The Postwar Literature on Externalities: An Interpretative Essay

E. J. Mishan

The Postwar Literature on Externalities: An Interpretative Essay

By E. J. Mishan

The London School of Economics
The American University

EXTERNAL EFFECTS on firms—or externalities, as they are now inelegantly referred to—make their appearance in Marshall's *Principles* as external economies; *i.e.*, economies external to the firm but internal to the industry. Little attention was given to this concept until Pigou's celebrated *Economics of welfare,* where, developed and extended, it appears as one of the chief causes of divergencies between "private net product" and "social net product." Expressed more generally, externalities today provide the standard exception to the equation of optimality with universal perfect competition. In addition to the increasingly overt recognition of this qualifying or limiting proviso, interest in the externality concept, as a phenomenon in the context of partial equilibrium analysis, has grown steadily and picked up momentum in the post-war period. Its current popularity warrants the demarcation of a new field of specialization within the broader terrain of welfare economics.

I

Other than R. F. Kahn's emendation of A. C. Pigou's general proposition in his classic paper on Ideal Output [23, Kahn, 1935, pp. 1-35]—to the effect that competitive industries having external economies (diseconomies) *above the average for the economy as a whole* should be expanded (contracted)—developments in the subject area during the interwar period appear to have consisted, in the main, in clearing up confusions about the nature of the long-period supply curve of competitive industry. Since the Second World War, however, and especially during the last decade, contributions to the subject have been prolific—though not surprisingly in the opinion of those who discern a close association between the development of economic analysis and the economic problems of society. Nor is it altogether inexplicable that, although environmental spillovers have been prominent in the news over the last few years, the bulk of the recent literature has confined its investigations to inter-industry, inter-firm, and inter-person externalities. Economists respond to real world problems with a time lag, initially making use of more familiar, if less relevant, bits of apparatus.

Since one of the purposes of this interpretative survey is to acquaint the non-specialist with the significance of the advances made on this front, and to leave him with a picture of works in progress, a chronological account would seem to be less suitable than one that divides the subject into a number of broad aspects. This treatment is,

therefore, organized in the main around four topics: 1) the problems of definition, 2) the traditional doctrine in the light of later refinements, 3) the relation of external economies to public goods, and 4) the new concern with environmental spillovers.

II Concerning the Definition of Externalities

In popular expositions, an external effect is commonly defined in terms of the response of a firm's output, or a person's utility, to the activity of others. Insofar as the standard smoke and noise examples are cited, the correct impression is conveyed. This casual definition is unsatisfactory because the statement that a firm's output, or a person's utility, can be influenced by the activity of others, also holds true in the absence of external effects. Within the context of an interdependent system, e.g., the Walrasian general equilibrium system, an exogenous change in the behavior of individuals can alter the equilibrium set of product and factor prices and thereby alter the utility levels of persons and the output levels of firms and industries. In the presence of universally perfect competition, however, such exogenous changes entail equilibrium solutions that are all Pareto optimal—and these solutions, therefore, cannot be ranked in the absence of a social welfare function.

In light of the above proposition, one is compelled to recognize the distinction between, on the one hand, instances where the influence upon the utility and outputs of others are exerted "indirectly," *i.e.*, via relative prices only in a general interdependent system, and, on the other hand, those where such an influence is exerted on them "directly," *i.e.*, via the arguments of their utility or production functions.

There is general agreement on the sort of mathematical notation required to indicate the presence of an external effect. Thus $F^1(x_1^1, x_2^1, \ldots, x_m^1; x_n^2)$ will represent an external effect generated by entity 2 on entity 1. F^1 can stand for the utility level of person

1, in which case the x^1s are the amounts of some goods $X_1, X_2, \ldots, X_m$, utilized by him, x_n^2 being the amount of some good X_n (where X_n could, of course, be $X_1, X_2, \ldots$, or X_m) that is utilized by person 2, or produced by an industry 2. Again, F^1 can stand for the output of a firm or an industry, in which case the x^1s are the amounts of its inputs, while x_n^2 is the amount of the input or output of some other firm or industry. Alternatively, looking at the productive process from the standpoint of cost, F^1 can stand for the total cost of all the goods X_1 to X_m produced by firm 1, where cost depends not only on the amounts produced of these goods but, also, on x_n^2, the amount of good X_n produced by firm 2.

A consideration of this notation suggests that an external effect arises wherever the value of a production function, or a consumption function, depends directly upon the activity of others. What the notation alone does not succeed in conveying, however, is that the essential feature of the concept of an external effect is that the effect produced is not a deliberate creation but an *unintended* or *incidental* by-product of some otherwise legitimate activity. This feature influences the economist's and the public's attitude toward externalities and, consequently, also influences remedial policies.

In pointing up the allocative significance of an external effect, we shall find it convenient, for the present at least, to follow tradition in treating the subject within the context of partial equilibrium analysis: assuming, that is, that all optimal conditions are met in all sectors of the economy save in those under scrutiny.

The existence of an externality falling on entity 1 as a result of the marginal unit of entity 2s equilibrium output of X_n is indicated by the term $\partial F^1/\partial x_n^2 \neq 0$. Let p_n^2 and c_n^2 stand for price and marginal cost of any output of the good X_n chosen by entity 2. If the existing externality is ignored by entity 2 and a competitive equilibrium chosen so as to equate c_n^2 to p_n^2, then $(p_n^2 - c_n^2) +$

$F^1/x_n{}^2 \neq 0$—which can be abbreviated $(p^2 - c^2) + F^{21} \neq 0$. The externality then has allocative significance at the margin of the existing equilibrium, and corrective action is called for.

We may conclude that $(p^2 - c^2) + F^{21} = 0$ is a necessary condition for an optimal output of $x_n{}^2$. But it is not a sufficient one. If $F^{21} = 0$ at the competitive equilibrium $(p^2 - c^2) = 0$, it may also be true that external diseconomies are generated only by intra-marginal units of the $x_n{}^2$ output. If this is so, the equilibrium output of $x_n{}^2$ may not be optimal. For it is possible that although the external diseconomy is generated only by some initial units of $x_n{}^2$, or is invariant to output $x_n{}^2$, if it inflicts on third parties (such as entity 1) a total loss that is greater than the total benefit—conceived as the sum of factor and consumer surpluses—deriving from the production and direct use of the competitive equilibrium output $x_n{}^2$ then the optimal output of the externality-producing good is zero.[1]

Similarly, if $F^{21} \neq 0$, where $(p^2 - c^2) = 0$, the optimal output of $x_n{}^2$ can be zero.

To summarize: a) in order for an external diseconomy to have no allocative significance (in the sense that no output correction is called for), the condition $F^{21} = 0$ at the equilibrium output $(p^2 - c^2) = 0$ is not sufficient. It is necessary also that in this equilibrium the total losses inflicted by external diseconomies do not exceed total surpluses; b) in order for a certain marginal externality, $F^{21} \neq 0$ at output $(p^2 - c^2) = 0$, to warrant a correction of the output so as to meet the condition $(p^2 - c^2) + F^{21} = 0$, it is also necessary that at this latter output[2] the same total condition is met—

total surplus exceeds total externality-imposed loss.[3]

→»«←

Some further light is shed on the nature of an externality by the notion of "internalizing" it. If the effluent of an upstream firm damages the product of a downstream firm, a merger of the two firms will internalize the spillover—for the upstream branch of the new firm has now to adjust its output in the light of the damage its effluent causes to the downstream firm. More generally, however, the externality can be internalized into the economy if a market for a product not previously sold comes into being. For example, straw might be simply one of the joint by-products of the threshing of wheat which happens to have value for some of the poorer peasants who habitually gather it for fodder or for filling mattresses. It would, however, cease to be an external economy if commercial uses for it were discovered, and the demand for it grew so that a market for straw came into being. Both grain and straw would then become intentionally produced and jointly marketed, and the demand prices of both together would be equated to the marginal resource cost of wheat production. Again, if we can imagine a world in which the airspace were allocated as a form of mobile property rights among the inhabitants, a "factor" market might be established for the services of smoke-absorbing space. Smoke-absorbing services would then appear as an item along with factor-payments as an integral

---

[1] Assuming that, under the existing technology, the least cost method of dealing with the external diseconomy is used.

[2] Thus the optimal outputs remain positive though, in general, different from the equilibrium output $x_n{}^2$.

It is commonly assumed that the equilibrium amount of the affected good, say X_m, is also altered by the external effect absorbed from the production

of the good X_n. If however, the external effect on the production function of the X_m good is "separable," the equilibrium amount of X_m is not altered; that is, there is an effect on X_m's total cost, but not on its marginal cost [11, Davis and Whinston, 1962].

[3] This total condition may also be extended to cases where $\partial F_1{}^2/\partial x_n > 0$—where, that is, external economies prevail at the margin—since it is at least conceptually possible that, although external economies are exerted at the margin, external diseconomies are generated by some of the intra-marginal units of X_n.

part of the production costs of the goods in question.

Of course, the price placed by the market on the erstwhile spillover, regarded either as a good or a "bad," may well be judged too high, or too low, by reference to optimality conditions—in which case one may be tempted to assert that the external effect is only "partially internalized" into the economy.[4] But extending the definition of an externality to make it turn on an 'optimal' pricing of the products and factors in question would be too stringent a requirement—for then any item whose price might be altered to bring outputs closer to an optimum would be deemed to partake of the nature of externalities. Moreover, such a definition is unnecessary as irrespective of the resulting price, the market internalization of the externality implies that, once priced, it comes under the control of that person, firm, or industry, which, hitherto, could only be a passive recipient. That is to say, beginning from a situation in which the function describing the response of person, firm, or industry 1 is written as $F^1(x_1^1, x_2^1, \ldots, x_m^1; x_n^2)$, the internalizing of x_n^2 now brings the value of the function under the direct control of 1, so that this original function is now to be written as $F^1(x_1^1, x_2^1, \ldots, x_m^1, x_n^1)$. For the price of X_n is now determined by the market along with the prices of all other goods and factors; to the extent that the levels of outputs and utility are dependent on X_n, they are, for everyone, now affected only "indirectly" by price changes of X_n.

It follows, incidentally, that if every effect on social welfare arising in the production, and use, of all goods is entered into the price system, universal perfect competition would tend to a general equilibrium that would, indeed, be Pareto optimal. One could say more—provided only that relative prices remain unaltered in the

presence of exogenous changes, the comprehension of all effects on social welfare by the price system implies that all admissible investment projects produce *actual* Pareto improvements.[5]

III The Pre-War Controversy

Before moving on to more recent developments, let us briefly re-examine the early controversy associated with the names of A. C. Pigou, F. H. Knight, J. Viner, and Mrs. Joan Robinson. Not only is the outcome of the controversy of some interest in itself, it also provided us with the gratuitous term *pecuniary* external effects.

Recalling a useful division of externalities into two polar cases, those *internal* to an industry[6] (or activity), and those *external* to it, the controversy in question confined itself to the treatment of external diseconomies that are *internal* to the industry. Two related questions were at issue: 1) whether the upward-sloping supply curve of the competitive industry was an *average* curve, in which case the competitive equilibrium output would be too large, and 2) whether the traffic-congestion problem was an instance of external diseconomies within the industry or else of the misuse of a scarce resource—in particular, of the zero pricing of scarce land.

---

[4] An instance of this sort of reasoning is found in Mishan [34, 1965, pp. 6–8].

[5] This is true whether or not perfect competition prevails. By assumption no one can be made worse off either through a change in relative prices or through external diseconomies, and some people will be made better off provided the new investment meets proper benefit-cost criteria.

[6] The external diseconomy *internal* to an industry *A* might well be the result of an external economy absorbed by *A* from another industry *B*. J. Meade's example of apple-blossom providing food for bees [31, 1952, pp. 54–67] is such a case. The scarce apple-blossom would cause the average cost of honey to rise as labor and capital were increased in the honey-producing industry (*A*). While the optimal amount of apples in the *B* industry has to take account of the value of the blossom as bee-food in honey production, the optimal amount of honey in *A* is below the competitive equilibrium, and is obtained by taking a curve marginal to the rising average cost of honey [34, Mishan, 1965].

In summarizing the debate, H. Ellis and W. Fellner [17, 1943] made it clear that, contrary to an initial allegation by Pigou (but later corrected by him) the equilibrium output of a perfectly competitive industry is indeed optimal. Thinking in terms of a fixed factor, say land, and a variable factor, say labor, available in any amount at a fixed price, the rising supply curve of the product is average inasmuch as it includes, at each output, the rent of land. But these rents, which rise as output is expanded, are but transfer payments to the owners of higher quality intra-marginal units of land—or, for that matter, to the owners of a fixed amount of land of uniform quality—and cannot be reckoned, therefore, in the *opportunity* cost of increasing output. *Excluding* rent-payments, therefore, the supply curve can be regarded as a true marginal cost curve of the product, comprising as it does, the cost only of the additional labor required to produce additional units of the product. The competitive equilibrium is, therefore, optimal. Yet, if the equilibrium output is one for which price is equal to marginal cost *excluding* rent, it is at the same time one for which price is equal to average cost *including* rent.

As for the other issue, congestion on the roads, the Pigou explanation [42, 1946], reduced to its elements, is that each of the owners of the vehicles operating on a stretch of highway would have regard to the costs incurred only as it affected his own vehicles; each owner, that is, would ignore the costs simultaneously imposed on all the other vehicles by the addition of his vehicles. In order, then, to take account of these external diseconomies suffered by the intra-marginal vehicles Pigou constructed a curve that is marginal to the per-vehicle average cost curve, optimal traffic being reached when this marginal cost curve is equal to price, or marginal benefit. This optimal traffic flow can be attained by levying a toll on all vehicles equal to the difference,

at the optimal flow, between the average and the marginal cost.

Knight [25, 1924, pp. 582–606] took a different view. To him it was not so much an instance of the divergence between private and social net products but rather an instance of the wasteful exploitation of a scarce natural resource. If good land were free, farm produce would also be excessive. Put land under private ownership, however, and its price would be bid up as demand expanded. Rising rents, as output expands, would then be included in the average costs of the product; in equilibrium, each factor would then be paid its full marginal product. By strict analogy, if the road in question is placed under private ownership, a price will be imputed to it that is equal to the full earnings of the owner. In a competitive situation this price will be its true scarcity value.

Bearing in mind that under competitive conditions the private owner is deemed to maximize his earnings by "exploiting" the supply curve—that is, by taking a curve marginal to the average cost curve—we perceive where the two explanations converge. Pigou's external diseconomy explanation elicits the marginal cost *excluding* rent concept. Knight's solution—the proper pricing of a scarce resource, road space—elicits the average cost *including* rent concept. And, as we have seen, the curves produced by each are coterminous, being but different ways of regarding the correct competitive supply curve.[7]

The above resolution of the controversy presents no difficulty *provided that* the upward-sloping supply curve is the result of adding increments of a variable factor that is infinitely elastic in supply to a factor that is fixed in supply. In that case the area above the industry supply curve (and below the resulting equilibrium price) can be

---

[7] A more detailed treatment of the controversy will be found in Mishan [34, 1965].

imputed wholly to the fixed factor. If, how-ever, the upward-sloping supply curve of an industry, composed of equally efficient firms, is the result of inelasticity in the supply of *both* factors, plus the familiar assumption that the factors are combined there in proportions different from the average of all industries, i) the curve is still a *marginal* curve in the sense that—for production functions homogeneous of degree one—the long period marginal cost, in terms of either factor, is equal to the equilibrium price; ii) it is also an *average* curve in the sense that the minimum average inclusive cost for each firm is equal to the equilibrium price. But the area above the supply curve can no longer be identified with a rent or a "surplus" to either factor [36, Mishan, 1968, pp. 1269–82].

->>><<-

In his classic paper of 1931, "Cost Curves and Supply Curves," Viner [53, 1931, pp. 23–46] introduced some terminological innovations which have since become standard currency despite their being, in my view, superfluous and possibly confusing. The term external *pecuniary* diseconomies was proposed to cover the case of a rising supply price that is the result solely of changes in relative factor prices as output expands. But in the complete absence of external effects, rising supply price is an implication of any interdependent economic model having such familiar features as production functions homogeneous of degree one, imperfectly elastic factor supplies, and factor proportions differing from one product to another. Seen from this perspective there is nothing special about a rising supply curve, and no optimizing correction of equilibrium outputs need be sought under conditions of universal perfect competition. Therefore to invoke the term *pecuniary* external diseconomies to "explain" supply curves that are in fact already explained by this familiar interdependent economic sys-

tem simply in order to distinguish them from external diseconomies proper—which in the Viner article take on the appellation, external *technological* diseconomies—strikes one today as, perhaps, a verbal extravagance. Moreover, the use of *pecuniary* external *economy* to refer to a *reduction* in the average cost of industry A as it expands its purchases of materials or services from a falling cost industry B, will surely confuse most readers because this phenomenon is neither more nor less than the original Marshallian conception of external economies that are internal to the competitive industry A, and attributable to economies of scale in the B industry.[8] We shall, therefore, make no further reference in this paper to a distinction between *pecuniary* and *technological* external effects. We shall speak only of external effects proper.

However, as was pointed out in 1965 [34, Mishan, 1965, pp. 4–5], the original clarity of the externality concept has become blurred in consequence of the term being used over the years as a convenient peg on which to hang a variety of economic phenomena which might be used to justify intervention in the private enterprise sector of the economy. Thus, the growth in skill and technical expertise of an expanding industry, where economic advantages are generally assumed "irreversible" (a consideration pertinent to the infant-industry argument) have occasionally been referred to as external economies. P. W. Rosenstein-Rodan [45, 1943, pp. 202–11] used the term to indicate an alleged reduction in risks, and therefore costs, arising from the central planning of industries producing complementary goods. Again, T. Scitovsky [50, 1954, pp. 70–82] classified as "pecuniary

---

[8] If, however, industry A were a monopoly, and used its size and buying-power to obtain more advantageous terms from its sellers, the gains it would make in this way are not those arising from the technological advantages of greater output, but only a transfer of revenue from weaker firms to itself.

external economies" such diverse phenomena as consumers' and producers' surpluses, unexploited investment opportunities to be found in complementary industries (arising, in the main, from inadequate information and cooperation), in decreasing cost industries (arising from indivisibilities), or in domestic import-competing industries (for reasons connected with optimal-tariff arguments).

In the remainder of this paper all such extensions of the original concept are ignored.

IV The Renovated Traditional Doctrine

The received doctrine, largely associated with Pigou's monumental *Economics of welfare*, expressed simply and without qualification, is that the equilibrium output of a competitive industry which generates an external *diseconomy* having allocative significance, is in excess of its optimal output. If positive, the optimal output is that at which the market price, less the social value of the marginal external diseconomy, is equal to the marginal resource cost. Conversely, if the competitive industry generates an external economy that has allocative significance, its equilibrium output is below the optimal output obtained by equating to its marginal resource cost the market price *plus* the social value of the marginal external economy.

Moreover, assuming total conditions are met, the traditional remedy for external effects is the tax-subsidy one. For a good generating an external diseconomy, the required excise tax is equal to the value of the marginal external diseconomy at the optimal output. *Per contra*, for any good generating an external economy, an excise subsidy equal to the value of the marginal external economy at the optimal output should be offered to producers. Clearly, the effect of these measures is, in the former case, to reduce output below its competive equilibrium and, in the latter case, to ex-

tend output beyond its competitive equilibrium.

To this oversimplified version there will now be appended a number of qualifications and modifications, some obvious, some less so.

(a) The cost of reducing the economic loss inflicted by external diseconomies is always to be minimized in the light of existing technological opportunities. In Pigou's example of the damage done to the crops by the sparks of a railway engine,[9] the unavoidable damage will not be the value of the crops destroyed if it transpires that the farmers can grow crops elsewhere, or can at least produce something with the movable factors and inputs used in growing the crops. For in moving his labor and capital to an otherwise less suitable location, the loss is ultimately carried by the owners of land adjacent to the railway. Given, then, that all factors other than land are mobile enough to be employed elsewhere at the market prices, the economic loss attributable to the railway service is equal to no more than the loss of rent suffered by the owners of the land. It goes without saying that the economic loss, total and/or marginal, will be still smaller if the switch to a strain of spark-resistant crops, or the installation of spark-preventive gadgets, costs less than the value of the loss of rent.

(b) If there are reciprocal externalities, competitive industry x imposing externalities on competitive industry y, and y also imposing them on x, the optimal outputs of x and y may both differ from their equilibrium outputs. However, at the optimal output for each industry, the marginal resource cost is equal to the market price plus the algebraic value of the marginal external effect. Thus, if there is a reciprocal external diseconomy that has allocative significance, optimal outputs in both industries are

[9] Discussed by R. H. Coase in his 1960 paper [9] and again in Mishan [34, 1965].

smaller than equilibrium outputs. If there are reciprocal external economies that have allocative significance, optimal outputs are larger than equilibrium outputs. For expository purposes, however, we shall assume henceforth that, unless otherwise stated, external effects are uni-directional only.

(c) The analysis of external effects has always been conducted within a partial equilibrium framework.[10] This condition means, as indicated earlier, that the outputs arising from the application of the *social* marginal cost pricing-rule to externality-generating sectors are correct only if all the optimum conditions are already met in the rest of the economy. If the conditions there are not fully met, then, as observed in the General Theory of Second Best [26, Lipsey and Lancaster, 1957 pp. 1–11], there is no certainty that the outputs satisfying optimal conditions in the sector(s) under examination will move the economy (as a whole) closer to the Pareto Optimum.

Nevertheless, plausible conditions can be invoked to justify this partial procedure [19, Farrell, 1958; 21, Green, 1961; 33, Mishan, 1962; 13, Davis and Whinston, 1967]. If, for instance, the initial price-marginal cost ratios at the non-optimal equilibrium outputs of the externality-generating goods are sufficiently large—say they are outside a "band" within whose limits are contained all, or nearly all, the social marginal cost price ratios of the remaining goods—then "correcting" the outputs of the externality-generating goods by the social marginal cost price rule is likely to bring the economy closer to an over-all optimum.

(d) Another qualification of the received doctrine arises from the possibility that, again within a more general framework, an uncorrected externality also entails the *inefficient* production of a batch of goods. The correcting of such an external diseconomy

then involves a movement from an interior point toward a position on the production-possibility boundary. It is no longer certain in such cases that correcting the external diseconomy results in a smaller output.

Such factor efficiency conditions are not met in the case of those (uncorrected) external diseconomies that are *internal* to the industry inasmuch as scarce resources (such as land, or a road, in the Knight-Pigou controversy) are valued at zero. In the economy as a whole, therefore, factor rates of substitution are not everywhere equal, and the batch of goods chosen by society cannot be on the production-possibility boundary. Once the external diseconomy is corrected, however, the scarce factor in question is properly priced and both production and top-level optima are met—assuming optimum conditions obtain in all sectors other than that generating the external diseconomy. This movement from an interior to a boundary point evidently does not preclude the possibility of an increase in the external-diseconomy good in the optimal position.

In contrast, an external diseconomy arising from product x that is *external* to the x industry does not violate the factor efficiency conditions. While it can be true that the greater the amount of x produced, the greater the scalar reduction of the production function of the other good y, this effect serves only to describe the locus of the production possibilities as between x and y. The uncorrected equilibrium amount of x in this case violates only the top level optimum, and the movement to an optimal position along the production boundary entails a reduction in the amount of x. If, however, this external diseconomy on y is attributable *not* to product x itself but only to one, or some but not all, of the *factors* used in x (or, more generally, if not all the factors used in x exert proportional external diseconomies on y), such factors are in effect overpriced. Thus, notwithstanding that factor rates of substitution are common

---

[10] An outstanding exception being the general equilibrium model produced by Ayres and Kneese [1, 1969]. In its present stage of development its heuristic advantages are more prominent.

to all goods, the uncorrected position is not on the production boundary.[11] The movement toward optimum in this case also entails a movement from an interior point to one on the production boundary, and it becomes possible that correcting for x's external diseconomy results in a larger output of x. The converse is true, in such a case, for external economies.[12]

Apart from such possibilities arising from the infringement of the factor efficiency conditions, there is always a temptation to concoct perverse cases by invoking unfamiliar postulates. An example [43, Plott, 1966, pp. 84–87] would be that of an external diseconomy on y generated by a firm's use of a factor A that varies inversely with the firm's output of x. Optimal output of x then entails an expansion. Clearly, it is this rather implausible assumption that the externality-producing factor A is "inferior" which brings about this apparently perverse result —at least for a noncompetitive industry.[13] Provided only that the factors are "normal" (that is, not "inferior") the same correction

is called for whether the externality arises directly from the use of any factor or directly in the process of producing or consuming the good in question.

V External Economies and Public Goods (i)

Explication of such a relationship is primarily an exercise in taxonomy. No apology is offered for pursuing the matter however. The classification of concepts, and their relation to one another, are preconditions of effective economic analysis.

P. A. Samuelson's original conception [47, 1954, pp. 387–89]—to the effect that a public good is one that is enjoyed in common, or one where person 1's consumption does not interfere with person 2's consumption—is a beginning. It comes close to the heart of the matter without being entirely satisfactory. For one thing, there is no explicit reference to the amount of the public good consumed by each of the beneficiaries. For another, as Margolis pointed out at the time [28, 1955, pp. 347–49], the proffered definition does not seem to accord well with the more common examples of a public good, such as education, hospitals, highways, courts of law, and the police. It does not appear true that for the use of such goods one person involves no cost to others: there are capacity limitations, congestions, and rationing in all of them. Again, despite some hints [6, Buchanan and Kafoglis, 1963, pp. 403–14; 52, Vincent, 1969, pp. 976–84] and notational distinctions [37, Mishan, 1969, pp. 329–43], the nature of the suspected relationship between public goods and external effects has remained elusive.

Finally, the necessary optimal conditions —for, let us say, a three-person community, $v_d{}^1 = v_d{}^2 = v_d{}^3 = c \ldots (1)$, and $v_g{}^1 + v_g{}^2 + v_g{}^3 = c \ldots (ii)$[14], where c is the marginal resource cost and $v_d{}^1$ and $v_g{}^1$ are the marginal valuations of the ith person for the private and public good respectively—pro-

[11] If, in the uncorrected x output, the use of factor A alone exerts an external diseconomy on the output of y, then, although the factor rate of substitution in x, and y is equal to the inverse of the common factor-price ratio, more of both x and y can be produced by switching some of A from x to y in exchange for some of factor B from y to x.
It should be evident, however, that this conclusion depends upon our assumption of external effects moving in one direction only—here from x to y. Assuming, instead, reciprocal effects in a two-good competitive economy, and also that, before correction, factor A used in y has an external effect on x that is proportional to the effect on y of factor A in x, the equilibrium amounts of the two goods will be on the production boundary. Indeed, in a two-good economy, a competitive equilibrium with proportional reciprocal externalities can be optimal.
[12] Notwithstanding, which no striking, new policy implications emerge. The tax subsidy solution which realizes an optimal output simultaneously corrects both sorts of optimal infringement [34, Mishan, 1965].
[13] If the x-industry were perfectly competitive, moreover, the standard result would obtain. The long-period social marginal cost of x would then be higher and the optimal output produced by fewer firms smaller than the uncorrected output.

[14] Assuming, throughout, that there is always sufficient divisibility to warrant the equality signs.

posed by Samuelson to distinguish private goods from public goods respectively are not unambiguous. Granted equation (i) applies to private goods, so also will the expression $v_d{}^1 + v_d{}^2 + v_d{}^3 = \delta c$ (at least if the marginal cost is constant at the equilibrium output), where δc is the increment in the cost of producing the final three units of the good. On the other hand, though equation (ii) applies to public goods, if the short period marginal cost of the public good happens to be zero, it can also be true that $v_g{}^1 = v_g{}^2 = v_g{}^3 = c \; (= 0)$.

The difficulties of making a clear distinction between public goods and private goods, and of linking the former to external effects, appear to be attributable to a number of factors. 1) The use of such terms as "public goods" (or "collective goods"), on the one hand, and "private goods" on the other, terms which have conventional associations, as a means of making a conceptual and functional distinction, is troublesome. Possible confusion can be avoided by provisional use of "G goods" to indicate the functional category corresponding to something like public goods, and "D goods" for the functional category corresponding to something like private goods. 2) Because of the conventional association of such terms, it was not recognized that what might be a G good in the long period could also be a D good in the short period. 3) There was a failure to abstract initially from congestion costs or, more generally, from external effects in determining which category a specific good falls into. 4) There was also failure to make a distinction between two types of G goods, optional and nonoptional, and 5) the attempt to link the definitions of D goods and G goods with distinct optimal equation forms when, as indicated above, either equation form might, under certain circumstances, be applied to one or the other type of good.

Be that as it may, the difficulties can be resolved, first, by making explicit the conditions necessary to transform an external economy into a G good—ignoring, for the time being, possible congestion costs.

<center>❧</center>

Suppose person A buys an amount of X at a price p_x and, unintentionally, confers a benefit Y on person B. The amount of Y thus received by person B may be so great as to have a zero marginal valuation, in which case the marginal consumption of X by person A involves no marginal external economy on B, and the amount of X bought by A is optimal. Only if the spillover has a positive marginal value for B, will B's marginal valuation schedule for X (in consequence of the Y it generates for him), when added to A's schedule for X, result in the optimal amount of X being greater than A's initial purchase. If, however, the amount of Y absorbed by B is constrained by A's consumption of X, the marginal value of that amount of Y may be negative to person B—from being a good to person B, Y has become a "bad" or a diseconomy. Addition of their marginal valuation schedules in this instance implies that the optimal amount of X is *below* the amount initially chosen by A.[15]

Needless to remark, the analysis is essentially unchanged if, instead of just one person B, there are a number of other persons, $B_1, B_2, .., B_n$, all of them benefitting at least from the first units of the spillover Y. The optimal amount of X, that is, may still be the same as the initial amount bought by person A, or more than, or less than, this.

If, instead of the externality produced by A being Y, it is X itself, the very thing that A

---

[15] If there is a reciprocal spillover arising from B's purchase of X conferring Y on person A, as well as A's purchase of X conferring Y on B, the optimal amount of X for B is determined in a manner symmetric to that for A. Thus it is determined by reference to B's marginal valuation schedule for X plus A's resulting marginal valuation for X (as derived from A's marginal valuation schedule for the Y generated by B's purchases of X).

consumes,[16] then although X is available to anyone at the market price, A's amount of X offers to others a larger, equal, or smaller amount of the beneficial spillover X than that which they otherwise would buy. Whatever X is thereby received gratis is, however, a perfect substitute for the X that can be bought. Each beneficiary from A's purchase of X will therefore reduce, wholly, or to some extent, the amount of X he would otherwise have bought. Once more, then, the optimal amount of A's purchase of X is found by adding, to A's schedule for X, those also of the beneficiaries. Under the condition mentioned, this optimal amount may be the same as, or greater, or smaller, than, A's initial amount of X.[17] What is to be noticed, however, is that there is now an optimal amount of a single kind of good X, shared by several, and for this optimal amount of X the algebraic sum of each intra-marginal unit is positive, and greater, than its marginal resource cost.

If we now suppose that the amount of X bought by A suffices for a given number of beneficiaries, then, although we are still supposing A foots the bill, X may now be regarded as a "shared" good—in the provi-

sional sense that the benefits of this amount of X are simultaneously enjoyed by a number of persons.[18] A simple example would be that of a television aerial erected on A's roof to which B, living in the semi-detached house, could connect his television. Possibly others close by, or living in apartments below A, could do the same without inflicting any loss on A—the total numbers availing themselves of A's enterprise being limited by the costs of making the connection, their costs rising with the distance of their television sets from A's aerial.

The shared good X—which may be said to tend to a G good as the numbers sharing it increase—arises therefore as the special case of an external economy in which the spillover X is itself identical to the good X that generates it.

VI External Economies and Public Goods (ii)

(1) We may describe a G good as *optional* if the amount absorbed by any person can be reduced without incurring costs; otherwise the G good is *nonoptional*. The significance of this distinction is revealed by a consideration of the relevant optimal conditions.[19] Suppose the short-period marginal resource cost of the G good in question is zero. Whether it is optional or nonoptional the equation $v_g{}^1 + v_g{}^2 + v_g{}^3 = c = 0$ is necessary. For an optional G good, however,

---

[16] An example would be the purchase by A of private protection of his home, which would benefit also the homes of his neighbors $B_1, B_2, \ldots, B_n$.

[17] According to Buchanan and Kafoglis [6, 1963] the standard doctrine states that if private behavior exerts a "Pareto relevant" external economy, the market-generated supply of resources used in this behavior falls short of the social optimum. They then produce the example of the spillover from A's consumption of X being X also, and for B, the recipient of this spillover, a perfect substitute for the X he buys. However, since they did not discuss the possibility of B's having a negative marginal valuation of A's spillover of X (in virtue of B's being constrained to absorb the lot), their result, that an optimal amount of X could be *smaller* than the market amount, depended, in the non-reciprocal case, on the rather implausible assumption that the X received as spillover by B has a higher value to B than the identical X he buys on the market. In the reciprocal case, on the other hand, the smaller-optimal-output result is obtained by an implicit assumption of increasing returns to scale. Also see P. E. Vincent [52, 1969].

[18] Buchanan [7, 1965] determines the optimal number of people sharing a good by explicit reference to external diseconomies. The reader is reminded that, at this stage, we are still abstracting from congestion costs.

[19] These so-called optimal conditions are sometimes referred to as first order, or necessary, or marginal conditions, thereby distinguishing them from second order, or stability, conditions and also from the all-important total conditions

$$\sum_i^n V^i \geq K,$$

where V^i is the (discounted) net value to the ith person of the total output produced, and K is the (discounted) cost of the total output.

each person chooses to absorb an amount that yields him a zero marginal valuation. Consequently the equation $v_g{}^1 = v_g{}^2 = v_g{}^3 = c = 0$ is also realized by the optimal output of it.[20] For a nonoptional G good, on the other hand, although the amount each person is initially constrained to absorb may differ,[21] his resulting marginal valuation may be zero, positive, or negative, and the latter equation is not met.

Bearing in mind that the optimal equation of some D goods may be written as $v_d{}^1 + v_d{}^2 + v_d{}^3 = \delta c$, and also, as indicated above, some G goods may meet the condition $v_g{}^1 = v_g{}^2 = v_g{}^3 = c$, it should be clear why neither G nor D goods can be associated in all circumstances, and for both short and long periods, with a particular form of an optimal equation even in the absence of congestion costs.

(2) Let us define D goods as those whose resource costs are attributable to each of the beneficiaries—G goods being defined as those whose resource costs are not attributable. The appropriate optimal equation of this definition of a D good is that $v^1 = v^2 = v^3 = c$ ($c > 0$; notwithstanding that it may also be true that $v^1 + v^2 + v^3 = \delta c$). If this optimal condition cannot be met, the resource costs cannot be attributable. The only optimal condition which applies is then $v^1 + v^2 + v^3 = c$ ($c > 0$), and the good is a G good.

A conventional public good, say a particular highway system, which has non-attributable costs in a short period, whether zero or positive, is a G good. If now, from variable costs being zero—or being positive but independent of the number of users—it is discovered that some small user cost can be attributed to the vehicles, the highway system (though still a public good in the conventional sense) is clearly no longer a G good. For variable costs are now attributable to users, and the optimal condition, $v^1 = v^2 = v^3 = c$ ($c > 0$) now applies.

(3) It follows also that although the discovery that user costs are attributable makes the highway a D good for allocative purposes, in the long period it again has to be regarded as a G good. True, the long period, where all costs are variable, is to be associated with any sort of alteration—that affecting the size, shape, or quality, of the facility —in particular any alteration enabling it to cater to a larger number of persons. Indeed, this latter long-period alteration might be thought to afford a clear case of a cost-attributable good and, therefore, for allocative purposes, a D good. Such an interpretation is not warranted, for were we to extend the good to accommodate four persons instead of the former three,[22] if the new long-period optimal is written as $v^1 + v^2 + v^3 + v^4 = c'$ then, as compared with the original optimal condition, we could indeed equate v^4 with ($c' - c$). But if the contemplated extension in favor of this one additional person also has some effect on the benefits to the existing number of persons,

---

[20] This optimal condition, when valid for an optional G good, does not, however, entitle us to infer that the amounts of the good taken can be exchanged as between persons. Where the marginal resource cost is not zero, but positive, the equation $v^1 = v^2 = v^3 = c$ applies only to D goods. In that case, the exchange optimum is also met, which implies that the total output of the D good is so distributed among the n persons as to maximize the valuation of that output.

[21] For example, the seeding of clouds may cause some farms to catch more rain than others. In some cases the marginal valuation of the rain received by a farm may be negative. Alternatively, costs may be incurred in reducing the amount of rain absorbed.

[22] We are concerned here with the long-period alteration in the capacity of some facility, the existing distribution of the population being accepted as a datum, and *not* with the possibility of people moving closer in order to avail themselves of the facility in question (in which case the additional resource costs required by the facility are nil). An example of the former would be increasing the power of an existing broadcasting station to reach a larger number of people. An example of the latter would be the movement of some families to be within range of an existing transmitter.

the new optimal equation has to be written $\bar{v}^1 + \bar{v}^2 + \bar{v}^3 + \bar{v}^4 = \bar{c}'$, where $\bar{c}' > c'$, and $\bar{v}^i > / = / < v^i$ ($i = 1,2,3$). For example, in extending a highway system to pass by an additional number of residences, it is likely that the pre-existing number of highway-users also will derive some additional benefit. In that case the costs of producing the extra benefits are not attributable to the additional user, or users. For allocative purposes, then, the long-period adjustment requires the good be treated as a G good.

(4) Finally, congestion costs incurred in the use of the G goods have to be introduced. Though resource costs are not, by definition, attributable to persons in the case of G goods, congestion costs are attributable. As indicated earlier, extracting the maximum social valuation from the collective use of the existing facility requires that marginal congestion costs be used as a rationing device.[23]

[23] The optimal conditions determining the collective use of a G good in circumstances in which congestion occurs can be derived as a special instance of the more general formulation determining the output of a D good where external economies and diseconomies are exerted by persons on other persons (either directly through their consumption activities, or indirectly through the effect of their consumption activities on the production of the good in question):

$$\left(\sum_i^n V_k^{i1} + \sum_i^n V_k^{1i} \ (i \neq 1) \right)$$

$$= \left(\sum_i^n V_k^{i2} + \sum_i^n V_k^{2i} \ (i \neq 2) \right) = \cdots$$

$$= \left(\sum_i^n V_k^{in} + \sum_i^n V_k^{ni} \ (i \neq n) \right) = c_k$$

where c_k is the marginal resource cost of the kth good

v^{i1} is the algebraic effect of the ith person's goods on the value of person 1's marginal consumption of k. (This is sometimes assumed to be zero for all i ($i \neq 1$), only the positive term v^{11} remaining).

and v^{1i} is the algebraic effect on person 1's marginal consumption of k on the ith person's valuation of his own (i's) goods.

Each person properly evaluates his first summation term but ignores the second summation term which the economist must take into account.

It should now be manifest that if we restrict ourselves to long periods, at least, the definition of a G good—one having non-attributable resource costs—accords with a fair number of goods that are conventionally referred to as collective, or public, goods. External defense, police, street lighting, broadcasting, a bridge, a park, a highway or railway system, are examples that spring to mind.

We may, therefore, if we wish, define collective goods as those having no attributable resource costs *in the long period*. Thus, highways, railways, and bridges, which might in the short period have a small ratio of variable operating costs to inclusive costs, would still qualify as collective goods. Hospitals and schools, on the other hand, will not be collective goods in this definition in so far as an increase in long-period resource costs can be attributable wholly to an increase in the numbers admitted.

Originally we approached the concept of a collective good from the benefit side—as the limiting case of an external economy in which the spillover on others is identical to the good enjoyed by the generator himself. The alternative approach from the cost side can now be indicated. This also conceives the collective good as a limiting case—as a good for which the long-period ratio of attributable to total costs is zero.

For external diseconomies *internal* to the kth activity, all the external diseconomies generated by each person (whose second summation term is therefore negative) are deemed to be exerted on persons engaged solely in the kth activity. Congestion costs are, of course, a common form of external diseconomies internal to the activity. Insofar as congestion costs are generated in the collective use of a G good, the above formulation still applies except that the marginal resource cost, c_k, is equal to zero.

Finally, in the special case in which each person is associated with only one unit of the kth good—as in the example of private automobiles using a highway—the optimal traffic flow is n, where for the nth person

$$\sum_i^n V_k^{in} + \sum_i^n V_k^{ni} \ (i \neq n) = c_k.$$

(5) It remains to make explicit the close relationship between a collective good and an external economy. For the external benefits conferred on each of a number of people as a result of a person's consumption of a unit of a good also cannot be attributed to each beneficiary on any economic principle. Indeed, there is, in general, no formal difference; in either case the summation-over-persons form of the optimal equation is required. There is a difference only in motivation. The benefits generated by person I's consumption of a private good are unintentional whereas the benefits generated by the public good are clearly intentional. An appropriate notation for the values created by the former is

$$\sum v^{1i},$$

and for the latter is

$$\sum v^i.$$

A second informal difference, more apparent than real, arises from the custom of considering an optimal amount of but a single public good in contrast to the habit of considering the optimal amounts of, say, n persons each conferring external benefits on a number of others. This apparent difference would, of course, disappear if we supposed instead a large number of identical types of public goods all produced by a competitive industry. In that event the optimal condition for public goods would be of the form

$$\sum_{}^{m} v_A{}^i = \sum_{}^{n} v_B{}^i = \sum_{}^{q} v_C{}^i = \cdots = c,$$

where A, B, C, ..., are groups comprising m, n, q, ..., members, respectively, in different parts of the economy. If, instead, A, B, C, .., were individuals, each conferring external benefits on n people (including himself), the optimal equation would be

$$\sum_{}^{n} v^{Ai} = \sum_{}^{n} v^{Bi} = \sum_{}^{n} v^{Ci} = \cdots = c.$$

Finally, the capacity, size, or coverage, of a shared or collective good[24]—whether, that is, it serves a few people or many, or whether it serves a locality, a region, or the country at large—depends in the main on three features: 1) economies of scale, 2) costs of travel to, or in connection with,[25] the collective good, and 3) income *per capita* and population density. Thus, 1) the greater is the decline in the long period average cost of providing the public service, 2) the lower is the cost of connecting or moving to it, and 3) the greater is population density and *per capita* income—the larger will be the size and coverage of the public good.

Moreover, any growth of congestion costs, as a result of an increase in the number of users, or an increase in usage by the same number of people, provides an incentive to extend the long-period capacity of the public good. Any optimal extension of capacity in effect entails economies from the substitution of long-period resource costs for some, at least, of the short-period congestion costs.

VII *Solutions to the Externality Problem*

Let us now turn our attention to several of the more familiar methods proposed for correcting outputs for external diseconomies.[26]

1. *Outright Prohibition.* The economist is prone to think of this solution as naive. It would be prohibitively expensive, if not im-

---

[24] The term collective (or public) good is used in two senses in the literature: sometimes to designate the physical asset itself, say a bridge, and sometimes to designate the services provided by that asset. The context usually makes clear the sense in which the term is used.

[25] According as the potential beneficiaries have to travel to avail themselves of the collective service (as in the case, say, of a park or theatre), or to form a link with the generator of the service (as in the case, say, of a television transmitter).

[26] A comparable though not entirely similar commentary on the more commonly proposed solutions to the externality problem can be found in a recent paper by Davis and Kamien [10, 1969, pp. 78–86].

possible, it is argued, to eliminate entirely all trace of some of the pollutants that inflict losses on others. Moreover, the argument continues, optimality does not require that external diseconomies be eliminated, simply that their amounts be consistent with the optimal amounts of the goods that create them.

This sort of argument is not conclusive for all pollutants. First of all, prohibition need not imply prohibition of every trace of a pollutant; it may be directed against producing "discernible" or "dangerous" amounts of the pollutant. Second, as we shall remark later, the cost of discovering and maintaining an optimal amount of the pollution may itself be prohibitive. The community may then be faced with the choice of zero or unchecked pollution.

2. *The Tax/Subsidy Solution.* This is the classic solution, and the one until recently most favored by theorists. The chief obstacle here is, of course, the costs of collecting the necessary information and the costs of supervision, costs which would be particularly heavy for industries in which demand and supply conditions are apt to vary frequently.[27] It is alleged, moreover, that this solution, even if feasible, overlooks a particular contingency that can result in "overcorrection."

See Figure 1 where SS is the "private," or commercial marginal cost curve of the output of X, DD is the market demand curve, and the vertical distance between SS and the *social* marginal cost curve, S'S', is the unit cost of spillovers generated in the production of X.[28]

Optimal output, OQ, can be achieved by

²⁷ It is, of course, possible that the industry producing the external diseconomy is a non-competitive industry. A monopoly firm equating marginal cost to marginal revenue will in any case produce an output smaller than the marginal (private) cost-price output, which may then be closer to the optimal output. In this connection, see D. A. Worcester [56, 1969].

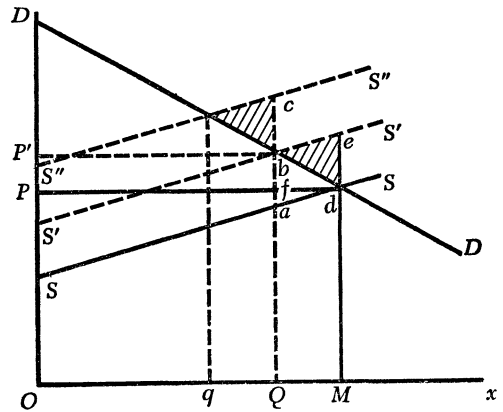²⁸ There is some slight geometric convenience in constructing S'S' parallel to SS.

FIGURE 1.

an excise tax equal, at Q, to the vertical distance between SS and S'S'. After the imposition of such a tax, however, the producers may regard S'S' as the new marginal cost curve.

At the new tax equilibrium, OQ, the marginal damage of the spillover effect is equal to *ab*. The victims of the spillover can then afford to pay producers as much as *cb*, equal to *ab*, to reduce output OQ by one unit, and so on for successive reductions. From such reasoning we construct a curve S"S" that is above S'S' by the same vertical distance at all points as S'S' is above SS. Clearly there can now be mutual agreement between producers and victims to reduce output to Oq, below the optimal output OQ.[29]

However, this possibility cannot be taken very seriously. If producers and their spillover victims can indeed reach voluntary agreement, they have more incentive to reach it before the excise tax is levied than afterward. The government, in any case, can always take measures to ensure that no further arrangements of this sort take place in order to prevent output being reduced below optimum.[30]

²⁹ See J. M. Buchanan and W. C. Stubblebine [5, 1962] and Turvey [51, 1963].

³⁰ A related objection to an effluent excise tax occurs in a paper by P. Bohm [4, 1970]. If the optimal excise tax increases with output, the firm (he

3. *Regulation.* Insofar as regulation of the production of goods that generate externalities is intended, much the same sort of information is required, and much the same sort of costs are incurred, as in the tax/subsidy solution. If, however, regulation is to be applied to the extent, and manner, of the usage of spillover-generating products, there will be additional costs of enforcement—the more so if flexibility is sought, and regulations are devised, to vary according to time, area, and circumstances.

4. *Voluntary Agreements.* If transactions costs in the broadest sense (to be defined presently) are *nil*, the initiative by either the producer or the recipient of the spillover in negotiating a mutually satisfactory agreement will bring about an optimal output. In Figure 1, for example, there will be an incentive to move from output *OM* to *OQ* since by so doing there will be a gain equal to the area of triangle *ebd* to be shared between the beneficiaries of *X* and the spillover victims. The maximum sum the spillover victims will pay to reduce the market output by *MQ* is given by the parallelogram area *abed*, while the loss to producers and consumers from reducing output by *MQ* is equal to the area of the triangle *abd*.

Such agreements, however, unless they

are between firms or industries (and supported by legal sanctions) are likely to be so expensive to negotiate and maintain as to be impractical.

5. *Preventive Devices.* For obvious reasons the professional economist is more likely to interest himself in optimal-output solutions than in the opportunities for installing preventive devices. This latter form of remedy, however, cuts across those mentioned above inasmuch as either government regulation or voluntary agreements can bring them into being. Whether there are opportunities for few or for many such devices, and whether they are less costly to the industry concerned than the alternative course of reducing the spillover-generating outputs, are, of course, empirical questions and ones to which economists are now turning.

VIII *The Abortive Consensus*

Post-war developments seemed about to culminate in a broad consensus in the early 1960s when increased attention to environmental spillovers compelled economists to re-examine some of their basic simplifications as well as the conclusions based on them. The more crucial propositions of this emergent consensus are summarized below.

1. On the assumption that the most economic way of dealing with an externality involves an output adjustment, the optimal output is uniquely determined. In this connection, it was also believed [9, Coase, 1960, pp. 1–44] that Pigou had failed to make explicit the duality of the tax-subsidy remedy. Whether the government offers an excise subsidy to the manufacturer to induce him to reduce the output of a good generating external diseconomies, or whether it imposes an excise tax on such a good, was believed to be a matter of indifference so far as allocation is concerned. Similarly, in the absence of government intervention, and assuming transaction costs are low enough, it was believed to be a matter of indifference from the point of

argues) might become aware of the relationship. Subtracting the schedule of optimal taxes from the demand price of the product would result in a downward-sloping net average revenue curve from which the firm could derive a marginal revenue curve. By equating marginal cost to this "marginal revenue" curve, the firm reduces its output below optimal.

However, the government is not obliged to impose a *uniform* effluent tax. It could as well make it clear that it would impose a *discriminating* tax, one equal at each unit of output to the marginal effluent and, therefore, at any output raising a total tax equal to the total loss inflicted by the effluent. Such a tax, already marginal, effectively precludes the industry from "exploiting" it by reducing its output. In addition, such a discriminating tax ensures that the total conditions are met. Thus, heavy effluent charges properly imposed on the initial units of the output could well prohibit production of the good.

view of allocation whether a manufacturer is compelled to compensate the victims or whether the victims offer to bribe the manufacturer.

2. Nor can the question of liability for the spillover properly be settled by a consideration of the equity involved. To use an example from Coase's 1960 paper [9, pp. 3–5], if the machinery of a confectioner disturbs the practice of a physician on the floor above, so also does the installation of vibration-reducing devices lower the profits of the confectioner. The interests of the two parties are mutually antagonistic, and with respect to equity the case is symmetric.[31]

3. However the matter is actually resolved—whether an excise tax or an excise subsidy is used, or whether the one party or the other is compensated—optimality is not at issue, only the distribution of welfare. This statement does not, of course, imply approval of all measures that realize an optimum position, since in moving from a non-optimal to an optimum position only a *potential* Pareto improvement, at best, is assured (gains exceed losses), and not an *actual* Pareto improvement. Thus a movement to an optimal position is quite consistent with one that makes the poor yet poorer.

4. In the absence of government intervention, whatever the legal position, the unfavored party has a clear interest in trying to bribe the other party to modify the "uncorrected" output. Successful mutual agreement between the parties, however, presupposes that the maximum possible amount of the shared gains, G, in moving to an optimal position, exceeds their combined transaction costs, T. Since the transactions costs, T, are real enough, inasmuch as they are ultimately the valuation of scarce resources, successful mutual agreement produces a *net* Pareto improvement—$(G - T) > 0$. Failure to reach mutual agreement, on the other

hand, can be regarded as *prima facie* evidence that $(G - T) < 0$; that is, a *net* potential Pareto improvement is not possible. Rationalizing the *status quo* in this way brings the economist perilously close to defending it.

Before subjecting the above propositions to scrutiny, it is as well to touch on an analytic difficulty of the partial analysis that seems to have been fudged.

The maximum social gain, G, from reducing the competitive output by MQ, in Figure 1, is generally calculated as follows: the gains to the spillover victims of reducing the output by MQ is equal to the area of the "parallelogram" *abcd*. From this gain, we subtract the loss from two other groups: consumers suffer a loss equal to the area of the "triangle" *fdb*, and producers suffer a loss equal to the area of the "triangle" *fda*. The residual gain—*abcd* less $(fdb + fda)$—is, of course, equal to the triangle, *dbe*.

However, if the supply curve is a long-period industry supply curve, one sloping upward in consequence of a relative price rise of the factor(s) used more intensively in this industry than in the economy as a whole, a zero Knightian profit is made by all firms in the industry in any long-period equilibrium. The area above the supply curve cannot then be identified with any surplus to the producers. Only if the upward slope of the long-period supply curve arises from the addition of increments of a constant-priced variable factor to the fixed amount of another factor—which may, however, include differences as between firms in the quality of the fixed factor—may the area between the supply curve and the price of the good be treated as a surplus. Moreover, it is not a surplus that accrues to the firms, but to the owners of the fixed factor whether of uniform quality or not.[32]

However, even if we suppose this long-period supply curve to slope upward as a

---

[31] This conclusion can be ascribed to the popularity in the literature of the two-firm or two-industry case.

[32] The rents earned by lands of superior quality, or location, being one of the earlier examples in the history of economic thought.

result only of a fixed amount of the scarce factor, we may note that the conclusion that mutual agreement between producers and spillover victims, if feasible, produces an optimal output, commonly ignores the consumer interest. This neglect of the consumer interest is another consequence of the popular preference for the two-firm model, and of the occasional simplification of a horizontal sales curve facing each of the firms.

Ignoring transaction costs, the most favorable conditions for negotiation between producers and spillover victims would seem to exist when producers can be taken as a corporation, and the supply curve then treated as a long-period marginal cost curve (excluding rent). It will facilitate the analysis still further if we suppose the corporation to act as a discriminating monopolist, appropriating all the consumer surplus, thus equating its private marginal cost to the demand curve.

IX Environmental Spillovers— Allocation (i)

The pertinent economic features of environmental spillovers, other than the observed fact that they appear to increase rapidly with economic growth, are 1) that their impact on the welfare of members of the public can be substantial, and 2) regarded as external diseconomies, they pose a problem not so much as between firms or industries, but as between, on the one hand, the producers and/or the users of spillover-creating goods and, on the other, the public at large. The implications of the latter feature are not diminished by the observation that, in important instances, the users of the spillover-creating goods and the affected public are all but indistinguishable —this being but a special case of external diseconomies internal to the activity in question.

A consequence of the first feature is that the so-called income-effects—or, more ac-

curately, *welfare effects,* as we shall call them—can no longer be treated as negligible. A consequence flowing from the second feature is that the transaction costs are likely to be inordinately large. These two consequences assume particular relevance when we recognize that an alteration in the law, say from tolerating to the prohibiting of certain spillovers, or the reverse, has significant effects not only on the distribution of welfare, but on the outcome of the allocative criterion. In particular, the notion of a Pareto Optimum, or, more accurately, since we are to restrict ourselves to partial economic analysis, a potential Pareto improvement, is no longer uniquely determined. Nor, for that matter, is a *net* potential Pareto improvement, $(G - T) > 0$, uniquely determined.

These propositions will now be demonstrated in connection with each of these two features in turn.

(1) If we assume that the welfare effects are positive, or "normal," a man who is prepared to spend up to $60,000 for a particular house with a view will experience a rise in his welfare if, unexpectedly, he finds he can buy it, for, say, $40,000. In consequence of this "surplus" of $20,000, the minimum price he will sell it for, after buying it for $40,000, will be more than $60,000, say $65,000. Invoking familiar Hicksian terminology, the difference of $5,000 in this case is equal to the difference between his *compensating variation* of $20,000 (the maximum sum he would pay—thus restoring his welfare to its original level W_0—in order to be allowed to buy his house for a price of $40,000) and his *equivalent variation* of $25,000 (the minimum sum he would accept to forego the opportunity of buying the house at $40,000, which sum raises his welfare to the level W_1 that he would have enjoyed had he indeed been permitted to buy the house at $40,000).

There is, however, another and possibly more potent factor in differentiating these magnitudes wherever the welfare involved

is substantial. The maximum sum he will pay for something valuable is obviously related to, indeed limited by, a person's total resources, while the minimum sum he will accept for parting with it is subject to no such constraint. To take an extreme example, a man may be ready to sacrifice every penny he can spare in order to pay for an operation that will save his life. This may amount to a present value of $10,000 or $10,000,000, but it will be a finite sum. On the other hand, there may be no sum large enough to compensate him for going without the operation, and so parting from this life.

Let individual B, with disposable income of $12,000 per annum, be exposed to aircraft noise which can be escaped with certainty only be relocating hundreds of miles away in some deserted area. Given the choice, he would, if hypersensitive enough to aircraft noise, pay as much as $5,000 per annum to be entirely free of it. At the same time, if the law compelled the airlines concerned to compensate all injured parties, his true minimum claim could be, say, $15,000 per annum.

Now instead of regarding the maximum and minimum sums as compensating-variation and equivalent-variation measures of a change in welfare under the existing law, we can regard each respectively as the compensating variation corresponding to two opposing states of the law. Thus, if the existing law, L, is tolerant of environmental spillovers, in particular aircraft noise, the compensating variation of a contemplated change banning all aircraft noise is a payment by B of $5,000, this being the sum which, given up in exchange for the ban, maintains B's welfare at the level W_0 which prevails under the existing L law. If, on the other hand, the existing law is $\bar{L}$, one that effectively bans all aircraft noise, B's level of welfare is W_1, which is higher than the level W_0 that prevails under the L law. The compensating variation of a contemplated

change introducing aircraft noise is then a receipt of $15,000 by B—this being the sum which, if the change occurs, will maintain his welfare at its original W_1 level.

Let A stand for the aircraft interests which extend, in this example, to all owning capital or employed in aircraft services as well as all the beneficiaries of air travel. The compensating variations of each of these persons will, in general, also vary according to which of the two kinds of law prevails. Let B stand for all those offended by aircraft noise. If the maximum sums that people are willing to pay to acquire a "good" (or to avoid a "bad") are prefixed by $+ve$ signs, while the minimum sums they are prepared to accept to forego a good (or to put up with a "bad") are prefixed by $-ve$ signs, the algebraic sum of all compensating variations indicates the social value of the change in question. In particular, if, under the existing law, the algebraic sum of a contemplated change is $+ve$, a potential Pareto improvement is possible. If, however, the algebraic sum is $-ve$, the existing unchanged situation is optimal; the change in question would only result in a potential Pareto loss.

Imagine now that a costless and perfectly accurate method of obtaining all the relevant data has been invented. The end-product of much research into the aircraft noise problem might then be summarized in the figures of Table I.

Reading along the first row we interpret as follows: given the existing law L, that is permissive of aircraft noise, the A group must be paid at least $55 million to secure agreement to change to $\bar{L}$ law, while the B

TABLE I

Existing Law	A	B	Total
L	$-$$55m	$+$$40m	$-$$15m
$\bar{L}$	$+$$45m	$-$$70m	$-$$25m

group will offer up to $40 million to have the L law changed to $\bar{L}$. Since the change-over would incur a potential Pareto loss of $15 million, the existing situation under L law is deemed Pareto Optimal. If, however, the existing law is $\bar{L}$ to start with, the second row indicates that the A group will pay up to $45 million to have the law changed to L. But this sum falls short by $25 million of the minimum compensation required by the B group to agree to the change. Again, therefore, the existing situation under the $\bar{L}$ law is Pareto Optimal.

We are to conclude, therefore, that irrespective of the existing distribution it is possible, if not likely, that for significant environmental spillovers, the arrangement that is optimal under one state of the law is not optimal under the other state of law. In our example, if aircraft are already allowed to fly unchecked (L law), then that situation appears as optimal. If, on the other hand, aircraft were banned under $\bar{L}$ law, that situation, too, would be optimal.[33] Under such conditions, how do we decide how to act?

The above analysis, applicable to indivisible economic arrangements as appear, say, in cost-benefit calculations, can easily be extended to economic arrangements having perfectly divisible external effects. Suppose the number of aircraft permitted to fly over a residential area is to be determined by reference only to optimality considerations— the exercise being to locate the point at which the marginal benefit of the aircraft group A is equal to the marginal loss suffered by the residential group B. Prior to

[33] The reader will observe that the above paradox (which depends on different compensating variations under different states of the law) has no affinity with that associated with the so-called "Kaldor-Hicks" and "Scitovsky" type welfare criteria. The latter paradox arises only from alterations in the set of relative prices (common to everyone) associated with the distributional changes whenever the community moves from producing one batch of goods to producing another.

calculating the optimal number of flights, the existence of L law, which permits unchecked flying over the area, will result in a higher level of welfare for the A group than if, instead, the $\bar{L}$ law prevails and no planes are permitted to fly, the reverse being the case for the residents comprising the B group.

Suppose the L law to be in force; then, prior to any agreement between A and B, the number of planes flying over the area is given by OM in Figure 2. The minimum compensation acceptable to the A group for reducing successive flights is given by the marginal curve $M - A_L$, while the maximum sums that the B group will pay for successive flight reductions is given by $B_L - O$, the two curves intersecting at Q_L. If, however, $\bar{L}$ is in force to start with, then, prior to any mutual agreement between A and B, the number of flights over the area is zero, and the minimum sums acceptable to the B group for each successive flight are given by the marginal curve $O - B_{\bar{L}}$, while the maximum sums that the A group is willing to pay for each additional flight are given by the marginal curve $A_{\bar{L}} - A_{\bar{L}}'$. The intersection of these two $\bar{L}$ curves is $Q_{\bar{L}}$.

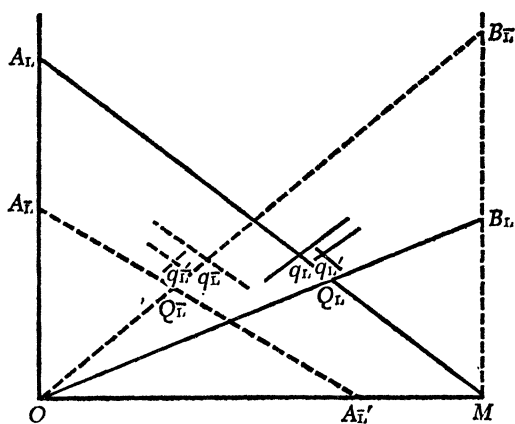Now in reaching agreement, beginning from either initial position—OM flights with

FIGURE 2.

L law, or zero flights with $\bar{L}$ law—one, or both, groups are made better off and, therefore (assuming normal welfare effects), one or both marginal valuation curves shift upwards, which implies that neither Q_L nor $Q_{\bar{L}}$ can be reached by bargaining alone. Nonetheless, if L law prevails, in which case flights are OM to begin with, and if we assume that in bargaining their way to an optimal position, all the gains go to the B group,[34] the increase in its welfare is still less than it would be, if instead $\bar{L}$ law initially prevailed and (without any payment from the B group) all flights were banned. Consequently the resulting maximum upward shift in B's marginal curve is still below the $O - B_{\bar{L}}$ curve, the optimal point being then q_L. However, if we suppose instead that A obtains some of the potential gains in the bargaining, B's marginal curve rises less and A's marginal curve rises somewhat, with the result that the revised optimal position, q_L' is to the right of q_L.

If, on the other hand, the $\bar{L}$ law prevails, and we assume first that all the gains in bargaining go to the A group, the increase in its welfare is less than it would be if the law were changed to L, and all flights were freely allowed. The resulting upward shift in A's original $A_{\bar{L}} - A_{\bar{L}}'$ curve is therefore below the $A_L - M$ curve, the optimal point being $q_{\bar{L}}$. However, according as the B group secures some of the gains in bargaining toward an optimal position, A's resulting marginal curve is somewhat lower while B's resulting marginal curve is above $O - B_L$. The revised optimal position $q_{\bar{L}}'$ being to the left of $q_{\bar{L}}$.[35]

---

[34] That is, the B group pays for each successive flight reduction no more than the sums traced out by the M-A_L curve.

[35] Another incidental implication of this sort of analysis, one making explicit allowance for welfare effects, is that an excise tax alone can no longer be counted on to realize an optimal position. In this connection see F. T. Dolbear, Jr. [15, 1967].

We may conclude, then, that however the bargaining goes, the resulting optimum output under the L law entails more flights than an optimum output under the $\bar{L}$ law.[36]

X Environmental Spillovers—Allocation (ii)

(2) Assuming that, whichever law prevails, the state does not oppose agreements tending to a Pareto improvement, any movement by any method toward such improvements involves a variety of costs, for which the term *transactions costs* is in common use. In general, the more favorable the law is in promoting mutual agreements of this sort, the lower will such transactions costs be. If, at first, we restrict ourselves to the method of voluntary agreement between two opposing groups in their attempt to reach a solution, either by curbing the activity of the offending industry, by installing preventive devices, or by moving the industry (or, alternatively, members of the B group) elsewhere—whichever method is the cheapest—the transactions costs, T, may be divided into three sub-categories: T_1, the initial costs leading to negotiations between the two groups; T_2, the costs of maintaining and, if necessary, revising, the agreement; and T_3, the capital expenditure, if any, required to implement the agreement.

The more important of these, the T_1 costs, can be broken down, for each group, into a number of phases: a) identifying the members of the group, b) persuading them to make, or to accept, a joint offer, c) reaching agreement within the group on all matters incidental to its negotiation with the other group, and d) negotiating with the other group.

It cannot be assumed, without investigation, that transactions costs would be any less under $\bar{L}$ law than they are under L law.

---

[36] The reverse being true for the improbable case of negative welfare effects.

What can be said, however, is that such costs, especially those subsumed under T_1, increase with the dispersion of the B group, and increase with the numbers involved, probably at an exponential rate. Whatever the magnitude of the T costs, however, relative to the maximum Pareto gains, G, three alternative cases exhaust the possibilities.

1. A net potential Pareto improvement, $(G - T) > 0$ emerges for the industry under either type of law—though if $\bar{L}$ law prevails, the optimal output, both of goods and pollution, will be smaller.

2. A net potential Pareto improvement for the industry emerges under neither type of law. By comparison with the (costless) potential optimum, therefore, we shall have "too much" pollution with the equilibrium output under the L law, and "too little" pollution under the zero output of the $\bar{L}$ law. Without further assumptions, however, it is not possible to say in general whether "too much" (under the L law) or "too little" (under the $\bar{L}$ law) is likely to be closer to the potential optimum, and, therefore, whether more is lost by adopting the L law rather than the $\bar{L}$ law.

3. Where the potential optimum position is closer to the initial $\bar{L}$ position than to the initial L position, a net Pareto improvement may take place only if the L law prevails. If the reverse is true, a net Pareto improvement may take place only if the $\bar{L}$ law prevails.[37] The former possibility is illustrated in Figure 3a; the latter in Figure 3b (welfare shifts being omitted so as not to encumber the diagrams).

Although it has been convenient to think of net potential Pareto improvements, through voluntary agreement between the

A and B groups, in terms of output adjustments, if any, the above three possibilities are equally valid if other, and cheaper, methods of effecting net potential Pareto improvements are contemplated—such as moving factories, or processes (or members of the B group) to other areas, installing any of a variety of preventive devices (or, in general, modifying the technology as to reduce spillovers), or government regulation of output either directly or through excise taxes.

In the absence of evidence indicating a clear connection between the magnitude of transactions costs and the type of law, there appear to be no firm allocative implications —save, perhaps, the advisability of thinking more closely of technical and institutional innovations that are likely to reduce trans-

FIGURE 3a.

FIGURE 3b.

[37] It is of incidental interest to note, however, that in the first case society would be better off if the $\bar{L}$ law prevailed (notwithstanding that it would not pay to move from the resulting zero output), since it would save the T costs incurred in moving to Q under the L law. As for the second case, society would be better off if the L law prevailed.
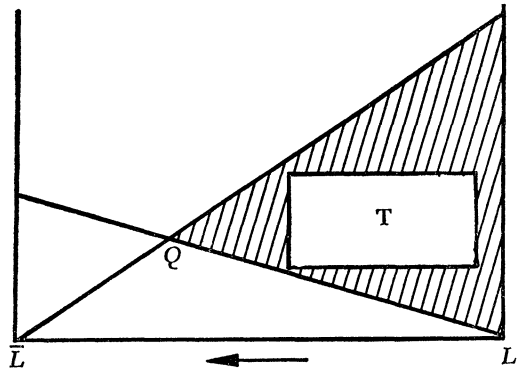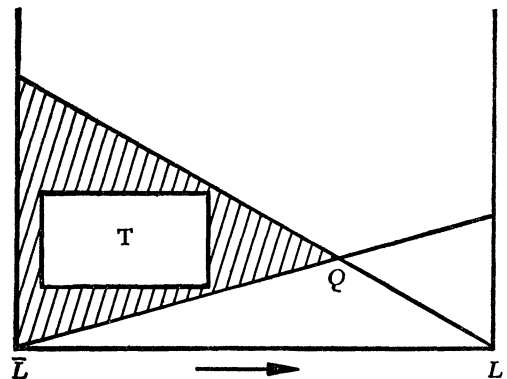
actions costs. Thus, with respect to the costs of government regulation of output, directly or through excise taxes, economists may like to remind themselves that the pursuit of the ideal is the enemy of the better. A roughly calculated excise tax imposed on a pollutant is likely to effect a distinct improvement, even if it were as much as 20% or so higher, or lower, than the "ideal" excise tax. As an immediate response to clear cases of excessive pollution, a roughly calculated tax is likely to be superior to not imposing a tax at all—or to procrastinating indefinitely while engaged in research to refine data and methods in the attempt to produce an ideal tax.

There are, nevertheless, a couple of considerations which appear to favor $\bar{L}$ law rather than L law. First, although the magnitude of transactions costs have been assumed independent of the type of law, the likelihood of a member of one group taking the initiative in approaching the other group is not independent of the law. If L law is in force, the possibility of some limited benefit for the person(s) taking the initiative on behalf of a large and widely dispersed B group has to be set against the certain loss of time and effort, and also against a large risk of incurring substantial and irrecoverable expenses in the attempt to complete phases (a) and (b). Under $\bar{L}$ law in contrast, the necessary initiative comes from industry. No personal risk is undertaken by one or more of the executives acting on behalf of shareholders—though, in any case, such an initiative would hardly stand out from the routine activities involving decision-taking by the managers of industry.[38]

[38] We have ignored welfare effects in order to avoid minor distractions from the shifting of the marginal valuation curves.

Under the $\bar{L}$ envisaged, it is not necessary for a plaintiff to incur any expense in pursuing a claim against pollution in the courts of law, since pollution—in the absence of explicit permission to the contrary—is illegal. Punitive action which includes cessation of the pollution-creating activity

Indeed, under $\bar{L}$ law, firms are unlikely to invest in plant and machinery for the manufacture of pollution-prone products unless they are fairly confident that—after all economic "preventive technology" has been employed—they can afford to meet claims for residual damage.

Second, under an L law, there is little incentive for industry to switch resources from promoting sales, or from research into product innovations or from cost-reducing technology, in favor of pollution-reducing technology. Assuming that firms allocate investable funds according to the equi-marginal principle, they will then, under the existing L law, misallocate resources because they tend to ignore all opportunities for social gains made by directing research funds into preventive technology.[39] Under an $\bar{L}$ law, in contrast, full liability for pollution damage enters directly into production costs, along with expenditures on productive services, and the consequent incentive to engage in such research is inescapable.

Opportunities for private industry, under an $\bar{L}$ law, of reducing the social costs of pollution are not, however, restricted to curbing outputs and engaging in research on pollution-reducing technology. There will be incentives for the polluting industries to investigate many other possible ways of reducing their liabilities for damage to the public. They may find it cheaper to concen-

is immediately taken by the public prosecutor unless the firm has a government permit issued periodically, which permit is never granted unless all claims to damages over the period in question are met. So severe a law will not prove costly to administer simply because businesses will almost certainly find it cheaper a) to move away from populous centers and/or b) to undertake further research into the technical changes necessary to reduce pollution as to be virtually undetectable.

[39] Any hope that funds from the B group will be offered to industry to engage in research so as to reduce widespread pollutants can be ruled out both because of the heavy risks of initiative by individual victims and because of the costs of transactions referred to.

trate (parts of) their plant in remote areas, to re-design and re-route highways, to re-route air flights over less populous areas, in the pursuit of which they create conditions conducive to separate "amenity areas" for the public.[40]

XI Environmental Spillovers—Equity

On the issue of the relative merits of L law and $\bar{L}$ law, there remain a number of considerations which may be subsumed under equity.

1. *Distribution.* If it can be shown a) that goods which generate spillovers also earn incomes for, and are purchased by, groups having above-average incomes, and b) that the bulk of modern spillovers fall more heavily on families with below-average incomes, then it may be asserted that, compared with $\bar{L}$ law, L law is a force acting to increase the regressive distribution of welfare. In the absence of any systematic research into the question, however, one can say only that it is not implausible to believe that the introduction of significant disamenities into a large area is likely to reduce the welfare of the more mobile rich less than that of the poor.[41]

---

[40] The conditions under which a separate areas' solution of group conflict is superior to the usual optimal solution that is constrained within a given area are discussed in Mishan [35, 1967].

[41] The fact that in any given neighborhood the rich will respond to local forms of pollution by moving from the locality in larger proportion than the poor certainly bears on the question of whether disamenities tend to fall in the first instance more heavily on the neighborhoods of the poorer groups in the economy. But even if it were the case that disamenities were introduced into initially unpolluted neighborhoods, rich and poor, in an entirely random fashion—which is implausible—it does not follow that the growth in pollution does not have regressive welfare effects. Thus if a man earning $100,000 per annum is willing to give up a maximum sum of $30,000 per annum to be rid of some particularly noxious spillover, but discovers that he is able to move out of the polluted area for a loss of about $10,000 per annum, he becomes better off than he would be if he remained in the area (to the tune of $20,000 per annum). An

2. *Malpractices.* If institutional innovations over time cause transactions costs to decline and initiative among the public to rise, there would be, under existing L law, a temptation for enterprising firms, and others in a position to do so, to produce unnecessary pollution in order to exact greater tribute from the public. This result can occur either prior to an initial agreement with the affected members of the public, or else subsequently—on the plea that market conditions have changed so radically that the existing agreement is irrelevant. Access to the detailed knowledge necessary to challenge businessmen's alleged expenditures on research and on consultations in attempting to meet public demands, or their subsequent allegations of changes in market conditions, is, if possible at all, likely to be costly and to lead to prolonged litigation.

3. *Culpability.* A part of the recent consensus was the belief that the conflict of interest entailed by an external diseconomy was symmetric in all relevant respects. The freedom of either group to pursue its interests or enjoyments necessarily interfered with the freedom of the other group. Thus, if the non-smokers' enjoyment is reduced by the smokers' freedom to smoke, so also, it is observed, is the smokers' enjoyment reduced by their abstaining for the greater comfort of the non-smokers. The question of who should compensate whom, it was occasionally stated, can be settled only arbitrarily or by reference to distributional implications.

But although they are indeed Pareto symmetric, such conflicts may not be ethically symmetric. In accordance with the

---

equally sensitive man earning only $10,000 per annum may be willing to sacrifice a maximum of $1,500 per annum to be rid of the pollution. But if the movement out of the area would involve him in a loss (or in the risk of a loss) of more than $1,500 per annum, he has to stay put and bear the full loss in his welfare.

classical liberal maxim, the freedom of a man to pursue his interests is qualified in so far as it tends to reduce the freedom, or the welfare, of others. It may then be argued that the freedom of the smoker to smoke in shared quarters is not on all fours with the freedom of the non-smoker to breathe fresh air, since the freedom to breathe fresh air does not, of itself, reduce the welfare of others. In contrast, the smokers' freedom to blow smoke into the air breathed by others does reduce their welfare. Similarly, it may be argued, the freedom to operate noisy vehicles, or pollutive plant, does incidentally damage the welfare of others, while the freedom desired by members of the public to live in clean and quiet surroundings does not, of itself, reduce the welfare of others. If such arguments can be sustained, there is a case in equity for the L law, and a case therefore for making polluters legally liable.[42]

4. *Amenity.* If, over time, transactions costs, and perhaps the costs of regulation also do not decline, the choice of L law or $\overline{L}$ law may imply that for a large class of spillovers the effective choice for society lies between "too much" spillover or "too little."[43] If the rate of growth of spillovers

equals or exceeds[44] the growth of Gross National Product, and if one assumed diminishing marginal utility of man-made goods and increasing marginal disutility of man-made "bads," the prevalence of L law will be a factor accelerating the rate at which *per capita* growth of real income approaches zero, and beyond.

5. *Posterity.* Indeed, for a range of spillovers, government regulation, intervention, or prohibition may be justified notwithstanding an apparent consensus among the groups immediately affected. The possibility that the damage being wrought by particular spillover effects is virtually irreversible has to be taken seriously in the new vision of our tiny and unique planet. In terms of man's life span, the continuing destruction of our limited resources of natural beauty, the poisoning of lakes and rivers, may be regarded as irrevocable. Consequently the losses to be suffered by future generations[45] has to be added to those carried by existing populations.

6. *Information.* If the pace of technological innovation extends the time lag between the immediate commercial exploitation of new products and processes, on the one hand, and, on the other, the knowledge of their long-term genetical and ecological effects, there is a presumption not only in favor of $\overline{L}$ law, but in favor also of direct prohibition of a number of hazardous polluting activities. There is a case, too, for public control over the adoption of new processes, and the marketing of new products, in particular, chemical products. The risks arising from insufficient knowledge of the long-term effects of any single innovation—or, indeed, the risks arising from in-

[42] This argument, if coaxed a little, might be made to take a finer turn: thus, if a switch in my demand from x to y causes either the price of x or y to rise, it obviously affects the welfare of others also. Nevertheless, considerations of equity need not ignore differences of magnitude. If the world were indeed such that a simple increase in my demand for notepaper inflicted injury on innocent families, it is likely—though the question of equity was far from clear—that broad agreement could be secured on the need to develop countervailing government mechanisms. If, however, the effects of changing tastes on relative prices were slight and random, and the costs of continually tracing them back to those responsible were prohibitive, there would be an explicit agreement or tacit understanding to ignore them for the undeniable conveniences offered by a comprehensive price system.

[43] As suggested, the adoption of $\overline{L}$ law will encourage the development of preventive technology more than will L law. The "too little" will not, then, be likely to last as long as the "too much."

[44] Which is more likely, since familiar growth industries (automobiles, motor-boats, motorized garden implements, chemicals, nuclear power, tourism, etc.) also appear prolific of spillover.

[45] Any discounting of the losses to be borne by *future* generations, moreover, cannot be justified on the usual arguments developed in the context of a single generation.

sufficient knowledge of the long-term effects of any of a number of existing products and processes—may well be thought slight. But even allowing for this more favorable contingency, as the number of such products spread over the globe—and today they tend to spread with incredible rapidity—the chance of some uncontrollable epidemic, or ecological catastrophe occurring becomes increasingly probable.

XII Epilogue

Many of the considerations brought forward in the last section do not, I recognize, lend themselves easily to analytic elegance. But with respect to environmental spillover —the most urgent economic problem of our fragile civilization—they are more pertinent than those arising from traditional allocative analysis. It is not, of course, hard to understand the somewhat exaggerated weight attached by economists to the allocative aspects of an economic problem as distinct, say, from those connected with equity. For the former aspects lend themselves nicely to formal theorizing and, with patience and a little finesse, impressive measures of social losses and gains can be foisted on credulous civil servants and a gullible public.

Yet the priority given to allocative aspects in real economic problems cannot, I think, be justified; certainly not by recourse to welfare economics. The more "affluent" a society becomes, the less important is allocative merit narrowly conceived. And in any society in the throes of accelerating technological change (one in which, of necessity, pertinent knowledge of the human, social, and ecological consequences of what we are doing is generally slight and partly erroneous) complacency on the part of any economist, guided in his professional decisions by considerations alone of allocative merit or economic growth potential, is both to be envied and deplored.

REFERENCES

1. AYRES, R. V. AND KNEESE, A. V. "Production, Consumption, and Externalities," *Amer. Econ. Rev.*, June 1969, 59(3), pp. 282-97.

2. BAUMOL, W. J. "External Economies and Second-Order Optimality Conditions," *Amer. Econ. Rev.*, June 1964, 54(3), pp. 358-72.

3. BOHM, P. *External economies in production.* Stockholm: Almquist & Wiksells, 1964.

4. ——, "Pollution, Purification and the Theory of External Effects," (Mimeograph, 1969). Also in French in *Annales de l'Insee,* 1970.

5. BUCHANAN, J. M. AND STUBBLEBINE, W. C. "Externality," *Economica,* Nov. 1962, 29, pp. 371-84.

6. BUCHANAN, J. M. AND KAFOGLIS, M. Z. "A Note on Public Goods Supply," *Amer. Econ. Rev.,* June 1963, 53(3), pp. 403-14.

7. BUCHANAN, J. M. "An Economic Theory of Clubs," *Economica,* Feb. 1965, 32 (125), pp. 1-14.

8. BURROWS, P. "On External Cost & the Visible Arm of the Law," *Oxford Econ. Pap.,* March 1970, 22(1), pp. 1-17.

9. COASE, R. H. "The Problem of Social Costs," *J. Law Econ.,* Oct. 1960, 3, pp. 1-44.

10. DAVIS, O. A. AND KAMIEN, M. I. "Externalities, Information and Alternative Collective Action" in *The analysis of public expenditures: The PPB system.* Washington: Government Printing Office, 1969, pp. 67-86.

11. DAVIS, O. A. AND WHINSTON, A. "Externalities, Welfare and the Theory of Games," *J. Polit. Econ.,* June 1962, 70 (3), pp. 241-62.

12. ——, "On Externalities, Information and the Government-Assisted Invisible

Hand," *Economica,* August 1966, *33* (131), pp. 303–18.

13. ——, "Piecemeal Policy in the Theory of Second Best," *Rev. Econ. Stud.,* July 1967, *34*(3), pp. 323–31.

14. ——, "On the Distinction between Public and Private Goods," *Amer. Econ. Rev.,* May 1967, *57*(2), pp. 360–73.

15. DOLBEAR, F. T., JR. "On the Theory of Optimal Externality," *Amer. Econ. Rev.,* March 1967, *57*(1), pp. 90–103.

16. DUESENBERRY, J. *Income, saving and the theory of consumer behaviour.* Harvard: Harvard University Press, 1949.

17. ELLIS, H. AND FELLNER, W. "External Economies and Diseconomies," *Amer. Econ. Rev.,* Sept. 1943, *23*(3), pp. 493–511. Reprinted in *Readings in price theory.* Chicago: Irwin, 1952.

18. EVANS, A. W. "Private Goods, Externality, Public Good," *Scottish J. Polit. Econ.,* Feb. 1970, *17*(1), pp. 79–89.

19. FARREL, M. J. "In Defence of Public Utility Pricing," *Oxford Econ. Pap.,* Feb. 1958, *10,* pp. 109–23.

20. GRAAFF, J. DE V. *Theoretical welfare economics.* Cambridge: University Press, 1957.

21. GREEN, H. A. J. "The Social Optimum in the Presence of Monopoly and Taxation," *Rev. Econ. Stud.,* 1962, 29(1), pp. 66–78.

22. HICKS, J. R. "The Four Consumer's Surpluses," *Rev. Econ. Stud.,* 1943, *11*(1), pp. 31–41.

23. KAHN, R. F. "Some Notes on Ideal Output," *Econ. J.,* March 1935, *45*(177), pp. 1–35.

24. KNEESE, A. V. *Approaches to regional water quality management.* Washington, D.C.: Resources for the Future, Inc., 1967.

25. KNIGHT, F. H. "Some Fallacies in the Interpretation of Social Cost," *Quart. J. Econ.,* August 1924, *37,* pp. 582–606.

26. LIPSEY, R. G. AND LANCASTER, K. "The General Theory of Second Best," *Rev. Econ. Stud.,* 1956–57, *24*(63), pp. 11–32.

27. LITTLE, I. M. D. *A critique of welfare economics.* Oxford: University Press, 1957.

28. MARGOLIS, J. "A Comment on the Pure Theory of Public Expenditure," *Rev. Econ. Statist.,* Nov. 1955, *37,* pp. 347–49.

29. MARSHALL, A. *Principles of economics* (8th ed.). London: Macmillan, 1925.

30. McGUIRE, M. C. AND AARON, H. "Efficiency and Equity in the Optimal Supply of a Public Good," *Rev. Econ. Statist.,* Feb. 1969, *51*(1), pp. 31–39.

31. MEADE, J. "External Economies and Diseconomies in a Competitive Situation," *Econ. J.,* March 1952, *62,* pp. 54–67.

32. MISHAN, E. J. "Welfare Criteria for External Effects," *Amer. Econ. Rev.,* Sept. 1961, *51*(4), pp. 594–613.

33. ——, "Second Thoughts on Second Best," *Oxford Econ. Pap.,* Oct. 1962, *14,* pp. 205–17.

34. ——, "Reflections on Recent Developments in the Concept of External Effects," *Can. J. Polit. Econ.,* Feb. 1965, *31,* pp. 3–34.

35. ——, "Pareto Optimality and the Law," *Oxford Econ. Pap.,* Nov. 1967, pp. 255–87.

36. —— "What is Producers' Surplus?" *Amer. Econ. Rev.,* Dec. 1968, *58*(5), pp. 1269–82.

37. ——, "The Relationship between Joint Products, Collective Goods, and External Effects," *J. Polit. Econ.,* May 1969, *72*(3), pp. 329–48.

38. MUSGRAVE, R. A. "Cost-Benefit Analysis and The Theory of Public Finance," *J. Econ. Lit.,* Sept. 1969, *7*(3), pp. 797–806.

39. OAKLAND, W. H. "Joint Goods," *Economica,* August 1969, *36*(143), pp. 253–68.

40. OLSON, M. *The logic of collective action.* Harvard University Press, 1965.

41. PAULY, M. V. "Clubs, Commonality and the Core: An Integration of Game Theory and the Theory of Public Goods," *Economica,* August 1967, *34* (135), 314–24.

42. PIGOU, A. C. *The economics of welfare* (4th ed.). London: Macmillan, 1946.

43. PLOTT, C. R. "Externalities and Corrective Taxes," *Economica,* Feb. 1966, *33* (129), pp. 84–87.

44. RIDKER, R. G. *Economic costs of air pollution.* New York: Praeger, 1967.

45. ROSENSTEIN-RODAN, P. W. "Problems of Industrialization of East and South East Europe," *Econ. J.,* June 1943, *53,* pp. 202–11.

46. SAMUELSON, P. A. "Contrast between Welfare Conditions for Joint Supply and for Public Goods," *Rev. Econ. Statist.,* Feb. 1969, *51*(1), pp. 26–30.

47. ———, "The Pure Theory of Public Expenditure," *Rev. Econ. Statist.,* Nov. 1954, *36,* pp. 387–89.

48. ———, "Aspects of Public Expenditure Theories," *Rev. Econ. Statist.,* Nov. 1958, *40,* pp. 332–38.

49. SCHUMACHER, E. F. "Clean Air and Future Energy," *Des Voeux Memorial Lecture,* 1967.

50. SCITOVSKY, T. "Two Concepts of External Economies," *J. Polit. Econ.,* April 1954, *62,* pp. 70–82.

51. TURVEY, R. "On Divergences between Social Cost and Private Cost," *Economica,* August 1963, *30,* pp. 309–13.

52. VINCENT, P. E. "Reciprocal Externalities and Optimal Input and Output Levels," *Amer. Econ. Rev.,* Dec. 1969, *59*(5), pp. 976–84.

53. VINER, J. "Cost Curves and Supply Curves," *Zeit. Nationalokonomie,* Sept. 1931, *3,* pp. 23–46. Reprinted in *Readings in price theory.* New York: The Blakiston Co., 1953.

54. WILLIAMS, A. "The Optimal Provision of Public Goods in a System of Local Government," *J. Polit. Econ.,* Jan. 1966, *74,* pp.

55. WELLISZ, S. "On External Diseconomies and the Government-Assisted Invisible Hand," *Economica,* Nov. 1964, *31,* pp. 345–62.

56. WORCESTER, D. A. "Pecuniary and Technological Externality, Factor Rents, and Social Costs," *Amer. Econ. Rev.,* Dec. 1969, 59(5), pp. 873–85.