent) game. Thus, a balanced temptation equilibrium strategy is in fact a contingent plan providing for retaliation in the case of the defection of some other player and has an effect tantamount to a threat. Friedman proved also the existence of stationary balanced temptation equilibria for time-dependent supergames, which possess, however, only the property of "local" Pareto efficiency, in the sense that Pareto optimality holds within each single time period only (1974; 1977). Morton Kamien and Nancy Schwartz analyzed an oligopoly with entry into the market under the assumption that a firm views its existing rivals in accordance with the classical
Cournot hypothesis and regards the timing of rival entry as represented by a random variable whose distribution is dependent on current industry prices (1975). Friedman (1979a, 1979b) used a supergame framework in which "firms-in-being" (Shubik, 1959a) are modelled as players that choose to be inactive and accept a payoff of zero during certain stages of the game. He proved the existence of "entry-exit" equilibria for a broad class of nonstationary, time-dependent oligopoly supergames which do not, however, display all Nash properties (a firm's "equilibrium" participation probability is a best reply only given its "equilibrium" constituent-game course of action, and vice versa, but both together are not necessarily best replies to the other oligopolists' equilibrium strategies).

Thomas Marschak and Selten (1974; 1978), in a most original attempt to cope with the analytical difficulties provided by oligopolistic interdependence both in a partial and a general equilibrium framework, have introduced an interesting new concept of reaction function ("convolutions"), which is to a great extent freed from the ad hoc flavor so typical of the other notions of reaction function (Friedman, 1968; Krelle, 1961; 1976) and, moreover, possesses an instructive interpretation in terms of supergame equilibria. In a situation described by a noncooperative Nash equilibrium, an oligopolist does not have any incentive to change his behavior if he assumes that in case he deviates the other competitors will stick to their strategies. But what if his appraisal of his competitors as rational players leads to the reasonable expectation that his own deviation would trigger a chain reaction of deviations on the part of the other agents? Taking this into account, it may well be that a deviation from a noncooperative equilibrium point is advantageous, whereas on the other hand, a defection from some other strategy combination not possessing the Nash equilibrium property, but preferred by all players to a noncooperative equilibrium outcome, might not pay for any oligopolist. Of course, considerations like this have abounded in the traditional oligopoly literature dealing with the question of how the oligopolists would go about stabilizing a joint-profit maximizing group behavior, which to many oligopoly theorists (e.g., Chamberlin, 1933; Fellner, 1949) seemed to be the only "rational" solution to "competition among the few." Also, the kinked demand curve literature (Paul M. Sweezy, 1939) trying to explain downward price rigidities relied on that kind of reasoning. But prior to Marschak and Selten (1974), none had given a precise treatment of the problem consistent with the accepted ideas about rational behavior. The basic assumption is that what matters to an oligopolist considering a change of his behavior is the situation that comes up after all of his competitors' responses have been completed, and any transitory profits accruing in the meantime are neglected. This assumption is in keeping with most of the traditional ideas about "conjectural variation" and the like, but distinguishes this approach from Friedman's usage of the reaction function concept.

3.6.3. Oligopolistic collusion

Most analyses of collusive oligopoly presume that the oligopoly market is played cooperatively and proceed to apply various cooperative solution concepts to quantity-variation and price-variation duopolies and general oligopolies. J. P. Mayberry, Nash, and Shubik (1953) examined the cooperative quantity duopoly in the light of the variable-threat cooperative Nash solution (also Shubik, 1959a); Schwödiauer applied the concept of bargaining set to side-payment characteristic functions derived from differentiated price-variation oligopolies (1970); and Telser analyzed cooperative oligopoly games.
in terms of the core (1972). Mamoru Kaneko studies a price oligopoly in which the firms have the same linear cost functions and sell the same homogeneous good (1978). The demand side is not given just by a demand schedule but is represented by utility-maximizing buyers assumed, however, to act as price takers and to buy from oligopolists setting the lowest price. Using these assumptions, a characteristic function is defined on the set of all agents (oligopolists and buyers), which has a non-empty core only for the case of monopoly, but a non-empty bargaining set also for the case of oligopoly. The bargaining set yields a price corresponding to joint profit maximization only in the case of duopoly. The lowest price given by the bargaining set approaches unit cost (i.e., the competitive solution) with an increasing number of sellers.

In contrast to the aforementioned contributions, Selten offers a model of cartel formation in oligopolies in which the proposition that few suppliers tend to maximize their joint profits whereas many suppliers are likely to behave non-cooperatively does not appear as an assumption but as a conclusion of the theory (1973). Working with a Cournot model with linear cost and demand functions, Selten models cooperative forms of behavior as moves in a noncooperative game. It is assumed that firms are free to form enforceable quota cartels, but each firm must decide without being informed about the corresponding decisions of the other oligopolists whether it wants to participate in cartel bargaining or not. Each of the firms which have chosen to participate must propose a cartel agreement in the form of a quota system that becomes binding if all members of the group have made the same proposal. Before the supply decision is taken, the outcome of the bargaining is made known to all firms in the market. Cartels may or may not include all oligopolists. Within the institutional framework described, it is advantageous to form a cartel; but if the number of competitors is sufficiently large, it may even be more advantageous to stay out of a cartel formed by others. The unique, in the technical sense noncooperative, equilibrium solution to the game exhibits a remarkable jump in the firms’ propensity to cartelize at a number of five oligopolists, a result which, of course, depends on the simple structure of the model. For \( n = 2, 3, 4 \) oligopolists, the probability that a cartel agreement is reached, if an equilibrium point is played, equals 1 and the outcome of the cartel bargaining is joint profit maximization. For \( n = 5 \) this cartel probability is approximately 1 percent or smaller; for \( n > 5 \) it is always smaller than 0.0001, making the solution virtually identical to the Cournot equilibrium. Thus, for the simple cost and demand structure underlying the model, four competitors are still “few,” but six are already “many.”

3.7. General Equilibrium with Strategically Active Agents

General equilibrium models with price-making firms at first circumvented the problem of mutual recognition of power among firms by simply assuming away all conceivable interactions via demand or supply relationships among monopolistic firms (Takashi Negishi, 1961; Arrow and Hahn, 1971). The dropping of this restrictive assumption suggests the application of the concept of noncooperative Nash equilibrium in a general equilibrium context. One approach, which is an immediate generalization of monopolistic general equilibrium models, is to combine Cournot and Walras equilibria by splitting the economy into an oligopolistic sector and a “competitive” one (with strategically passive, price-taking agents). Jaskold-Gabszewicz and Jean-Philippe Vial (1972), considered an economy where on the exchange side the agents are price-takers, while on the production side the firms be-
have like Cournot oligopolists. If the number of firms is increased by a replication device, prices prevailing on the exchange markets become less and less sensitive to the supply of any one of the firms, and each firm tends to choose its equilibrium strategy closer and closer to the competitive supply. Marschak and Selten have shown how general equilibria of the Cournot-Nash type may be defined in a natural way even when full allowance is made for the selling of commodities within the oligopolistic sector and the distribution of profits to share-holding consumers (1974). Proofs of existence of such pure-strategy Nash equilibria require, however, rather strong conditions, which do not easily follow from the standard assumptions about preferences, endowments, and technologies. In particular, in order to permit the application of the Kakutani fixed-point theorem, the set of optimal choices of each oligopolistic firm given the behavior of the other agents must be assumed to be convex. As the examples given by Roberts and Sonnenschein demonstrate (1977), we cannot expect to obtain upper-hemicontinuous, convex-valued reaction correspondence—the relevant generalization of the continuous reaction functions known from partial equilibrium oligopoly theory—even for highly simplified and idealized economies. On neither of the examples they produce do the reaction curves display the required convex-valuedness and, in both cases, no equilibrium exists, although the characteristics of firms and consumers fulfill all the standard assumptions. Moreover, the non-existence of equilibrium is robust in the sense that small perturbations in the data of the economies cannot restore equilibrium. The difficulties pointed out by Roberts and Sonnenschein affect the existence of noncooperative equilibria in pure strategies (1977). From a game-theoretical standpoint, it would therefore be natural to look for equilibria in mixed strategies implying expected-profit-maximizing probability distributions over output levels or prices. Novshek and Sonnenschein (1978) study noncooperative equilibria of private ownership economies satisfying all the standard Debreu-type conditions except for two: the firms' technology sets are non-convex so that the efficient scale of firms' productions is bounded away from zero, and the number of firms is not given a priori but determined endogenously. Firms choose production levels given the actions of other firms, where payoffs are defined relative to the demand function of the price-taking consumer sector (a selection from the appropriate Walrasian demand correspondence) and mixed strategies are permitted. A Cournot-Nash equilibrium is defined by the conditions that no firm in the market can increase expected profit by changing its quantity strategy, and no firm absent from the market can enter and achieve a positive expected profit. The existence of Cournot-Nash equilibria is shown for sufficiently small optimum scales of firms; in these equilibria, the equilibrium strategies of firms with positive expected profits are in fact pure quantity strategies, while for so-called marginal firms (with zero expected profits), they are mixed strategies of the form "produce a certain quantity with some probability q and stay out of the market with probability 1 - q." These Cournot-Nash equilibria are shown to converge to the traditional Walrasian equilibria if the efficient scale of production in each industry approaches zero. The Novshek-Sonnenschein model is an ingenious extension and sharpening of the Chamberlinian notion of large-group equilibrium to a general equilibrium setting. Another way to escape the difficulties posed by the likely non-convex-valuedness of the reaction correspondences in oligopolistic general equilibrium models is offered by the notion of convolution and consists in abandoning the concept of Nash equilibrium.
for its behavioral "unreasonableness" in favor of the strategy combinations at which all the agents are stable with respect to some suitably constructed rationality-preserving response function. Marschak and Selten study several variants of oligopolistic economies as games of limited information employing what they call a "convolution" concept (1974).

Game-theoretical equilibrium concepts suggest themselves not only for the analysis of economies that are "intrinsically" monopolistic or oligopolistic, but also for so-called fix-price economies in which the compatibility of individual plans is brought about via quantity rationing mechanisms leading to situations, which from a Walrasian point of view, might be considered persistent disequilibria (see, e.g., Drèze, 1975; Jean-Pascal Benassy, 1977; Jean Michel Grandmont, 1977; 1978; Grandmont, Laroque, and Yves Younes, 1978). If the rationing schemes employed are "manipulable" in the sense that the transactions assigned to the agents on the long-side of a market are sensitive to their offered volumes of trade (as is the case, e.g., for proportional rationing schemes), the hypothesis that the traders formulate their trade offers ("effective demands") in the terminology introduced by Robert Clower (1965) and Benassy (1977) for each market separately and without taking into account their final consequences appears rather implausible. More attractive seems the modelling of the problem as a noncooperative game where each agent would know the rationing scheme and would send to the market trade offers which, given the other agents' trade offers, would result in a vector of final transactions most preferred by him, i.e., a noncooperative Nash equilibrium in the space of trade offers (Böhm and P. Levine, 1976; Walter Heller and Ross Starr, 1976), which might also be termed a "rational-expectations" or "self-fulfilling expectations" equilibrium. However, in order to ensure the existence of non-trivial Nash equilibria (in pure strategies), i.e., Nash equilibria where trade actually takes place, additional assumptions are required, in particular, on the properties of rationing schemes (the strictly proportional rationing schemes, e.g., must be ruled out).

Martin Shubik, in cooperation mainly with Pradeep Dubey and Lloyd Shapley, has presented a series of models describing economies with trade, with and without production, and with several types of money (commodity money, fiat money, bank money, accounting money) and other financial institutions (shares, bankruptcy rules, etc.) as games in extensive form (Shubik, 1971/72; 1973; Shubik and W. Whitt, 1973; Shapley, 1976; Shapley and Shubik, 1977; Shubik and Wilson, 1977; Dubey and Shubik, 1977a; 1977b; 1978; 1979; J. Evers and Shubik, 1976). In these models, commodities are typically exchanged in markets where buyers bid money and sellers offer quantities of the respective commodities; each buyer then receives the proportion of the aggregate amount of the commodity equal to the proportion of his bid to the aggregate bids, and vice versa for sellers. The extensive-game formulation of the model ensures a full description of the temporal and informational structure implied by the assumed production and exchange moves and the various institutional assumptions, and renders the behavior of all agents well defined also in disequilibrium states of the economy. The equilibrium concept employed is that of a noncooperative Nash equilibrium in pure strategies. Under a variety of institutional assumptions, the existence of non-trivial noncooperative equilibria is proven. The conditions under which such noncooperative equilibria converge to Walrasian competitive equilibria for replicated economies (or, coincide with them in case of non-atomic measure spaces of agents) are
investigated (see also G. Jaynes, M. Okuno, and David Schmeidler, 1978; Postlewaite and Schmeidler, 1978). If, e.g., borrowing from an outside central bank is permitted, the bankruptcy rules have to provide for sufficiently high penalties against those who fail to repay their loans, in order to ensure that the distribution of goods and prices associated with any competitive equilibrium of a large economy corresponds to some of its noncooperative equilibria (Dubey and Shubik, 1979). The attraction of these fully game-theoretical models of general economic equilibrium lies in three things. First, they are capable of distinguishing between feasible and equilibrium actions of economic agents (whereas in the Walrasian analysis only equilibrium outcomes are feasible). Second, they can determine in a completely endogenous way the oligopolistic character or competitiveness of the agents' behavior (without splitting the economy in an ad hoc manner in a competitive and an oligopolistic sector). Finally, and most importantly, they are open to the introduction of new institutional arrangements, which allow the comparative evaluation of the role played, in particular, by various conceivable monetary and financial institutions.

Conclusions

Space does not permit a long set of conclusions or speculations about the future of game theory in economics. One thing that is certain, however, is that game theory has a definite role to play in the analysis of economic and social institutions. This development, we feel, holds great promise for economics, since it may one day allow a more institutionally flexible analysis of economic problems than is currently found in neoclassical analysis, which is tied to one and only one institutional framework—competitive markets. By allowing the analyst the ability to study both the endogenous creation of social institutions (see Schotter, 1980), and their positive (Shubik, 1971/72) and comparative properties (see Hurwicz, 1973), game theory has carved out a natural place for itself within economic theory, a place that is extremely consistent with the intention of its originators 36 years ago.

References


BAUMOL, WILLIAM J.; BAILEY, ELIZABETH E. AND WILLIG, ROBERT D. "Weak Invisible Hand Theo-


COUROT, ANTOINE A. *Recherches sur les principes mathematiques de la theorie des richesses.* Paris:
Librairie de Sciences Politiques et Sociale, M. Riviere & cie, 1838.


——— "Non-Cooperative Equilibria in Time-De-
Journal of Economic Literature, Vol. XVIII (June 1980)

...Supergames,” Econometria, March 1974, 42(2), pp. 251-37.


“Non-Cooperative Equilibria for Exit Super-

“On Entry Preventing Behavior and Limit
Price Models of Entry,” in Applied game theory. Edited by STEVEN BRAMS, ANDREW SCHOTTER,
and GERHARD SCHWÖDLAUER. Wien-Würzburg:
Physica-Verlag, 1979b, pp. 28-63.

Gale, Douglas. “The Core of a Monetary Economy
without Trust,” J. Econ. Theory, Dec. 1978, 19(2),
p. 458-91.

Gardner, Roy. “Shapley Value and Disadvanta-
geous Monopolies,” J. Econ. Theory, Dec. 1977,
16(2), pp. 513-17.

Gately, Dermot. “Sharing the Gains from Re-
gional Cooperation: A Game Theoretical Applica-
tion to Planning Investment in Electric Power,”

Gates, D. J.; Rickard, J. A. and Wilson, D. J. “A
Convergent Adjustment Process for Firms in Com-
1349-63.

Rickard, J. A. and Wilson, D. J. “Conver-
gence of a Market Related Game Strategy,” J.

Gibbard, Allan. “Manipulation of Voting Schemes:
A General Result,” Econometria, July 1973, 41(4),
p. 581-601.

Gillies, Donald B. “Solutions to General Non-
Zero-Sum Games,” in Vol. IV. Annals of mathemat-
ics studies. No. 40. Contributions to the theory
of games. Edited by ALBERT W. TUCKER and R.
DUNCAN LUCE. Princeton, N.J.: Princeton Uni-

Grandmont, Jean-Michel. “The Logic of the Fix
169-86.

Laroque, Guy and Younès, Yves. “Equi-
librium with Quantity Rationing and Recontract-
ing,” J. Econ. Theory, Oct. 1978, 19(1), pp. 84-
102.

Green, Jerry R. “The Stability of Edgeworth's Re-
contracting Process,” Econometria, Jan. 1974,
42(1), pp. 21-34.

“Satisfactory Mechanisms for Environments
with Consumption Lower Bounds.” Unpublished
manuscript, 1976.

Laffont, Jean Jacques. “Revelation des Prefer-
1976.

AND LAFFONT, JEAN JACQUES. “Characteriza-
tion of Satisfactory Mechanisms for the Revelation
of Preferences for Public Goods,” Econometria,

Groves, Theodore. The allocation of resources un-
der uncertainty: The informational incentive roles
of price and demand in a team. Ph.D. dissertation,
unpublished. University of California, Berkeley,
1970.

“Incentives in Teams,” Econometria, July

“Efficient Collective Choice when Compensa-
tion is Possible,” Rev. Econ. Stud., April 1979,
46(2), pp. 227-41.

AND LEdYARD, John O. “Some Limitations of
Demand Revealing Processes,” Public Choice,
Spring 1977, 29-2, Special Supplement, pp. 107-
24.

AND LOEB, MARTIN. “Incentives and Public
211-16.

Guesnerie, Roger. “Monopoly, Syndicate, and
Shapley Value: About Some Conjectures,” J. Econ.

AND ODDOU, Claude. “A Game Theoretical
Approach to a Second Best Problem with an Ap-
plication to the Theory of Local Public Goods.” M.

Hahn, Frank H. “The Stability of the Cournot Oli-
gopoly Solution,” Rev. Econ. Stud., Oct. 1962,

Hammond, Peter J. “Straightforward Individual In-
centive Compatibility in Large Economies,” Rev.

Hansen, Terje and Jaskold-Gabszewicz, Jean.
“Collusion of Factor Owners and Distribution of
pp. 1-18.

Harsanyi, John C. “Games with Incomplete In-
formation Played by 'Bayesian' Players,” 3 parts,
456-502.

AND SELTEN, Reinhard. “A Generalized
Nash Solution for Two-Person Bargaining Games
1972, 18(5), Part 2, pp. 80-106.

Hart, Sergiu. “Formation of Cartels in Large
Markets,” J. Econ. Theory, April 1974, 7(4), pp. 455-
66.

“Asymptotic Value of Games with a Contin-
pp. 57-80.

“Values of Large Market Games,” in Applied
Game theory. Edited by STEVEN BRAMS, ANDREW
Schotter, and GERHARD SCHWÖDLAUER. Wien


LAUNHARDT, WILHELM. Mathematische begründung der Volkswirtschaftslehre. Leipzig: Englemann, 1885.


AND SHUBIK, MARTIN. "Price Duopoly and..."


“Implementations which are Immune to Collusion.” Unpublished manuscript, 1978b.


NASH, JOHN F. “Equilibrium Points in n-Person


SCHMEIDLER, DAVID. “The Nucleolus of a Charac-


SHITOVITZ, BENJAMIN. "Oligopoly in Markets with


