COLLECTED PAPERS OF KENNETH J. ARROW

The Economics of Information

The Belknap Press of Harvard University Press
Cambridge, Massachusetts 1984

16 The Property Rights Doctrine and Demand Revelation under Incomplete Information

The conditions under which the price system might not achieve optimal resource allocation have gradually been refined since they were first given reasonably accurate expression by Young (1913) in a review of Pigou's *Wealth and Welfare* (the first edition of what became *The Economics of Welfare*). Landmarks on the way to better understanding were the work of Knight (1924) and that of Scitovsky (1954).

The basic thesis is that the optimal resource allocation will not be achieved by a competitive market system if there are technological externalities. These are goods (or bads) for which no market can be formed. The usual reason given is that the good is not property in law or in practice, the latter covering the cases in which the act of enforcing property rights is itself costly and may therefore not be worthwhile. An excellent survey of market failure and its implications is to be found in Bator (1958).

A conclusion which is usually drawn from the presence of externalities not mediated through competitive markets is that the state has to intervene in some form to improve resource allocation, whether in the form of taxes and subsidies or other regulatory forms. Such were the recommendations of Pigou (1952) after he had absorbed Young's critique; an enthusiastic formulation is that of Baumol (1952).

A recent counterattack has been that of the so-called property rights school, starting with the well-known and important paper of Coase (1960)

Reprinted from *Economics and Human Welfare*, ed. M. Boskin (New York: Academic Press, 1979), pp. 23–39.

and including the work of Buchanan, McKean, and others, as surveyed by Furubotn and Pejovich (1972). They start with the position, common to all, that market failures are associated with lack of definition of property rights. However, they then argue that in principle clear definition of property rights is sufficient to ensure efficiency. This position goes well beyond the standard neoclassical position that competitive markets suffice for efficiency. As is well known, defined property rights are only one of the necessary conditions for competitive markets; large numbers (actual or potential) of buyers and sellers, concavity of the production possibility sets, and informed buyers are others. Scitovsky (1951, chaps. 15 and 16) has given the classic characterization of competitive markets as means of achieving efficiency.

The property rights theorists do not usually set out their underlying assumptions with the utmost clarity; but it appears that the basic postulate is the same one that underlies the theory of cooperative games, in the original formulation of von Neumann and Morgenstern (1947, chaps. 5, 6, and 10) and virtually all later developments (see, for example, Luce and Raiffa, 1957, chaps. 6, 8, and 9). That is, whatever else may be true about the outcome of the bargaining process, it will certainly be Pareto-optimal. The argument is obvious. Suppose A and B are both possible outcomes of the game, achievable by suitable choices of strategies by the players. Suppose the players can bargain about the choices of strategies, including possible side payments, and suppose that every player prefers A to B. Then clearly they will not stop at B, since, if nothing else is achievable, they can all improve by going to A.

What is not always recognized is that this argument depends crucially on the unstated assumption that every player knows every other player's payoff (utility, profit, whatever) as a function of the strategies played. In the case of bargaining over externalities, the strategies might be offers and counteroffer strategies, that is, plans beforehand to make a counteroffer as a function of the initial offer. If player I misperceives the payoff function of player II, then he or she may make an offer judged to be Pareto-superior to the initial position but not in fact superior in player II's payoff function. Thus, getting stuck at a Pareto-dominated point is no longer impossible.

In the traditional smoke case, suppose the landowners in the neighborhood of a factory own the property rights to clean air. (The opposite assignment of property rights leads to a similar analysis, and discussion is omitted.) The factory owner must buy out the rights from all landowners before he can emit smoke. Each landowner has a reservation price for permitting smoke. Efficiency implies that the smoke be emitted if and only if the pure profits of the factory owner exceed the sum of the reservation prices of the landowners. It would appear that a mutually advantageous bargain will achieve the efficient allocation.

But the factory owner does not know the reservation prices of the landowners. If the land is used for residential purposes, then the reservation price is determined by the indifference surfaces of the residents, clearly private information. If the land is used for production, then the reservation price depends on the effect of smoke on productivity, and it is an essential virtue of decentralized private enterprise that businesses do not know in detail each others' production functions. By the same token, the landowners do not know the value of smoke emission to the factory owner; and, what is at least as important, no landowner knows the reservation prices of the other landowners. (A landowner might infer the reservation price of another landowner who was visibly identical to him in circumstances, by assuming it was the same as his own; but to the extent that landowners vary in land use and location relative to the factory, their mutual knowledge of reservation prices becomes correspondingly hazy.)

Consider, then, how bargaining might proceed. The factory owner might make an offer to each landowner. Since he does not know the reservation prices of the landowners, the offer might well be rejected by some as below the reservation price. This, by itself, is only a difficulty in the adjustment process, for the factory owner has now acquired information about the reservation prices. But a more serious problem is that a landowner might reject the offer even if it is above his reservation price, to convey the idea that it is still higher, for he knows that the factory owner cannot be sure of the deception. In turn, knowing this possibility means that the factory owner cannot draw any inference from a refusal.

Suppose instead that the landowners initiate the offers. Independently of the others, each landowner states a price at which he is willing to yield his rights to clean air. Even if, and perhaps especially if, each landowner has a fairly good idea of the factory owner's profits and of the reservation prices of others, he will be tempted to set a price much higher than his true reservation price. Each one will attempt to garner for himself the entire surplus in the economy, the excess of the producer's profits over the sum of the reservation prices (assuming it is positive and large relative to any one landowner's reservation price). Indeed, precisely this situation actually occurs in land assembly, where a large plot of contiguous land is needed, and the odd shapes of some department stores and office buildings testify mutely to failures to achieve efficient resource allocation.

The second case brings out more fully the essential identity between the achievement of efficiency through bargaining and the well-known "free-rider" problem in efficient allocation of resources to public goods. In the second case, each landowner may be tempted to ask for much more than his reservation price, relying on the others not to ask too much; similarly, in the public goods case, each may be tempted to understate his benefits if his taxes are related to his statement.

This connection reminds us of the recent literature which suggests the possibility of overcoming the free-rider problem by appropriate incentives (Clarke, 1971; Groves and Loeb, 1975; Green and Laffont, 1977a,b). These incentives have been referred to as achieving "demand revelation" (the term is due to Tideman and Tullock, 1976, who give a most spirited interpretation and an argument for practical implementation).

The dominant theme in this work has been the search for strategy-proof incentive structures. The agents are supposed to send messages which reflect or purport to reflect their valuations. In addition to the initial resource allocation problem, a system of rewards and penalties as functions of the messages is created; the resulting game is supposed to be such that the message sent by an agent reflects only his valuations and is independent of his guesses as to the evaluations of others. If the preference orderings of the agents over the possible outcomes are unrestricted a priori, then indeed there is no possible mechanism which will guarantee strategy-proof behavior, as brilliantly shown by Gibbard (1973) and, independently, by Satterthwaite (1975). However, strategy-proof procedures may be possible if the orderings are restricted, for example, by the assumption that there are no income effects, so that any rewards or penalties are additive to the results of the bargaining procedure. This assumption is in fact the basis of most of the demand-revelation literature and will be made here.

(There is a close relation between the existence of strategy-proof procedures and the existence of social welfare functions. In fact, as shown by Maskin, 1976, chap. 3, the two problems are essentially equivalent, in that an a priori restriction on the agents' preferences which permits one problem to be solved also permits the other to be solved.)

The demand-revelation procedures do indeed yield efficient decisions, with, however, a curious qualification. The decision made is indeed efficient, but the rewards and penalties do not in general add up to zero; to ensure against infeasibility, it is in general necessary to permit resources to go to waste, in the sense that the sum of the penalties exceeds the sum of the rewards. Green and Laffont (1977a) have shown that this lack of balance is necessary.

I wish to suggest a different approach to demand revelation, which achieves efficiency and avoids the waste of resources in the incentive payments. As might be expected, it makes stronger assumptions, in this case assumptions about the expectations that each agent has about the others' valuations.

In the next section I will exhibit one possible formalization of the bargaining process which would lead to efficient allocation when the agents know each others' utility functions. The third section illustrates how efficiency fails to be achieved when this knowledge is absent. In the fourth section the general approach to demand revelation under incomplete information will be discussed. In the fifth section I show how this approach can be used to develop a demand-revealing game for which truthful revelation is a local equilibrium point (in the sense of Nash) and for which all the side-payments balance. The final section is supplementary; it shows that the utilitarian criterion ensures that the collective decision rule has a responsiveness property assumed in the fifth section.

A Formalization of Bargaining

We are accustomed to proofs of the optimality of allocation resulting from the market under a competitive price regime. The rules of the system in equilibrium are well defined, and the propositions clearly stated. What is meant by an assertion, whether by Coase or by von Neumann and Morgenstern, that unrestricted bargaining with well-defined property rights will lead to efficient allocation? What are the rules of the game and how do we determine its outcome?

We have to describe bargaining as a series of permitted moves and then an equilibrium concept which determines when no agent will wish to change his or her strategy. This turns out to be a difficult task, and most proposed solutions have some unsatisfactory qualities.

In game-theoretic language, the problem is to devise a noncooperative game whose equilibrium point (in the sense of Nash) is an efficient allocation of resources. Thus, each agent will have the right to change unilaterally, but the game is so devised that at some efficient allocation no agent will in fact find it preferable to change.

I confine discussion to the case of two agents, Estragon (E) and Vladimir (V). To simplify further, assume that there is only one private good and one public good (or externality-producing activity). Suppose first that the public good is resource-using. Let x_E be the amount of the private good going to Estragon, x_V the amount of the private good going to Vladimir, and y the

amount of the public good. Total initial resources are w, and there is no production. An allocation, then, is feasible if and only if

$$(16\text{-}1) \qquad x_E + x_V + y = w.$$

Assume that the two individuals can have any feasible allocation they agree to. If they do not agree, suppose that there is no public good ($y = 0$) and that each keeps his initial stock of the resource for private use, $x_E = w_E$ and $x_V = w_V$, where w_i is the initial resource holding of agent i ($i = E, V$). We shall refer to the allocation $(w_E, w_V, 0)$ as the disagreement payoff and denote it by A^d.

Let $U^i(x_i, y)$ be the utility of agent i if he receives x_i of the private good and the amount of the public good is y. An efficient resource allocation then satisfies the well-known Samuelson condition that the sum of the marginal rates of substitution of the public good for the private for the two individuals equals 1:

$$(16\text{-}2) \qquad \frac{\partial U^E}{\partial x_E} \Big/ \frac{\partial U^E}{\partial y} + \frac{\partial U^V}{\partial x_V} \Big/ \frac{\partial U^V}{\partial y} = 1.$$

Equations (16-1) and (16-2) constitute two equations in the three unknowns, x_E, x_V, y, and define the Pareto frontier. However, because of initial holdings, the contract curve is restricted in addition to the allocations which satisfy the conditions of *individual rationality,*

$$(16\text{-}3) \qquad U^i(x_i, y) \geqq U^i(w_i, 0), \qquad (i = E, V);$$

that is, no agreement will make any individual worse off than he could be without agreement. Note that the disagreement payoff A^d satisfies these conditions by definition.

Consider now the following simple procedure. Estragon proposes an allocation, that is, x_E, x_V, and y, satisfying the feasibility condition (16-1). Vladimir can accept or reject it. Suppose that Estragon knows Vladimir's utility function. Vladimir will accept if and only if the proposed allocation is individually rational from his point of view. Then clearly Estragon's optimal strategy will be to choose that allocation which maximizes his utility subject to the constraints

$$(16\text{-}4) \qquad x_E + x_V + y = w, \qquad U^V(x_V, y) \geqq U^V(w_V, 0).$$

In general, the optimal allocation to Estragon subject to (16-4) will be better for him than the disagreement point $(w_E, 0)$, and therefore he will not choose an allocation which will be rejected by Vladimir, while among those which

Vladimir will accept, Estragon will, of course, choose the best from his viewpoint.[1]

The outcome of this procedure is clearly Pareto-efficient by definition, since it maximizes the utility of one agent for a given level of utility for the other. Hence, there is a game whose equilibrium is Pareto-efficient. The outcome, it must be noted, is not the competitive equilibrium.

The Failure of Bargaining under Privacy

One of the virtues of decentralization is the respect for privacy. The utility functions of different individuals cannot easily be known to each other. Indeed, since they have meaning only in terms of observable behavior, there may be no way of transmitting utility functions from one agent to another. There is in general no way of forcing an individual to reveal his utility function; if he knows that this knowledge will be used in some allocation process, what he will transmit will or at least can be designed to affect that allocation favorably to him.

Let us analyze the outcome of the procedure of the last section with one additional complication: Estragon does not know for sure what Vladimir's utility function is. More specifically, suppose that it could be one of two utility functions, $U_a^V(x_V,y)$ or $U_b^V(x_V,y)$, and Estragon does not know which.

1. Wilson (1978) has given a procedure for any number of economic agents which will achieve efficient allocation by means of a noncooperative game; the procedure above is the special case of Wilson's when there are but two agents.

The procedure given in the text is not quite standard in game theory. The game has been described in so-called extensive form. It is usual to reduce games to normal form, in which each agent describes his potential behavior at any point for every possible history of the game up to that point. These descriptions or *strategies* are thought of as chosen simultaneously at the beginning of the game. Then the equilibrium point is a choice of strategy by each agent such that neither could gain by changing if the other player does not change. In the game described above, Vladimir's strategy, in this sense, would be a statement describing which allocations he would accept if offered by Estragon. But then if he should choose a strategy of accepting only one allocation, Estragon's optimal strategy in reply would be to make that offer if it is at least as good for him as the disagreement allocation; and given that offer, Vladimir should announce that he would accept it if it is at least as good for *him* as the disagreement allocation. Hence, any individually rational allocation could be achieved as an equilibrium in this sense.

In the text, and in the work of Wilson (1978), the last player is not allowed to formulate a strategy in advance but rather is required to accept or reject previous offers on the basis of a comparison with the disagreement allocation. This concept of equilibrium, in which the second player is optimizing given all previous history, has been termed a *perfect* equilibrium by Selten (1975); it is the game-theoretic counterpart of the principle of optimality in dynamic programming.

Let us go further, and represent his uncertainty by an assignment of probabilities; Vladimir's utility function is U_a^V with probability p_a, U_b^V with probability p_b, $p_a + p_b = 1$. These probabilities may be regarded as objective facts or as subjective probabilities of Estragon's.

Now consider Estragon's choice problem. If he chooses a feasible allocation (x_E,x_V,y) which Vladimir would accept for either utility function, then the probability of acceptance would be one. Clearly, among such allocations, Estragon would choose the best from his point of view. Hence, one possible candidate for his choice would be

$$(16\text{-}5) \qquad A^* = (x_E^*,x_V^*,y^*) \text{ maximizes } U^E(x_E,y) \text{ subject to } x_E + x_V + y = w,\; U_a^V(x_V,y) \geqq U_a^V(w_V,0),\; U_b^V(x_V,y) \geqq U_b^V(w_V,0).$$

Another possibility would be to consider a broader class of feasible allocations, those which Vladimir would choose if his utility function were U_a^V. Again, among these, Vladimir would choose that one which maximizes his utility. This allocation might happen to satisfy also the condition that $U_b^V(x_V,y) \geqq U_b^V(w_V,0)$, in which case it would be the allocation A^*. If we disregard this possibility, we have another candidate for Estragon's choice,

$$(16\text{-}6) \qquad A^a = (x_E^a,x_V^a,y^a) \text{ maximizes } U^E(x_E,y) \text{ subject to } x_E + x_V + y = w,\; U_a^V(x_V,y) \geqq U_a^V(w_V,0).$$

Note that Estragon will prefer this policy to any other which is rational for Vladimir under U_a^V because, if rejected, the payoff to Estragon is independent of the offer. Since the probability of acceptance is p_a, the expected utility of Estragon is

$$(16\text{-}7) \qquad p_a U^E(x_E^a,y^a) + p_b U^E(w_E,0).$$

Symmetrically, of course, Estragon can choose an allocation which would be acceptable to Vladimir if his utility function were U_b^V. He would choose the best:

$$(16\text{-}8) \qquad A^b = (x_E^b,x_V^b,y^b) \text{ maximizes } U^E(x_E,y) \text{ subject to } x_E + x_V + y = w,\; U_b^V(x_V,y) \geqq U_b^V(w_V,0),$$

and the expected return to him would be

$$(16\text{-}9) \qquad p_a U^E(w_E,0) + p_b U^E(x_E^b,y^b).$$

Finally, Estragon could, if it were desirable, choose an allocation which Vladimir would reject whether his utility function were U_a^V or U_b^V. But Estragon would know that the disagreement allocation would result, an allocation he could also obtain by offering it, since it would satisfy all the individual rationality conditions for Vladimir. In that case, he would do at least as well by choosing A^*, so that we can assume that Estragon will always choose an allocation which satisfies Vladimir's individual rationality condition for at least one possible utility function.

Estragon then chooses that one of the three allocations A^*, A^a, A^b which makes his expected utility as large as possible; the expected utility of A^* to him is $U^E(x_E^*,y^*)$, since there is no uncertainty in that case, while the expected utilities of A^a and A^b are given by (16-7) and (16-9), respectively. Given all values of the utilities, the choice depends on p_a, being A^a if p_a is sufficiently large, A^b if it is sufficiently small; there may also be an intermediate range in which A^* is chosen.[2]

If A^a is offered and in fact Vladimir's utility function is U_a^V, the allocation will be Pareto-efficient. But there is a probability p_b that the disagreement allocation will be the equilibrium. Hence the system cannot guarantee efficiency; indeed, p_b might easily be nontrivial and the inefficiency of the disagreement point very considerable. Similarly, if A^b is offered, there is a probability p_a of winding up at the disagreement allocation. Finally, A^* is inefficient whichever utility function Vladimir has; hence, if it is offered, there will be inefficiency with probability 1.

Thus, it can be seen that a procedure which would achieve a Pareto-efficient allocation if each agent knew the other's utility function will have a positive probability of falling short of efficiency if this knowledge is absent.

Bargaining as a Game of Incomplete Information

The game analyzed in the last section is one of incomplete information in the sense introduced in an important series of papers by Harsanyi (1967–68). It can be reduced to a game in standard form by introducing, for each

2. For those interested in the detailed results, it will be easy to verify the following statements. Define $V^a = U^E(x_E^a,y^a) - U^E(w_E,0)$, $V^b = U^E(x_E^b,y^b) - U^E(w_E,0)$, $V^* = U^E(w_E^*,y^*) - U^E(w_E,0)$. Let $\overline{V}$ be half the harmonic mean of V^a and V^b. Then there are two cases: (1) if $V^* > \overline{V}$, then Estragon chooses allocation A^a if $p_a > V^*/V^a$, A^b if $p_a < 1 - (V^*/V^b)$, and A^* in the intermediate range; (2) if $V^* < \overline{V}$, then A^* is not chosen for any value of p_a, while A^a is chosen if $p_a > V^a/(V^a + V^b)$, and A^b otherwise.

agent, a chance move which determines his utility function, the outcome of the move being revealed to him but not to other agents. (In the particular game just discussed, uncertainty about Estragon's utility function is irrelevant, since Vladimir has to respond passively to Estragon's proposal; but in general, allocation games will be more symmetric.) The probabilities of different utility functions are part of the rules of the game and are known to both (or, more generally, all) agents. It might be asked why Vladimir, who comes to know his own utility function, has to know the ex ante probabilities; in the particular game in fact he does not have to know, but more generally, he should know what probabilities about his utility functions are in Estragon's mind.

The question to be posed is the following: Can we find a set of rules, a game, in which each agent is to announce his or her utility function, the allocation of resources is a function of the announcements, and the rewards are such as to induce each to announce truly? The answer to be demonstrated here is that such a game can be devised if (1) income effects are neglected, and (2) the probabilities of different possible utility functions for the agents are known and are independent of each other.[3]

It should be remarked that there has been a shift from revelation of indifference maps to revelation of utility functions. The reason is that once probabilities are introduced, the aim of each agent in the bargaining game is to maximize *expected* utility according to the conventional Bernoulli–Ramsey–von Neumann-Morgenstern hypothesis. The results will not be invariant under monotone transformations of the utility functions, since such transformations alter attitudes toward risk bearing.

The absence of income effects means that the utility function is linear in income (and therefore, in particular, implies risk neutrality toward income). This is a serious limitation, though one that is common to the entire demand-revelation literature.

More specifically, we model the making of a public decision, a variable x in some domain. (The decision might include the allocation of private goods, but it is the choice of public goods and externalities that is most in mind here.) Agent i has a utility function for the decision; the form of this

3. After working out the ideas to be presented, I found that the representation of incentives for satisfactory bargaining as a game of incomplete information had already been used by d'Aspremont and Gerard-Varet (1975). The ideas are essentially the same; their utility functions, arising out of a pollution problem, are more specialized, and this fact permits some simplification.

function is known to all up to some parameters which are private. Let ρ_i be the parameters of the ith agent; his utility function then is $U^i(x,\rho_i)$. Individual i has a net increment of income t_i (which may be negative); then his utility is $U^i(x,\rho_i) + t_i$. Since there could be infinitely many parameters, the restriction of the utility functions U^i to a class of known form in unknown parameters is no real restriction.

Each individual makes an announcement of his parameter values, which may or may not be true. For example, in a pollution problem, the manufacturer may announce the unit cost of antipollution measures to him while the neighbor announces the damage due to unit pollution. If the issue is whether or not to build a bridge, each agent may be asked to announce the value to him of the bridge. Let r_i be the announcement by agent i.

There is specified a rule for making the decision as a function of the announcements, $x(r_1, \ldots, r_n)$, where n is the number of agents. At the moment we leave the rule unspecified; it might be derived on the hypothesis of maximizing the sum of the agents' utilities on the assumption that their announcements are true, a case that will be considered in the last section.

Each agent i has a probability distribution over the true parameter values of other agents. Let ρ^i be the parameters of all individuals other than individual i: $\rho^i = (\rho_1, \ldots, \rho_{i-1}, \rho_{i+1}, \ldots, \rho_n)$. Suppose the ith agent assumes that everyone else will tell the truth (he does not, of course, know what those parameters are but knows their distribution). If he announces r_i, then the social decision is $x(r_i,\rho^i)$. In the absence of any money transfers, his utility will be $U^i(x(r_i,\rho^i),\rho_i)$; but, since he does not know ρ^i, he would be induced to maximize

$$(16\text{-}10) \qquad E_{\rho^i}[U^i(x(r_i,\rho^i),\rho_i)],$$

where, as the notation indicates, the expectation is taken over ρ^i. One would like to see that $r_i = \rho_i$ for all ρ_i; that is, we would like to have individual i be induced to reveal his private parameters truthfully if everyone else did so. If this held for all i, truth telling would be an equilibrium point of the game.

There is no reason so far for truth telling, as we well know from the literature on the free-rider problem in the theory of public goods. The demand-revelation literature suggests that we modify the payoff to individual i by assessing a change in income as a function of his announcement. Specifically, we will prescribe a function $T_i(r_i)$ for each individual, which is the amount he pays if he makes the announcement r_i. This income has to go somewhere; it would be inefficient to throw it out if positive, and there is no source to supply it if negative. Hence, we add a specification that the income dispensed by one agent is paid out to all other agents. That is, we specify functions, $T_{ij}(r_j)$, the amount paid by individual j to individual i if j's announcement is r_j. These functions are defined only if $i \neq j$, of course. By definition,

$$(16\text{-}11) \qquad \sum_{i \neq j} T_{ij}(r_j) = T_j(r_j).$$

From the viewpoint of individual i, then, his utility is decreased by his payments, $T_i(r_i)$, and increased by the payments made to him by others, $\Sigma_{j \neq i} T_{ij}(r_j)$. Therefore, once the functions T_i, T_{ij} satisfying (16-11) are selected, and each agent chooses his announcement r_i, the net payoff to individual i is

$$(16\text{-}12) \qquad U^i(x(r),\rho_i) - T_i(r_i) + \sum_{j \neq i} T_{ij}(r_j),$$

where $r = (r_1, \ldots, r_n)$. Each agent must choose an announcement r_i for each possible value of the parameters ρ_i; that is, a *strategy* for i is a function $r_i(\rho_i)$. Once the strategies are selected by the agents, the expected utility to agent i is the expected value of (16-12) over the parameters of other agents, ρ^i. Let r^i be the announcements of agents other than i, and $r^i(\rho^i)$ the strategies of agents other than i. Then agent i's expected utility is

$$(16\text{-}13) \qquad E_{\rho^i}\{U^i[x(r_i(\rho_i),r^i(\rho^i),\rho_i]\} - T_i(r_i(\rho_i)) + E_{\rho^i}\left[\sum_{j \neq i} T_{ij}(r_j(\rho_j))\right].$$

An equilibrium point is a specification of a strategy $r_i(\rho_i)$ for each agent, such that if every agent $i \neq j$ chooses $r_j(\rho_j)$, then (16-13) is maximized for agent i by the choice of the strategy $r_i(\rho_i)$. This is the same as saying that, for each particular value of ρ_i, the corresponding value of the strategy $r_i(\rho_i)$ maximizes (16-13).

The Demand-Revelation Game

A truth-telling strategy is given by $r_i(\rho_i) \equiv \rho_i$ (that is, identically in ρ_i). A truth-telling equilibrium is one in which each agent is playing a truth-telling strategy. Our problem is to find functions T_i, T_{ij} so that truth telling is an equilibrium.

Define $F^i(r_i,\rho_i)$ to be the expected payoff to agent i, when all other agents are telling the truth. This is defined by setting $r_j(\rho_j) \equiv \rho_j$ in (16-13); for a given ρ_i, we also replace $r_i(\rho_i)$ by r_i:

$$(16\text{-}14) \qquad F^i(r_i,\rho_i) = E_{\rho^i}[U^i(x(r_i,\rho^i),\rho_i)] - T_i(r_i) + E_{\rho^i}\left[\sum_{j \neq i} T_{ij}(\rho_j)\right].$$

At an equilibrium, agent i chooses r_i to maximize $F^i(r_i,\rho_i)$. Assume now differentiability of all relevant functions, and also assume that the optimum choice is not a boundary. Also, for simplicity, assume that ρ_i, and therefore r_i, are one-dimensional variables; this restriction is solely expository and can be removed with only notational changes. Then r_i satisfies the condition $\partial F^i/\partial r_i = 0$. This condition in fact defines r_i for each value of ρ_i and therefore defines the strategy $r_i(\rho_i)$. Since the last term in (16-14) is independent of r_i, optimal behavior for agent i is defined by

$$\text{(16-15)} \qquad E_{\rho^i}[U^i_x(x(r_i,\rho^i),\rho_i)x_{r_i}(r_i,\rho^i)] - T'_i(r_i) = 0.$$

To have a truth-telling equilibrium, the solution of (16-15) in r_i should be ρ_i for all ρ_i. That is, (16-15) should hold with r_i replaced by ρ_i for all ρ_i:

$$T'_i(\rho_i) = E_{\rho^i}[U^i_x(x(\rho_i,\rho^i),\rho_i)x_{r_i}(\rho_i,\rho^i)]$$

for all ρ_i, or, if we replace the variable ρ_i by the variable r_i,

$$\text{(16-16)} \qquad T'_i(r_i) = E_{\rho^i}[U^i_x(x(r_i,\rho^i),r_i)x_{r_i}(r_i,\rho^i)]$$

for all r_i.

Equation (16-16) supplies a complete solution of the problem, if one exists. It defines the function $T_i(r_i)$, up to an irrelevant constant of integration. The right-hand side is well defined by the conditions of the problem. The transfer function T_{ij} does not affect the behavior of any individual and therefore can be chosen arbitrarily, subject to the condition that the transfer to all other individuals equals the payment required; see (16-11).

The right-hand side has a simple interpretation. If the true value of the parameter ρ_i is in fact equal to the announcement, then a perturbation dr_i in the announcement will change the social decision by $x_{r_i}dr_i$ and will therefore change the individual's utility by $U^i_x x_{r_i} dr_i$. This product depends, however, on the true values of other individuals' parameters ρ^i; hence, from agent i's point of view, he must consider the expected marginal utility. The payment function T_i has to be fixed to just offset any resulting incentive to change from r_i to $r_i + dr_i$.

Thus, each agent will find that truth telling satisfies the first-order conditions for a maximum. We must, however, check to see that the second-order conditions are also satisfied.

The derivative $\partial F^i/\partial r_i = F^i_{r_i}(r_i,\rho_i)$ is given by the left-hand side of (16-15). By our construction of T_i, we know that $F^i_{r_i}(r_i,\rho_i) = 0$ when $r_i = \rho_i$, that is,

$$F^i_{r_i}(r_i,r_i) \equiv 0$$

identically in r_i. Total differentiation with respect to r_i yields

$$\text{(16-17)} \qquad F^i_{r_ir_i}(r_i,r_i) + F^i_{r_i\rho_i}(r_i,r_i) \equiv 0.$$

In order that truth telling be optimum for individual i, given that all others are telling the truth, it is necessary that the first- and second-order conditions for a maximum in r_i be satisfied at $r_i = \rho_i$, that is, that $F^i_{r_i}(\rho_i,\rho_i) = 0$, which is guaranteed by construction, and that $F^i_{r_ir_i}(\rho_i,\rho_i) < 0$. Since these conditions must hold for all ρ_i, we can replace ρ_i by r_i and require that they hold for all r_i. From (16-17), then, we require that $F^i_{r_i\rho_i}(r_i,r_i) > 0$ for all r_i.

In (16-14), note that the second and third terms are independent of ρ_i. If we differentiate (16-14) with respect to ρ_i and that derivative with respect to r_i and set $\rho_i = r_i$, we find that the second-order condition for individual i to have a truth-telling optimum is

$$E_{\rho i}[U^i_{x\rho_i}(x(r_i,\rho^i),r_i)x_{r_i}(r_i,\rho^i)] > 0.$$

This means that, on the average, for any given values of the utility function parameters of the others, if an individual is telling the truth, a shift in his parameter changes the amount of the public good, through the collective decision rule, and the marginal utility of the public good to him in the same direction. This is a minimum condition for the collective decision rule to represent positively the desires of the economic agents.

We can restate our results formally.

DEFINITION 1. *A collective decision rule is a function which determines the amount of the public good for any specification of the utility function parameters of the economic agents.*

DEFINITION 2. *A collective decision rule is said to be responsive if, for any given individual, a shift in his parameter changes on the average his marginal utility for the public good and the amount of the public good in the same direction. More precisely, the expected value of the product of these changes, averaged over the utility function parameters of others, is positive.*

THEOREM 1. *Suppose*

(a) *there are no income effects so that agent i has a utility function in the public good and other income of the form* $U^i(x,\rho_i)$ + *(other income);*

(b) *the form of the utility function is known to all, but the value of* ρ_i *is known only to the agent i;*

(c) *each agent has a probability distribution over the utility function parameters of others which is independent of his own parameter value;*

(d) *the collective decision rule is responsive.*

Consider the following social decision procedure. Each individual chooses a strategy which associates to each possible value of his parameter an announced value of the parameter. The public good quantity is determined by the collective decision rule as a function of the announced values of the parameters. Finally, there are specified functions $T_{ij}(r_j)$ $(i \neq j)$ which specify amount of income paid by agent j to agent i as a function of j's announcement r_j of his utility parameter. These functions satisfy the condition

$$(16\text{-}18) \qquad \sum_{i \neq j} T_{ij}(r_j) = \int E_{\rho^j}[U^j_x(x(r_j,\rho^j),r_j)x_{r_j}(r_j,\rho^j)]dr_j.$$

Then the truth-telling strategies, $r_i(\rho_i) \equiv \rho_i$, form a local equilibrium point of the social decision procedure.

Remark 1. By a local equilibrium point is meant that, given that others tell the truth, agent i will find that for each value of ρ_i, the action $r_i = \rho_i$ is a local maximum. Examination of the second-order conditions cannot by itself show that the action is globally optimal for the agent.

Remark 2. Condition (16-18) is simply a restatement of (16-16), with i replaced by j and condition (16-11) substituted in. The constant of integration in (16-18) can be chosen arbitrarily.

The Utilitarian Decision Rule

I have so far only required that the collective decision rule be responsive. A natural condition is that the collective decision rule be defined by the condition that the amount of the public good be such as to maximize the sum of individuals' utilities. This rule is especially plausible if we maintain the assumption that utility is linear in money, for then the sum of utilities is the money value of net benefits. This rule will in fact imply that the collective decision rule is responsive if the utility functions are concave in the public good.

We are requiring that $x(r_1, \ldots, r_n)$ be defined as the value of x which maximizes the sum of utilities on the assumption that announcements are true:

$$(16\text{-}19) \qquad x(r) \text{ maximizes } \sum_{i=1}^{n} U^i(x,r_i),$$

where $r = (r_1, \ldots, r_n)$. Then $x(r)$ satisfies the identity in r,

$$\sum_{i=1}^{n} U^i_x(x(r),r_i) \equiv 0.$$

Partial differentiation with respect to r_i yields

$$\left[\sum_{i=1}^{n} U^i_{xx}(x(r),r_i)\right] x_{r_i}(r) + U^i_{xr_i}(x(r),r_i) \equiv 0.$$

Multiply through by $x_{r_i}(r)$:

$$\left[\sum_{i=1}^{n} U^i_{xx}(x(r),r_i)\right] [(x_{r_i}(r)]^2 + U^i_{xr_i}(x(r),x_{r_i}(r) \equiv 0.$$

But, because U^i is assumed concave in x for each i, the first term is in general negative, so that the second term must be positive. But this certainly implies that x is a responsive collective choice rule.[4]

THEOREM 2. *If the collective choice rule x(r) is defined so as to maximize the sum of the agents' utilities, then it is responsive.*

References

Arrow, K. J., and Scitovsky, T., eds. (1969). *Readings in Welfare Economics.* Homewood, Ill.: Richard D. Irwin.

Bator, F. (1958). The Anatomy of Market Failure, *Quarterly Journal of Economics* 72, 351–379.

Baumol, W. J. (1952). *Welfare Economics and the Theory of the State.* Cambridge, Mass.: Harvard University Press.

Clarke, E. (1971). Multipart Pricing of Public Goods, *Public Choice* 11, 17–33.

Coase, R. H. (1960). The Problem of Social Cost, *Journal of Law and Economics,* 3, 1–44.

d'Aspremont, C., and Gerard-Varet, L. (1975). Incentives and Incomplete Information. CORE discussion paper (unpublished).

Furubotn, E., and Pejovich, S. (1972). Property Rights and Economic Theory: A Survey of Recent Literature, *Journal of Economic Literature* 10, 1137–1162.

Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result, *Econometrica* 41, 587–602.

Green, J., and Laffont, J. (1977a). Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods, *Econometrica* 45, 427–438.

Green, J., and Laffont, J. (1977b). Révélation des préférences pour les biens publics, *Cahiers du Séminaire d'Econometrie* 19, 83–103.

Groves, T., and Loeb, M. (1975). Incentives and Public Outputs, *Journal of Public Finance* 4, 211–226.

4. Laffont and Maskin (1978), in some unpublished work, have shown that the utilitarian rule actually implies that truth telling is a global equilibrium point in this game; that is, if everyone else is telling the truth, then truth telling is a global maximum strategy. They have also shown the close connections between the revelation mechanism developed here and the Groves mechanism and its generalization.

Harsanyi, J. (1967–68). Games with Incomplete Information Played by "Bayesian" Players, I–III, *Management Science* 14, 159–182, 320–334, 486–502.
Knight, F. H. (1924). Some Fallacies in the Interpretation of Social Cost, *Quarterly Journal of Economics* 38, 582–606. (Reprinted in Arrow and Scitovsky, 1969, pp. 213–227.)
Laffont, J., and Maskin, E. (1978). A Differential Approach to Expected Utility Maximizing Mechanisms, Notes, Groupe de Travail, Economie de l'Information, Ecole Polytechnique, Paris (unpublished).
Luce, R. D., and Raiffa, H. (1957). *Games and Decisions.* New York: Wiley.
Maskin, E. (1976). Social Choice on Restricted Domains, Ph.D. dissertation in Applied Mathematics, Harvard University.
Pigou, A. C. (1952). *The Economics of Welfare,* 4th ed. London: Macmillan.
Satterthwaite, M. A. (1975). Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions, *Journal of Economic Theory* 10, 187–217.
Scitovsky, T. (1951). *Welfare and Competition.* Chicago, Ill.: Richard D. Irwin.
Scitovsky, T. (1954). Two Concepts of External Economies, *Journal of Political Economy* 62, 143–151. (Reprinted in Arrow and Scitovsky, 1969, pp. 242–252.)
Selten, R. (1975). Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games, *International Journal of Game Theory* 4, 25–55.
Tideman, N., and Tullock, G. (1976). A New and Superior Process for Making Social Choices, *Journal of Political Economy* 84, 1145–1160.
von Neumann, J., and Morgenstern, O. (1947). *Theory of Games and Economic Behavior,* 2nd ed. Princeton, N.J.: Princeton University Press.
Wilson, R. (1978). A Competitive Model of Exchange, *Econometrica* 46, 577–585.
Young, A. (1913). Pigou's *Wealth and Welfare, Quarterly Journal of Economics* 27, 672–686.

17 Allocation of Resources in Large Teams

For a number of years I have given a course on the economics of information and organization, both at Harvard and at Stanford. One year I examined in detail the work of Roy Radner on teams with quadratic payoffs and normal disturbances, under alternative assumptions about the conditions of communication. I realized that his results had strong implications for the value of communication, for his work showed that the differences between two communication systems, one clearly stronger than the other, vanished if the number of units became large. There were some other implications that were unacceptable. I developed in class a new model that avoided some of the difficulties in Radner's. At this point Radner came to visit for a year at Harvard, and we jointly developed the chapter that follows.

One of the oldest themes in economics is the use of the market in the coordination of widespread and diverse but interdependent activities. The underlying situation may be taken to be one of optimal resource allocation, that is, the maximum achievement of some objective subject to constraints on the resources used. Under suitable hypotheses of convexity and differentiability, it is well known that optimality conditions can be stated in terms of

This chapter was written with Roy Radner. Reprinted from *Econometrica,* 47 (1979):361–385.