JOURNAL OF ECONOMIC THEORY 4, 180–199 (1972)

Fundamental Nonconvexities in the Theory of Externalities*

DAVID A. STARRETT

Harvard University, Cambridge, Massachusetts 02138

Received February 19, 1971

It has long been recognized that externalities among economic agents will lead to inefficiency of decentralized competitive allocation. To correct these inefficiencies, various modifications of the competitive system have been suggested; one scheme pioneered by Meade (1952) would set up a system of artificial markets for externalities (like smoke, nectar, etc.), while the other first suggested by Pigou (1920) would impose a system of taxes on polluters or subsides on pullutees (with an analogous structure for external economies). As is well known, the first system suffers from a thiness of markets [typically there will be only one buyer and one seller of an externality; see Arrow (1969)], while the second requires information which the market alone does not provide, so neither one could work without some government administration. But assuming that they are administered properly, the difference between them in economic effect is that in the first case, any payments made by a polluter are automatically received by a corresponding pollutee while in the second, they are not. (If pollutees are compensated in the tax scheme, there is no difference.)

In the cases discussed so far, the pollutee has the "right" to a clean environment in the sense that he must be compensated for any pollution foisted upon him. While our common-law upbringing makes this scheme sound reasonable, other payment systems could also be used. In particular, the pollutee could pay polluter to clean up his pollution. We can refer to this plan as a "polluter's rights" market scheme. And of course, there would be a corresponding plan for subsidizing polluters.

All of these schemes would seem to satisfy the same set of necessary conditions. (The arguments here are essentially those of Coase, 1960).

* I have written a less general and rigorous but more policy-oriented version of this paper in collaboration with Richard Zeckhauser (1971). Readers who get bogged down in the theorems are referred to that paper. I also benefitted from discussions with Kenneth Arrow, Peter Diamond, and Leo Hurwicz. This research was supported in part by NSF grant GS 2797.

In each case, the marginal benefit from pollution (or any other externality) is equated to the sum of the marginal losses of those affected. For the tax or subsidy schemes, this is done by setting the tax or subsidy equal to the sum of marginal losses and letting the profit-maximizing polluter equate his margin to the price. For the market-type schemes the polluter participates in a market for "pollution rights" with each pollutee. The net price he pays (or receives) is the sum of the prices on all these markets, so the same efficiency conditions will be satisfied.

Since the Samuelson necessary conditions for efficiency are satisfied by all of our schemes, we should be able to apply a standard theorem of economic analysis to assert that they are equivalent up to a redistribution of income. However, we should not accept this result so readily. Redistribution of income among households is one thing; redistribution among firms is something else entirely. To get from the market scheme with pollutees rights to the one with polluter's rights clearly requires a transfer of income from the pollutee to the polluter (we shall see below just how much must be transferred). But the competitive firm (in long-run equilibrium with free entry) operates with no pure profits—no economic rent, so if the pollutee is a competitor, we should expect him to go out of business when income is transferred away. For our schemes to be really equivalent, it is necessary that there be considerable economic rent to the firm.

The purpose of the present paper is to analyze the apparent paradox just raised. In the first section, we show that *under the usual convexity assumptions*, there is indeed considerable economic rent in the system—enough to guarantee the equivalence of all our allocation schemes. In the following section we show that for a large class of externality problems, this result breaks down because of logical nonconvexities in production sets. These nonconvexities and their implications for handling the externality problem are discussed in depth. To avoid the additional complication of income effects and minimize notational requirements, we will restrict the discussion to a production model. The type of externality situations we have in mind are: water pollution affecting industrial water use and air or noise pollution affecting worker productivity. For a brief discussion of consumption externalities see Section 4.

1. An Equivalence Theorem

We borrow directly Arrow's abstract treatment of externalities (see Arrow, 1969); the only major restriction of this formulation is that all externalities must be associated with measurable commodities. This is not as restrictive as it sounds, since we allow new (possibly artificial) commodities to be defined as needed (for example, nectar, in the bees-and-honey case discussed by Meade). We will also restrict ourselves to a static framework with no uncertainty.

Let x_{jk} be the net output of the k-th commodity by the j-th firm. This list must include all variables of the firm which have a potential external effect on others. In general the production possibilities of firm j will depend not only on his own net outputs but on those of everyone else as well. Hence, the production possibilities can be represented by sets $X^j(x_{11}, x_{12}, ..., x_{mn})$, where m is the number of potential firms, n the number of potential commodities[1], and j an index over potential firms (n and m finite). The economy-wide net output set (Z) is then the direct sum of the X^j:

$$Z = \left\{ z_k \,\middle|\, z_k = \sum_j x_{jk}, \ x_{jk} \in X^j, \text{ all } j \right\}.$$

Without prejudging consumer preferences, we wish to characterize the undominated points of Z in terms of supporting price and tax systems. The standard analysis without externalities proceeds from here by examining and characterizing maximal points in Z. This would be inappropriate here, since some of the commodities in our list are not "goods" but "bads" and some may be of ambiguous value depending on who you are. It is important to realize that we are operating in an extended commodity space, one in which any production variable which imposes an externality on households constitutes a separate dimension.

While we wouldn't expect maximality, we might at least want to be on the boundary of Z. Otherwise (assuming convexity which we shall be in this section), we are interior in Z and could move at least slightly in *any* direction of net-output space; intuitively, an incremental move which provides more of goods and less of bads could be distributed so as to make someone better off without making anyone worse off.[2] In the absence of externalities, such a distribution is guaranteed by the innocuous assumption of individual nonsatiation; here a nonsatiation assumption will also work, but it is not nearly so innocuous. Let us call an interior point of Z which is Pareto efficient, a point of *local Pareto satiation.*

[1] The definition of a commodity is somewhat subtle; two items are the same commodity if they are perfect substitutes in consumption. If households can tell the difference between the products of different firms, then there is no distinction between sets Z and X.

[2] Of course, if the externality variables such as smoke are left out of the commodity space, we would have no such expectation.

Unfortunately, local Pareto satiation is possible even if (taking the behavior of others as given) no individual consumer can ever be satiated. As an example, think of the "keep up with the Jones" effect. Each individual would be happier consuming more *ceteris paribus*, but if everyone consumes more, then everyone feels worse off (or at least no better) because of jealousies. Somewhere before this point is reached, will be a point of local Pareto satiation. Intuitively, if externalities in consumption are of second order importance compared to the direct benefits from consumption, local Pareto satiation (LPS) could be ruled out. Probably some theorems could be worked out in this direction, but we will not pursue the matter further here. It is worth noting one condition that trivially rules out LPS: the existence of a commodity which everyone finds desirable and with which no one associates any consumption externalities (labor could perhaps serve as such a commodity, although critics would probably argue that people are jealous of their industrious neighbors). Clearly, it would also be sufficient to assume that there are no externalities in consumption.

Rather than carry the axiomatization back further, we assume:

P.1 *The economy admits no point of* LPS.[3]

Given P.1, we can restrict attention to the boundary points of Z. If Z is convex, then each boundary point is supported by a "budget" hyperplane, and we can proceed to derive efficient price and tax/subsidy schemes in a standard way. Toward this end we assume *in this section only*

P.2 X^j *is convex, each* j.

(Note that P.2 implies Z convex.) The standard procedure for getting from supporting hyperplanes on Z to profit-maximizing price systems is to exploit the independence properties of firm's production sets. It is precisely because, with externalities, these sets are not independent that the "new welfare economics" breaks down in this case. However, Arrow (1969) has shown that one can artificially induce independence of production sets and still derive a certain type of price systems. Here we will follow his treatment closely, although without using calculus.

Suppose we distinguished commodities not only according to which firm produces them but also by which firm is being affected (we will use the letter y to denote net outputs so distinguished). Thus y_{ijk} now stands for the net-output of commodity k by the firm j as observed by firm i.

[3] If we are willing to assume full convexity of preference, the word local can be replaced by the word global in this postulate.

Let Y^i denote the production set of firm i, defined on variables y^i_{jk}.[4] Now if we ignore the obvious requirements: $y^i_{jk} = y^s_{jk}$, all i, s, production sets become independent, since the level of y^i_{jk} cannot affect the possibilities of any firm except i. For future reference, notice that this independence is not bought cheaply; indeed, to obtain it, we have to define new commodities, one for each potential external effect. (y_{ijk}, $i \neq k$ represents the external effect on i from j's production of k.)

We now define a series of net-output sets. The significance of each will be explained after they are all defined.

$$Z = \left\{ z_k \,\middle|\, z_k = \sum_j x_{jk}\,,\; y_{ijk} = x_{jk}\,,\; y^i_{jk} \in Y^i,\ \text{all } i \right\},$$

$$X = \{x_{jk} \mid y_{ijk} = x_{jk}\,,\; y^i_{jk} \in Y^i,\ \text{all } i\},$$

$$Y = \{y_{ijk} \mid y^i_{jk} \in Y^i,\ \text{all } i\},$$

$$\bar{Y}^i = \{y^{ii}_k \mid (y^{ii}_1, \ldots, y^{ii}_n, \ldots, \bar{y}^i_{jk}, \ldots) \in Y^i\},$$

where $\bar{y}$ is some reference program.

Z is as before except that the new artificial index has been incorporated. X is the nm-dimensional economy-wide net-output set with commodities distinguished according to the firm which produced them. Y is the nm^2-dimensional net-output set distinguished also according to which firm is being affected, and without feasibility requirements. Because of the induced independence of production sets, $Y = \chi_i\, Y^i$. $\bar{Y}^i$ is the decision set for firm i, given that all other firms operate according to a reference program $\bar{y}_{ijk}$. Clearly,

$$\bar{Y}^i = Y \cap \{y_{stk} \mid y_{stk} = \bar{y}_{stk}\,,\ \text{if } s \neq i \text{ or } t \neq i\} \equiv Y \cap \bar{\Lambda}^i.$$

As we shall show below, supporting prices on the sets Y^i define a competitive equilibrium with "artificial markets" while prices on the sets $\bar{Y}^i$ provide a set of optimal taxes or subsidies.

[4] A word on notation: large letters will always stand for sets, small letters, for elements of those sets. Subscripts will always represent matrix indices while superscripts are identifiers of the variable to which they are attached. For net output variables with three indices, the first will always refer to the firm which this variable affects; the second, to the firm which produced it; and the third, to the specific commodity involved. For example, $y_i{}^j{}_k$ is a vector of net outputs by firm j, the first index (i) running over the firm affected by this variable, and the last (k) running over commodities.

PROPOSITION 1. *Assume* P.1 *and* P.2. *Then* $\bar{z}$ *is boundary in* Z *if and only if there exists a price vector* q (*non-zero, but not necessarily non-negative!*) *such that*

$$\sum_k q_k z_k \leqslant \sum_k q_k \bar{z}_k, \text{ all } z \in Z. \tag{1}$$

Proof. The "if" part is trivial; the "only if" part is a direct application of the supporting hyperplane theorem. q is a vector of social opportunity costs, reflecting rates of social transformation between commodities. When (1) is satisfied, we say that q supports $\bar{z}$ in Z.

PROPOSITION 2. q *supports* $\bar{z}$ *in* Z, *if and only if*

$$\{p_{jk} \mid p_{jk} = q_k, \text{ all } j\} \text{ supports } x_{jk} \text{ in } X. \tag{2}$$

Proof. Suppose (2) does not hold. Then there exists x_{jk} in X such that

$$\sum_{jk} p_{jk} x_{jk} > \sum_{jk} p_{jk} \bar{x}_{jk}.$$

But then,

$$\sum_k q_k \sum_j x_{jk} > \sum_k q_k \sum_j \bar{x}_{jk}$$

which directly contradicts (1). The converse is proved by reading backwards.

Define a new set:

$$\Omega = \{y_{ijk} \mid y_{ijk} = y_{sjk}, \text{ all } s \text{ and each } j, k\}.$$

Note that Ω is a linear subspace of our grand Euclidean space and also that $X = Y \cap \Omega$. In our next proposition we want to take the price system (2) on $Y \cap \Omega$, and extend it to all of Y. Fortunately, this is exactly the situation for which the Hahn–Banach Theorem applies[5]; but to use it, we need an interior point condition. Hence, the following lemma:

LEMMA 1. *If* $\bar{x}_{jk}$ *is interior in* X, *then* $\bar{y}_{ijk}$ *is interior in* Y (*where* $\bar{y}_{ijk} = \bar{x}_{jk}$, *all* i).

Proof. Suppose $\bar{y}_{ijk}$ is boundary. Then there is a ω_{ijk} with

$$\bar{y}_{ijk} + t\omega_{ijk} \notin Y,$$

any $t > 0$. Now, by the independence property of Y, this implies

[5] For a discussion of the Hahn–Banach Theorem and the version of it we are using here, see the Appendix.

$\bar{y}^i_{jk} + t\omega^i_{jk} \notin Y^i$ for some i. Now, letting $\omega_{jk} = \omega^i_{jk}$, we argue that $\bar{x}_{jk} + t\omega_{jk} \notin X$ for any $t > 0$, since otherwise $\bar{x}_{jk} + t\omega_{jk}$ must be possible for every firm including (certainly) the i-th. This contradiction establishes the lemma.

PROPOSITION 3. *Given* P.1 *and* P.2, *q supports $\bar{z}$ in Z if and only if there are price vectors s^i_{jk} (not all zero) which support $\bar{y}^i_{jk}$ in Y^i (each i) and satisfy the conditions,*

$$\sum_i s_{ijk} = q_k \text{ (all } j, k). \tag{3}$$

Proof. By Lemma 1 any interior point of X is interior in Y[6], so the Hahn–Banach Theorem guarantees that the functional p_{jk} (of Proposition 2) supporting $\bar{x}_{jk}$ in X can be extended to a functional s_{ijk} supporting $\bar{y}_{ijk}$ in Y. By the independence property of Y, s^i_{jk} supports $\bar{y}^i_{jk}$ in Y^i (each i). Furthermore, the functional s_{ijk} agrees with p_{jk} on the subspace spanned by X. Hence,

$$\sum_{ijk} s_{ijk} x_{jk} = \sum_{jk} p_{jk} x_{jk}, \text{ all } x_{jk} \text{ or}$$

$$\sum_{jk} x_{jk} \left(\sum_i s_{ijk}\right) = \sum_{jk} p_{jk} x_{jk}, \text{ all } x_{jk}.$$

Sequentially plugging in basis elements, we obtain

$$\sum_i s_{ijk} = p_{jk} = q_k \text{ (each } j, k).$$

Furthermore, if s^i_{jk} supports $\bar{y}^i_{jk}$ (each i), then conditions (3) guarantee that p_{jk} supports $\bar{x}_{jk}$ in X. Proposition 3 now follows from Proposition 2. ||

COROLLARY 1. *Given* P.1 *and* P.2, *$\bar{z}$ is an "efficient" production allocation if and only if there are price vectors s^i_{jk} (not all zero) which support $\bar{y}^i_{jk}$ in y^i (each i) and satisfy* (3) *for some final demand price vector q.*

Proof. The corollary follows immediately from Propositions 1 and 3 (and P.1). ||

We will now argue that, given q, the matrix s_{ijk} defines *both* a competitive equilibrium in the artificial markets scheme *and* a system of

[6] If X does not have interior, the theorem can still be proved by reducing the dimensionality of the space to the smallest subspace containing X; X (being convex) *must* have interior in this space.

supporting taxes. Suppose that we have established artificial markets in the externalities y_{ijk} $(i \neq j)$, in addition to the traditional markets for y_{iik}. Suppose further that prices s_{ijk} have been proposed for these markets. Then, assuming that the firm takes price as given, Proposition 3 tells us that each firm i finds $\bar{y}^{i}{}_{jk}$ profit-maximizing. Clearing of all markets is guaranteed by conditions $\bar{y}_{ijk} = \bar{x}_{jk}$, all i. Firm i actually participates on several markets for $y^{ii}{}_{k}$. Before selling to consumers, he may have to buy rights from affected parties; the total price paid to others for this purpose is given by $\sum_{t \neq i} s_t{}^i{}_k$. Since the net price is $s^{ii}{}_k$, the effective price to consumers must be $s^{ii}{}_k + \sum_{t \neq i} s_t{}^i{}_k$. Hence, (3) assures us that q_k is the vector of final demand prices, as required.

To derive a set of supporting taxes, recall that $\bar{Y}^i = Y \cap \bar{A}^i$. Now we apply another simple principle: If a convex set is supported by a hyperplane and we restrict attention to a subspace containing the point of support, then within that subspace, the restricted hyperplane still supports the restricted set. Therefore, $\bar{y}^{ii}{}_k$ is supported in Y^i at "prices" $s^{ii}{}_k$. This means that if each firm assumes that everyone else will follow the program $\bar{y}$ and chooses his decision variables only (the $\bar{y}^{ii}{}_k$), then $\bar{y}^{ii}{}_k$ is profit maximizing for him at net prices, $s^{ii}{}_k$. If q_k are prices to consumers, net prices $s^{ii}{}_k$ are achieved by taxing firm i at the rate $\sum_{t \neq i} s_t{}^i{}_k$, since from (3):

$$s^{ii}{}_k = p_{ik} - \sum_{t \neq i} s_t{}^i{}_k = q_k - \sum_{t \neq i} s_t{}^i{}_k . \tag{4}$$

Note that the net price of good k may differ from firm to firm. It will be the same for a pair of firms only if in their production of k, each confers the same external effects on third parties.

Our equivalence is now complete: the tax which firm i pays on good k is exactly equal to the total payments i would have made to all others firms on the externality markets for k. The only difference between the two schemes is that recipients of externalities are compensated in the artificial markets framework while with taxes, they are not; with full convexity, we have shown that this payment is a pure rent item to the firm. However, one further point is worth making. In the artificial markets scheme, the government obtains information and calls out prices, but never plays a direct role on any market; but in the tax scheme, the government receives taxes and pays subsidies directly. In the latter case, there is no guarantee that the government budget will balance automatically. (Indeed, if all externalities are diseconomies, it will have revenues but no expenditures.) Thus, a tax plan implicitly assumes that lump sum taxes or subsidies can be used to balance the government budget. If lump sum taxes are infeasible, then a tax scheme will be "second best," and the analysis is quite different [see Diamond and Mirrlees (1971)].

Notice that Proposition 3 guarantees only $s_{ijk} \neq 0$, which certainly implies $s^i{}_{jk} \neq 0$ (some i) but does not imply $s^i{}_{jk} \neq 0$ (all i). Thus, from what we know so far, it is possible that some firms in the artificial markets scheme face zero prices. Furthermore, even if $s^i{}_{jk} \neq 0$ (all i), it doesn't follow that $s^{ii}{}_k \neq 0$ (all i); so even if all firms face nonzero prices in the artificial markets plan, they need not in the tax/subsidy scheme.

Unfortunately, we cannot in general rule out the case of zero prices for some firms. The reason is that it is sometimes Pareto efficient to have a firm operating in the interior of its production set, and no nonzero set of prices can support that situation. (This possibility was first pointed out, and an example given, by Murakami and Negishi, 1964.) Operating at an interior point can be easily ruled out in the absence of externalities by assuming that each firm produces something desirable to consumers; but if the desirable goods produced give off external diseconomies as by-products, a stronger assumption is clearly needed. Further work on this aspect of the problem is in progress. [See Starrett (1970) for some preliminary results.]

So far, our schemes have been entirely of the pollutee's rights type. But we can handle the polluter's rights cases with no additional effort! We can convert the polluttee's rights market system into a pulluter's rights system by a simple translation of the commodity space. Instead of measuring the zero level of pollution by its physical zero level, we could make the zero level any arbitrary benchmark. For example, suppose that z is the pollution variable and N the benchmark. Then, by the new measuring stick, the pollution level is $z - N$; if z is less than N, this is a negative number, otherwise positive. In our general notation, we associate a benchmark $N_{ijk}(i \neq j)$ for each external effect and measure the level of the external effect from this benchmark.

Now, translations in the commodity space do not affect the equilibrium derived above in any significant way.[7] The only effect of such a translation is to subtract a constant ($p_z N$) from the profits of *every* production possibility for the pollutee and add the same constant to profits of the polluter. "Every" is the crucial word in the previous sentence: The profits differential is the same for any firm (say the pollutee) whether it decides to operate or shut down; if it shuts down, then under the first scheme, it

[7] This is a special case of the general proposition that maxima and minima are covariant under linear transformations. Samuelson ("Foundations of Economic Analysis," Harvard University Press, Cambridge, Mass., 1947) used this same proposition to argue that economic theory is not sensitive to the way in which we measure commodities.

receives $p_z N$ while under the second, it receives nothing. On the other hand, if it operates with pollution level z, it gets private profit π plus $p_z z$ under scheme one and $\pi + p_z(z - N)$ under scheme two; the profit differential is the same in both cases $(p_z N)$.

Since a transfer which is completely independent of behavior cannot affect profit maximizating choice in any way, any price quantity combination which is an equilibrium under scheme one must also be an equilibrium under scheme two. The second scheme should be interpreted as a polluter's rights scheme; the polluter has a right to pollute at the benchmark level (N); if he pollutes less, he must be compensated for the implied cleanup, while if he pollutes more, he must pay for the privilege. The benchmark levels (N_{ijk}) are clearly somewhat arbitrary, being determined by a "political" decision as to what pollution levels polluters are entitled to.

Thus, under the assumption of full convexity, all of the allocation schemes we have discussed are equivalent in the sense that given a vector of final demand prices, any allocation which is an equilibrium in one of the schemes is also an equilibrium in any of the others. This is true even though the actual profits earned by the various firms will differ markedly from case to case.

The fact that profits will differ suggests that the assumption of full convexity must logically imply the presence of considerable economic rent in the system. We already saw an example of this rent in our study of the tax system, where it was shown that any compensation paid to pollutees constituted a pure economic rent. Actually, the size of the rent may be much larger. Indeed, the profits in any artificial markets equilibrium must always be larger than what could be earned by shutting down production and "selling" the benchmark externality rights. But now something is clearly wrong since the benchmarks are quite arbitrary and could be set as large as we like, which would seem to imply that economic rent is arbitrarily large. The culprit turns out to be our assumption of full convexity, so we turn now to a discussion of nonconvexities.

2. Fundamental Nonconvexities

To illustrate the basic nonconvexity, consider a single product firm affected by an externality, and plot its output (b) as a function of the externality (z), holding all inputs at fixed levels. The resulting function must be downward sloping, since z is a diseconomy; but it must also be nonnegative, since the firm could always choose to produce nothing regardless of the externality level. Such a curve clearly cannot be concave over the entire z axis. Two possible shapes for the curve are shown in

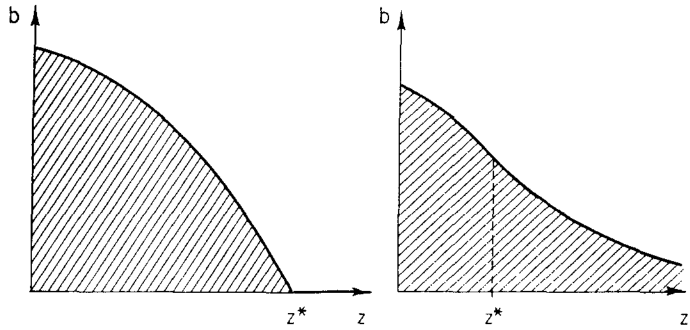

FIGURE 1

Fig. 1. In case (1) marginal losses to z increase right up to the point where the firm quites, after which they are zero; in case (2) they increase up to a point, and then decrease toward zero as the affected firm becomes "saturated" in z.

Once nonconvexities are introduced, the arguments of Section 1 naturally break down. Indeed, we can easily see that there will be (in general) no equilibria on artificial markets. At any price for externality rights, the pollutee will seek to sell an arbitrarily large number of rights (since this will increase his profits without bound). Intuitively speaking, the firm is being offered a positive price for accepting externalities; since the externality can do him at most finite damage, and since (under the assumption of competitive behavior) he believes that he can purchase whatever amounts he likes at the quoted prices, he has (theoretically) an infinite supply of rights.

However, if the price is zero, he will supply no rights, while the polluter will want to buy some positive number. Here, we have a classic case of nonexistence of equilibrium in the presence of nonconvexities.

Actually, we have rejected the artificial markets schemes a little too quickly; there may be logical or practical limits to the amount of pollution (or other externality) which society can produce in a given year. If so, then we can effectively limit the commodity space to this feasible region. In terms of markets, this would mean that a pollutee could not offer more than a certain number (say, N) of externality rights in a given year. Now, if N is less than z^* in Fig. 1 (whichever case applies), then all sets are convex in the region of feasibility, and results of Section 1 apply. (Of course, those results may also apply in cases of external economies.) In this case, all the allocation schemes are equivalent as long as the benchmark (N) is below the feasibility limit (z^*). With this restriction, all externality payments are pure rent items to both parties. Of course, in the long run, the presence of economic rent will cause entry and exist from the industries involved. For a discussion and analysis of long-run equilibrium conditions in the presence of externalities, see Starrett (1971).

Optimal tax-subsidy schemes may still exist even when there are no artificial-market equilibria. In the tax scheme, a firm faces prices for its decision variables only, and makes it schoices, taking the behavior of others as given; it never bids directly for "externalities." The decision set Y^i which (by *a priori* arguments only) must be nonconvex is replaced by the decision set $\bar{Y}^i$ which may still be convex.

But will the "as if" decision set $\bar{Y}^i$ be convex? There is at least one reason for believing it won't. Suppose that we accept the idea of constant returns overall. To the decentralized firm, this means that if he doubles all his inputs, and all externality levels double, then he will be able to double all his outputs. But if this is so, then by doubling his inputs, with externality levels fixed, he should be able to more than double his outputs (when the externalities are "bads"). Clearly, such a possibility would deny even "as if" convexity.

There are two rebuttals to this argument. First, we might restrict the analysis to the short run in which case there will be many "fixed goods" to the firm as well as fixed bads. It seems reasonable to assume that the effect of diminishing returns to fixed goods will outweight the increasing returns to fixed bads. Indeed, convexity of the $\bar{Y}_i$ in the short run seems quite acceptable.

A second argument can be given for the long run situation. We argue that most external diseconomies in production behave like a "public bads" at a point of space. The familiar examples of air, water, and noise pollution as well as physical congestion, all satisfy this condition. I feel that this observation is no fluke. In cases where the externality is appropriable, as in the case of physical garbage, there are incentives for private enterprise to engage in disposal operations which insulate firms from any detrimental effects.

But if externalities are public bads, doubling the size of the firm by replication "doubles" the external effect, though the amount of the externality produced is unchanged. Measuring externalities by their production yardstick, therefore, implies constant returns to fixed bads in the long run. This assumes that there are no additional reasons for believing in increasing or decreasing returns. In the sequel, we assume also that there are no additional reasons for nonconvexities in the $\bar{Y}^i$:

P.3 *$\bar{Y}^i$ is convex.*

As an example of a production set satisfying P.3, consider the following:

$$(-a, b) \in \bar{Y}^i \quad \text{if} \quad b \leqslant a/(\bar{z} + 1) \quad \text{and} \quad a, b \geqslant 0.$$

Y^i exhibits a smooth nonconvexity of the second type, but $\bar{Y}^i$ is convex and in fact exhibits constant returns to scale for each $\bar{z}$. (See Fig. 2.)

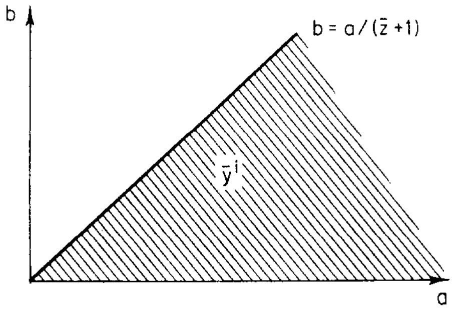

FIGURE 2

It is worth noting that if we believe in *constant* returns to fixed bads the artificial markets scheme will not work no matter where we set the benchmark externality levels. Intuitively, the argument is as follows: With constant returns and competitive pricing on all private goods, the externality recipient will just break even in the long run, assuming that he makes no net payments on externality "markets". But he always has the option of making positive profits by going out of business and "selling" some externality rights (as long as the price of rights is positive). Since the more he sells the better, he will want to sell right up to any feasibility limit we set; hence, the artificial market will always be unstable. We make these arguments precise in the following proposition.

PROPOSITION 4. *Suppose $\bar{Y}^i$ is a convex cone for every reference program and each i. Assume in addition that y^{ivw} represents an external diseconomy variable to firm i, of which firm v will produce positive amounts unless taxed. Suppose further that we set benchmark externality levels outside the upper feasibility limits as outlined in the 4-th paragraph of this section. Then there exists no long run equilibrium on artificial markets.*

Proof. Suppose that $\hat{y}_{ijk}$, $\hat{s}_{ijk}$ constitute an equilibrium pair of quantity and price vectors. First, we argue that $\hat{s}^{ivw} > 0$ at equilibrium[8] since otherwise firm i will demand $y^{ivw} = 0$ while firm v will supply $y^{ivw} > 0$. Now we argue that firm i's equilibrium demand will always be $y^{ivw} = N^{ivw}$:[9]

[8] With our orientation, increases in y^{ivw} damage firm i and require positive compensation; hence $s^{ivw} \geqslant 0$.

[9] We have chosen an orientation in which pollution is measured by its production yardstick. Therefore, in spite of the fact that pollution acts as a "net input" to the pollutee, it is measured positively.

The equilibrium point must always be more profitable than the position with firm i out of business and externality levels at N_{ijk}. Writing down this condition, we have

$$\sum_{j,k} \hat{y}^i_{jk}\hat{s}^i_{jk} \geqslant \sum_{\substack{k\\ j\neq i}} N^i_{jk}\hat{s}^i_{jk}$$

or

$$\sum_{k} \hat{y}^{ii}_{k}\hat{s}^{ii}_{k} \geqslant \sum_{\substack{k\\ j\neq i}} (N^i_{jk}-\hat{y}^i_{jk})\,\hat{s}^i_{jk} \tag{5}$$

Now the left side of (5) is zero under our assumption of constant returns to fixed bads. Hence, since every term in the sum on the right hand side of (5) is nonnegative (by feasibility of $\hat{y}_{ijk}$), every such term must be zero. It follows that $\hat{y}^{ivw}$ must be equal to N^{ivw} since we already know that $\hat{s}^{ivw} > 0$. But N^{ivw} has been constructed so as to be infeasible on the supply side. Therefore, no equilibrium can exist on artificial markets. ||

Clearly one implication of Proposition 4 is that given constant returns to fixed bads, there can be no region of full convexity in the externality variables (such a region would correspond to the interval(s) $(0, z^*)$ in Fig. 1). This observation seems interesting in its own right, as it does not seem intuitively obvious that there should be any connection between the shape of $\bar{Y}^i$ and convexity of Y^i.

3. Finding Optimal Taxes

Assuming that the "myopic" decision sets $\bar{Y}^i$ are convex, an efficient set of taxes will generally exist for any given vector of final demand prices. Given a price vector (q), the necessary conditions for efficiency (3) must naturally be satisfied, where the s_{ijk}'s are now interpreted as the marginal gains and losses of the respective firms. Assuming that we know all marginal losses associated with the efficient program, taxes can be computed from (4). Now if firms face these taxes and maximize profits (under the assumption that others will behave according to the optimal program), the necessary conditions for profits-maximization are just (3) (each firm equates its marginal benefits for each commodity to the net price), so that marginal benefits must be those of the efficient program. Thus, the efficient program (for given q) satisfies necessary conditions for individual profits maximization in the proposed tax scheme. But if the sets $\bar{Y}^i$ are convex, the necessary conditions are sufficient for profits-maximization, so that the efficient program could in fact be chosen at the quoted net prices.

While it is true that every efficient program is supported by a set of taxes (given P.3), it is unfortunately *not* true that every tax equilibrium supports an efficient program. This may seem surprising since an artificial markets equilibrium is always efficient if it exists. Competitive equilibrium is efficient with or without convexity; we can show this in the present context by observing that the "if" part of Corollary 1 does not depend on P.2.

This result does not carry over to the tax scheme. Even though every firm maximizes *globally* within its decision set, the tax equilibrium need not be even a local optimum, much less a global one. This is most easily seen in a simple example: (For convenience, we drop the net output orientation in this example; a will be measured positively as an input rather than negatively as a "net output".)

Two firms produce the same output (b), using the same input (a) but the output of one is externally detrimental to the other. The production possibly sets are defined as follows:

$$\text{firm 1: } b^1 \leqslant 2a^1,\ b^1,\ a^1 \geqslant 0,$$

$$\text{firm 2: } b^2 \leqslant a^2/(b^1 + 1). \quad b^2,\ a^2 \geqslant 0.$$

Note that our example satisfies the conditions of constant returns to fixed bads, and smooth nonconvexity to variable bads.

Suppose we have 7/2 units of a; then our efficiency problem is simply:

$$\max\ b^1 + b^2 \text{ subject to } a^1 + a^2 \leqslant 7/2 \quad \text{(and nonnegativity).}$$

Given "as if" convexity, the necessary conditions for an interior maximum will define a set of equilibrium taxes. But given global nonconvexity, this solution will be a local minimum. To find a solution satisfying necessary conditions, we set up the Lagrangian problem (with q_a the price of a in terms of b), and solve using calculus. The resulting solution is $a^1 = \frac{1}{2}$, $b^1 = 1$, $a^2 = 3$, $b^2 = 3/2$, $q_a = \frac{1}{2}$, yielding total output of 5/2. Clearly we could do better than this by shutting down the second firm whereupon the first firm could produce $b^1 = 7$ with the available input.

But the calculus solution does have a set of supporting taxes. Firm 2 will not be taxed, since it imposes no externalities; hence it faces consumer prices. In our previous notation, $s^{22}{}_a = \frac{1}{2}$, $s^{22}{}_b = 1$, at which prices firm 2 finds the solution $a^2 = 3$, $b^2 = 3/2$ profits-maximizing (given $b^1 = 1$). The output of firm 1 must be taxed at the rate of marginal losses to firm 2. These are given by

$$s^{21}{}_b = \partial b^2/\partial b^1 = -a^2/(b^1 + 1)^2 = -3/4.$$

Hence, firm 1 faces net prices of $s^{11}{}_a = \frac{1}{2}$ and $s^{11}{}_b = 1 - s^{21}{}_b = 1/4$. And at these prices, he finds the solution $a^1 = \frac{1}{2}$, $b^1 = 1$ profits-maximizing.

One can object to this tax equilibrium on the grounds that it is not stable. But this is beside the point, since there is a third stable and inefficient equilibrium with firm 1 out of business, $a^2 = b^2 = 7/2$, and the tax rate equal to 7/2. In either case our tax equilibrium corresponds to an interior point of the feasible production set. Note that this set (which corresponds to Z in the general analysis) is convex in the example, being simply the possibilities of firm 1. Thus, even convexity of Z would not be sufficient to guarantee efficiency of the tax equilibrium.

Of course, the true optimum also satisfies the set of necessary conditions and can be supported by a set of taxes. With firm 2 out of business, there there are no marginal losses due to externalities, so the tax rate is zero. The price of a rises to two, the marginal product of a in firm 1. At these prices firm 2 would not want to enter business; so the new solution is an equilibrium.

Our example illustrates a general principle: Given the nonconvexity inherent in situations involving external diseconomies, there will be multiple tax equilibria; obviously, only one of these equilibria is optimal. This fact was certainly recognized by Pigou, but it has been neglected in the recent literature, since most authors assume full convexity, and under that assumption, the global optimum is the only local optimum.

How are we to know which taxes are optimal? Of course, if we know the optimal solution to begin with, we would automatically know the appropriate taxes; but this is not very helpful, since if we knew the optimal allocation, we could simply legislate it. Ordinarily, we don't know the efficient allocation and hope to use the tax mechanism to find it. Presumably, we would use an iterative scheme of the following sort. We start with arbitrary y_{ijk} and ask firms what marginal losses they would suffer due to externalities if the proposed allocation went into effect. From this information, we use the necessary conditions (3) to compute an initial set of taxes, given the social opportunity costs q_k. A new program is now generated by firms seeking to maximize profits at the "prices" $s^{ii}{}_k$ and assuming that others will operate according to the old program. We then replace the old program with this new one and iterate. With $\bar{Y}^i$ convex, the convergence properties of this process are analogous to those of a standard tâtonnement process. For an example of such an analysis for a tax system similar to ours and a brief history of adjustment processes of this type see Aoki (1970). Clearly, if we have convergence, the resulting allocation must satisfy necessary conditions (3).

But even assuming convergence, all we know is that we arrive at a tax equilibrium; there is no presumption that we arrive at the *right* tax equilibrium. For example, suppose that we are trying to control air pollution by imposing taxes. If we start from a situation of unbridled pollution, we may find that most affected parties have either moved away from the sources or are relatively saturated with pollution. If this is the case, we may well find a tax equilibrium in which tax rates are very low (indicating that marginal benefits and losses are equal and low), and pollution levels relatively high. On the other hand, society might be much better off with substantially lower pollution levels. Clearly, the decentralized tax scheme will never uncover this possibility. Only a global cost-benefit analysis is sure to uncover it. For a further discussion of the appropriate cost-benefit analysis, see Starrett and Zeckhauser (1971).

4. The Case of Consumption Externalities

Note that so far the analysis has been restricted to allocation in production without regard to the distribution side of things. We did this for the sake of simplicity in notation and exposition; transfers of economic rent among firms will have no effect on production allocation, but similar transfers among households create income effects which do change the equilibrium. However, aside from this additional complication, the analysis of externalities in consumption is very similar to our preceding analysis of the production side. We will not present that analysis here, but merely indicate the nature of the results. For further discussion, see Starrett and Zeckhauser (1971).

With full convexity, either a tax or a market scheme will be Pareto efficient in handling the distribution role. However, inasmuch as these schemes have different implications for the distribution of income, they are not equivalent; only an appropriate redistribution of income will convert one into the other.

External diseconomies will introduce nonconvexities into consumer indifference surfaces just as they introduce nonconvexities into production surfaces. While a household cannot shut itself down in the same sense that a firm can, it can escape the effect of an externality such as pollution by moving to the country. Furthermore, if the externality affects him through his consumption of particular goods, he can avoid the externality by ceasing his consumption of those goods; for example, if I care about river pollution only as it affects my pleasure from swimming, I can escape the pollution by swimming elsewhere or not at all. Either of these situations imply a nonconvexity (for a demonstration, see Starrett and Zeckhauser, 1971), and the subsequent analysis of tax and markets schemes is just as in Section 2.

5. Externalities and Strategy

Some time ago, Buchanan and Stubblebine (1962) argued that a tax subsidy scheme would fail to allocate resources efficiently (in the sense of pareto). However, their argument was quite different from ours and seems to be unconvincing if we stick to the usual definitions of Pareto efficiency. They argued that if an upstream firm was paying taxes not received by some downstream firm, then there would always be a trade which the two of them could make that would make them both better off. However, the required trade involves a reduction in the pollution level *and therefore a lower tax payment*; since a lower tax payment means that there will be less purchasing power to distribute among others, it is not true that the trade leaves everyone better off. Our pair of firms has simply exploited its position vis-à-vis the rest of the world.

What the Buchanan–Stubblebine argument does show is that it may be difficult to enforce the tax-subsidy allocation even if we do know the appropriate rates: Pairs of individuals will have an incentive to deviate from the plan by making private deals on the side. To work, the tax/subsidy scheme requires that all agents take price *and the behavior of others* as given. The problem is not that the Pigovian allocation is inefficient, but rather that people have incentives to act in noncompetitive ways. For market schemes this has long been recognized; externality markets will generally be "thin" so that participants will have some influence over the prices charged. For tax schemes, the difficulties are compounded. If there are few firms affected by a particular externality, then each will have some bargaining influence over that level, and we have to face the Buchanan–Stubblebine difficulty. But if, on the other hand, there are many affected firms (so that no direct bargains will be useful), each has an incentive to overstate his marginal losses as an alternative method of lowering the level of pollution. What is needed is a method of determining a firm's marginal losses without having to trust the firm. "For an analysis of several possibilities, see Starrett (1971)."

Historical Note

In an earlier paper, Professor Baumol (1964) discusses a relationship between externalities and nonconvexity. His point is quite different from the one we are making and seems to be less general. His argument can be paraphrased as follows: Suppose that my utility (W) depends on my consumption (c_1) and yours (c_2). Ordinarily, we would say that there is an externality present if $\partial W/\partial c_2 \neq 0$. However, Professor Baumol defines externality to exist only if $\partial^2 W/\partial c_1 \, \partial c_2 \neq 0$, that is if increases in your consumption affect my *marginal* benefits. He goes on to use $\partial^2 W/\partial c_1 \partial c_2$ or a measure of the size of externalities. Under that definition it is clear that if externalities are large enough (relative to direct effects), the Hessian of W will be negative and the utility function must exhibit a nonconvexity over some ranges.

But under the usual definition of externalities, external economies can be as large as we like without implying nonconvexity. For example, if the welfare function is of the form $W = c^{1-\alpha} c_2{}^{\alpha}$, we would want to say that the externality is stronger and stronger relative to the direct effect as α approaches 1. However, for all α between zero and one, W naturally exhibits full convexity.

Hence, the question of whether strong externalities imply nonconvexities of preference seems to reduce to the question of whether externalities are of the Baumol type, or of the more general type. This would appear to be an empirical question. As mentioned earlier, one can demonstrate logically the presence of nonconvexity in situations involving consumption externalities, but only if consumers have the option of escaping the effects of the externality (by, say, moving or installing air-conditioning). The argument is quite independent from that of Baumol.

APPENDIX

The version of the Hahn–Banach Theorem which we use in Proposition 3 follows directly from standard versions but is rarely stated in texts. The standard theorem is reproduced here from Royden (1963).

THEOREM (Hahn–Banach). *Let p be a real-valued function defined on the vector space T satisfying $p(s + t) \leqslant p(s) + p(t)$ and $p(as) = ap(s)$ for each $a \geqslant 0$. Suppose that f is a linear functional defined on a subspace S and that $f(s) \leqslant p(s)$ for all s in S. Then there is a linear functional F defined on T such that $F(t) \leqslant p(t)$ for all t, and $F(s) = f(s)$ for all s in S.*

To prove our version of the theorem, we follow the normal procedure for proving the separating hyperplane theorem interpreting the function p as a support function of the convex set Y. If zero is not a point of X interior to Y, then we rescale the axes so as to translate some point of X which is interior to Y into zero. Then the support function is defined by:

$$p(y) = \inf \left\{ \lambda \,\middle|\, \frac{1}{\lambda} y \in Y, \lambda \geqslant 0. \right\}$$

Royden shows (p. 176, 18. Lemma) that this support function satisfies the conditions for p in the Hahn–Banach Theorem, and in addition that

$$\{y: p(y) < 1\} \subset Y \subset \{y: p(y) \leqslant 1\}.$$

Let $f(x)$ be a linear function which defines a supporting hyperplane to X at $\bar{x}$. Since f is determined only up to a positive constant, we can normalize so that $f(\bar{x}) = 1 = p(\bar{x})$. To apply the theorem above, we need to show that $f(x) \leqslant p(x)$ all x in the subspace spanning X. But this is trivial, since at every boundary point, x, of X, $f(x) \leqslant p(\bar{x}) = 1$ (recall that f defines supporting hyperplane at $\bar{x}$), and both f and p are linear homogeneous.

We now let X be the set S of the theorem and apply the Hahn–Banach Theorem to extend the functional f defined as the space spanning X to a functional F defined on the space spanning Y. Then $F(y) \leqslant p(y)$ all y; so in particular, $F(y) \leqslant p(y) \leqslant p(\bar{x})$, all $y \in Y$. But since F agrees with f on the subspace spanning X, $F(y) \leqslant F(\bar{x})$ all $y \in Y$, and F defines a supporting hyperplane to Y at $\bar{x}$; furthermore, F agrees with f on X.

References

1. M. Aoki, Two planning processes for an economy with production externalities," HIER Discussion Paper No. 157, Harvard University, Cambridge, Mass., 1970.
2. K. J. Arrow, The organization of economic activity: issues pertinent to the choice of market versus non-market allocation *in* "The Analysis and Evaluation of Public Expenditures: The PPB System, pp. 47–64. Joint Economic Committee of the Congress of the United States, Washington, D. C., 1969.
3. Baumol, W. J., External economies and second-order optimality conditions. *American Economic Review* **54** (1964), 358–72.
4. J. M. Buchanan and W. Stubblebine, Externality, *Economica*, N.S. **29** (1962), 374–84.
5. R. H. Coase, The problem of social cost, *J. Law Econ.* **3** (1960), 1–44.
6. P. A. Diamond and J. Mirrlees, Optimal taxation and public investment, to appear in the *Amer. Econ. Rev.* (March-June 1971), to appear.
7. J. Meade, External economies and diseconomies in a competitive situation, *Econ. J.* **62** (1952), 54–67; "Readings in Welfare Economics" Arrow and Scitovsky, Eds.), pp. 185–98, Irwin Press, Homewood, Ill., 1969.
8. Y. Murakami and T. Negishi, A note on a formulation of external economy. *Int. Econ. Rev.* **5** (1964), 328–34.
9. A. C. Pigou, "The Economics of Welfare" Macmillan, New York, 1920 (first edition), 1932 (fourth edition).
10. H. L. Royden, "Real Analysis," Macmillan, New York, 1963.
11. D. Starrett, On a fundamental non-convexity in the theory of externalities, HIER Discussion paper No. 115, Harvard University, Cambridge, Mass., 1970.
12. D. Starrett, Identifying production and cost relationships in the presence of externalities," HIER Discussion Paper No. 186, Harvard University, Cambridge, Mass., 1971.
13. D. A. Starrett and R. Zeckhauser, Treating external diseconomies—Markets or taxes? Discussion Paper No. 3, John F. Kennedy School of Government, Harvard University, Cambridge, Mass. 1971.