Journal of Public Economics 5 (1976) 131–138. © North-Holland Publishing Company

REGULATION OF EXTERNALITIES

Theodore C. BERGSTROM*

Washington University, St. Louis, MO 63130, U.S.A.

Received October 1973, revised version received March 1975

The problem of efficient allocation in the presence of externalities can be usefully viewed as having two parts. Standards for aggregate levels of externalities must be determined. From among many possible patterns of individual activities which achieve the same standards, some single allocation must be selected. In the first part of this paper, we argue that with a proper specification of property rights, decentralized markets can be used to efficiently meet specified standards. More centralization seems to be required in the determination of standards. Possible methods of determining and enforcing aggregate levels of externalities will be considered in the latter portion of the paper.

The formal models treated here do not incorporate production explicitly. That the introduction of production does not alter any of our results significantly is demonstrated in Bergstrom (1973a). Consider an economy with n consumers. Each consumer i has preferences represented by a utility function $u_i(x_i, y_i, z)$ with domain $C_i \subset E^m \times E^k \times E^k$. The vector x_i is an m-vector of private goods made available to consumer i. The vector y_i is the k-vector of nonprivate activities pursued by consumer i, and the vector z is the sum of the vectors of nonprivate activities by all consumers in the economy.

If, for example, consumer i takes pleasure in releasing pollutant j but views increases in the total amount of pollutant j in the environment as unpleasant, then u_i is a monotone increasing function of y_{ij} and a monotone decreasing function of z_j, where y_{ij} and z_j represent quantities of this pollutant. Alternatively, consumer i may regard the performance of task k as unpleasant, but

*This paper is based on research undertaken while the author was a visitor at CORE in Louvain, Belgium. The research was carried out within the project 'Analyse économique de la lutte contre la pollution des eaux', under contract between the Université Catholique de Louvain and the Belgium Ministry of Scientific Policy; the latter project is itself part of the 'Premier programme national de Recherche et de développement sur l'environnement physique et biologique', administered by Service de la Politique et de la Programmation Scientifiques. Many of the ideas in the paper evolved from discussions with Professor Henry Tulkens of CORE. The errors are mostly originals.

prefers that the total extent to which task k is pursued by members of the community be large rather than small. Then, where y_{ik} and z_k are measures of the performance of this task, u_i will be monotone decreasing in y_{ik} and monotone increasing in z_k.

Consider an economy in which there is no production of private goods but there is a fixed stock represented by a nonnegative vector $w \in E^m$ which may feasibly be divided to give any nonnegative allocation vector $x = (x_1, \ldots, x_n) \in E^{mn}$ such that $\sum x_i = w$. Suppose that in such an economy a central authority decides to enforce a specified level $\mathring{z}$ for the sum of all vectors of nonprivate activities. So long as z is fixed at $\mathring{z}$, each u_i depends only on x_i and y_i. The feasible allocations for which $\sum y_i = \mathring{z}$ are just those allocations such that $\sum x_i = w$ and $\sum y_i = \mathring{z}$. Such allocations will be called *$\mathring{z}$-feasible.*

From the characterization of $\mathring{z}$-feasible allocations and from the fact that with fixed $\mathring{z}$ each consumer's utility depends only on his own vectors of private consumption and of nonprivate activities, it is apparent that there is an isomorphism between the problem of choosing from among alternative $\mathring{z}$-feasible allocations and the much studied problem of allocation in a pure exchange economy with only private goods. Corresponding to the notions of a Pareto optimum for an exchange economy, we define a *conditionally efficient allocation relative to* $\mathring{z}$ to be an allocation which is $\mathring{z}$-feasible and to which no other $\mathring{z}$-feasible allocation is Pareto superior.

The isomorphism to an exchange economy can be further developed if we define and interpret vectors of 'initial holdings' not only of private goods but also of nonprivate activities. In particular, let $(\mathring{x}, \mathring{y}) = (\mathring{x}_1, \ldots, \mathring{x}_n, \mathring{y}_1, \ldots, \mathring{y}_n)$ be specified so that $\sum \mathring{x}_i = w$ and $\sum \mathring{y}_i = \mathring{z}$. The vector $(\mathring{x}_i, \mathring{y}_i)$ will be called the initial holdings of consumer i. The economic interpretation of initial holdings vectors in nonprivate activities will be developed after we introduce two more definitions.

Corresponding to competitive equilibrium in exchange theory we define a *market equilibrium relative to* $(\mathring{x}, \mathring{y})$ to be an allocation $(\bar{x}, \bar{y})$ and price vectors $\bar{p}$ and $\bar{q}$ such that: (i) for all i. $(\bar{x}_i, \bar{y}_i)$ maximizes $u_i(x_i, y_i, \mathring{z})$ subject to $(x_i, y_i, \mathring{z}) \in C_i$ and $\bar{p}x_i + \bar{q}y_i \leqq \bar{p}\mathring{x}_i + \bar{q}\mathring{y}_i$; (ii) $\sum y_i = \sum \mathring{y}_i = \mathring{z}$ and $\sum x_i = \sum \mathring{x}_i = w$.

Corresponding to the core in exchange theory, we define the *core relative to* $(\mathring{x}, \mathring{y})$ to be the set of $\mathring{z}$-feasible allocations (x, y) such that there exists no coalition $K \subset \{1, \ldots, n\}$ which can use its own resources to improve on (x, y). By this we mean that for no coalition $K \subset \{1, \ldots, n\}$ does there exist a collection of vectors (x_i', y_i') for each $i \in K$ such that $\sum_K x_i' = \sum_K \mathring{x}_i$, $\sum_K y_i' = \sum_K \mathring{y}_i'$ and $u_i(x_i', y_i', \mathring{z}) \geqq u_i(x_i, y_i, \mathring{z})$ for all $i \in K$ with strict inequality for some $i \in K$. We are now able to interpret the initial holdings $\mathring{y}_i$ of nonprivate activities. If, for example, the jth nonprivate activity is the emission of a pollutant, we can imagine an institutional arrangement in which anyone who is to emit y_{ij} units of this pollutant must hold a total of y_{ij} 'pollution tickets'. Consumer i initially holds $\mathring{y}_{ij}$ tickets, where $\sum_i \mathring{y}_{ij} = \mathring{z}_j$. No new tickets may be printed but consumers are allowed to trade these tickets just as they are allowed to trade from their initial holdings of private goods. Thus a system of property rights in polluting activities is established and enforced in much the same way as with property rights in private goods.

If, say, the kth nonprivate activity represents provision of a public service which, though privately odious in its performance, is socially useful, then initial holdings $\mathring{y}_{ik}$ in the kth nonprivate activity represent obligations to perform this activity at the level $\mathring{y}_{ik}$. Again these obligations may be bought and sold (typically at a negative price). But to sell an obligation at a negative price is to pay someone else to perform the obligation. Thus in market equilibrium such initial holdings represent obligations to bear a certain share of the cost of inducing the performance of activity k at level $\mathring{z}_k$.

The definition of a *core relative to* $(\mathring{x}, \mathring{y})$ must be interpreted to mean that strategies which are independently available to a coalition are restricted to those in which its members jointly meet the social obligations and restrictions implicit in their initial holdings of rights and obligations. This seems a quite natural extension of the role of property rights in the usual exchange model.

In view of the isomorphism between the model of this paper and the usual exchange economy, we can draw on well-known results to conclude that if preferences are convex, locally nonsatiated, and represented by continuous utility functions, then, subject to a few familiar technical assumptions, a market equilibrium exists relative to $(\mathring{x}, \mathring{y})$; such a market equilibrium is conditionally efficient relative to $\mathring{z} \equiv \sum \mathring{y}$ and is in the core relative to $(\mathring{x}, \mathring{y})$; and every allocation which is conditionally efficient relative to $\mathring{z}$ is a market equilibrium relative to some $(\mathring{x}, \mathring{y})$ where $\sum \mathring{x}_i = w$ and $\sum \mathring{y}_i = \mathring{z}$. If also there are many similar consumers (in a certain well-defined sense), then any allocation which is in the core relative to $(\mathring{x}, \mathring{y})$ is also 'nearly' a market equilibrium relative to $(\mathring{x}, \mathring{y})$. Furthermore where there are 'many' consumers these results remain largely intact, even if the convexity assumption is weakened to allow nonconvexities of preference which are 'not too large'.[1]

These results enable one to make a case that the goal of efficient implementation of a specified level of nonprivate activities would be reasonably well served by a central authority whose primary role is to establish and enforce a system of property rights with respect to nonprivate activities as well as private goods, while remaining largely passive in other respects.

The selection of a level of nonprivate activities to be enforced is less amenable to decentralized decision making. While a single standard $\mathring{z}$ must be determined, the utilities of many consumers are simultaneously influenced by this decision. Somehow the divergent interests of different consumers must be resolved. One possibility is to view the choice of $\mathring{z}$ as the outcome of a political process, such as majority voting. This approach, though perhaps of considerable descriptive

[1]Such results for large but finite economies can be found in Bergstrom (1973b) and Arrow and Hahn (1971).

value, will not guarantee selection of efficient allocations. Here we confine our attention to decision methods which lead to at least approximate efficiency.

Consider an economy in which there is a central authority which markets property rights in nonprivate activities. The authority may either sell new pollution rights or buy existing pollution rights and destroy them. Similarly, it may sell or buy obligations to perform socially desired activities. Individuals may make transactions either with each other or with the central authority. Where initially held property rights in nonprivate activities total $\mathring{z} = \sum \mathring{y}_i$ and where the sum of all such rights held after the transactions of the central authority are performed is z, the total revenue of the central authority is $q(z-\mathring{z})$, where q is the vector of market prices for rights in nonprivate activities. Total revenue of the central authority is divided among individuals. The amount received by consumer i is determined by a function $s_i(q, z-\mathring{z})$, where

$$\sum_{i=1}^{n} s_i(q, z-\mathring{z}) = q(z-\mathring{z}).$$

Each consumer then has a budget constraint: $px_i+qy_i \leqq p\mathring{x}_i+q\mathring{y}_i+s_i(q, z-\mathring{z})$. He could determine those values of x_i, y_i and z which maximize $u_i(x_i, y_i, z)$ subject to this constraint. In general, different individuals would choose different values for z. However, for suitably chosen functions $s_i(q, z-\mathring{z})$, it can be shown that prices exist under which all demands are mutually consistent. This is the case in Lindahl equilibrium.

A Lindahl equilibrium is an allocation $(\bar{x}, \bar{y}, \bar{z}) = (\bar{x}_1, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_n, \bar{z})$ such that for some price vectors $\bar{p}, \bar{q}_1, \ldots, \bar{q}_n$ and $\bar{q}$, where $\bar{q} = \sum \bar{q}_i$: (i) for all i, $(\bar{x}_i, \bar{y}_i, \bar{z})$ maximizes $u_i(x_i, y_i, z)$ subject to $\bar{p}x_i+\bar{q}y_i \leqq \bar{p}\mathring{x}_i+\bar{q}\mathring{y}_i+\bar{q}_i(z-\sum \mathring{y}_i)$; (ii) $\sum \bar{x}_i = w$ and $\sum \bar{y}_i = \bar{z}$.

Here unanimous agreement is achieved on the vector $\bar{z}$ when there is an appropriate division of the revenue and costs from the marketing activities of the central authority. In a previous paper, Bergstrom (1970), I have proved that Lindahl equilibrium exists under very general circumstances, the only really restrictive assumption being convexity of preferences. Also, Lindahl equilibrium is Pareto optimal and any Pareto optimum is a Lindahl equilibrium for some allocation of initial holdings. Where there are some unpleasant externalities, Lindahl equilibrium will be in the alpha core only if the core is rather specially defined. This situation is explained in Bergstrom (1975a). However, Lindahl equilibrium will always be in the *core relative to* $(\mathring{x}, \mathring{y})$ as defined above.

In subsequent discussion, we will view Lindahl equilibrium as a possible objective for central planning. As a starting point alternative to complete decentralization, let us consider a highly idealized version of totally centralized planning. Suppose that there is a central authority with unlimited ability to gather information at no cost, with powers of computation enabling it to find maxima whenever they exists, and with the ability to enforce any feasible activities on the part of individual agents. Such a central authority might 'simply' gather full information on all preferences and technical capabilities, and compute the set of all Pareto optimal allocations together with the corresponding requisite activities for individual agents. The authority must use some criterion to choose a particular Pareto optimum and then enforce individual activities leading to that outcome. At almost every step in this process assumptions are made which are quite obviously untenable. We will examine some of the difficulties and consider possible ways of coping with them.

The most direct means of gaining information about preferences and technical capabilities would be simply to ask all individuals and firms to tell the central authority all about their preferences and technologies. Here there are two important difficulties. In the first place, it is unlikely that individuals will be so introspective or firms so prescient as to be able to decide without a great deal of costly effort how they would behave in choice situations which are very different from those in their recent experience. At best we might hope that an individual could tell rather easily what he would choose in a budget situtation which is not very different from those of his recent experience. Similarly, without expensive research efforts, it may be that a firm could reasonably estimate its profit maximizing response only to small changes in factor prices. A second difficulty is that if the individual agents know that information about their preferences or technologies will be used in determining the policy of the central authority, it will be in the interest of any individual to make that statement which affects the central authority's policy in the way which is most favorable to him. In these circumstances there is no reason to suppose that the statement which he chooses to make is a true description of his preferences or of his technology. For example, the well-known free rider problem in the 'public goods' literature consists of the observation that if an individual is to be taxed for the support of a public good at a rate equal to his stated marginal rate of substitution for that good, it is generally in his interest to understate his true marginal rate of substitution.

The extreme difficulty of obtaining detailed information about the preferences and technical capabilities of each individual and firm in the economy suggests that we seek ways to obtain and employ rough approximations. It has been suggested by Bergstrom (1975a) and Kurz (1974) that, if there is sufficient regularity of preferences and of technology in a large economy, useful inferences about preferences or technology of reasonably homogeneous groups of individuals and firms may be obtained from detailed study of samples drawn from the population at large. If the sample constitutes a small fraction of the entire population, then at a cost which is small relative to the size of the economy, substantial amounts of resources could be used to induce each sampled individual to reveal information carefully and honestly. Sometimes it may be possible to choose a sample which, though representative of a population which will be affected by the information obtained, will itself not be directly influenced by the uses of that information. For example, citizens of one municipality might be asked to reveal their preferences for local public goods, with the knowledge that although this information will be used to influence tax and expenditure policies in other cities, their responses will not influence policies in their home community. Thus there would be little direct incentive to mislead the interviewer.

Even where a large amount of resources is available to reward truthful answers, it remains a difficult problem to find ways of rewarding truth telling if there is no independent knowledge of what is true. One possibility would be to use elaborate devices of cross-examination to detect the inconsistencies into which prevarication is likely to lead. Rewards for participating in the sample would presumably depend on how 'consistent' one's answers are. Another interesting strategy for eliciting honest responses is to arrange rewards so that sampled individuals are engaged in what Schelling (1963) calls a game of 'tacit coordination'. This possibility is discussed in detail in Bergstrom (1975b). Other possibilities are discussed by Kurz (1974).

We construct a procedural paradigm which a central authority might follow in guiding an economy toward an efficient allocation using limited information. For the time being, assume convexity of preferences and of production sets. Recalling our earlier discussion, we know that there exists a Lindahl equilibrium. Though such an equilibrium is Pareto optimal and has the core property, we are not assured that decentralized forces will lead to a Lindahl equilibrium. Suppose that the central authority employs a research laboratory which gathers information from a sample of the population as suggested above and uses this information to estimate economy-wide aggregate excess demand functions for nonprivate activities and for private goods. These estimated demand functions are then used to estimate Lindahl equilibrium prices and quantities. If the estimates of the central authority were precisely accurate, the central authority could simply issue property rights for the indicated consumptions of private goods and levels of nonprivate activities and announce the computed equilibrium price vector. The resulting situation would turn out to be a market equilibrium and no changes would result from voluntary decentralized activity. But, of course, we cannot expect complete accuracy from the central authority. It therefore becomes very important that the information gained by the central authority be employed in such a way as to allow corrections to be readily made and to ensure that damage caused by inaccurate results be relatively small and not too unevenly distributed.

To this end, we suggest the following procedure. To facilitate discussion, suppose that there are several private goods, but that the only nonprivate activity is emission of a single pollutant. Recalling the definition of Lindahl equilibrium, the laboratory would compute an allocation $(\bar{x}_1, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_n, \bar{z})$, and prices $\bar{p}, \bar{q}_1, \ldots, \bar{q}_n$, and $\bar{q} = \sum \bar{q}_i$ such that if its estimates were correct, then: (i) for all i, $(\bar{x}_i, \bar{y}_i, \bar{z})$ maximizes $u_i(x_i, y_i, z)$ subject to $\bar{p}x_i + \bar{q}y_i \leqq \bar{p}\mathring{x}_i + \bar{q}\mathring{y}_i + \bar{q}_i(z - \sum_{i=1}^n \mathring{y}_i)$; (ii) $\sum \bar{x}_i = w$ and $\sum \bar{y}_i = \bar{z}$.

Observe that if the total quantity of pollution is fixed at $\bar{z}$, the budget constraint of each consumer i is equivalent to $\bar{p}x_i+\bar{q}y_i < \bar{p}\mathring{x}_i+\bar{q}\hat{y}_i$, where

$$\hat{y}_i = y_i+\frac{\bar{q}_i}{\bar{q}}\left(\bar{z}-\sum_{i=1}^{n}\mathring{y}_i\right).$$

Thus it is apparent from our earlier discussion that the allocation $(\bar{x}, \bar{y})$ and the price vector $(\bar{p}, \bar{q})$ constitute a market equilibrium relative to the initial holdings vector $(\mathring{x}, \hat{y}) = (\mathring{x}_1, \ldots, \mathring{x}_n, \hat{y}_1, \ldots, \hat{y}_n)$. In fact, if the laboratory had been entirely accurate in its computations, the central authority could simply issue pollution tickets in the quantity $\hat{y}_i$ to each i, leave initial ownership of private goods unchanged and announce that with this distribution of ownership, the price system $(\bar{p}, \bar{q})$ supports a conditional market equilibrium. If one believes that unrestricted bargaining leads to an allocation close to a market equilibrium in the resulting exchange economy, then he would expect the allocation $(\bar{x}_1, \ldots, \bar{x}_n, \bar{y}_1, \ldots, \bar{y}_n, \bar{z})$ to emerge without further central interference.

Where there is likelihood of error in the laboratory's calculations, the case for so restricting the interference of the central authority seems very attractive. If the economy can be expected to select a conditional market equilibrium allocation, then even if the estimates of the central authority are inaccurate, the resulting allocation will be conditionally efficient subject to the choice $\bar{z} = \sum_{i=1}^{n} \hat{y}_i$. The cost of erroneous computation of equilibrium values will be limited to the cost of using the wrong value for z. In case a proper choice of z is made, the outcome will be Pareto optimal. Distributional inequities due to imprecise estimates of equilibrium values will be limited in the sense that individuals retain ownership of their initial holdings of private goods and suffer 'only' from an improper allocation of pollution tickets. Of course, if $\bar{z} < \sum_{i=1}^{n} \mathring{y}_i$, then there is nothing to insure that $\hat{y}_i$ will be nonnegative. Negative values of $\hat{y}_i$ must be construed as reducing the wealth of consumer i by $\bar{q}\hat{y}_i$. If a consumer initially holds little or no rights to pollute and is incorrectly supposed to place a high value on the reduction of pollution, then $\bar{q}\hat{y}_i$ may represent a large 'undeserved' reduction in his income.

When the market has had some time to adjust to the change in endowments of pollution imposed by the central authority, it is time for the laboratory to renew its calculations. This is true in part because the laboratory may be able to improve its predictions of individual demand behavior from observations of actual behavior in the market, and in part because individual agents may have better information about their own preferences and technology as a result of actual experience with alternative prices, wealth and pollution levels. In a dynamic setting, of course, capital accumulation, population change, and technical progress provide additional reasons for continual revision.

This paradigm suggests the usefulness of allowing an artificial 'laboratory economy' to interact sequentially with the real economy, where information generated by each economy is used to guide the other. The informational and computational demands of the schemes are, of course, formidable even where relatively modest demands of precision are made. Recent progress with algorithms for computing equilibrium values for systems of equations [see, for example, Scarf (1973)] would suggest that equilibrium could be approximated for reasonably complicated systems of demand equations.

References

Arrow, K. and F. Hahn, 1971, General competitive analysis (Holden Day, San Francisco).
Bergstrom, T., 1970, Collective choice and the Lindahl allocation method, mimeo (Washington University, St. Louis, MO).
Bergstrom, T., 1973a, The use of markets to control pollution, Recherches Economique de Louvain.
Bergstrom, T., 1973b, Equilibrium in a finite economy with bounded non-convexity, mimeo (Washington University, St. Louis, MO).
Bergstrom, T., 1975a, The core when strategies are restricted by law, Review of Economic Studies, forthcoming.
Bergstrom, T., 1975b, On the free rider problem, in: S.Y. Lin, ed., Economics of Externalities, forthcoming.
Kurz, M., 1974, Experimental approach to the determination of demand for public goods, Journal of Public Economics 3, 329–348.
Scarf, H., 1973, The computation of economic equilibrium (Yale University Press, New Haven).
Schelling, T., 1963, The strategy of conflict (Oxford University Press, New York).