Journal of Public Economics 3 (1974) 71–82. © North-Holland Publishing Company

SOCIAL COSTS AND BENEFITS AND THE TRANSFER PRICING PROBLEM

Joshua RONEN
New York University, New York, U.S.A.

First version received July 1972, revised version received July 1973

1. Introduction

This paper points out a strong similarity between the problems of dealing with externalities and the intra-firm transfer pricing problem. A solution suggested for the latter problem by Ronen and McKinney (1970) is applied to the social cost problem. Several systems have been suggested to deal with the problem of divergence between social costs and social benefits, e.g., Buchanan and Stubblebine (1962), Mishan (1971), and Turvey (1963). Many of them rest on governmental intervention via regulation or various tax methods. None of these systems, however, provides incentives to individual firms, either producing or suffering harmful side effects,[1] to supply information necessary for a social arrangement that minimizes resource misallocation. The framework discussed in this paper – which was suggested for the transfer pricing problem by Ronen and McKinney (1970) – leads to such incentives.

2. The parallel between the transfer pricing problem and the social cost problem

The transfer pricing problem arises in the context of an individual firm; the problem of divergence between social costs and private costs is this problem's counterpart at the economy level. At the individual firm level, the problem arises in the context of decentralization through divisions that operate as profit centers.[2] The need to establish the proper system of transfer prices for the decentralized profit centers is predicated primarily upon three requirements [Ronen and McKinney (1970), pp. 99–100]: First, these prices must allow central management to evaluate as accurately as possible the performance of the profit centers in terms of their separate contributions to corporate profits. Second, the system of transfer pricing must motivate profit center managers to pursue

[1]The term 'side effect' was used by Demsetz (1964) rather than either 'external effect' or 'neighborhood effect' to avoid the connotations implied by these terms.

[2]For an elaborate specification of the transfer pricing problem, see, for example, Ronen and McKinney (1970), Solomons (1965), Cook (1955), Hirshleifer (1956 and 1957), and Gould (1964).

their own self-interests in a manner which is conducive to the success of the company as a whole, and third, the system must serve as a stimulus to managers to increase their efficiency without restricting the divisions' autonomy as profit centers.

These objectives are strikingly similar to those of dealing with side effects; the same economic cost and benefit evaluations underlie the centralization/decentralization decisions within a firm, and the centralization/decentralization decisions within the economy as a whole [Coase (1960), pp. 16–17]. The high administrative costs of organizing transactions within the firm when operations are centralized, or within the economy where the government extensively intervenes, induce a firm, among other things, to decentralize, and the government to let the free market mechanism bring about a desirable equilibrium. But at the same time decentralization may lead to less than optimal value accumulation whenever there are interdependencies among the autonomous units.

At the individual firm's level, when the profit centers are not economically independent (as would be the case whenever the external market for the intermediate product is not perfectly competitive) decisions taken by the autonomous profit centers in pursuit of their self-interest will most likely result in less than optimal profits for the firm as a whole. Analogously, in an economy where producers are interdependent, decentralization may lead to less than optimal social product if the individual firms maximize private profits that diverge from social profits. Just as the intrafirm's dysfunctional decisions can be avoided by recentralization (i.e., by making the relevant decisions by central management) all industries could be integrated to eliminate divergences at the economy-wide level between private profits and public benefits.[3]

However, for both the firm and the economy levels, recentralization is not the only solution. Since recentralization could be associated with high costs, other schemes that were similar between the two levels (individual and economy-wide) were developed. The suggested schemes for the externality problem include, for example, government regulation, extending the role of the firm through integrating the entities affected by the diseconomy-creating activity [Coase (1937)], and the solution that combines the extension of the firm with combination/sales devices [Demsetz (1964)]. Also, it may be that the most efficient alternative is not to take into account some side effects since the market solutions are too costly.

These same schemes apply with slight modification to the decentralized firm's problem. The economic analysis necessary to determine the optimal amount of externality to be produced is identical in the transfer pricing [Hirshleifer (1956)], and the social cost situation [Mishan (1971) and Turvey (1963)]. The harmful side effect imposed by one firm on others corresponds to a "noxious"

[3]Thus Dr. Rodan Rosenstein (1943, p. 204) advocated that 'the whole of the industry to be created is to be treated and planned like one huge firm or trust'.

intermediate product that is transferred from one division to another within a decentralized firm in the context of transfer pricing.

None of these schemes (except for voluntary exchange) for dealing with the social cost problem provides incentive for the firms to communicate information that facilitates efficient allocation.[4] The isomorphism between the interfirm social cost and the intrafirm transfer pricing problem should lead us to attempt to apply schemes that provide such incentives in the transfer pricing case to the social cost case. A scheme that accomplishes this for the intrafirm transfer pricing problem was suggested by Ronen and McKinney (1970). It is briefly discussed below and then applied to the social cost case.

3. The analysis in the transfer pricing context

For the purpose of the discussion, I assume, with no loss of generality, a simple case of a decentralized firm with a manufacturing division and a distribution division. The manufacturing division transfers some intermediate product to the distribution division which, in turn, sells the final product in a perfectly competitive final output market.[5]

To maximize profits for the firm as a whole without unduly restricting the autonomy of divisional managers,[6] we suggested (refer to figs. 1 and 2) that the manufacturing division communicate to central management its marginal manufacturing costs curve MMC by stating how much it would produce at various transfer prices. From this, central management derives the average cost function $P'(Q)$ which is then given to the distribution division designating the actual transfer price it will be charged for alternative quantities (see AMC in fig. 2). Similarly, central management obtains from the distribution division its demand schedule ($MR - MDC$ in fig. 1), showing how much that division would purchase at various transfer prices (this equals the marginal revenue from the sale of the final product MR less the marginal distribution cost MDC). From

[4]Indeed, the problem of providing information that facilitates the choice of an appropriate policy is of primary importance, as argued by Demsetz (1966, p. 68): 'the costs and benefits of a prospective change in resource allocation cannot be treated as given datum. Marginal costs and benefit curves associated with a prospective realignment of resources are not known by the Government. Each affected individual knows his benefit or cost, and in the absence of high exchange costs this information will be transmitted to others in the form of market negotiations. The primary problem of the Government is the estimation problem. The compensation principle by its assumption that costs and benefits are known, begs the most difficult question posed by a prospective change'.

[5]To simplify the analysis, it is assumed that the divisions are technologically independent; i.e., the level of production in one division does not affect the cost of the other and a common level of output is to be reached by the two divisions (either because there exists no market for the intermediate product or because the marginal costs of either division rise sharply when dealing with an outside market).

[6]'. . . in that they are not permitted to act as monopolistic buyers or sellers where a perfectly competitive market for the intermediate product does not exist externally.' See Hirshleifer (1956).

this, an average revenue function $P^*(Q)$ (which is the final product demand D less average distribution cost ADC; see fig. 1) is given to the manufacturing division as its demand schedule designating the transfer price offered to the manufacturing division for any quantity supplied.

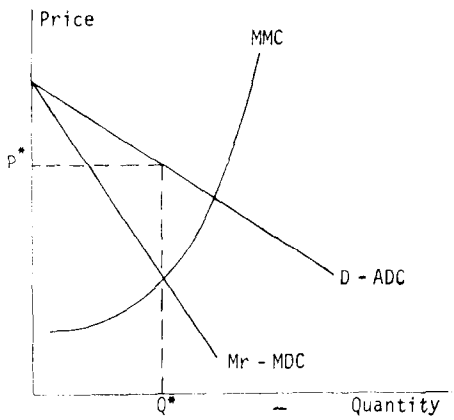

Fig. 1. Manufacturing division.

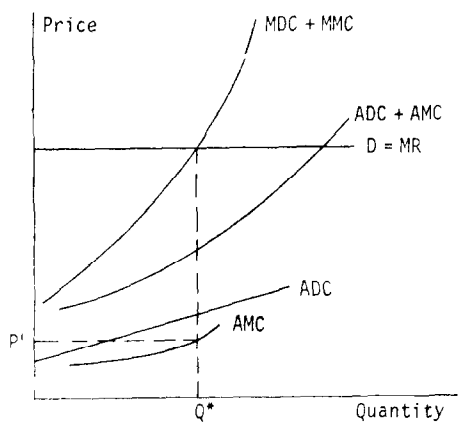

Fig. 2. Distribution division.

When the manufacturer faces the $(D-ADC)$ curve as his demand curve he derives a curve marginal to it $(MR-MDC)$ and chooses to produce the quantity optimal for the firm Q^* where MMC equals $MR-MDC$. Through a similar process, the distribution division will choose the same level of output. The distribution division is charged with the unit price $P'(Q^*)$ for the quantity (Q^*) transferred. The manufacturing unit is credited with $P^*(Q^*)$ per unit,

consisting of the payment from the distribution division plus a subsidy from central management of $[P^*(Q^*)-P'(Q^*)]\cdot Q^*$. Thus, both divisions without undue restriction of their autonomy will want to produce at the same level and will maximize their own profits as well as the firm's in so doing. Furthermore, each division's reported profit equals its contribution to the firm's profit, i.e., the amount by which the firm's profit is reduced if the division is abandoned. (In this simple case, each division's contribution to the firm's profit is identical to the total firm's profit.)

Note that the system provides incentive for communicating correct demand and supply information by the divisions to top management. Each division manager will find that by distorting the information supplied the quantity demanded from or supplied to him will differ from the quantity that maximizes his profits. For example, if we assume that the manufacturer overstates his marginal cost, and the distributor provides his correct demand curve, the latter will inevitably indicate willingness to buy a lesser quantity than the one that equates the marginal revenue of the manufacturing division with its (true) marginal cost (i.e., lesser than the seller would want to supply). This would indicate that the supplier overstated his costs. Buying from him the lesser quantity would leave him with less profit than he could earn without overstating his costs. Similarly, if the buyer understates his demand, lesser quantity will be supplied to him than he would like to buy and his profits will accordingly shrink. If both the supplier overstates his costs and the buyer understates his demand, both will find themselves trading lesser quantities than would maximize their profits.

The divisions need to merely communicate information explicitly about their costs and demand functions. This information should be available in some form since it is also essential for each division's management to make informed decisions. The function of central management is restricted in this scheme to transferring information between the divisions and paying the difference between the seller's revenues and the buyers' purchase outlays. The divisions can adapt to environmental changes by periodically reporting to central management changes in costs and demand. Whether this scheme should be implemented depends on whether the benefits of intervention by central management (which consist of eliminating the reduction in the firm's profits that result from dysfunctional decisions by the divisions in pursuit of their own self-interest) exceed the cost of communicating the information and processing it in order to effect the appropriate transfer pricing policy.

4. The analysis of the social cost problem

To illustrate the analysis of the social cost problem, I assume two entities (although the conclusion would be applicable to any number of entities) A and B. A is the firm, person or group (or persons or firms) which engages in a

diseconomy creating activity causing damages to B, another firm, person, or group. The magnitude of the damage naturally depends not only on the scale of A's activities, but also on the way B adjusts to it. The optimal allocation of resources is obtained when A's gain from the harmful activity less B's loss resulting from it is maximized (when all alternative actions are considered, including the discontinuation of the harmful activity).

Suppose that A and B are firms; then the effects the harmful activity has on their profits measure their gains or losses. In addition, assuming no serious market imperfections, these changes in profits would be the appropriate basis for the determination of a social optimum. If we also assume that each, in pursuit of its own self-interest, seeks to maximize its profits, that each knows about the available alternative activities and that it is willing to negotiate, then both will achieve the optimum without governmental intervention. They will either merge and thus internalize the harmful activity or they will reach the desired level of activity by having B pay A to modify the nature or scale of its harmful activity. When there is liability for damage, A will compensate B for the optimal amount of damage imposed by A. All these solutions are parallel to those applicable to the transfer pricing case between two divisions of a decentralized firm if these divisions were free to merge, to agree on a mutually optimal transfer price for the intermediate product or if central management were to force A to pay B the amount of optimal loss imposed as a 'transfer price' for the noxious 'intermediate product' transferred. Thus, central management can (in the transfer pricing case) specify the quantity of the intermediate product to be produced and transferred or apply a dual subsidy system to avoid suboptimization. And, similarly, the government may appropriately intervene in the social cost case and regulate the nature and the scale of the harmful activity or apply a corrective subsidy system. However, there are dissimilarities between the social tax-subsidy solution as traditionally known, and the transfer pricing system described above.

According to the traditional tax-subsidy solution at the social level the required excise tax for a good generating an external diseconomy is equal to the value of the marginal diseconomy at the optimal output, whereas the excise subsidy should equal the value of the marginal external economy at the optimal output for a socially beneficial good. But these measures are dysfunctional in that the suggested tax will reduce output below its competitive equilibrium and the subsidy will extend output beyond the optimal level.[7]

A related objection to the excise tax-subsidy solution is raised by Bohm (1970) as quoted by Mishan (1971, pp. 15–16):

> If the optimal excise tax increases with output, the firm (he 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

[7]For a discussion of this issue, see Mishan (1971), Buchanan and Stubbelbine (1962), and Turvey (1963).

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, there, 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.

It is interesting to notice the similarity between Bohm's suggestion at the social level and Hirshleifer's (1956) solution for the transfer pricing problem in that divisional managers are not permitted to act as monopolistic buyers or sellers where a perfectly competitive market for the intermediate product does not exist externally. Unfortunately, this solution, for both the transfer pricing and the social cost cases, may unduly restrict the autonomy of the manager (of the division and the firm, respectively). In addition, it does not provide incentives to communicate correct information about the cost and demand or information about the contribution of the division and of the firm to the overall profits of the firm as a whole and of the economy, respectively.[8]

It can be shown, however, that the Ronen and McKinney (1970) subsidy solution for the transfer-pricing problem can apply to the social cost case as illustrated in fig. 3.

Fig. 3 illustrates the situation faced by A. The scale of the harmful activity conducted by A as represented by the horizontal axis as well as the scale of B's losses resulting from the activity as represented in fig. 4 are assumed to be continuously variable. In fig. 3, the line MG(A) represents the marginal gain to A from the harmful activity, i.e., the net gain that results from producing the goods and services that necessitate engaging in a harmful activity [the revenue minus all the costs (including the private costs of the harmful activity) related to the product necessitating the activity harming B].

The area under this line gives the total gain to A from the harmful activity. The line ML(B) in fig. 4 represents the marginal loss to B (reduction in profit) resulting from the side effects of A's activity. The total area under this line reflects the direct loss to B as reduced by the best adjustment possible to A's activity plus the cost of making that adjustment. If A and B cannot negotiate and if no restrictions whatever are imposed on A, A would choose to engage in the harmful activity at a level OC rather than the socially optimal level OD (with net social gain of OEF) which would be obtained if A and B could merge, as illustrated in fig. 5. If A and B were to negotiate, or if A is liable to compensate B for actual damages, the optimal level OD will be produced [Mishan (1971, p. 15) or Turvey (1963, p. 311)]. If the net gain attributable to negotiation (CEH) is larger than its costs, negotiation would be socially desirable.

[8]These criticisms are discussed in more detail in Ronen and McKinney (1970).

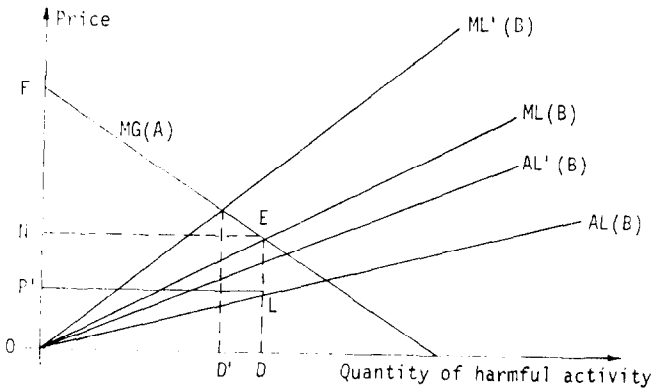

Fig. 3. Situation as viewed by A.

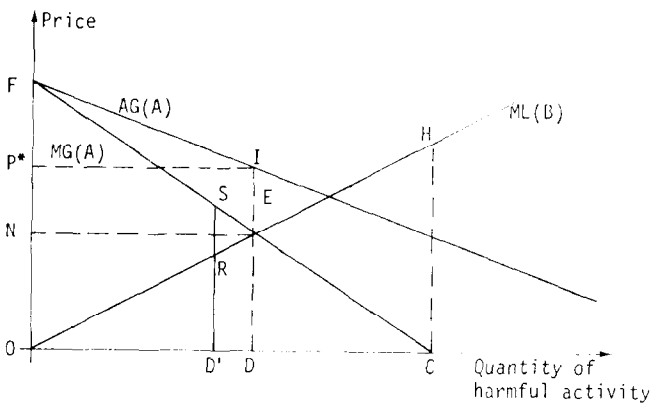

Fig. 4. Situation as viewed by B.

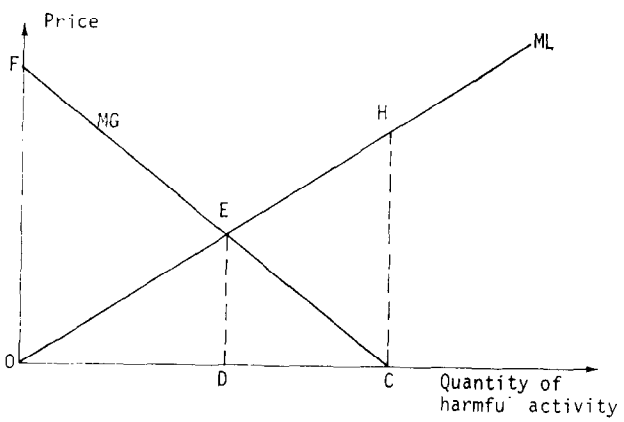

Fig. 5. A and B merged.

If the cost of government intervention through any of the alternative schemes is less than the lower of CEH or the cost of negotiation, it would be desirable from a social standpoint if the government uses the least costly scheme to induce A to produce the optimal activity level OD. The Ronen and McKinney (1970) scheme can be used by the government as well as the alternative schemes. Its cost should be compared to that of other methods of intervention (such as outright regulation, establishing liability for damages, and other tax subsidy systems). In comparing the costs, however, the benefits of our system that are absent in other schemes should be considered. As indicated, these benefits include providing much of the firm's autonomy and, in addition, obtaining measures of the contribution of the firm – inasmuch as the harmful activity is concerned – to overall social gain.

5. Application of the Ronen and McKinney solution to the social cost problem[9]

To simplify the application, I make the same assumptions as in the transfer pricing context; i.e., firms A and B (refer to figs. 3, 4, and 5) are technologically independent and a common level of harmful activity is to be reached by the two firms. The harmed firm B can be viewed as transferring an intermediate product to the harming firm A. The intermediate product is B's acceptance of A's harmful activity which is assumed to be continuously variable. The higher the level of the harmful activity accepted, the higher is B's marginal loss and A's marginal gain. Under the suggested scheme, B would be required to communicate to the governmental regulatory agency its marginal loss schedule ML(B) by stating how much harmful activity it would accept at various transfer prices. From this the agency derives the average loss function $P'(Q)$ [AL(B) in fig. 3] which is then given to A designating $P'(Q)$, the transfer price that will be charged for the acceptance of alternative quantities of harmful activity. The prices to A are then taken from this supply schedule.[10]

Similarly, the agency also obtains from A a schedule showing the prices it would be willing to pay for B's acceptance of alternative quantities of harmful activity. This schedule will reflect A's marginal gain from engaging in the harmful activity [MG(A) in figs. 3 and 4]. The agency derives the average gain from these data constructing thus a demand function $P^*(Q)$ [AG(A) in fig. 4] that is given to B designating $P^*(Q)$, the transfer price offered to B for the acceptance of any quantity of harmful activity produced by A.

For any particular quantity of harmful activity created (such as OD), A is charged $P'(OD)$ per unit. B is credited with $P^*(OD)$ per unit, consisting of the payment from A plus a subsidy from the agency of $[P^*(OD) - P'(OD)] \cdot OD$.

[9] The analysis is shown for the case of the harmful side effect activity. It can be symmetrically applied to the case of a beneficial side effect activity.

[10] These schedules reflect the actual prices paid or credited and are in no way fictitious.

These rules are illustrated in figs. 3 and 4. B faces the situation in fig. 4. The agency supplies him with the $AG(A)$ curve which determines his transfer price. He derives the curve marginal to it $[MG(A)]$, equates his marginal loss to the marginal gain and chooses to accept the quantity OD of harmful activity where $MG(A) = ML(B)$. Analogously, A faces the situation in fig. 3; the agency supplies him with the $AL(B)$ curve that determines his transfer price paid. He equates the marginal curve to this, $ML(B)$, with his marginal gain curve $MG(A)$ and chooses to pay for acceptance of OD units of harmful activity where $ML(B) = MG(A)$. Thus, both A and B want to produce and to accept the same level of harmful activity and maximize their profit and social gain by so doing.

The total receipts of the governmental agency from A would be $ODLP'$ (see fig. 3) and the agency will pay B a total of $ODIP^*$ (see fig. 4). As a result, a net subsidy amounting to the difference between the total net gain to A from his activity $ODEF$ and the total loss to B from the activity ODE will be paid to B by the agency. This net subsidy equals OEF which is the net social gain from the harmful activity.

Each firm's contributions to the social product as a result of either engaging in or accepting the harmful activity (that is, the amount by which the social product is reduced if either A does not choose to engage in the activity, or B does not accept the activity, thereby forcing A to cease it) is in fact the total contribution of the activity to the social product. The apparent profits to B from the harmful activity equal the total contribution to the social product as do those of A.[11]

As in the intrafirm transfer pricing case, it can be easily shown that both A and B would have incentives to supply correct information under this scheme. A will find that by providing incorrect information the quantity of harmful activity accepted by B will differ from the quantity that maximizes its gains. Similarly, B will find that the quantity with respect to which A would be willing to pay for its acceptance differs from what would maximize its own gains if it supplies incorrect information. For example, suppose that B is tempted to overstate its marginal loss from the harmful activity and supplies to the agency the schedule $ML'(B)$ instead of the correct schedule $ML(B)$ (see fig. 3). The agency will construct the average schedule $AL'(B)$ and provide it to A as the supply schedule for accepting the harmful activity. A will, in turn, derive the marginal schedule $ML'(B)$, equate it with $MG(A)$ and want to produce only OD' units of harmful activity. B will find that by having to accept only OD' of harmful activity, its gain is reduced by the area RES (see fig. 4). It can avoid this loss by supplying correct information about its marginal loss from the harmful activity.

[11] These rules and results can be extended to more complex situations as shown in Ronen and McKinney (1970).

6. Information requirements and conclusion

The information requirements for the administration of such a policy are identical to those arising in the transfer pricing case. Information about the marginal loss and marginal gain resulting from the activities should be obtained and communicated.

In particular, the information that firms can probably produce and communicate at relatively low cost consists of (a) the cost to a firm (lost profits) caused by a harmful side activity of another firm, (b) the cost to a firm (lost profits) that can result from either eliminating or reducing the damage inflicted on others by its own harmful activity.[12]

This communication should preferably proceed on a systematic basis since changing factors in the environment may affect the loss and the gain curve.[13] Evidently, the information must be periodically communicated unless the cost of doing so exceeds the benefit. The benefit can be measured by the net gain resulting from governmental interference, negotiation, or any other social arrangement that becomes desirable as a result of obtaining the information. The cost involves estimation of the loss by B and of the gain by A. Since presumably either A or B is the best qualified to estimate his gains and losses resulting from the activity, respectively, it is reasonable to assume that information on these parameters can be most cheaply provided by A and B. Certainly, whether in any particular case the benefit of communicating the information does not exceed the costs is an empirical question. But the proposition that the benefits do exceed the costs in the majority of the cases has a great intuitive appeal.

[12]If A and B are people rather than firms, their gain and loss must be measured as the amount of money they respectively would pay to indulge in and prevent A's activity. Alternatively, it would be quantified as the amount they respectively would require to refrain from or to endure A's activity. These amounts would be different unless the marginal utility of income is constant [Turvey (1963)].

[13]Communicating this information systematically becomes crucial in a changing environment because the cost of obtaining such information under these conditions becomes prohibitive. Thus, in discussing the traditional tax-subsidy solution, Mishan (1971, p. 15) argues: '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.'

References

Buchanan, J.M. and W.C. Stubblebine, 1962, Externality, Economica 29, 371–384.
Bohm, P., 1969, Pollution, purification and the theory of external effects, mimeo. Also in French in: Annales de l'INSEE, 1970.
Coase, R.H., 1937, The nature of the firm, Economica.
Coase, R.H., 1960, The problem of social cost, Journal of Law and Economics, 1–44.
Cook, P.W., 1955, Decentralization and the transfer-price problem, Journal of Business, 87–94.
Demsetz, H., 1964, The exchange and enforcement of property rights, Journal of Law and Economics, 11–26.

Demsetz, H., 1966, Some aspects of property rights, Journal of Law and Economics 9, 61–70.
Gould, J.R., 1964, Internal pricing in firms when there are cost of using and outside market, Journal of Business, 61–67.
Hirshleifer, J., 1956, On the economics of transfer pricing, Journal of Business, 72–84.
Hirshleifer, J., 1957, Economics of the divisionalized firm, Journal of Business, 96–100.
Mishan, E.J., 1971, The postwar literature on externaiities: An interpretative essay, Journal of Economic Literature, 1–28.
Ronen, J. and G. McKinney, 1970, Transfer pricing for divisional autonomy, Journal of Accounting Research, 99–112.
Rosenstein, R., 1943, Problems of industrialization of eastern and south eastern Europe, Economic Journal, 202–211.
Solomons, D., 1965, Divisional performance: Measurement and control (Financial Executive Research Foundation, New York).
Turvey, R., 1963, On divergence between social costs and private costs, Economica, 309–313.