DISCUSSION PAPER NO. 87

"INFORMATION, INCENTIVES, AND THE INTERNALIZATION OF PRODUCTION EXTERNALITIES"

by

Theodore Groves*

May 25, 1974

* Associate Professor of Managerial Economics
Graduate School of Management

1. Introduction

The coordination of economic agents' decisions having external effects (hereafter referred to as external decisions) has been a central issue of the theory of externalities. Two distinct approaches to this problem have been followed by most of the literature - one emphasizing the incentives agents have to coordinate or internalize their external decisions on their own and the other proposing different explicit mechanisms for achieving the coordination. In both approaches the problems of what information agents must have to achieve the coordination and how this information is to be acquired often have been ignored, assumed away, or treated in an informal manner making it difficult or impossible to verify the claims made.

The first approach emanated from the seminal paper of Coase [6]. The Coasian tradition has emphasized that, in the absence of transaction costs, since a lack of coordination of external decisions is inefficient in the Pareto sense, rational agents will, on their own, seek out each other in ways not offered by traditional markets and conclude such agreements that lead to efficiency. Thus, if one agent is damaged by a decision of another agent, the damaged party would be expected to offer the damaging agent a deal, commonly called a "bribe", to reduce the level of the damages. If the agents are sufficiently skilled in their negotiations to leave no opportunity to mutually improve unexplored, efficiency will result. A change in legal liability requiring damaging agents to compensate the damaged parties would, income effects aside, not affect the resulting coordinated decisions if the compensation rules are efficient. Thus, if income effects are negligible, bribes and damage compensation are symmetric; both lead to the same coordinated efficient decisions; only the distribution of benefits is affected. 1/

However, even though economic agents might have an incentive to negotiate efficient agreements, precisely how the negotiation process occurs and what information is required are questions which remain, especially when there are many agents and many external decisions to coordinate. Also, once transaction costs are allowed, the particular mechanism is of crucial importance in order to compare the benefits of coordination with the cost of achieving it. There is no reason to believe that a lack of coordination and a reliance on traditional market incentives to motivate decisions including those with external effects imply that the transaction costs associated with any mechanism capable of achieving efficiency are greater than the benefits to be realized. It may be that an appropriate mechanism for coordination has not yet been considered.

The second approach to the coordination of external decisions has focussed on various explicit mechanisms for achieving coordination. Although much of the discussion has been cast in the context of government rules such as tax and subsidy formulae, 2/ direct regulation, and the creation of additional markets for, say, pollution rights, 3/ the particular coercive authority of government has not been the central issue. The focus of this Pigouvian tradition has been to devise explicit behavioral rules for the economic agents, including an additional agent (e.g. "government" or "central planning board"), such that if all agents follow the prescribed rules, efficient decisions will result.

Two important questions that invariably arise with such formulations and that are often not fully explicated or explored are the informational requirements of the decision rules and the incentives the economic agents have to follow the prescribed decision rules. For example, in a typical two firm model where one firm's output adversely affects the other firm's cost (e.g. a negative externality such as pollution produced in fixed proportions with output), an appropriate tax may be levied on the output of the externality producing firm so that individual firm profit maximization will lead to optimal (joint-profit maximizing) decisions. However, since the appropriate tax depends on substantial information regarding the technological, cost, and demand conditions facing the two firms, the authority responsible for setting the tax rate must acquire this information in some manner. If, as seems most reasonable, the tax authority must depend, at least in part, on information supplied by the firms, they typically would have considerable incentives to distort or communicate false information to the taxing authority. The damaged firm would typically have an incentive to overstate the damages inflicted on it and the damaging firm would typically have an incentive to overstate the benefits of its activity which damages the other firm.

This paper follows the spirit of the Pigouvian tradition in formulating explicit mechanisms to coordinate external decisions. Considered is an n-firm production model in which some decision of the firms have external effects on other firms. In Section 2, a general scheme for coordinating the external decisions is formulated and the twin problems of the informational requirements of any mechanism and the incentives the agents have to communicate the correct information are posed. In Section 3 a specific mechanism is proposed to coordinate the external decisions that takes into account both the informational limitations and the agents' incentives to communicate correct information. Additionally, the mechanism is interpreted as a scheme both for the voluntary coordination or internalization of external decisions and for imposed coordination by a government relying on governmental coercive authority. As a voluntary scheme the mechanism may be viewed in the Coasian tradition of emphasizing the mutual self-interest of the agents to coordinate their external decisions.

2. The Coordination of External Decisions of Many Firms

Consider a collection of n firms that are interdependent in some aspects of their operations. In order to isolate for attention the interdependencies, we distinguish among three types of decisions each firm may take. First, denote by x_i a vector of firm i's "local" input and/or technique decisions; that is, each element of x_i is the quantity of some input or the level of some variable specifying a choice of technique for firm i and is such that the decision has no direct effect on any of the other firms. Second, denote by y_i a vector of firm i's "local" output decisions; each element of y_i is the quantity of some output of firm i and is also such that it has no direct effect on any of the other firms. Third, denote by z_i a vector of all those decisions of firm i that do have a direct effect on at least one other firm. Letting $z = (z_1,\ldots,z_n)$ denote the n-tuple of all firms' external decisions, the external effects of these decisions is represented by including the entire n-tuple z in every firm's production relation, and revenue and cost functions.

Each firm i has a production relation $f_i(x_i,y_i;z) \leq 0$ that defines the feasible decision choices (x_i,y_i,z_i) of the firm given the external decisions z_j of each of the other firms. Examples of external decisions affecting the production relation of firm i might be the level of pollution discharged by firm j into a river whose water is used as an input by firm i, or the quantity of some input that is public in the sense that firm j cannot exclude firm i from also consuming the input.

In addition to the production relation, associated with each firm i is a revenue function $P_i(y_i,z)$ and a cost function $C_i(x_i,z)$. The external decisions z are included as arguments of these functions to capture various types of possible pecuniary externalities.[4/] For example, if the output of firm j affects the price received by firm i, then z_j would include that output variable. Also, if firm j imposes external costs on firm i and is legally liable to firm i for these costs, then the cost and revenue functions of firms i and j would include the legally required compensation from j to i.

Assuming the firms are profit maximizing and choose their local decisions (x_i, y_i) after the level of all external decisions z are known, we may confine our attention to each firm's profit function, as a function of the external decisions z_j that is, to the functions $\pi_i(z)$ defined by:

$$(2.1) \quad \pi_i(z) = \max_{(x_i, y_i)} [P_i(y_i, z) - C_i(x_i, z)] \quad \text{subject to } f_i(x_i, y_i; z) \leq 0$$

We assume henceforth that the functions $\pi_i(\cdot)$ exist.

In the absence of any coordination of the n firms' choices of their external decisions z, the n profit functions $\pi_i(\cdot)$ define the payoff functions of an n-person non-cooperative game. A natural though not unique solution concept for this game is the Nash Equilibrium, viz. $\tilde{z} = (\tilde{z}_1, \ldots, \tilde{z}_n)$ is a Nash Equilibrium if, for every i, $\tilde{z}_i$ maximizes $\pi_i(\tilde{z}/z_i)$ over all possible z_i, where $\tilde{z}/z_i \equiv [\tilde{z}_1, \ldots, \tilde{z}_{i-1}, z_i, \tilde{z}_{i+1}, \ldots \tilde{z}_n)$.

In general, a Nash Equilibrium, even if one exists, will not maximize the joint profits of the n firms, $\sum_i \pi_i(z)$, and thus there are potential benefits to the firms individually and collectively from voluntarily co-ordinating their choices of the external decision. Additionally, if the n firms are perfectly competitive and the external effects of the decisions z

do not extend beyond the group of n firms, then external decisions that maximize joint profits are efficient from a social welfare point of view and hence there are social benefits to be gained by coordinating the choices of the external decisions through, perhaps, governmental intervention.[5/] In either case, the goal of coordination adopted here is to choose external decisions that maximize the n firms' joint profits. Henceforth, it is assumed that there exists an n-tuple of external decisions, denoted by z^*, that maximizes joint profits:

(2.2) There exists a z^* that maximizes $\sum_i \pi_i(z)$ over all possible z.

To accomplish the task of coordinating the external decisions, an additional agent, hereafter called the Center, is introduced into the model. The Center may be thought of as an agent hired by the n firms to help coordinate the choices of external decisions in the case of voluntary coordination or as a government agency in the case of imposed coordination.

Since the Center is assumed to be dependent on information acquired from the n firms, an elementary process of communication between the firms and the Center is specified. Each firm i is required to send the Center a message, denoted m_i, that is chosen from a "language" set, denoted M, of all possible messages.[6/] Upon receipt of the n-tuple of messages $m = (m_1, \ldots m_n)$, the Center selects an n-tuple of external decisions in accordance with a rule, denoted $\hat{z}(\cdot)$; that is $z = \hat{z}(m)$ are the external decisions selected by the Center if it receives the messages m.

Although there are many ways to coordinate the external decisions, the process envisioned here requires that the actual external decisions z taken by the firms are sufficiently visible or capable of being monitored at negligible cost to permit the Center to announce the external decisions chosen, $\hat{z}(m)$, to the firms and verify subsequently whether or not they have taken these decisions. Sufficiently high penalties are assumed to be assessed against a firm that takes a decision z_i different from the one, $\hat{z}_i(m)$, selected and announced by the Center. Thus any self-interested firm may be assumed to take the selected decisions. This assumption is analogous to the assumption that no firm will release more pollution than a system of standards permits or than it has purchased rights for in a system of auxilliary markets for purchase and sale of pollution rights. It is also similar to the implicit assumption of models with taxation schemes that the required tax will in fact be paid. Obviously the costs of enforcing compliance with such types of rules may not be negligible. However, consideration of enforcement costs are ignored in this paper in order to focus on the prior issue of the selection of the optimal external decisions. This is the analogous problem to the selection of an optimal pollution standard, an optimal quantity of pollution rights to create, or an optimal tax.

Now, in order for the Center's coordination to be successful, the language set M and the Center's rule $\hat{z}(\cdot)$ must be selected with reference to the objective of choosing the joint profit maximizing or optimal external decisions z^*. Since the optimal decisions z^* depend on the firms' profit $\pi_i(\cdot)$, the language set M must be sufficiently large so that for every allowable [7/] collection of profit functions $\pi_i(\cdot)$, $i = 1,\ldots,n$, there are messages m_i^* in M such that if sent, the Center's rule $\hat{z}(\cdot)$ will select the optimal decisions z^* for the particular collection of profit functions;

that is:

(2.3) Given $(\pi_1,\ldots,\pi_n)$, there exists for every i, $m_i^* \in M$ such that $z^* = \hat{z}(m^*)$ maximizes $\sum_i \pi_i(z)$.

It should be noted that the relation between the allowable profit functions $\pi_i(\cdot)$ and the messages m_i^* defines the concept of "truth" with reference to the decision rule $\hat{z}(\cdot)$. Truthful reporting or sending true messages consists of sending messages that yield optimal decisions.

A simple example may help clarify these concepts. Suppose the language set M is a set of all allowable profit functions $\pi_i(\cdot)$; that is, every message m_i in M is an allowable (profit) function of the decisions z. Further, suppose that the Center's rule $\hat{z}(\cdot)$ for selecting the external decisions is defined by:

(2.4) $\hat{z}(m)$ = that z which maximizes $\sum_i m_i(z)$.

Such a rule is optimal if the actual profit functions $\pi_i(\cdot)$ are identical with the messages m_i. Thus "truth" in this case consists of sending the Center the firm's actual profit function $\pi_i(\cdot)$, i.e. $m_i^* = \pi_i$, since $\hat{z}(m^*)$ maximizes $\sum_i m_i^*(z) = \sum_i \pi_i(z)$.

Now given a general language set M and the Center's rule $\hat{z}(\cdot)$, a firm may or may not have any reason to send the correct or truthful message m_i^* corresponding to its profit function π_i. Furthermore since the Center does not know the firm's actual profit function it is unable to ascertain whether or not the message it receives is the truthful message. The essence of the incentive problem is to provide the firms with some reason to choose truthful messages over all other messages, even when the Center is unable to verify subsequently whether or not it was sent these messages.

From the point of view of any firm, its interest in communicating its truthful message depends on the consequences to it of not sending this message. In order to provide an incentive to the firms, the Center is assumed to have the authority to levy charges or make payments to the firms in addition to its role in choosing the external decisions $\hat{z}(m)$. However, since the Center's only information consists of the messages m received from the firms, the charges or payments selected by the Center can depend, at most, on this information. Thus, the Center selects charges or payments, hereafter called <u>transfers</u>, in accordance with n rules $T_i(\cdot)$, $i = 1,\ldots,n$, each of which is a real-valued function of the n-tuple of messages m. Given the messages m, the transfer $T_i(m)$ is a charge levied against the $i^{\underline{th}}$ firm if it is negative and a payment to the $i^{\underline{th}}$ firm if it is positive.

Given the Center's decision rule $\hat{z}(\cdot)$ and transfer rules $T_i(\cdot)$, the outcome or payoff to each firm is the firm's after-transfer profits realized when the messages m are sent:

$$(2.5) \qquad \omega_i(m;T_i) \equiv \pi_i[\hat{z}(m)] + T_i(m), \; i = 1,\ldots,n.$$

Since both the decisions $\hat{z}(m)$ and the transfers $T_i(m)$ depend on the messages sent by all the firms, the consequence to firm i of sending a message m_i other than its truthful message m_i^* depends on the messages of the other firms and is given by:

$$(2.6) \qquad \omega_i(m;T_i) - \omega_i(m/m_i^*;T_i) \text{ where } m/m_i^* \equiv (m_1,\ldots,m_i^*,\ldots,m_n)$$

The incentive problem in this framework is then to find transfer rules $T_i(\cdot)$ such that each firm maximizes its own after-transfer profit (payoff) by sending its truthful message m_i^*, <u>regardless of the messages sent by the other firms</u>.[8/] Formally, we call a collection of n transfer functions $\hat{T} = \{\hat{T}_i, i = 1,\ldots,n\}$ an <u>optimal incentive structure</u> relative to the decision rule $\hat{z}(\cdot)$ and language set M, if

(2.7) m_i^* maximizes $\omega_i(m/m_i;\hat{T}_i)$ over M for any

$$m\backslash m_i \equiv (m_1,\ldots,m_{i-1},m_{i=1},\ldots,m_n) \text{ where } m_i \in M,\ j \neq i.$$

Summarizing, the problem of coordinating the external decision choices subject to the informational limitations of the Center and the incentives of the n firms is to choose (1) a language set M and a decision rule $\hat{z}(\cdot)$ such that (2.3) holds, and (2) an optimal incentive structure $\hat{T}$ relative to the decision rule $\hat{z}(\cdot)$ and language set M.[9/]

Several properties of an optimal incentive structure deserve emphasis. First, the only information required by the Center to make the transfers (and choose the decisions as well) is the messages received from the firms and furthermore the Center does not have to know whether or not it was sent the correct messages. An optimal incentive structure provides a rationale for it to assume the messages are the correct ones. Second, the transfers made according to an optimal incentive structure do not depend on the external decisions that the firms actually make. As noted above, it is assumed that the Center can monitor the firms' external decisions at negligible cost and severely penalize any firm that does not take the decisions selected by the Center. Nevertheless, no firm need worry that its transfer given by an optimal incentive structure will be affected by another firm's failure to implement the selected decision. Third, under an optimal incentive structure, the best message for any firm is completely independent of the messages the other firms are sending to the Center. Thus, a firm needs to know only what its own truthful message is and that its transfer will be computed according to an optimal incentive structure. 10/

A fourth property of an optimal incentive structure that is less desirable is that there is no guarantee that the Center's budget exactly balances; that is, that the sum of all the transfers will be identically zero, even when all firms respond to the incentives and send their truthful messages. This issue is discussed at length in the next section. As shown there, however, the Center's net balance (the sum of all charges less the sum of all payments) can be guaranteed to always be non-negative so that the financial feasibility of the Center can be assured. Also, it can be shown that in special cases, 11/ if the profit functions $\pi_i(\cdot)$ are known to have a special structure, it is possible to ensure a zero net balance for the Center. For example, for some special cases, the language set M may be taken to be a Euclidean space and an optimal incentive structure $\hat{T}$ can be exhibited with the property that the sum of all transfers $\Sigma \hat{T}_i(m)$ is a polynomial function of the vectors m_i of finite degree K less than the number of firms n. Also, any optimal incentive structure $\hat{T}$ may be modified without altering its incentive properties by adding to each transfer rule $\hat{T}_i(\cdot)$ a function of the messages, say $R_i(\cdot)$, that is constant in the i^{th} firm's message m_i, i.e.

(2.8) $R_i(m/m_i) = R_i(m/m_i')$ for all $m_i, m_i' \in M$.

Thus, since $\sum \hat{T}_i(m)$ is a polynomial function of degree K less than n, it is easy to find polynomial functions $R_i(\cdot)$ constant in m_i such that

(2.9) $\sum_i [\hat{T}_i(m) + R_i(m)] = 0$ for every $m \in M^{(n)}$.

3. A General Solution of the Coordination Problem

Given the model of Section 2 it is not difficult to solve the coordination problem posed subject to the informational limitations of the Center and the incentives of the firms. 12/ For a very general solution, the set Z of external decisions z is assumed to be a compact space and the language set M the collection of all upper semi-continuous real valued functions $m_i(\cdot)$ of z. The firms' actual profit functions π_i are assumed to be members of the set M also. Given M, the Center's rule $\hat{z}(\cdot)$ for choosing the external decisions z is defined by:

(3.1) For every $m \varepsilon M^{(n)}$, $\hat{z}(m)$ maximizes $\sum_i m_i(z)$.

Note that although each message m_i is a function of z, $\hat{z}(m)$ is an n-tuple of external decisions.

Given the language set M and the rule $\hat{z}(\cdot)$, the firms' actual profit functions $\pi_i(\cdot)$ are truthful messages, that is, $m_i^* = \pi_i$. These are, of course, not the only "truthful" messages since, for example, the addition of a constant to $m_i(\cdot)$ will not change the value $\hat{z}(m)$. Truthful messages, it will be recalled, are defined as any messages m_i^* such that

(3.2) $\hat{z}(m^*)$ maximizes $\sum_i \pi_i(z)$.

where $\pi_i(\cdot)$ is the i^{th} firm's actual profit function.

The rule $\hat{z}(\cdot)$ is easily interpreted as a rule which selects external decisions that maximize the joint reported profits of the n firms. If each firm reports its actual profit function, the rule will select the optimal external decisions. It is the purpose of an optimal incentive structure to provide the firms with a reason to tell the truth or report their actual profit functions.

Consider an incentive structure $\hat{T} = \{\hat{T}_i, i = 1,\ldots,n\}$ defined by:

$$(3.3) \quad \hat{T}_i(m) \equiv \sum_{j\neq i} m_j[\hat{z}(m)] - R_i(m), \; i = 1,\ldots,n$$

where $R_i(\cdot)$ is any real-valued function of the messages m that is constant in its i^{th} component, m_i. Note that although m_j is a function of z, $R_i(m)$ is a real number that depends only on the functions m_j and not on the values of these functions at $\hat{z}(m)$. For example, $R_i(m)$ might be defined by:

$$(3.4) \quad R_i(m) = \sum_{j\neq i} m_j(\bar{z}) \text{ where } \bar{z} \text{ is some fixed value of } z.$$

To show $\hat{T}$ is an optimal incentive structure, it is only necessary to recall that $\hat{z}(m/m_i^*)$ maximizes the joint profits $\pi_i(z) + \sum_{j\neq i} m_j(z)$. Thus, from the definition of $\omega_i(\cdot,\hat{T}_i)$ [see (2.5)],

$$\omega_i(m/m_i^*;\hat{T}_i) + R_i(m/m_i^*) = m_i^*[\hat{z}(m/m_i^*)] + \sum_{j\neq i} m_j[\hat{z}(m/m_i^*)] \geq m_i^*(z) + \sum_{j\neq i} m_j(z)$$

for all z.

In particular, the inequality holds for $\hat{z}(m/m_i)$ for all m_i. Thus

$$\omega_i(m/m_i^*;\hat{T}_i) + R_i(m/m_i^*) \geq \omega_i(m/m_i;\hat{T}_i) + R_i(m/m_i)$$

or, since $R_i(m/m_i)$ is constant in m_i,

$$\omega_i(m/m_i^*;\hat{T}_i) \geq \omega_i(m/m_i;\hat{T}_i) \quad \text{for all } m_i \text{ in } M,$$

which is the requirement for an optimal incentive structure.

The optimal incentive structure $\hat{T}$ is interpreted most easily by defining the transfers independent of the particular rule used to determine the level of the external decisions z. Specifically, for every level z, define transfer rules T_i' of z and the messages m by:

$$(3.5) \quad T_i'(z,m) \equiv \sum_{j \neq i} m_j(z) - R_i(m), \; i = 1,\ldots,n$$

where R_i is, as above, constant in m_i. Note that $T_i'(z,m)$ is also constant in m_i. That is, the message m_i affects firm i's transfer only through the selection of z.

When the Center uses the rule $\hat{z}(\cdot)$ to select the level of z, the transfers defined by T_i' are the same as those given by the optimal incentive structure $\hat{T}$:

$$(3.6) \quad T_i'[\hat{z}(m), m] = \hat{T}_i(m) \quad \text{for all } m \in M^{(n)}, \; i = 1,\ldots,n.$$

To interpret the transfer rules T_i', recall that the Center interprets each firm's message m_i as the firm's profit function. Thus, T_i' transfers to firm i the full amount of "reported" profits of the other firms less an amount, $R_i(m)$ that is independent of firm i's message. The role of the function $R_i(\cdot)$ is examined below in connection with the budget balance question.

Now, each firm's after-transfer profits in terms of the transfer rules T_i' are given by a function $\omega_i'(\cdot)$ of the decisions z, the messages m, and the rule T_i':

$$(3.7)\quad \omega_i'(z,m;T_i') = \pi_i(z) + T_i'(z,m).$$

Assuming the external decisions are real variables and all profit functions (the actual π_i and reported m_j) are differentiable, the marginal profitability of any external decisions z_{jk} (the $k^{\underline{th}}$ external decision of the $j^{\underline{th}}$ firm) and hence the value to the $i^{\underline{th}}$ firm of the marginal unit of z_{jk} is:

$$(3.8)\quad \frac{\partial \omega_i'}{\partial z_{jk}} = \frac{\partial \pi_i}{\partial z_{jk}} + \sum_{\ell \neq i} \frac{\partial m_\ell}{\partial z_{jk}} \quad .$$

However, when the Center receives the messages m_j from the firms and endeavors to maximize joint-(reported) profits, the marginal joint-profitability of the external decision z_{jk} as perceived by the Center is $\sum_{\ell=1}^{n} \frac{\partial m_\ell(z)}{\partial z_{jk}}$. Thus, if firm i reports truthfully, i.e. sends $m_i^* = \pi_i$, the Center will value the marginal unit of z_{jk} at every level of z the same as firm i. Also, when all firms communicate truthfully, each firm's after transfer profit $\omega_i'(z,m^*;T_i')$ is maximized at the same quantity $\hat{z}(m^*) = z^*$ -- the true joint profit maximizing quantity -- although there is no reason for the firms' profits to be all equal at this quantity since the amounts $R_i(m^*)$ need not be identical for all i.

Summarizing this interpretation, the optimal incentive structure may be viewed as a scheme to induce each firm to evaluate each external decision in terms of its true marginal joint-profitability $\sum_{\ell} \partial\pi_{\ell}(z)/\partial z_{jk}$.

As noted in Section 2, the Center's choice of transfer rules T_i is not forced to satisfy a budget constraint:

$$(3.9) \quad \sum_i T_i(m) = 0.$$

Thus, under any particular optimal incentive structure $\hat{T}$, the Center may run a surplus or a deficit. The magnitude of the surplus or deficit, for an optimal incentive structure of the form $\hat{T}$ given by (3.3), depends on the functions $R_i(\cdot)$ chosen, for the <u>net surplus</u> for any such $\hat{T}$ may be defined as:

$$(3.10) \quad \text{Net surplus} \equiv -\sum_i \hat{T}_i(m) = \sum_i R_i(m) - (n-1) \sum_i m_i[\hat{z}(m)].$$

Although in special cases, as noted in Section 2, it is possible to find functions $R_i(\cdot)$ that will ensure a zero net surplus, in general such functions do not exist. However, the importance of this difficulty depends on the interpretation of the model.

If the Center is viewed as a government agency imposing the coordination on the n firms with the authority to levy taxes, the budget balancing property is not of crucial significance, especially since the Center can at least guarantee that its surplus is non-negative. For example, consider the functions $R_i^o(\cdot)$ defined by:

$$(3.11) \quad R_i^o(m) \equiv \max_z \sum_{j \neq i} m_j(z), \; i = 1,\ldots,n.$$

The transfer functions $\hat{T}_i^o(\cdot)$ with this specification of $R_i(\cdot)$ are then defined by:

$$(3.12)\quad \hat{T}_i^o(m) = \sum_{j\neq i} m_j[\hat{z}(m)] - \operatorname*{Max}_{z} \sum_{j\neq i} m_j(z),\ i = 1,\ldots,n,$$

and it is obvious by inspection that the transfer of each firm is non-positive, i.e. each firm is charged or taxed. Thus, the Center's net surplus is always non-negative. This specific optimal incentive structure may be interpreted as assessing each firm for the full impact that its existence has on the optimal joint profits of all the other firms.

If the Center is interpreted as an agent hired by the n firms in an attempt to voluntarily coordinate the choice of the external decisions, the budget balance issue is more important. Even though the Center can be guaranteed a non-negative net surplus and the external decisions that maximize the joint <u>before-transfer</u> profits will be selected by the Center, the firms' <u>after-transfer</u> profits <u>may</u> be lower than what they would have been in the absence of the coordination of the external decisions. In assuring the Center a non-negative surplus, in some cases it could accumulate a positive surplus larger than the total joint-profits foregone by uncoordinated decision making by the n firms. Although the surplus could always be redistributed back to the firms in such a way that they would all be better off than with no coordination , the optimal incentive property of the scheme would be destroyed since any firm would then take into account the effect of its message on its share of the Center's surplus.

One method for avoiding this problem in cases of repeated decision periods is to require only that the Center balance its budget in the long run, permitting it to run surplusses or deficits in any one period. If this could be accomplished, then total joint-profits over many periods will be received as total after-transfer profits by the n firms. It might be additionally hoped that each firm in the long run would receive greater aggregate profits than it would have in the absence of the coordination. If this could be assured, it would not be unreasonable to expect that all the firms could agree to participate in such a voluntary coordination arrangement.

A way to formalize a long run budget balance requirement for the Center is to require that the Center's budget be balanced in expectation. To be specific, suppose in each period the n firms' true profit functions $\pi_i(\cdot)$, $i = 1,\dots,n$, are given as the realization of some fixed probability law. Further, relaxing the assumption of the Center's total ignorance, assume that the Center knows the probability distribution of the n profit functions, here a distribution over $M^{(n)}$, the n-fold product of the language set M. Given this distribution, the Center can select the functions $R_i(\cdot)$ in such a way that its expected net surplus is zero; i.e. such that

$$(3.13) \quad \sum_i E[\hat{T}_i(\pi)] = 0$$

where E is the expected value operator with respect to the probability distribution over $M^{(n)}$.

An interesting example of such functions $R_i(\cdot)$ are:

$$(3.14) \quad R_i^1(m) = \sum_{j \neq i} E[m_j(\hat{z}(m)) \mid m \backslash m_i], \; i = 1,\dots,n,$$

or the sum of the conditional expected reported profits of all the firms excepting firm i, conditioned on the reported messages of these firms. The transfer functions $\hat{T}_i^1$ for this specification of $R_i(\cdot)$ are then

$$(3.15)\quad \hat{T}_i^1(m) = \sum_{j\neq i} \{m_j[\hat{z}(m)] - E[m_j(\hat{z}(m)) \mid m\setminus m_i]\},\ i = 1,\ldots,n,$$

and may be interpreted as transferring to firm i the full amount by which its message m_i changes the Center's expectation of all the other firms', $j \neq i$, reported profits.

It is easy to verify that these transfer rules provide the Center with a zero expected net surplus. However, in any one period, the net surplus is likely to be different from zero. An initial reserve could be provided the Center to cover deficits uncovered by previous surplusses. If either consistent deficits or surplusses are realized over time, the Center could reasonably infer that its probability distribution was biased and that some adjustment would be necessary. The optimality of the transfer rules $\hat{T}_i^1(\cdot)$ would not be destroyed as long as such revisions occured sufficiently infrequently so that no firm would be likely to consider the effect of its current message on its future transfers as affected by any revisions of the probability distribution.

Since the transfer rules $\hat{T}_i^1(\cdot)$ defined in (3.15) balance the Center's budget in expectation, the expected total of all firms' after-transfer profits is greater than their expected profits if their decisions were uncoordinated. However, any one firm might expect to do worse under these rules than with no coordination at all.

For example, suppose the conditional distribution of m_i given all the other m_j is concentrated in a small subset of M; that is, knowing all the other firms' profit functions gives the Center a very close, but not exact, idea of what firm i's profit function must be. In this case, the transfer rule $\hat{T}_i^1(\cdot)$ defined by (3.15) will define a transfer to (from) firm i that is very close to zero. Thus the after-transfer profit of firm i will be very similar to its before-transfer profit $\pi_i(\hat{z}(m))$ where $\hat{z}(m)$ is the optimal external decisions selected by the Center.

In addition, suppose that firm i's own external decisions z_i (the i^{th} component of the n-tuple z) has a large negative impact on the other firms' joint profits. A major effect of the coordination might well be to choose a level of $z_i = \hat{z}_i(m)$ that significantly reduces firm i's (before-transfer) profits $\pi_i(z)$ from what firm i could achieve if it were not a party to the coordination effort. Since firm i, under the transfer rule $\hat{T}_i^1(\cdot)$ would not share in the increases in joint-profits realized through the coordination, it would likely not agree to participate in a voluntary coordination effort with rules such as these.

Thus, an optimal incentive structure, in addition to the zero expected net surplus property, should have the property that each firm can at least expect to benefit from the coordination. It would be unlikely that a voluntary coordination effort could otherwise be acceptable to all the firms. However, it is possible to find such an optimal incentive structure under the same assumptions as made for $\hat{T}^1$ above.

Suppose that, in the absence of any coordination, the external decisions chosen by the n firms would be $\bar{z}(m)$ if m is the n-tuple of the firms' profit functions. For example, $\bar{z}(m)$ might be the (a) Nash Equilibrium of the n-person noncooperative game defined by the n payoff functions $m_i(\cdot)$, $i = 1,\ldots,n$ of $z = (z_1,\ldots,z_n)$, if one exists. Next consider the transfer functions $\hat{T}_i^2(\cdot)$ defined by:

$$(3.16)\quad \hat{T}_i^2(m) \equiv \sum_{j\neq i} \{m_j[\hat{z}(m)] - E[m_j(\hat{z}(m)) \mid m\setminus m_i]\}$$

$$+ E\{m_i[\bar{z}(m)] - m_i[\hat{z}(m)] + w_i \sum_j [m_j(\hat{z}(m)) - m_j(\bar{z}(m))] \mid m\setminus m_i\}$$

$i = 1,\ldots,n$, where the w_i are fixed positive weights summing to unity. The first term is, of course, identical to the transfer specified above by $\hat{T}_i^1$ [see (3.15)]. Since $\hat{T}_i^1$ is an optimal incentive structure and the second term does not depend on m_i (having been "expected out"), the incentive structure $\hat{T}^2$ is also optimal. Further, taking the expected value of the entire expression and summing over all i demonstrates that under $\hat{T}^2$ the Center's expected net surplus is also zero.

These transfer functions may be interpreted as follows: the first term transfers to (from) firm i the full amount by which its message m_i changes the Center's expectations of all the other firms', $j \neq i$, reported profits. The last term consists of two parts; the first part transfers to (from) firm i any loss (gain) the Center expects the firm to suffer (receive) as a result of the coordination. The second part transfers to firm i a fixed share of the total expected gain in joint-profits resulting from the Center's coordination. It should be noted that the expectations defining the transfer $\hat{T}_i^2(m)$ are taken only with respect to the message m_i and are conditioned on the actual messages received from the other firms, m_j, $j \neq i$. Since the fixed weights w_i, $i = 1,\ldots,n$, are arbitrary in part, they could be selected through some type of initial bargaining process among the n firms.

With the transfer rules $\hat{T}_i^2$, every firm can expect over the long run to receive greater (after-transfer) profits than it would in the absence of any coordination. The expected after-transfer profits of firm i are:

$$(3.17)\quad E[\pi_i(m;\hat{T}_i^2)] = E[\pi_i[\bar{z}(m)] + w_i \sum_j [m_j(\hat{z}(m)) - m_j(\bar{z}(m))]]$$

or are equal to the expected non-cooperative profits of the firm plus a share of the expected total gain in joint profits realized from the coordination. Thus, a voluntary coordination effort among the n firms to hire an agent (the Center) that would use these transfer rules might be expected to be agreed to by every firm.

As a postscript, it should be emphasized that the assumption of knowledge by the Center of the probability distribution of the firms' profit functions is a very strong assumption and is somewhat contrary to the spirit of the prior discussion in this paper where the only information the firms had was their own profit function and the Center was entirely dependent on the firms for its information. For this reason, the results for the interpretation of the Center as an agent hired in a voluntary effort of the firms to coordinate their decisions must be, perhaps, viewed cautiously.

More generally, finding incentive compatible rules for voluntary coordination of external decisions seems much more difficult than finding them for imposed coordination. The key source of the difficulty is that an optimal incentive structure for voluntary coordination must compensate firms whose before-transfer profits are decreased from what they would be under no coordination if they are expected to agree to participate in the coordination effort. No compensation is necessary for the other firms since their before-transfer profits are increased by the coordination. In fact, to maintain the budgetary viability of the Center, the gainers from the coordination must provide the compensation paid to the losers. This means that optimal incentive rules for voluntary coordination cannot treat all firms symmetrically. Yet the Center, unless it has independent information regarding which firms will gain and which will lose such as the independent knowledge of the probability distribution of the firms profit functions, must rely exclusively on information provided it by the firms themselves.

An imposed coordination scheme can avoid this problem and treat all firms symmetrically since losers are not provided the option of not participating in the scheme. Furthermore, if the government is the Central agent imposing the coordination, the budgetary feasibility problem is not essential, especially since it can assure a non-negative net surplus with an optimal incentive structure such as $\hat{T}^o$ defined by (3.12).

4. Appendix: On the Uniqueness of the Truthful Message Equilibrium

In some comments on the previous sections of this paper, Professor Trout Rader raised an interesting question regarding the uniqueness of the equilibrium consisting of the truthful messages $m_i^* \equiv \pi_i$ [see (2.7)]. In particular, considering the n person game defined by the payoff functions $\hat{\omega}_i(m;T_i)$ [see (2.6) and (3.3)] where the language set M is a set of real-valued functions $m_i(\cdot)$ of the external decisions z, Rader pointed out that there are in general many non-cooperative equilibria and that there is no assurance any particular non-cooperative equilibrium $\bar{m} = (\bar{m}_1,\ldots,\bar{m}_n) \in M^{(n)}$ will yield external decisions $\bar{z} = \hat{z}(\bar{m})$ that maximize true joint profits $\sum_i \pi_i(z)$. Furthermore, Rader suggested that under stronger conditions on the functions ω_i and the message space M a uniqueness theorem could be proved.

Under the restrictions assumed in Section 3 that the set of decisions Z is compact, that every element $m_i(\cdot)$ of the language set is an upper semi-continuous function of z and that the true profit functions $\pi_i(\cdot)$ belong to this set, two results can be proved. First, as pointed out by Rader, a non-cooperative equilibrium $\bar{m}$ may exist that does not maximize true joint profits. Second, however, the n-tuple of truthful messages $m^* = (m_1^*,\ldots,m_n^*)$ is the unique (up to the addition of arbitrary constants to the functions $m_i^* \equiv \pi_i$) non-cooperative equilibrium with the additional property that each firm's truthful message m_i^* maximizes its after-transfer profits $\hat{\omega}_i(m/m_i;T_i)$ <u>for every (n-1)-tuple of the other firms' messages</u> $m\setminus m_i \equiv (m_1,\ldots,m_{i-1},m_{i+1},\ldots,m_n)$. Thus, since the addition of a constant to the truthful message $m_i^* \equiv \pi_i$ does not change the values of the external decisions $\hat{z}(m^*)$ maximizing reported joint-profits, any non-cooperative equilibrium with this additional property yields external decisions $\hat{z}(m^*)$ maximizing true joint profits.

Concerning the first result an intuitive counter-example will be given. Suppose the external decision z to be coordinated is the level of some real variable that cannot assume negative values. Under the further restriction that each firm's true profit function π_i and message m_i are strictly concave and differentiable functions of z, a zero level of the decision z will maximize reported joint profits if the sum of the reported marginal profitabilities evaluated at zero is non-positive: i.e.

$$(4.1) \quad z = \hat{z}(m) = 0 \text{ maximizes } \sum_i m_i(z) \quad \text{if } \sum_i \partial m_i(z)/\partial z \Big|_{z=0} \leq 0.$$

Now suppose that the sum of the true marginal profitabilities is strictly greater than zero: i.e.

$$(4.2) \quad \sum_i \partial\pi_i(z)/\partial z \Big|_{z=0} > 0$$

so that the true joint profit maximizing level of z is strictly positive. Now consider any n-tuple of messages, $\bar{m} = (\bar{m}_1, \ldots, \bar{m}_n)$ such that, for every i

$$(4.3) \quad \frac{\partial\pi_i(z)}{\partial z}\Bigg|_{z=0} + \sum_{j\neq i} \frac{\partial\bar{m}_j(z)}{\partial z}\Bigg|_{z=0} < 0.$$

That is, each message $\bar{m}_j$ sufficiently understates the marginal profitability of the decision z such that, even if a firm i reports the truth (sends $m_i^* \equiv \pi_i$ instead of $\bar{m}_i$) the level maximizing reported joint profits (here, $\pi_i(z) + \sum_{j\neq i} \bar{m}_j(z))$ is zero. It follows that $\bar{m}$ is a non-cooperative equilibrium that yields the zero decision, $\hat{z}(\bar{m}) = 0$, which is not true joint profit-maximizing. However, it should be noted that since the true messages $m_i^* \equiv \pi_i$ maximize after-transfer profits $v_i(m/m_i,\hat{T}_i)$ for every $m\backslash m_i$ (as proved in Section 3), including $\bar{m}\backslash\bar{m}_i$, the message m_i^* would be no worse for the firm to send than the message $\bar{m}_i$. Furthermore, in the event that the other firms are, in fact, reporting truthfully, the truthful message $m_i^* \equiv \pi_i$ is better, in general, than $\bar{m}_i$. Thus, the true message m_i^* dominates any other message, even though some other message $\bar{m}_i$ may be no worse given some particular messages of the other firms.

To prove the second result, suppose $m^o = (m_1^o,\ldots,m_n^o)$ is any non-cooperative equilibrium with the additional property that m_i^o is best for any $m\backslash m_i$:

(4.4) m_i^o maximizes $v_i(m/m_i;\hat{T}_i)$ for every $m\backslash m_i \in M^{(n-1)}$.

To be shown is that $m_i^o(z) = \pi_i(z)$ + constant, or, equivalently, that

(4.5) $m_i^o(z^1) - m_i^o(z^2) = \pi_i(z^1) - \pi_i(z^2)$ for every z^1 and z^2.

Suppose the contrary and without loss in generality that

(4.6) for some z^1 and z^2, $m_i^o(z^1) - m_i^o(z^2) > \pi_i(z^1) - \pi_i(z^2)$.

It can be shown that there exist messages from the other firms $j \neq i$ such that, if firm i's message is $m_i^* \equiv \pi_i$, the decision z^2 will maximize reported joint-profits, whereas if the firm's message is m_i^o, the decision z^1 will maximize reported joint profits. Furthermore, after-transfer profits of firm i when m_i^* is sent are greater than when m_i^o is sent, thus contradicting (4.4). A brief sketch of the proof follows.

Let P be any number satisfying

$$(4.7)\quad m_i^o(z^1) - m_i^o(z^2) > P > \pi_i(z^1) - \pi_i(z^2).$$

Define the quantities C and D by:

$$(4.8)\quad C \equiv \min\{m_i(z^1), \pi_i(z^1), m_i(z^2) + P, \pi_i(z^2) + P\}$$

$$D \equiv \max\{\max_z \pi_i(z), \max_z m_i(z)\} + \varepsilon, \text{ for some small } \varepsilon > 0.$$

The quantity D is well-defined since π_i and m_i^o are upper semi-continuous functions. Now let $\hat{m}\setminus\hat{m}_i$ be any (n-1)-tuple in $M^{(n-1)}$ such that

$$(4.9)\quad \sum_{j\neq i} \hat{m}_i(z) = \begin{cases} -C & \text{if } z = z_1 \\ P - C & \text{if } z = z_2 \\ -D & \text{otherwise} \end{cases}$$

The existence of such an $\hat{m}\setminus\hat{m}_i$ is clear. It is straightforward to verify that when firms $j \neq i$ send the messages $\hat{m}_j$, if firm i sends $m_i^* \equiv \pi_i$, z^2 is chosen and, if m_i^o is sent, z^1 is chosen i.e.

$$(4.10)\quad \hat{z}(\hat{m}/\pi_i) = z_2 \quad\text{and}\quad \hat{z}(\hat{m}/m_i^o) = z_1$$

Then using the definition of $\omega_i(m;\hat{T}_i)$ and P it follows that

$$(4.11)\quad \omega_i(\hat{m}/\pi_i;\hat{T}_i) - \omega_i(\hat{m}/m_i^o;\hat{T}_i) = \pi_i(z^2) + P - \pi_i(z^1) > 0.$$

Thus, m_i^o does not maximize $\omega_i(\hat{m}/m_i;\hat{T}_i)$ contradicting (4.4).

FOOTNOTES

1/ The symmetry property of the Coase Theorem has been a subject of much controversy in the literature. See Buchanan and Stubblebine, [4], Wellisz [21], Calabresi [5], Demsetz [8], Gifford and Stone [9], Kamien, Schwartz and Dolbear [14], Bramhall and Mills [3], Tybout [20], and Marchand and Russell [15].

2/ The literature on the classical remedy of corrective taxes and/or subsidies is voluminous. The classic work is by Pigou [17]; a later classic paper is by Meade [16].

3/ Cf. Dales [7]. More generally, the possibility of augmenting the economy with auxilliary markets was proposed by Arrow [1]; see, however, Starrett [19] for a discussion of the difficulties.

4/ See Scitovsky [18] for the distinction between "technological" and "pecuniary" externalities.

5/ Governmental coordination, of course, would not be socially desirable in cases of pecuniary externalities. In fact, voluntary coordination of the external decisions in cases of pecuniary externalities should be prohibited from a social welfare point of view.

6/ The formalization of a "language" and the communication process is based on Hurwicz [13].

7/ The allowable collection of profit functions are those admitting a solution to the joint profit maximization.

8/ Thus, the incentive problem requires that a truthful message from a firm be defined independently of the messages of the other firms. Furthermore, the n-tuple of truthful messages m^* must be a stronger equilibrium than a Nash Equilibrium of the game defined by the n payoff functions $\omega_i(m; T_i)$, $i = 1,\ldots,n$.

9/ Implicitly assumed in this model is that the Center automatically follows the rules selected; in other words, there is no consideration given to the Center's incentives to follow the rules. See Alchian and Demsetz [2] for a discussion of this issue.

10/ The firms, therefore, do not even have to know what the Center's rules are.

11/ Cf. Groves and Loeb [12] for a detailed example.

12/ The general incentive problem in this form was posed and solved by Groves [10]. In Groves and Loeb [12] the solution was applied to the case of a public input. Ledyard has extended these results to a general equilibrium model with public (consumption) goods; cf. Groves and Ledyard [11].

REFERENCES

[1] ARROW, K. "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Non-Market Allocation," in The Analysis and Evaluation of Public Expenditures: The PPB System, Joint Economic Committee of the Congress of the United States, Washington, D.C., 1969, pp. 47-64

[2] ALCHIAN, A. and H. DEMSETZ. "Production, Information Costs, and Economic Organization" American Economic Review, 62 (5), 1972, pp. 777-795.

[3] BRAMHALL, D. and E. MILLS. "A Note on the Asymmetry Between Fees and Payments," Water Resources Research, 2, 1966, pp. 615-616

[4] BUCHANAN, J. and W. STUBBLEBINE. "Externality," Economica, 29, 1962, pp. 371-384

[5] CALABRESI, G. "Transaction Costs, Resource Allocation and Liability Rules - A Comment," Journal of Law and Economics, 11, 1968, pp. 67-73

[6] COASE, R. "The Problem of Social Cost," Journal of Law and Economics, 3, 1960, pp. 1-44

[7] DALES, J. Pollution, Property, and Prices, University of Toronto Press, Toronto, 1968

[8] DEMSETZ, H. "When Does the Rule of Liability Matter," Journal of Legal Studies, 1, 1972, pp. 13-28

[9] GIFFORD, A. and C. STONE. "Externalities, Liability and the Coase Theorem: A Mathematical Analysis," Western Economic Journal, 11, 1973, pp. 260-269

[10] GROVES, T. "Incentives in Teams," Econometrica, 41, 1973, pp. 617-631

[11] GROVES, T. and J. LEDYARD. "Incentives and Public Goods: A General Equilibrium Analysis," in preparation.