"The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-market Allocation"

Kenneth J. Arrow

Kenneth Arrow, a recent Nobel Laureate in Economics, is Professor of Economics at Stanford University. This paper was published by the Joint Economic Committee of Congress in 1969.

Introduction

The concept of public goods, has been developed through a process of successive refinement over a long period of tire. Yet surprisingly enough there does not seem to exist anywhere in the literature a clear general definition of this concept or the more general one of "externality." The accounts given are usually either very general and discursive, difficult of interpretation in specific contexts, or else they are rigorous accounts of very special situations. What exactly is the relation between externalities and such concepts as "appropriability" or "exclusion"?

Also, there is considerable ambiguity in the purpose of the analysis of externalities. The best developed part of the theory relates to only a single, question: the statement of a set of conditions, as weak as possible, which insure that a competitive equilibrium exists and is Pareto efficient. Then the denial of any of these hypotheses is presumably a sufficient condition for considering resort to non-market channels of resource allocation-usually thought of as Government expenditures, taxes, and subsidies.

At a second level the analysis of externalities should lead to criteria for non-market allocation. We are tempted to set forth these criteria in terms analogous to the profit-and-loss statements of private business; in this form, we are led to benefit-cost analysis. There are, moreover, two possible aims for benefit-cost analysis; one, more ambitious but theoretically simpler, is specification of the non-market actions which will restore Pareto efficiency; the second involves the recognition that the instruments available to the Government or other non-market forces are scarce resources for one reason or another, so that all that can b achieved is a "second-best."

Other concepts at seem to cluster closely to the-concept of public goods are those of "increasing returns" -and-" market failure." These are related to Pareto inefficiency on the one hand and to the existence and optimality of competitive equilibrium on the other; sometimes the, discussions in the literature do not adequately distinguish these two aspects. I contend that market failure is a more general category than externality; and both differ from increasing returns in a basic sense, since market failures in general and externalities in particular are relative to the mode of economic organization, while increasing returns are essentially a technological phenomenon.

Current writing has helped bring out the point that market failure is not absolute; it is better to consider a broader category that of transaction costs, which in general impede and in particular cases completely block the formation of markets. It is usually though not always emphasized that transaction costs are costs of running the economic system. An incentive for vertical integration is replacement of the costs of buying and selling on the market by the costs of intra-firm transfers; the existence of vertical integration may suggest that the costs of operating competitive markets are not zero, as is usually assumed in our theoretical analysis. Monetary theory, unlike value theory, is heavily dependent on the assumption of positive transaction costs; the recurrent complaint about the difficulty of integrating these two branches of theory is

certainly governed by the contradictory assumptions made about transaction costs. The creation of money is in many respects an example of a public good. ,

The identification of transaction costs in different contexts and under different systems of resource allocation should be a major item on the research agenda of the theory of public goods and indeed of the theory of resource allocation in general. Only the most rudimentary suggestions are made here. The "exclusion principle" is a limiting case of one kind of transaction cost, but another type, the costliness of the information needed to enter and participate in any market, has been little remarked. Information is closely related on the one hand to communication and on the other to uncertainty. '

Given the existence of Pareto inefficiency in a free market equilibrium, there is a pressure in the market to overcome it by some sort of departure from the free market; i.e., some form of collective action. This need not be undertaken by the Government. I suggest that in fact there is a wide variety of social institutions, in particular generally accepted social norms of behavior, which serve in some means as compensation for failure or limitation of the market, though each in turn involves transaction costs of its own. The question also arises how the behavior of individual economic agents in a social institution (especially in voting) is related to their behavior on the market. A good deal of theoretical literature has arisen in recent years which seeks to describe political behavior as analogous to economic, and we may hope for a general theory of socioeconomic equilibrium. But it must always be kept in mind that the contexts of choice are radically different, particularly when the hypotheses of perfectly costless action and information are relaxed. It is not accidental that economic analysis has been successful only in certain limited areas. Competitive Equilibrium and Pareto Efficiency

A quick review of the familiar theorems on the role of perfectly competitive equilibrium in the efficient allocation of resources will be useful. Perfectly competitive equilibrium has its usual meaning: households, possessed of initial resources, including possibly claims to the profits of firms, choose consumption bundles to maximize utility at a given set of prices; firms choose production bundles so as to maximize profits 'at the same set of prices; the chosen production and consumption bundles must be consistent with each other in the sense that aggregate production plus initial resources must equal aggregate consumption. The key points in the definition are the parametric role of the prices for each individual and the identity of prices for all individuals. Implicit are the assumptions that all prices can be known by all individuals and that the act of charging prices is not itself a consumer of resources.

A number of additional assumptions are made at different points in the theory of equilibrium, but most are clearly factually valid in the usual contexts and need not be mentioned. The two hypotheses frequently not valid are (C), the convexity of household indifference maps and firm production possibility sets, and (M), the universality of markets. While the exact meaning of the last assumption will be explored later at some length, for the present purposes we mean that the consumption bundle which determines the utility of an individual is the same as that which he purchases at given prices subject to his budget constraint, and that the set of production bundles among which a firm chooses is a given range independent of decisions made by other agents in the economy.

The relations between Pareto efficiency and competitive equilibrium are set forth in the following two theorems:

1. If (M) holds, a competitive equilibrium is Pareto-efficient. This theorem is true even if (C) does not hold.
2. If (C) and (M) hold, then any Pareto-efficient allocation can be achieved as a competitive equilibrium by a suitable reallocation of initial resources.

When the assumptions of proposition 2 are valid, then the case for the competitive price system is strongest. Any complaints about its operation can be reduced to complaints about the distribution of income, which should then be rectified by lump-sum transfers. Of course, as Pareto already emphasized, the proposition provides no basis for accepting the results of the market in the absence of accepted levels of income equality.

The central role of competitive equilibrium both as a normative guide and as at least partially descriptive of the real world raises an analytically difficult question: does a competitive equilibrium necessarily exist?

3. If (C) holds, then there exists a competitive equilibrium. This theorem is true even if (M) does not hold.

If both (C) and (M) hold, we have a fairly complete and simple picture of the achievement of desirable goals, subject always to the major qualification of the achievement of a desirable income distribution. The price system itself determines the income distribution only in the sense of preserving the status quo. Even if costless lump-sum transfers are possible, there is needed a collective mechanism reallocating income if the status quo is not regarded as satisfactory.

Of course (C,) is not a necessary condition for the existence of a competitive equilibrium, only a sufficient one. From proposition 1, it is possible to have an equilibrium and therefore efficient allocation without convexity (when (M) holds). However, in view of the central role of (C) in these theorems, the implications of relaxing this hypothesis have been examined intensively in recent years by Farrell (19$9), Rothenberg (1960), Aumann (1966), and Starr (1969). Their conclusions may be summarized as follows: Let (C') be the weakened convexity assumption that there are no indivisibilities large relative to the economy.

4. Propositions 2 and 3 remain approximately true if (C.) is replaced by (C.).

Thus, the only non-convexities that are important for the present purposes are increasing returns over a range large relative to the economy. In those circumstances, a competitive equilibrium cannot exist.

The price system, for all its virtues, is only one conceivable form of arranging trade, even in a system of private property. Bargaining can assume extremely general forms. Under the assumptions (C') and (M), we are assured that not everyone can be made better off by a bargain not derived from the price system: but the question arises whether some members of the economy will not find it in their interest and within their power to depart from the perfectly competitive price system. For example, both Knight (1921, pp. 190-194) and Samuelson (1967, p. 120) have noted that it would pay all the firms in a given industry to form a monopoly. But in fact it can be argued that unrestricted bargaining can only settle down to a resource allocation which could also be achieved as a nerfectlv competi tive equilibrium, at least if the bargaining itself is costless and each agent is small compared to the entire economy. This line of

argument originated with Edgeworth (1881., pp. 20-43) and has been developed recently by Shubik (1959), Debreu and Scarf (1963), and Aumann (1964) .

More precisely, it is easy to show:

5. If (M) holds and a competitive equilibrium prevails, then no set of economic agents will find any resource allocation which they can accomplish by themselves (without trade with the other agents) which they will all prefer to that prevailing under the equilibrium. Proposition 5 holds for any number of agents. A deeper proposition is the following converse:
6. If (C) and (M) hold, and if the resources of any economic agent are small compared with the total o f the economy, then, given any allocation not approximately achievable as a competitive equilibrium, there will be some set of agents and some resource allocation they can achieve without any trade with others which each one will prefer to the given allocation.

These two propositions, taken together, strongly suggest that when all the relevant hypotheses hold, (a) a competitive equilibrium, if achieved, will not be upset by bargaining even if permitted, and (b) for any bargain not achievable by a competitive equilibrium there is a set of agents who would benefit by change to another bargain which they have the full power to enforce.

The argument that a set of firms can form a monopoly overlooks the possibility that the consumers can also form a coalition, threaten not to buy, and seek mutually advantageous deals with a subset of the firms; such C deals are possible since the monopoly allocation violates some marginal equivalences.

In real life, monopolizing cartels are possible for a reason not so far introduced into the analysis: bargaining costs between producers and consumers are high, those among producers low-a point made most emphatically by Adam Smith (1937, p. 128); "People of the same trade seldom meet together, even for merriment or diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices." It is not the presence o f bargaining costs her se but their bias that is relevant. If all bargaining costs are high, but competitive pricing and the markets are cheap, then we expect the perfectly competitive equilibrium to obtain, yielding an allocation identical with that under costless bargaining. But if bargaining costs are biased, then some bargains other than the competitive equilibrium can be arrived at which will not be upset by still other bargains if the latter but not the former are costly.

Finally, in this review of the elements of competitive equilibrium theory, let me repeat the obvious and well-known fact that in a world where time is relevant, the commodities which enter into the equilibrium system include those with future dates. In fact, the bulk of meaningful future transactions cannot be carried out on any existing present market, so that assumptiom (M), the universality of markets, is not valid.

Imperfectly Competitive Equilibrium

There is no accepted and well-worked out theory corresponding to the title of this section. From the previous section it is clear that such a theory is forcibly needed in the presence of increasing returns on a scale large relative to the economy (hereafter,

the phrase "increasing returns" will always be understood to include the prepositional phrase just employed), and is superfluous in its absence.

There are two approaches to a theory of general equilibrium in an imperfectly competitive environment; most writers who touch on public policy questions implicitly accept one or the other of these proto theories without always recognizing that they have made a choice. One assumes all transactions are made according to the price system, i.e., the same price is charged for all units of the same commodity; this is the monopolistic competition approach. The alternative approach assumes unrestricted bargaining; this is the game theory approach. The first might be deemed appropriate if the costs of bargaining are high relative to the costs of ordinary pricing, while the second assumes costless bargaining.

It cannot be too strongly emphasized that neither approach is, at the present stage, a fully developed theory, and it is misleading to state any implications about the working of these systems. Chamberlin's (1933), purpose was certainly the incorporation of monopoly into a general equilibrium system, together with a view that the commodity space should be viewed as infinite-dimensional, with the possibility of arbitrarily close substitutes in consumption; Triffin (1941) emphasized this aspect, but the only completely worked-out model of general monopolistic equilibrium is that of Negishi, (1960-61), and he made the problem manageable by regarding the demand functions facing the monopolists as those preceived by them, with only loose relations to reality. Such a theory would have little in the way of deducible implications (unless there were a supplementary psychological theory to explain the perceptions of demand functions) and certainly no clear welfare implications.

Of course, whatever a monopolistic competitive equilibrium means, it must imply inefficiency in the Pareto sense if there are substantial increasing returns. For a firm can always make zero profits by not existing; hence, if it operates, price must at least equal average cost which is reater than marginal cost. Kaldor (1935) and Demsetz (1964), however, have argued that in the "large numbers" case, the welfare loss may be supposed very small. I would conjecture that this conclusion is true, but it is not rigorously established, and indeed the model has never been formulated in adequate detail to discuss it properly.

With unrestricted bargaining it is usual to conclude that the equilibrium, whatever it may be, must be Pareto-efficient for, by definition, it is in the interest of all economic agents to switch from a Paret0-inefficient alocation to a suitably chosen Pareto-efficient one. This argument seems plausible, but is not easy to evaluate in the absence of a generally acceptedconcept of solution for game theory. Edgeworth (1881) held the outcome of bargaining to be indeterminate within limits, and von Neumann and Morgenstern (1944) have generalized this conclusion. But when there is indeterminacy, there is no natural or compelling point on the Pareto frontier at which to arrive. It is certainly a matter of common observation, perhaps most especially in the field of international relations, that mutually advantageous agreements are not arrived at because each party is seeking to engross as much as possible of the common gain for itself. In economic affairs a frequently cited illustration is the assembly of land parcels for large industrial or residential enterprises whose value (net of complemen tary costs) exceeds the total value of the land in its present uses. Then each owner of a small parcel whose acquisition is essential to the execution of the enterprise can demand the entire net benefit. An agreement may never be reached or may be long delayed; at positive discount rates

even the latter outcome is not Pareto-efficient. It is to avoid such losses that the coercive powers of the state are invoked by condemnation proceedings.

There is, however, another tradition within game theory which argues for the determinacy of the outcome of bargaining. Zeuthen (1930, ch. IV) had early propounded one such solution. After von Neumann and Morgenstem, Nash (1950, 1953) offered a solution, which Harsanyi (1956) later showed to be identical with that of Zeuthen. Nash's analysis of bar gaining has been extended by Harsanyi (1959, 1963, 1966) ; variant but related approaches have been studied by Shapley (1953) and Selten (1964) .The analysis has proceeded at a very general level, and its specific application to resource allocation has yet to be spelled out. In the simplest situation, bargaining between two individuals who can cooperate but cannot injure each other except by withholding cooperation and who can freely transfer benefits between them, the conclusion of the theories is the achievement of a joint optimum followed by equal splitting of the benefits of cooperation net of the amounts each bargainer could obtain without cooperation. Thus, in a land assembly, if the participation of all parcels is essential, each owner receives the value of his parcel in its present (or best alternative) use plus an equal share of the net benefits of the project. Without further analytic and empirical work it is not easy to judge the acceptability of this conclusion.

An elementary example may bring out the ambiguities of allocation with unrestricted bargaining. Since the perfectly competitive equilibrium theory is satisfactory (in the absence of marketing failures and costs) when increasing returns on a substantial scale are absent, the problem of imperfectly competitive equilibrium arises only when substantial increasing returns are present. In effect, then, there are small numbers of effective participants. Suppose there are only three agerits. Production is assumed to take place in coalitions; the output of each coalition depends only on the number of members in it. If the average output of the members of a coalition does not increase with the number of members, then the equilibrium outcome is the perfectly competitive one, where each agent produces by himself and consumes his own product. If the average output of a coalition increases with the number of members, then clearly production. will take place in the three-member coalition; but the allocation is not determined by the threats of individuals to leave the coalition and go on their own, nor by threats of pairs to form coalitions (for any one member can claim more than onethird of the total output and still leave the other two more than they could produce without him). But perhaps the most interesting case is that where the average output is higher for two individuals than for either one or three; i.e., increasing returns followed by diminishing returns. For definiteness, suppose that one agent can produce one unit, two agents can produce four units, and all three together can produce five units. Clearly, Pareto efficiency requires the joint productive activity of all three. Since each pair can receive four units by leaving the third agent out, it would appear that each pair must receive at least four units. But this implies that the total allocated to keep the three-man coalition together must be at least six, more than is available for distribution.

(Theories of the Nash-Harsanyi type arrive at solutions in cases like this by assuming that the economic agents foresee these possible instabilities and recognize that any attempt by any pair to break away from the total coalition can itself be overturned. If each is rational and assumes the others are equally rational, then they recognize, in the completely symmetric situation of the example, that only a symmetric allocation is possible.)

The point of this lengthy discussion of possible game theory concepts of equilibrium is to suggest caution in accepting the proposition that bargaining costs alone prevent

the achievement of Pareto efficiency in the presence of increasing returns, as Buchanan and Tullock (1962, p. 88) and Demsetz (1968, p. 61) assert.

Risk and Information

The possible types of equilibria discussed in the previous two sections are not, in principle, altered in nature by the presence of risk. If an economic agent is uncertain as to which of several different states of the world will obtain, he can make contracts contingent on the occurrence of possible states. The real-world counterparts of these theoretical contingent contracts include insurance policies and common stocks. With these markets for contingent contracts, a competitive equilibrium will arise under the same general hypotheses as in the absence of uncertainty. It is not even necessary that the economic agents agree on the probability distribution for the unknown state of the world; each may have his own subjective probabilities. Further, the resulting allocation is Pareto-efficient if the utility of each individual is identified as his expected utility according to his own subjective probability distribution.

But, as Radner (1968) has pointed out, there is more to the story. Whenever we have uncertainty we have the possibility of information and, of course, also the possibility of its absence. No contingent contract can be made if, at the time of execution, either of the contracting parties does not know whether the specified contingency has occurred or not. This principle eliminates a much larger number of opportunities for mutually favorable exchanges than might perhaps be supposed at first glance.

A simple case is that known in insurance literature as "adverse selection:" Suppose, for example, there are two types of individuals, A and B, with different life expectancies, but the insurance company has no way to distinguish the two; it cannot in fact identify the present state of the world in all its relevant aspects. The optimal allocation of resources under uncertainty would require separate insurance policies for the two types, but these are clearly impossible. Suppose further that each individual knows which type he belongs to. The company might charge a rate based on the probability of death in the two types together, but the insurance buyers in the two types will respond differently; those in the type with the more favorable experience, say A, will buy less insurance than those in type B, other things (income and risk aversion) being equal. The insurance company's experience will be less favorable than it intended, and it will have to raise its rates. An equilibrium rate will be reached which is, in general, between those corresponding to types A and B separately but closer to the latter. Such an insurance arrangement is, of course, not Pareto-efficient. It is not a priori obvious in general that this free market arrangement is superior to compulsory insurance even though the latter is also not Pareto-efficient because it typically disregards individual differences in risk aversion.

As the above example shows, the critical impact of information on the optimal allocation of risk bearing is not merely its presence or absence but its inequality among economic agents. If neither side knew which type the insured belonged to, then the final allocation would be Pareto-efficient if it were considered that the two types were indistinguishable; but in the above example the market allocation is Pareto-efficient neither with the types regarded as indistinguishable nor with them regarded as distinguishable.

There is no particular case of the effect of differential information on the workings of the market economy (or indeed any complex economy) which is so important as to deserve special comment: one agent can observe the joint effects of the unknown state of the world and of decisions by another economic agent, but not the state or the decision separately. This case is known in the insurance literature as "moral hazard,"

but because the insurance examples are only a small fraction of all the illustrations of this case and because, as Pauly (1968) has argued, the adjective "moral" is not always appropriate, the case will be referred to here as, the "con founding of risks and decisions." An insurance company may easily observe that a fire has occurred but cannot, without special investigation, know whether the fire was due to causes exogenous to the insured or to decisions of his (arson, or at least carelessness). In general, any system which, in effect, insures against adverse final outcomes automatically reduces the incentives to good decision making.

In these circumstances there are two extreme possibilities (with all intermediate possibilities being present): full protection against uncertainty of final outcome (e.g., cost-plus contracts for production or research) or absence of protection against uncertainty of final outcome (the one-person firm; the admiral shot for cowardice "*pour encourager les autres*"). Both policies produce inefficiency, though for different reasons. In the first, the incentive to good decision making is dulled for obvious reasons; in the second, the functions of control and risk bearing must be united, whereas specialization in these functions may be more efficient for the workings of the system.

The relations between principals and agents (e.g., patients and physicians, owners and managers) further illustrate the confounding of risks and decisions. In the. professions in particular they also illustrate the point to be emphasized later: that ethical standards may to a certain extent overcome the possible Pareto inefficiencies.

So far we have taken the information structure as given. But the fact that particular information structures give rise to Pareto inefficiency means that there is an economic value in transmitting information from one agent to another, as well as in the creation of new information. J. Marschak (1968), Hirshleifer (unpublished), and others have begun the study of the economics of information, but the whole subject is in its infancy. Only a few remarks relevant to our present purpose will be made here.

(1) As both communications engineering and psychology suggest, the transmission of information is not costless. Any professor who has tried to transmit some will be painfully aware of the resources he has expended and, perhaps more poignantly, of the difficulties students have in understanding. The physical costs of transmission may be low, though probably not negligible, as any book buyer knows; but the "coding" of the information for transmission and-the limited channel capacity of the recipients are major costs.

(2) The costs of transmitting information vary with both the type of information transmitted and the recipient and sender. The first point implies a preference for inexpensive information, a point stressed in oligopolistic contexts by Kaysen (1949, pp. 294-295) and in other bargaining contexts by Schelling (1957) . The second point is relevant to the value of education and to difficulties of transmission across cultural boundaries (so that production functions can differ so much across countries).

(3) Because the costs of transmission are nonnegligible, even situations which are basically certain become uncertain for the individual; the typical economic agent simply cannot acquire in a

meaningful sense the knowledge of all possible prices, even where they are each somewhere available. Markets are thus costly to use, and therefore the multiplication of markets, as for contingent claims as suggested above, becomes inhibited.

Externalities Illustrated

After this long excursus into the present state of the theory of equilibrium and optimality it is time to discuss some of the standard concepts of externality, market failure, and public goods generally.. The clarification of these concepts is a long historical process, not yet concluded, in which the classic contributions of Knight (1924), Young (1913 pp. 676-684), and Robertson (1924) have in more recent times been enriched by those of Meade (1952), Scitovsky (1954), Coase (1960), Buchanan and Stubblebine (1962), and Demsetz (1966) . The concept of externality and the extent to which it causes non-optimal market behavior will be discussed here in terms of a simple model.

Consider a pure exchange economy. Let x_{ik} be the amount of the k^{th} commodity consumed by the i^{th} individual (i =1, . . . , n; k =1, . . . , m) and x_k be the amount of the k^{th} commodity available. Suppose in general that the utility of the ill individual is a function of the consumption of all individuals (not all types of consumption for all individuals need actually enter into any given individual's utility function); the utility of the i^{th} individual is $U_i(x_{11}, , , , , , x_{mm})$. We have the obvious constraints.

(1) $$\sum_i x_{ik} \leq x_k$$

Introduce the following definitions:

(2) $$x_{jik} = x_{ik}$$

With this notation a Pareto-efficient allocation is a vector maximum of the utility functions $U_j(x_{j11}, . . . , x_{jmn})$, subject to the constraints (1) and (2) . Because of the notation used, the variables appearing in the utility function relating to the j^{th} individual are proper to him alone and appear in no one else's utility function. If we understand now that there are n 2m commodities, indexed by the triple subscript jik, then the Pareto-efficiency problem has a thoroughly classical form. There are n^2m prices, P_{jik}, attached to the constraints (2), plus m prices q_k , corresponding to constraints (1) . Following the maximization procedure formally, we see, much as in Samuelson [1954], that Pareto efficiency is characterized by the conditions:

(3) $$\lambda_j (fU_j / fx_{ik}) = P_{jik}$$

and

(4) $$\sum_j P_{jik} = q_k \ .$$

where .λ is the reciprocal of the marginal-utility of income for individual j. (These statements ignore corner conditions; which can easily be supplied.)

Condition (4) can be given the following economic interpretation: Imagine each individual i to be a producer with m production processes, indexed by the pair (i.k) . Process (i,k) has one input namely commodity k, and n outputs, indexed by the triple (j,i,k). In other words, what we ordinarily call individual i's consumption is regarded

as the production of joint outputs, one for each individual whose utility is affected by individual i's consumption.

The point of this exercise is to show that by suitable and indeed not unnatural reinterpretation of the commodity space, externalities can be regarded as ordinary commodities, and all the formal theory of competitive equilibrium is valid, including its optimality.

It is not the mere fact that one man's consumption enters into another man's utility that causes the failure of the market to achieve efficiency. There are two relevant factors which cannot be discovered by inspection of the utility structures of the individual. One, much explored in the literature, is the appropriability of the commodities which represent the external repercussions; the other, less stressed, is the fact that markets for externalities usually involve small numbers of buyers and sellers.

The first point, Musgrave's "exclusion .principle," (1959, p. 86) is so well known as to need little elaboration. Pricing demands the possibility of excluding nonbuyers from the use of the product, and this exclusion may be technically impossible or may require the use of considerable resources. Pollution is the key example; the supply of clean air or water to each individual would have to be treated as a separate commodity, and it would have to be possible in principle to supply to one and not the other (though the final equilibrium would involve equal supply to all). But this is technically impossible.

The second point comes out clearly in our case. Each commodity (j,i,k) has precisely one buyer and one seller. Even if a competitive equilibrium could be defined, there would be no force driving the system to it; we are in the realm of imperfectly competitive equilibrium.
In my view, the standard lighthouse example is best analyzed as a problem of small numbers rather than of the difficulty of exclusion, though both elements are present. To simplify matters, I will abstract from uncertainty. so that the lighthouse keeper knows exactly when each ship will need its services, and also abstract from indivisibility (since the light is either on or off). Assume further that only one ship will be within range of the lighthouse at any moment. Then exclusion is perfectly possible; the lighthouse need only shut off its light when a nonpaying ship is coming into range. But there would be only one buyer and one seller and no competitive forces to drive the two into a competitive equilibrium. If in addition the costs of bargaining are high, then it may be most efficient to offer the service free.

If, as is typical, markets for the externalities do not exist, then the allocation from the point of view of the "buyer" is determined by a rationing process. We can determine a shadow price for the buyer: this will differ from the price, zero, received by the seller. Hence, formally, the failure of markets for externalities to exist can also be described as a difference of prices between buyer and seller.

In the example analyzed, the externalities related to particular named individuals; individual i's utility function depended on what a particular individual, j, possessed. The case where it is only the total amount of some commodity (e.g., handsome houses) in other people's hands that matters is a special case, which yields rather simpler results. In this case, $\partial U_j / \partial x_{ik}$ is independent of i for $i \neq j$, and hence, by (3), P_{jik} is independent of i for $i \neq j$. Let

$$P_{iik} = P_{ik}, P_{jik} = \overline{P}_{jk} \text{ for i ? j}$$

Then (4) becomes,

$$P_{ik} + \sum_{j \neq i} \overline{P}_{jk} = q_k$$

or

$$(P_{ik} - \overline{P}_{ik}) + \sum_{j} \overline{P}_{jk} = q_k$$

from which it follows that the difference, P_{ik}- P_{ik} is independent of I There are two kinds of shadow prices, a price ¢;k, the price that individual I is willing to pay for an increase in the stock of commodity k in any other individual's hands, and the premium, P_{ik},- P_{ik} he is willing to pay to have the commodity in his possession rather than someone else's. At the optimum, this premium for private possession must be the same for all individuals.

Other types of externalities are associated with several commodities simultaneously and do not involve named individuals, as in the case of neighborhood effects, where an individual's utility depends both on others' behavior (e.g., esthetic, criminal) and on their location. There is one deep problem in the interpretation of externalities which can only be signaled here. What aspects of others' behavior do we consider as affecting a utility function? If we take a hard-boiled revealed preference attitude, then if an individual expands resources in supporting legislation regulating another's behavior, it must be assumed that that behavior affects his utility. Yet in the cases that students of criminal law call "crimes with out victims," such as homosexuality or drug-taking, there is no direct relation between the parties. Do we have to extend the concept of externality to all matters that an individual cares about? Or, in the spirit of John Stuart Mill, is there a second-order value judgment, which excludes some of these preferences from the formation of social policy as being illegitimate

Market Failure

The problem of externalities is thus a special case of a more general phenomenon, the failure of markets to exist. Not all examples of market failure can fruitfully be described as externalities. Two very important examples have already been alluded to; markets for many forms of risk-bearing and for most future transactions do not exist and their absence is sure suggestive of inefficiency.

Previous discussion has suggested two possible causes for market failure (1) inability to exclude; (2) lack of necessary information to permit mark transactions to be concluded.

The failure of futures markets cannot be directly explained in the terms. Exclusion is no more a problem in the future than in the present Any contract to be executed in the future is necessarily contingent on son events (for example, that the two agents are still both in business), b~ there must be many cases where no informational difficulty is present. The absence of futures markets may be ascribed to a third possibility: (3) supply and demand are equated at zero; the highest price at which anyone would buy is below the lowest price at which anyone would sell.

This third case of market failure, unlike the first two, is by itself in r way presumptive of inefficiency. However, it may usually be assumed that its occurrence is the result of failures of the first two types on complementary markets. Specifically, the demand for future steel may be low because of uncertainties of all types; sales and technological uncertainty for the buyer's firm, prices and existence of competing goods, and the

quality specification of the steel. If, however, adequate markets for risk-bearing exist, the uncertainties could be removed, and the demand for future steel would rise.

Transaction Costs

Market failure has been presented as absolute, but in fact the situation is more complex than this. A more general formulation is that of transaction costs, which are attached to any market and indeed to any mode resource allocation. Market failure is the particular case where transaction costs are so high that the existence of the market is no longer worthwhile The distinction between transaction costs and production costs is that tl former can be varied by a change in the mode of resource allocation, while the latter depend only on the technology and tastes, and would be tl same in all economic systems.

The discussions in the preceding sections suggest two sources of transa tion costs. (1) exclusion costs; (2) costs of communication and information, including both the supplying and the learning of the terms on which transactions can be carried out. An additional source is (3) the costs of disequilibrium; in any complex system, the market or authoritative allocation, even under perfect information, it takes time to compute the optimal allocation, and either transactions take place which are inconsistent with the final equilibrium or they are delayed until the computations are completed (see T. Marschak, 1959).

These costs vary from system to system; thus, one of the advantages of a price system over either bargaining or some form of authoritative allocation is usually stated to be the economy in costs of information and communication. But the costs of transmitting and especially of receiving a large number of price signals may be high; thus, there is a tendency not to differentiate prices as much as would be desirable from the efficiency viewpoint; for example, the same price is charged for peak and offpeak usage of transportation or electricity.

In a price system, transaction costs drive a wedge between buyer's and seller's prices and thereby give rise to welfare losses as in the usual analysis. Removal of these welfare losses by changing to another system (for example, governmental allocation on benefit-cost criteria must be weighed against any possible increase in transaction cost (for example, the need for elaborate and perhaps innumerable studies to determine demand functions without the benefit of observing a market).

The welfare implications of transaction c s would exist even if they were proportional to the size of the transaction, but in fact they typically exhibit increasing returns. The cost of acquiring a piece of information, for example, a price, is independent of the scale of use to which it will be put.

Collective Action: The Political Process

The State may frequently have a special role to play in resource allocation because, by its nature, it has a monopoly of coercive power, and coercive power can be used to economize on transaction costs. The most important use of coercion in the economic context is the collection of taxes; others are regulatory legislation and eminent domain proceedings.

The State is not an entity but rather a system of individual agents, a widely extensive system in the case of a democracy. It is appealing and fruitful to analyze its behavior in resource allocation in a manner analogous to that of the price system. Since the same agents appear in the two systems, it becomes equally natural to assume they have the same motives. Hotelling (1929, pp. 54-55) and Schumpeter (1942, ch. XXII

) had sketched such politicoeconomic models, and von Neumann and Morgenstem's monumental work is certainly based on the idea that all social phenomena are governed to essentially the same motives as economics. The elaboration of more or less complete models of the political process along the lines of economic theory is more recent, the most prominent contributors being Black (1958), Downs (1957), Buchanan and Tullock (1962), and Rothenberg (1965) .

I confine myself here to a few critical remarks on the possibilities of such theories. These are not intended to be negative but to suggest problems that have to be faced and are raised by some points in the preceding section.

1. If we take the allocative process to be governed by majority voting, then, as we will know, there are considerable possibilities of paradox. The possible intransitivity of majority voting was already pointed out by Condorcet' (1785) .(cf, instead of assuming that each individual votes according to his preferences it is assumed that they bargain freely before voting (vote selling), the paradox appears in another form, a variant of the bargaining problems already noted in section
2. If a majority could do what it wanted, then it would be optimal to win with a bare majority and take everything , but any such bargain can always be broken up by another proposed majority.

Tullock (1967) has recently argued convincingly that if the distribution of opinions on social issues is fairly uniform and if the dimensionality of the space of social issues is much less than the number of individuals, then majority voting on a sincere basis will be transitive. The argument is not, however, applicable to income distribution, for such a policy has as many dimensions as there are individuals, so that the dimensionality of the issue space is equal to the number of individuals.

This last observation raises an interesting question. Why, in fact, in democratic systems has there been so little demand for income redistribution? The current discussion of a negative income tax is the first serious attempt at a purely redistributive policy. Hagstrom (1938) presented a mathematical model predicting on the basis of a self-interest model for voters that democracy would inevitably lead to radical egalitarianism. -

Political policy is not made by voters, not even in the sense that they choose the vector of political actions which best suits them. It is in fact made by representatives in one form or another. Political representation is an outstanding example of the principal-agent relation. This means that the link between individual utility functions and social action is tenuous, though by no means completely absent. Representatives are no more a random sample of their constituents than physicians are of their patients.

Indeed, the question can be raised: to what extent is the voter, when acting in that capacity, a principal or an agent? To some extent, certainly, the voter is cast in a role in which he feels some obligation to consider the social good, not just his own. It is in fact somewhat hard to explain otherwise why an individual votes at all in a large election, since the probability that his vote will be decisive is so negligible.

Collective Action: Social Norms

It is a mistake to limit collective action to State action; many other departures from the anonymous atomism of the price system are observed regularly. Indeed, firms of any

complexity are illustrations of collective action, the internal allocation of their resources being directed by authoritative and hierarchical controls.

I want, however, to conclude by calling attention to a less visible form of social action: norms of social behavior, including ethical and moral codes. I suggest as one possible interpretation that they are reactions of society to compensate for market failures. It is useful for individuals to have some trust in each other's word. In the absence of trust it would become very costly to arrange for alternative sanctions and guarantees, and many opportunities for mutually beneficial cooperation would have to be, foregone. Banfield (1958) has argued that lack of trust is indeed one of the causes of economic underdevelopment.

It is difficult to conceive of buying trust in any direct way (though it can happen indirectly, for example, a trusted employee will be paid more as being more valuable); indeed, there seems to be some inconsistency in the very concept. Nonmarket action might take the form of a mutual agreement. But the arrangement of these agreements and especially their continued extension to new individuals entering the social fabric can be costly. As an alternative, society may proceed by internalization of these norms to the achievement of the desired agreement on an unconscious level.
There is a whole set of customs and norms which might be similarly interpreted as agreements to improve the efficiency of the economic system (in the broad sense of satisfaction of individual values) by providing commodities to which the price system is inapplicable.

These social conventions may be adaptive in their origins, but they can become retrogressive. An agreement is costly to reach and therefore costly to modify; and the costs of modification may be especially large for unconscious agreements. Thus, codes of professional ethics, which arise out of the principal-agent relation and afford protection to the principals, can serve also as a cloak for monopoly by the agents.

References

Aumann, R. J. 1964. “Markets with a Continuum of Traders”. *Econometrica* 32:39-50.

Aumann, R. J. 1966. “The Existence of Competitive Equilibria in Markets with a Continuum of Traders”. *Econometrica* 34:1-17.

Aumann, R. J. and Maschler, M. 1964. “The Bargaining Set for Cooperative Games”. In M. Dresher, L. S. Shapley, A.W.Tucker (eds). *Advances in Game Theory. Annals of Mathematics Study.* Princeton, New Jersey: Princeton, University Press, 52:443-476.

Banfield, E. C. 1958. *The Moral Basis of a Backward Society*. The Free Press.

Black, D. 1958. *The Theory of Committees and Elections*. Cambridge, U.K.: Cambridge University Press.

Buchanan, J. and W. C. Stubblebine. 1962. “Externality”. *Economica* 29:371-384.

Buchanan, J. and G. Tullock. 1962. *The Calculus of Consent*, Ann Arbor, Michigan: University of Michigan Press.

Chamberlin, E. H. 1933. *The Theory of Monopolistic Competition*. Cambridge, Mass. Eighth edition, 1965.

Coase, R. H. 1960. “The Problem of Social Cost”. *Journal of Law and Economics* 3:1-44.

Condorcet, Marquis de. 1785. *Essai sur 1'application de 1'analyse de la probabilité des décisions rendues de la pluralité des voix*. Paris.

Debreu, G. and H. Scarf. 1963. “A Limited Theorem on the Core of an Economy”. *International Economic Review* 4:235-246.

Demsetz, H. 1964. “The Welfare and Empirical Implications of Monopolistic Competition”. *Economic journal* 74:623-641.
Demsetz, H. 1966. “Some Aspects of Property Rights”. *Journal of Law and Economics* 9:6170.
Demsetz, H. 1968. “Why Regulate Utilities?” *Journal of Law and Economics* 11:55-66.
Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper.
Edgeworth, F. Y. 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London: C. Kegan Paul & Co.
Farrell, M. J. 1959. “The Convexity Assumption in the Theory of Competitive Markets”. *Journal of Political Economy* 67:377-391.
Hagstrom, K. G. 1938. “A Mathematical Note on Democracy”. *Econometrica* 6:381-383.
Harsanyi, J. C. 1956. “Approaches to the Bargaining Problem Before and After the Theory of Games: A Critical Discussion of Zeuthen's, Hicks', and Nash's Theories.” *Econometrica* 24:144-157.
Harsanyi, J. C. 1959. “A Bargaining Model for the Cooperative n-Person Game”. In A. W. Tucker and R. D. Luce (eds.) *Contributions to the Theory of Games IV. Annals of Mathematics Study*, Princeton, New Jersey: Princeton University 40:325-355.
Harsanyi, J. C. 1963. “A Simplified Bargaining Model for the n-Person Cooperative Game”. *International Economic Review* 4:194-220.
Harsanyi, J. C. 1966. “A General Theory of Rational Behavior in Game Situations”. *Econometrica*. 34:613-634.
Hotelling, H. 1929. “Stability in Competition”. *Economic Journal* 39:41-57.
Kaldor, N. 1935. “Market Imperfection and Excess Capacity”. *Economica,* N.S. 2:35-50.
Kaysen, Carl. 1949. “Base Point Pricing and Public Policy”. *Quarterly Journal of Economics* 63:289-314.
Knight, F. H. 1921. *Risk, Uncertainty, and Profit.* Boston and New York: Houghton Mifflin. Reprinted by London School of Economics and Political Science 1948.
Knight, F. H. 1924. “Some Fallacies in the Interpretation of Social Cost”. *Quarterly Journal of Economics* 38:582-606.
Luce, R. D. 1954. “A Definition of Stability for n-Person Games”. *Annals of Mathematics* 59:357-366.
Luce, R. D. 1955a. “K-Stability: A New Equilibrium Concept for n-Person Game Theory”. In *Mathematical Models of Human Behavior.* Stamford, Corm.: Dunlap and Associates, pp.32-44.
Luce, R. D. 1955b. “K-Stability of Symmetric and Quota Games”. *Annals of Mathematics* 62:517-555.
Marschak, T. 1959. “Centralization and Decentralization in Economic Organizations”. *Econometrica* 27:399-430.
Meade, J. E. 1952. “External Economies and Diseconomies in a Competitive Situation”. *Economic Journal* 62:54-67.
Musgrave, R. A. 1959. *The Theory of Public Finance: A Study in Public Economy*. New York: McGraw-Hill Book Company.
Nash, J. F. Jr. 1950. “The Bargaining Problem”. *Econometrica* 18:155-162.
Nash, J. F. Jr. 1953. “Two Person Cooperative Games”. *Econometrica* 21:128-140.
Negishi, T. 1960-1. “Monopolistic Competition and General Equilibrium”. *Review of Economic Studies* 28:196-201.
von Neumann, J., and O: Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton, N. J.: Princeton University Press. Second edition. 1947.
Pauly, M. V. 1968. “The Economics of Moral Hazard: Comment”, *American Economic' Review* 58:531-537.

Radner, R. 1968. “Competitive Equilibrium under Uncertainty”. *Econometrica* 36:31-58.
Robertson, D. H. 1924. “Those Empty Boxes”. *Economic Journal* 34:16-30.
Rothenberg, J. 1960. “Non-Nonvexity, Aggregation, and Pareto Optimality”. *Journal of Political Economy* 68:435-468.
Rothenberg, J. 1965. “A Model of Economic and Political Decision-Making”. In J. Margolis (ed.) *The Public Economy of Urban Communities*. Washington, D.C.: Resources for the Future.
Samuelson, P. A. 1954. “The Pure Theory of Public Expenditures”. *Review of Economics and Statistics* 36:387-389.
Samuelson, P. A. 1967. “The Monopolistic Competition Revolution”. In R. E. Kuenne (ed.) *Monopolistic Competition Theory: Studies in Impact*. New York, London, and Sydney: Wiley, pp. 105-138.
Schelling, T. 1957. “Bargaining, Communication, and Limited War”. *Journal of Conflict Resolution* 1:19-36.
Schumpeter, J. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper. Third Edition, 1950.
Scitovsky, T. 1954. “Two Concepts of External Economies”. *Journal of Political Economy* 62:143-151.
Selten, R. 1964. “Valuation of n-Person Games”. In M. Dresher, L. S. Shapley, A. W. Tucker (eds.) *Advances in Game Theory*. Princeton, New Jersey; Princeton University Press, 52:577-626.
Shapley, L. S. 1953. “A Value for n-Person Games”. In H. W. Kuhn and A. W. Tucker (eds.) *Contributions to the Theory of Games II. Annals of Mathematics Study*. Princeton, New Jersey: Princeton University Press, 28:307-317.
Shapley, L. S., and M. Shubik. 1967. “Ownership and the Production Function”. *Quarterly Journal of Economics* 81:$8-111.
Shubik M. 1959. “Edgeworth Market Games”. In A. W. Tucker and R. D. Luce (eds.) *Contributions to the Theory of Games IV. Annals of Mathematics Study*. Princeton, New Jersey: Princeton University Press, 40:267-278.
Smith A. 1937. *An Enquiry Concerning the Causes of the Wealth of Nations*. New York: Modern Library reprint of 1776 edition.
Starr, R. 1969. “Quasi-Equilibria in Markets with Nonconvex Preferences”. *Econometrica* 37:25-38.
Triffin, R. 1941. *Monopolistic Competition and General Equilibrium Theory*. Cambridge, Mass.: Harvard University Press.
Tullock, G. 1967. *Toward a Mathematics of Politics*. Ann Arbor, Michigan: University of Michigan Press.
Young, A. A. 1913. “Pigou's Wealth and Welfare”. *Quarterly Journal of Economics* 27: 672-686.
Zeuthen, F. 1930. *Problems of Monopoly and Economic Welfare*. London, George Routledge & Sons Ltd.