

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Analysis of Public Output

Volume Author/Editor: Julius Margolis, ed.

Volume Publisher: NBER

Volume ISBN: 0-87014-220-8

Volume URL: <http://www.nber.org/books/marg70-1>

Publication Date: 1970

Chapter Title: Political and Economic Evaluation of Social Effects and Externalities

Chapter Author: Kenneth Arrow

Chapter URL: <http://www.nber.org/chapters/c3349>

Chapter pages in book: (p. 1 - 30)

# POLITICAL AND ECONOMIC EVALUATION OF SOCIAL EFFECTS AND EXTERNALITIES

KENNETH J. ARROW  
*Harvard University*

## Introduction

The concept of public goods has been developed through successive refinement over a long period of time. Yet, surprisingly enough, nowhere in the literature does there appear to be 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 to interpret 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 problem: 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 nonmarket 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 nonmarket 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 nonmarket actions which will restore Pareto efficiency. The second involves the recognition that the instruments available to the government or other nonmarket forces are scarce resources for one reason or another, so that all that can be achieved is a "second best."

Other concepts that 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 the replacement of the costs of buying and selling on the market by the costs of intrafirm 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 the costliness of the information needed to enter and participate in any market, another type of cost, has received little attention. 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 will be pressure in the system 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 of 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 the theoretical literature of recent years seeks to describe political behavior as analogous to economic, and we may hope for a general theory of socio-economic 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 at this point. 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.<sup>1</sup> 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 does not itself consume resources.

A number of additional assumptions are made at different points in the theory of equilibrium, but most clearly are 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

<sup>1</sup> Sometimes this is stated to permit an excess of supply over demand, with a zero price for such free goods; but this can be included in the above formulation by postulating the existence of production processes (disposal processes) which have such surpluses as inputs and no outputs.

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 [1959], 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 nonconvexities 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 perfectly competitive 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 of 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 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 of bargaining costs per 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 assumption *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 needed perforce 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 seen 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 such a choice. One assumes that 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 were high relative to the costs of ordinary pricing, while the second assumes costless bargaining.<sup>2</sup>

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 considered 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 perceived 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 greater 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.<sup>3</sup>

<sup>2</sup> Within the framework of each prototheory attempts have been made to modify it in the direction of the other. Thus, price discrimination is a modification of the price system in the pure theory of monopoly, though I am aware of no attempt to study price discrimination in a competitive or otherwise general equilibrium context. Some game theorists (Luce [1954, 1955a, b], Aumann and Maschler [1964]) have attempted to introduce bargaining costs in some way by simply limiting the range of possible coalitions capable of making bargains.

<sup>3</sup> Suppose that the degree of increasing returns is sufficient to prevent there being more than one producer of a given commodity narrowly defined, but not to prevent production of a close substitute. Is this degree of returns sufficiently substantial to upset the achievement of an approximately perfect competitive equilibrium, as discussed in the last section?



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 Pareto-inefficient allocation to a suitably chosen Pareto-efficient one. This argument seems plausible, but is not easy to evaluate in the absence of a generally accepted concept 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 complementary costs) exceeds the total value of the land in its present uses. Then the owner of each 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, Chapter IV] had early propounded one such solution. After von Neumann and Morgenstern, Nash [1950, 1953] offered a solution, which Harsanyi [1956] later showed to be identical with that of Zeuthen. Nash's analysis of bargaining 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. With-

out 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 agents. 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 one-third 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 agents 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.<sup>4</sup>

(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.)

<sup>4</sup> The general principle illustrated by this example has been briefly alluded to by Shapley and Shubik [1967, p. 98, fn. 5].

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 as distinguishable.

There is one 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 "confounding 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 out-

come (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. 294–95] 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 nonoptimal 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 ( $i = 1, \dots, n$ ;  $k = 1, \dots, m$ ) and  $\bar{x}_k$  be the amount of the  $k$ th commodity available. Suppose in general that the utility of the  $i$ th 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  $U_i(x_{11}, \dots, x_{mn})$ . We have the obvious constraints:

$$(1) \quad \sum_i x_{ik} \leq \bar{x}_k$$

Introduce the following definitions:

$$(2) \quad x_{jik} = x_{ik}.$$

With this notation a Pareto-efficient allocation is a vector maximum of the utility functions  $U_j(x_{j11}, \dots, x_{jmn})$ , 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:

$$(3) \quad \lambda_j(\partial U_j / \partial x_{ik}) = p_{jik},$$

and

$$(4) \quad \sum_i p_{jik} = q_k,$$

where  $\lambda_j$  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 it to some and not to others (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 keeper need only shut off the 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 condition 3,  $p_{jik}$  is independent of  $i$  for  $i \neq j$ . Let,

$$p_{iik} = p_{ik}, p_{jik} = \bar{p}_{jk} \text{ for } i \neq j.$$

Then condition 4 becomes,

$$\bar{p}_{ik} + \sum_{j \neq i} \bar{p}_{jk} = q_k,$$

or,

$$(p_{ik} - \bar{p}_{ik}) + \sum_j \bar{p}_{jk} = q_k,$$

from which it follows that the difference,  $p_{ik} - \bar{p}_{ik}$ , is independent of  $i$ . There are two kinds of shadow prices, a price  $\bar{p}_{ik}$ , 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,  $p_{ik} - \bar{p}_{ik}$ , 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., aesthetic, 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 expends 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 without 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 infringements of individual freedom?

## 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 surely suggestive of inefficiency.

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

The failure of futures markets cannot be directly explained in these 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 some events (for example, that the two agents are still both in business), but there must be many cases where no informational difficulty is presented. 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 no 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 existed, 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 of 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 the 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 the same in all economic systems.

The discussions in the preceding sections suggest two sources of transaction costs: (1) exclusion costs; and (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 off-peak 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 costs (for example, the need for elaborate and perhaps impossible studies to determine demand functions without the benefit of observing a market).

The welfare implications of transaction costs 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, e.g., 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 that they have the same motives. Hotelling [1929, pp. 54–55] and Schumpeter [1942, Chapter XXII] had sketched such politico-economic models, and von Neumann and Morgenstern's monumental work is certainly based on the idea that all social phenomena are governed by 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 discussion.

1. If we take the allocative process to be governed by majority voting, then, as we well know, there are considerable possibilities of paradox. The possible intransitivity of majority voting was already pointed out by Condorcet [1785]. If, instead of assuming that each individual votes according to his preferences, it is assumed that all 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, Chapter III] 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. Hagström [1938] presented a mathematical model predicting on the basis of a self-interest model for voters that democracy would inevitably lead to radical egalitarianism.

2. 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, e.g., 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*, Vol. 32, pp. 39-50.
- Aumann, R. J., 1966. "The Existence of Competitive Equilibria in Markets with a Continuum of Traders," *Econometrica*, Vol. 34, pp. 1-17.
- Aumann, R. J., and Maschler, M., 1964. "The Bargaining Set for Cooperative Games," in Dresher, M., Shapley, L. S., and Tucker, A. W. (eds.), *Advances in Game Theory. Annals of Mathematics Study*, Vol. 52, Princeton, pp. 443-476.
- Banfield, E. C., 1958. *The Moral Basis of a Backward Society*, Glencoe, Ill.
- Black, D., 1958. *The Theory of Committees and Elections*, Cambridge, U.K.
- Buchanan, J., and Stubblebine, W. C., 1962. "Externality," *Economica*, Vol. 29, pp. 371-384.
- Buchanan, J., and Tullock, G., 1962. *The Calculus of Consent*, Ann Arbor, Michigan.
- Chamberlin, E. H., 1933. *The Theory of Monopolistic Competition*, Cambridge, Mass., eighth edition.
- Coase, R. H., 1960. "The Problem of Social Cost," *Journal of Law and Economics*, Vol. 3, pp. 1-44.
- Condorcet, Marquis de, 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris.
- Debreu, G., and Scarf, H., 1963. "A Limit Theorem on the Core of an Economy," *International Economic Review*, Vol. 4, pp. 236-246.
- Demsetz, H., 1964. "The Welfare and Empirical Implications of Monopolistic Competition," *Economic Journal*, Vol. 74, pp. 623-691.
- Demsetz, H., 1966. "Some Aspects of Property Rights," *Journal of Law and Economics*, Vol. 9, pp. 61-70.
- Demsetz, H., 1968. "Why Regulate Utilities," *Journal of Law and Economics*, Vol. 11, pp. 55-66.
- Downs, A., 1957. *An Economic Theory of Democracy*, New York.
- Edgeworth, F. Y., 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*, London.
- Farrell, M. J., 1959. "The Convexity Assumption in the Theory of Competitive Markets," *Journal of Political Economy*, Vol. 67, pp. 377-391.
- Hagström, K. G., 1938. "A Mathematical Note on Democracy," *Econometrica*, Vol. 6, pp. 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*, Vol. 24, pp. 144-157.

Harsanyi, J. C., 1959. "A Bargaining Model for the Cooperative N-Person Game," in Tucker, A. W., and Luce, R. D. (eds.), *Contributions to the Theory of Games IV. Annals of Mathematics Study*, Princeton, pp. 325-355.

Harsanyi, J. C., 1963. "A Simplified Bargaining Model for the N-Person Cooperative Game," *International Economic Review*, Vol. 4, pp. 194-220.

Harsanyi, J. C., 1966. "A General Theory of Rational Behavior in Game Situations," *Econometrica*, Vol. 34, pp. 613-634.

Hotelling, H., 1929. "Stability in Competition," *Economic Journal*, Vol. 39, pp. 41-57.

Kaldor, N., 1935. "Market Imperfection and Excess Capacity," *Economica*, Vol. 2, pp. 33-50.

Kaysen, Carl, 1949. "Basing Point Pricing and Public Policy," *Quarterly Journal of Economics*, Vol. 63, pp. 289-314.

Knight, F. H., 1921. *Risk, Uncertainty, and Profit*, Boston and New York.

Knight, F. H., 1924. "Some Fallacies in the Interpretation of Social Cost," *Quarterly Journal of Economics*, Vol. 38, pp. 582-606.

Luce, R. D., 1954. "A Definition of Stability for N-Person Games," *Annals of Mathematics*, Vol. 59, pp. 357-366.

Luce, R. D., 1955a. " $\Psi$ -Stability: A New Equilibrium Concept for N-Person Game Theory," in *Mathematical Models of Human Behavior*, Stamford, Conn., pp. 32-44.

Luce, R. D., 1955b. "K-Stability of Symmetric and Quota Games," *Annals of Mathematics*, Vol. 62, pp. 517-555.

Marschak, J., 1968. "Economics of Inquiring, Communicating, Deciding," *American Economic Review Papers and Proceedings*, Vol. 58, pp. 1-18.

Marschak, T., 1959. "Centralization and Decentralization in Economic Organizations," *Econometrica*, Vol. 27, pp. 399-430.

Meade, J. E., 1952. "External Economies and Diseconomies in a Competitive Situation," *Economic Journal*, Vol. 62, pp. 59-67.

Musgrave, R. A., 1959. *The Theory of Public Finance: A Study in Public Economy*, New York.

Nash, J. F., Jr., 1950. "The Bargaining Problem," *Econometrica*, Vol. 18, pp. 155-162.

Nash, J. F., Jr., 1953. "Two Person Cooperative Games," *Econometrica*, Vol. 21, pp. 128-140.

Negishi, T., 1960-61. "Monopolistic Competition and General Equilibrium," *Review of Economic Studies*, Vol. 28, pp. 196-201.

von Neumann, J., and Morgenstern, O., 1944. *Theory of Games and Economic Behavior*, Princeton, second edition.

- Pauly, M. V., 1968. "The Economics of Moral Hazard: Comment," *American Economic Review*, Vol. 58, pp. 531-537.
- Radner, R., 1968. "Competitive Equilibrium Under Uncertainty," Vol. 36, pp. 31-58.
- Robertson, D. H., 1924. "Those Empty Boxes," *Economic Journal*, Vol. 34, pp. 16-30.
- Rothenberg, J., 1960. "Non-Convexity, Aggregation, and Pareto Optimality," *Journal of Political Economy*, Vol. 68, pp. 435-468.
- Rothenberg, J., 1965. "A Model of Economic and Political Decision-Making," in Margolis, J. (ed.), *The Public Economy of Urban Communities*, Washington, D. C.
- Samuelson, P. A., 1954. "The Pure Theory of Public Expenditures," *Review of Economic Statistics*, Vol. 36, pp. 387-389.
- Samuelson, P. A., 1967. "The Monopolistic Competition Revolution," in Kuenne, R. E. (ed.), *Monopolistic Competition Theory: Studies in Impact*, New York, London, and Sydney, pp. 105-138.
- Schelling, T., 1957. "Bargaining, Communication, and Limited War," *Journal of Conflict Resolution*, Vol. 1, pp. 19-36.
- Schumpeter, J., 1942. *Capitalism, Socialism, and Democracy*, New York, third edition.
- Scitovsky, T., 1954. "Two Concepts of External Economies," *Journal of Political Economy*, Vol. 62, pp. 143-151.
- Selten, R., 1964. "Valuation of N-Person Games," in Dresher, M., Shapley, L. S., and Tucker, A. W. (eds.), *Advances in Game Theory. Annals of Mathematics Study*, Vol. 52, Princeton, pp. 577-626.
- Shapley, L. S., 1953. "A Value for N-Person Games," in Kuhn, H. W., and Tucker, A. W. (eds.), *Contribution to the Theory of Games II. Annals of Mathematics Study*, Princeton, Vol. 28, pp. 307-317.
- Shapley, L. S., and Shubik, M., 1967. "Ownership and the Production Function," *Quarterly Journal of Economics*, Vol. 81, pp. 88-111.
- Shubik, M., 1959. "Edgeworth Market Games," in Tucker, A. W., and Luce, R. D. (eds.), *Contribution to the Theory of Games IV. Annals of Mathematics Study*, Princeton, Vol. 40, pp. 267-278.
- Smith, A., 1937. *An Enquiry Concerning the Causes of the Wealth of Nations*, New York.
- Starr, R., 1969. "Quasi-Equilibria in Markets with Nonconvex Preferences," *Econometrica*, Vol. 37, pp. 25-38.
- Triffin, R., 1941. *Monopolistic Competition and General Equilibrium Theory*, Cambridge, Mass.
- Tullock, G., 1967. *Toward a Mathematics of Politics*, Ann Arbor, Mich.
- Young, A. A., 1913. "Pigou's Wealth and Welfare," *Quarterly Journal of Economics*, Vol. 27, pp. 672-686.
- Zeuthen, F., 1930. *Problems of Monopoly and Economic Warfare*, London.



## COMMENT

by SIDNEY S. ALEXANDER, *Massachusetts Institute of Technology*

Instead of commenting in detail on Arrow's paper, I will address myself directly to the original question. The topic announced is "The Political and Economic Evaluation of Social Effects and Externalities." The key word is evaluation, and Arrow has very carefully avoided treating evaluation. What he has done is to sketch with great virtuosity the relationship of the existence of markets, and the convexity of utility and production possibility sets, to Pareto-efficiency. In particular he has indicated how Pareto-efficiency can fail to be achieved if externalities exist, and how externalities may be regarded as equivalent to the failure of certain markets to exist.

The essence of an externality is that there is an effect on some person of a transaction to which he is not a party. Since the attainment of Pareto-efficiency through the operation of a competitive market is brought about by each party to a transaction adjusting to the price parameters, the fact that there are people affected by the transaction but not parties to it means that the appropriate adjustments will not be made, and the points attained in such a process will not in general be Pareto-efficient. As I construe the topic, the question is how the effects of economic and political transactions which fall upon those who are not parties to the transactions are to be evaluated with a view toward possible public action.

To this question Arrow has furnished only an implicit response, but a very clear one, and one that he shares with the majority of social scientists. It is, in effect, that all value judgments but one are outside the business of the social scientist: the one exceptional value judgment that can be embedded in economic analysis is the so-called "pig-principle," that more is better, so an economist's concern is with efficiency.

Except for an occasional lapse Arrow refers to Pareto-*efficiency* rather than to Pareto-*optimality*. It is there that he parts company with me and with evaluation, for evaluation is normative and efficiency is descriptive. A Pareto-efficient point is one that is undominated from a preference point of view; a Pareto-optimal point is one that is undominated from a welfare point of view. A point is Pareto-efficient if nobody can be *put into a preferred position* by moving to some other

attainable point without putting someone else into a less preferred position. A point is Pareto-optimal if nobody can be *made better off* by moving to some other attainable point without making somebody else worse off. The set of Pareto-efficient points will be different from the set of Pareto-optimal points if to be in a preferred position is not the same as to be better off. To say that the two states are the same is a normative statement equating a positive operational state of being preferred to a normative state of being better off. Indeed it specifies a particular form of utilitarianism, a normative doctrine. To talk about Pareto-optimality rather than Pareto-efficiency, then, is to leap the gulf from the *is* to the *ought*, and Arrow has elected not to take the jump but to stay on the *is* side of that gulf. But evaluation lies on the other side.

Arrow recognizes with Pareto that a Pareto-efficient point has a normative claim to make only if the income distribution is appropriate, or, more accurately, if there is an appropriate initial distribution of control over resources. But what makes such a distribution appropriate or inappropriate, good or bad, neither Pareto nor Arrow says. In fact, they both imply that it is something which, if it exists at all, is given from outside the realm of their inquiries. Any complaints can be met by lump-sum transfers, Arrow tells us (p. 4). Whether such complaints are well or poorly grounded, or how to tell whether such a complaint is well or poorly grounded, Arrow does not say. But that is one of the most important problems of evaluation of the operation of the market mechanism, the topic of this discussion. Faced with this problem Little took recourse to Superman to evaluate different income distributions.<sup>1</sup>

If we are to evaluate social effects and externalities we must evaluate income distributions as well, and much more. In fact we must evaluate Pareto-efficiency. My message to you is that you too can be a Superman, that you can make judgments of better or worse in social affairs, not infallible judgments it is true, but to the extent that you err such error can be detected only by another judgment of the same sort.

What do we do when we judge between two possible states of the world? In general, we look to two characteristics, welfare and justice, corresponding to two principal strains in ethical theory, utilitarianism and deontology respectively. Utilitarianism judges between states of the world according as it finds more happiness, by some measure deemed suitable, in one state than in another. From a deontological point of

<sup>1</sup> I. M. D. Little, *A Critique of Welfare Economics*, 1950, p. 87.

view, one state of the world is better than another when it is more the way things should be by some prior standard. Pareto-efficiency makes its normative claim through an implicit utilitarianism that says that people are better off when they have their wants satisfied to a greater extent. The income distribution actually found in practice in our economy is largely based on a deontological notion of justice, that one should not take away from a man that which is his. This may indeed have a utilitarian basis in the long run, but again it may not. When we do use the social coercion of the state to take by taxation from a man what would otherwise be his, we usually justify that action either in terms of distributive justice or in terms of increasing the over-all happiness suitably measured.

A conceptually simple test for both justice and happiness together can be based on a fundamental principle of morality expressed in the question "how would you feel if you were in his shoes?" If we have to make sense of saying that one state of the world is better than another, we can, I submit, do it in the following way. If a reasonable man completely knowledgeable of the conditions in both states is given the choice of taking a chance with equal probability of being anyone concerned in either state A or state B and he chooses state A, we can say that state A is better than state B.<sup>2</sup> This is not given as a *definition* of the word "better" but rather as a proposed positive test of the normative concepts of better or worse in social affairs, it being presumed that "better" has a primitive normative meaning in our language. Whether state A is preferred by the test man because there is more happiness in it or more justice we need not here inquire. The test would presumably test both simultaneously, with trade-offs implicit.

It might be argued on the basis of the reigning attitudes in the social sciences that this "place-taking" test probably appeals to the ethical intuitions of many of us, but the resulting agreement is only a conjunction of personal preferences of those who share common standards. If I say state A is better than state B, it is commonly argued, I am merely saying that I personally would rather take the chance of being anyone in state A than of being anyone in state B. I am saying something more than that, however, and that something is critical. It is the implication that anyone else who thought clearly and was knowledgeable would have a similar preference. I am making an impersonal claim

<sup>2</sup> This test, and the philosophy underlying it, is more fully discussed in my essay "Public Television and the Ought of Public Policy," *Washington University Law Quarterly*, Winter 1968 (to be reprinted in a forthcoming Brookings Institution volume).

as an agent for all reasonable moral men and not just expressing my personal taste.

We need have no deep metaphysical support for this usage. It is simply proposed as what, upon reflection, we mean when we say state of the world A is better than state of the world B.

How do I know that this is a good test? The same way I know that an operational test is a good test, say of the specific gravity of some substance. These are methodological judgments appropriate to the subject in hand. If you agree that this is the right test to use we can proceed with our inquiry. If not, we have to escalate the question to the next higher methodological level of what is a good test, or the meta-methodological level of what is a good method for choosing methods—and the sequence has no end. No discussion at all can take place except among people who agree on something.

There can be hardly any question that if a place-taking test were applied to our current social arrangements as compared with possible alternatives the test man would almost certainly demand a state with more equalization of opportunities than our society affords. How much, if at all, other desirable features of the society would have to be traded off for greater equality of opportunity requires a lot of study. But it would be the greatest of coincidences if the current balance was the best possible world, since there is, apparently, no equivalent of a market process to lead to an optimal trade-off between all desirable characteristics. This is a subject which is simply avoided in social science discussions because of the presumed impossibility of making normative judgments. Deference to that impossibility is perfectly illustrated in Arrow's paper where no attempt is made to treat income distribution except as a boundary condition to be governed from some external position from which complaints are made, or distributions regarded as acceptable.

Suppose that an ideal distribution of income has been achieved, relative to the attainable Pareto-efficient points, and the corresponding Pareto-efficient point attained. Thus, we may assume Arrow's conditions *C* and *M* to hold, so that each person is compensated for the smoke nuisance just up to the point which makes him choose the amount of smoke he actually gets and each smoke emitter is charged accordingly. Conditions 3 and 4 (pp. 4–5) will accordingly hold. Would that indeed be heaven, or at least the Good Life? Everybody would be getting what he wants to the extent that is jointly possible given the resource endowment, the conditions of production, and the best possible distribution of ownership of resources. While that would undoubtedly

be a far better world than we now have, it need by no means be the Good Life. Whether it would or would not be a Good Life depends on what sort of people it makes of us, and what we get out of the satisfaction of our wants. One does not have to go very deep into philosophy to observe rather tritely with Frank Knight and others that want-satisfaction is not a particularly worthy ethical ideal. The wealthier people in our society do not live such worthy lives that we can follow Hayek in regarding them as pioneers of the Good Life.

That the value of a life depends upon the values served and realized rather than on the extent of want-satisfaction is certainly a trite observation, but the current state of thought in the social sciences so clearly ignores this point that I must remind you of it however trite it may be. If you feel that this is preacher talk and unsuitable for serious social scientists, I must disagree, and give my reasons. We concentrate on Pareto-efficiency because there are many elegant things we can say about it, and because it lies closest to what we can reasonably hope that our economy can achieve through its current institutions. If, then, we delude ourselves into thinking that that which we can analyze is that which we ought to analyze when we seek to evaluate our social institutions, we are making the mistake illustrated by the vaudeville routine of looking for the lost quarter at the corner under the lamp post where the light is good rather than in the darkness of the middle of the block where the quarter was lost.

Arrow recognizes externalities where there is a failure of a market to exist but where the wants do exist so that if a market existed there would be a tendency for the appropriate economic agent to adjust his demand for the "commodity" to its price. There are some externalities, and indeed I think the most important ones, where even if a market should be brought into existence there would be no appropriate adjustment because the element does not enter into the utility function of the person affected. In some cases this might just be a matter of ignorance; if a man knew what was happening to him he would be interested in paying for changing the amount of that particular activity, or in Arrow's term, "commodity." But the pig will not pay to become Socrates. The external effects of our economic, political and social activities which make us what we are are not such that an appropriate balance would be struck if a market were created or established for these "commodities." The value of a good society is not to be measured by how much people in our present society are ready to pay for it. There is a difference between what serves a man's welfare and what enters his utility function if the latter is to be derived from

his pattern of preference in the manner called "revealed preference."

Perhaps much of the difference between what a man wants and what serves his welfare is just a matter of information. This was certainly argued by Socrates when he claimed that no man knowingly did wrong, that evil came from ignorance. It was also the basis of Mill's argument for poetry's superiority over pushpin in that those who really knew both preferred poetry. But the argument, I think, is unsound. There are, I believe, many of us who knowing the better choose the worse. Here perhaps Dostoevski saw more clearly than John Stuart Mill what the problem is. So, while the imperfection may in part be one of lack of information, there may indeed be a more deep-seated root of evil in man's nature and it may take more than better information to bring us closer to the saints through want-satisfaction. Is it just ignorance that makes a racist, or is he responding to other environmental forces?

But whether the trouble comes from lack of information or lack of desire for the good, the practical problem is the same. There are effects of economic and political transactions whose impact on individuals not party to the transactions are individually negligible but cumulatively determine the very nature of those individuals. These cumulative forces operate outside an optimizing mechanism. They constitute a fundamental challenge to social policy. I maintain that, the positive facts being available, better guidance could be obtained from the place-taking test than from a Pareto-efficiency test. The place-taking test might indeed be a test for Pareto-optimality, in that it is an attempt to measure welfare rather than want-satisfaction.

But the positive facts are not generally available, and their lack is the principal obstacle to the recognition of sound policy. How, for example, can the system be altered to achieve greater equality of opportunity? What would be the effects on output or on other aspects of our life? But recognition of policy and its achievement are two different things. Even were we to gain perfect knowledge of the operation of our social and economic system and have it all in the computer, would our political process, there so well described, realize an optimal point? Surely that self-interest which drives an economic system to a Pareto-efficient point may not drive our political system to a Pareto-optimal one.

If one wishes to operate on that system, one needs a model far more complex than our elegant model of a perfectly competitive market. But complex as the understanding of the positive features of our society are, the normative is relatively simple. The way to evaluate is to evaluate, to judge between two social states the best we can. That means funda-

mentally to consider in which state a man would prefer to live who was reasonable, moral, well informed as to the nature and consequences of various actions and of how it is to live under the two different regimes. That is hardly the end of evaluation but it is a beginning, and a beginning that leads to something other than want satisfaction as the ideal and the relegation of ideal income distribution to some offstage force which is to supply us with an answer that is beyond our capacity to formulate for ourselves.