DISCUSSION PAPER NO. 166

INCENTIVE COMPATIBLE CONTROL OF DECENTRALIZED ORGANIZATIONS

by

Theodore Groves

August 1975

Prepared for: Conference on Directions in Decentralized Control, Many-Person Optimization and Large-Scale Systems, September 1-3, 1975
Cambridge, Massachusetts

ORSA/TIMS Joint Meeting, Team Theory Session, Las Vegas, November 17-19, 1975

INCENTIVE COMPATIBLE CONTROL OF DECENTRALIZED ORGANIZATIONS

I. The Basic Organizational Decision Problem

Many organizational decision problems may be usefully modeled as the programming problem:

$$P: \quad \max_{x} F(x) \tag{1.1}$$

$$\text{subject to } G(x) \le 0$$

where $x \in \mathbb{R}^N$, $F : \mathbb{R}^N \to \mathbb{R}$, and $G : \mathbb{R}^N \to \mathbb{R}^K$.

However, for large organizations in which the numbers of decisions N and constraints K are very large it is often either infeasible or prohibitively expensive to accumulate under a single individual's control complete information regarding the functions $F(\cdot)$ and $G(\cdot)$ and for a single individual to solve problem P. In such cases, what one individual is unable to do may be possible for many, working together, to accomplish. The organizational design problem, in part,[1] is concerned with how to organize a multi-person organization to solve such a problem.

However, for multi-person organizations the problem of organizational control arises. This problem may be described, following Arrow [1], in two parts: (1) the definition of the operating rules or the specification of rules of behavior for the different agents to follow in making the decisions assigned to their control, and (2) the definition of the enforcement rules or the specification of rules for evaluating the individual agents that provide appropriate incentives for them to follow the prescribed operating rules.

Much research has been devoted to the elaboration of operating rules for such large organizations, including work in team theory,[2] decentralization theory,[3] and decomposition algorithms.[4] Far less work has addressed the specification of enforcement rules or the incentive problem.[5] In this paper I describe a general approach for solving the incentive problem in the organizational decision problem context or more generally for solving the control problem and summarize some of the major results obtained by myself and others following this approach.[5a]

II. The Canonical Divisional Form of the Decision Problem

Given the organizational decision problem P and any specified number of organization members, an organizational form[6] is defined by an assignment of the N decisions to the various members and a specification of the *a priori* information available to each member--that is, the information each member has prior to any communication among the members. In this paper we assume the organization consists of I+1 members (agents)-- I divisional managers, i=1,...,I, and a headquarters manager or Center, i=o--; that each agent i controls the decisions x_i where $x = (x_o, x_1, \ldots, x_I) \in \mathbb{R}^N$; and that with the assignment of decisions the problem P can be written as:

$$\text{D:} \quad \max_{\{x_i\}} \Sigma_{i=1}^{I} f_i(x_i, x_o) + f_o(x_o) \tag{2.1}$$

$$\text{subject to} \quad g_i(x_i, x_o) \le 0 \quad , \; i = 1, \ldots, I$$

$$g_o(x_o) \le 0$$

where $x_i \in \mathbb{R}^{N_i}$, $\Sigma_{i=o}^{I} N_i = N$, $g_i : \mathbb{R}^{N_i + N_o} \to \mathbb{R}^{K_i}$, $i = 1, \ldots, I$, $g_o : \mathbb{R}^{N_o} \to \mathbb{R}^{K_o}$, and $\Sigma_{i=o}^{I} K_i = K$. We also assume *a priori* information is dispersed or decentralized such that each agent i knows only the functions $f_i(\cdot)$ and $g_i(\cdot)$.

Although any problem of form P can be written in the canonical form D by merely setting $N_o = N$ and $K_o = K$, i.e. assigning all decisions to the Center, the a priori information assumption clearly restricts the model. Given a particular structure of the a priori information it may not be possible to assign the decisions to the agents in a manner implied by form D. Also of considerable importance is the requirement that if any of the N decisions enters as a non-trivial argument of more than one of the functions f_i or g_i, then it must be assigned to the Center, i.e. be a component of the vector x_o. Although a priori information may be completely decentralized, this requirement limits the degree of decentralization of decision making.

Despite the restrictions imposed on the model by canonical form D, the model is quite general and, by augmenting the number of decisions and constraints, may be applicable for situations not appearing to satisfy the restrictions at first glance. For example, consider the widely discussed simple additive decomposition model:[7]

$$A: \quad \max_{\{x_i\}} \Sigma_{i=1}^{I} F_i(x_i) \qquad (2.2)$$

$$\text{subject to} \quad \Sigma_{i=1}^{I} G_i(x_i) \leq K$$

where x_i is an activity vector of division i, K is a vector of resources available to the organization, $G_i(x_i)$ is the resource requirements of activities x_i, and $F_i(x_i)$ is the i^{th} division's profitability at the level x_i. Typically it is assumed that a priori the Center knows only the available resources K and each division i knows only the functions $G_i(\cdot)$ and $F_i(\cdot)$. Since the decisions x_i and x_j enter the same constraint, the requirement of form D would imply that all decisions be assigned to the Center, i.e. that is, $x_o = (x_1,\ldots,x_I)$. But this would violate the a priori information assumption since the Center does not know $F_i(\cdot)$ or $G_i(\cdot)$.

However, by adding I new decisions, x_{oi}, $i = 1,\ldots,I$, (where each x_{oi} is of the same dimensionality as the vector K) and adding the constraint $\Sigma_{i=1}^{I} x_{oi} \leq K$, it is easy to show that problem A is equivalent to:

$$\text{B:} \quad \max_{\{x_i, x_o\}} \Sigma_{i=1}^{I} F_i(x_i) \equiv \Sigma_{i=1}^{I} f_i(x_i) \tag{2.3}$$

$$\text{subject to} \quad \begin{array}{l} G_i(x_i) - x_{oi} \equiv g_i(x_i, x_{oi}) \leq 0 \\ \Sigma_{i=1}^{I} x_{oi} - K \equiv g_o(x_o) \leq 0 \end{array}$$

where $x_o = (x_{o1},\ldots,x_{oI})$. Furthermore, problem B is precisely of form D and satisfies the a priori information restrictions. The additional I decisions are, of course, allocations of the available resources K to the I divisions.

III. Defining the Control Problem

As mentioned above, the multi-person organizational control problem consists of two parts: (1) the choice of operating rules and (2) the choice of enforcement rules. However, given any set of enforcement rules or, more specifically, rules for evaluating the I divisional managers, the divisional manager's operating rules are automatically defined by assuming the managers strive to take whatever decisions will maximize their evaluation measures. Thus, the fundamental control problem is to find enforcement rules for the divisional managers and operating rules for the Center such that, when the managers maximize their evaluation measures and the Center follows its operating rules, all the agents' decisions solve the decision problem D.

Now under the _a priori_ information assumptions for Problem D, since the Center knows only the functions $f_o(\cdot)$ and $g_o(\cdot)$, in order to compute evaluation measures and its own decisions x_o nonarbitrarily, it must acquire some information from the divisions. This requirement complicates the control problem since the evaluation rules must not only induce the managers to take optimal decisions x_i but also to communicate optimally. This is frequently referred to as the "revelation problem" or the problem of inducing the managers to communicate "truthfully" or "honestly"--i.e. to send messages the operating rules of the Center require for optimal decisions to be taken.

Thus, to define the control problem a communication process must be formalized. In this paper we consider a very general abstract communication process.[8] Let M denote an abstract set called a _language_; each element m_i in M denotes a possible message division manager i can send the Center. An element m_i may denote a single message sent at one time to the Center or an entire sequence of messages sent in a lengthy iterative communication procedure.[9] In any case, the Center acquires an I-tuple $m = (m_1,\ldots,m_I)$ of divisional messages which it uses to compute its decisions x_o.

Concerning the sequencing of message, decision, and evaluation operations, in this paper we assume that the messages m_i from the divisional managers are sent first; next, the Center computes its decisions x_o; then, the division managers compute their decisions x_i; and finally, the Center computes the divisional evaluation measures. The reason for specifying that the Center makes its decision first is that this permits the division managers to base their own decisions on the value of x_o selected by the Center.[10]

At the time the Center takes its decisions x_o, its only information is the I-tuple of divisional messages m and its <u>a priori</u> knowledge of $f_o(\cdot)$ and $g_o(\cdot)$. Thus, any decision rule for the Center is a function $x_o(\cdot)$ of $m \in M^I$ (suppressing $f_o(\cdot)$ and $g_o(\cdot)$ as arguments). However, at the time the Center computes the evaluation measures all agents have taken their decisions and hence it is reasonable to allow the division's evaluations to depend on the <u>realized</u> values of the division's contributions to total payoff, $f_i(x_i,x_o)$, $i = 1,\ldots,I$, as well as the other information of the Center--the messages $m \in M^I$ and the functions $f_o(\cdot)$ and $g_o(\cdot)$.

Although the evaluation measure for any division i may depend on the realized contributions to total payoff of other divisions, $f_j(x_j,x_o)$, $j \neq i$, accountants have emphasized that for motivational and other reasons a manager's evaluation should be based on controllable performance only--that is, on performance attributable or responsive to the particular manager's decisions and not other managers' decisions.[11] While this is a vague prescription and is achievable in the strictest sense only when the Center has no decisions to make (i.e. the divisions are completely independent),[12] we will interpret the dictum by restricting the evaluation measures to depend only on the individual divisions' realized contributions and the joint message of all division managers. Thus, any <u>evaluation measure</u> for division i is a real-valued function E_i of the joint message $m = (m_1, \ldots m_I) \in M^I$ and the i^{th} division's realized contribution to total payoff $f_i(x_i, x_o)$. Not allowing a division's evaluation measure to depend on other divisions' realized contributions eliminates using "profit-sharing" as an evaluation measure.[13]

Summarizing, a <u>control mechanism</u> is defined to be a triple $C = \{M, x_o(\cdot), \langle E_i(\cdot)\rangle_{i=1}^{I}\}$ consisting of:

a) a language M,

b) a decision rule for the Center $x_o : M^I \to \mathbb{R}^{N_o}$, and

c) I divisional evaluation measures

$E_i : \mathbb{R} \times M^I \to \mathbb{R}$ where the first argument of $E_i(\cdot)$ is the i^{th} division's realized contribution to total payoff, $f_i(x_i, x_o)$.

Given a control mechanism $C = \{M, x_o(\cdot), \langle E_i(\cdot)\rangle_i\}$, the i^{th} division's evaluation measure depends on its own decisions x_i and the I-tuple of all divisions' messages m ; $E_i[f_i(x_i, x_o(m)); m]$. The i^{th} manager is assumed to choose his message m_i and decision x_i in an effort to maximize this measure.

Now, <u>a priori</u> the i^{th} manager knows $f_i(\cdot)$ and $g_i(\cdot)$. Additionally we assume he knows the control mechanism C or at least the language M, decision rule $x_o(\cdot)$, and his own evaluation measure $E_i(\cdot)$. Further we assume that once the Center chooses its decisions $x_o(m) = x_o$ all divisions are informed of the Center's choice. (This information is, of course, not available to the divisions when they choose their messages.)

Since an arbitrary evaluation measure E_i will depend on the messages of all the divisions, m_j, $j = 1,\ldots,I$, the i^{th} division manager's best message m_i and decision x_i may depend on which messages the other divisions choose. However, for an optimal control mechanism, we require that each division best decisions x_i depend only on the decisions taken by the Center, $x_o(m)$, and his best message, m_i, be independent of the other divisions' messages. Thus, for an optimal control mechanism, a division manager needs no information about the other divisions' messages (nor, of course, about their decisions either).[14]

Now it is possible to find control mechanisms for which a divisions' best message is not unique--even though they all lead to optimal decisions. Although choosing one of these best messages rather than another will have no effect on the particular division, it may affect the choice of the Center's decisions (in cases of multiple optima) and the value of the other divisions' evaluation measures. Since there would be no way to know which among multiple best messages a division manager would send nor to induce him to choose one rather than another, such control mechanisms are capricious in terms of the evaluations of the divisions. To eliminate such capriciousness, we require, for an optimal control mechanism, that if there are multiple best messages for the divisions then all lead the Center to pick the same decisions and lead to the same value of the evaluation measures for all divisions.

Summarizing, then, we define an optimal control mechanism by:[15]

Definition: An optimal control mechanism, denoted $\hat{C}$, is a control mechanism $\{\hat{M},\hat{x}_o(\cdot),\langle\hat{E}_i(\cdot)\rangle_i\}$ such that:

a) it is <u>decisive</u>: for each division $i = 1,\ldots,I$, there exists a decision rule $\hat{x}_i(\cdot)$ (a function of x_o) and a message $\hat{m}_i \in \hat{M}$ that maximize

$$\hat{\omega}_i[x_i(\hat{x}_o(m)),m] \equiv \hat{E}_i[f_i(x_i(\hat{x}_o(m)),\hat{x}_o(m)),m]$$

subject to $g_i(x_i(\hat{x}_o(m)),\ \hat{x}_o(m)) \leq 0$

for every $m \setminus m_i \equiv (m_1,\ldots,m_{i-1},m_{i+1},\ldots,m_I) \in \hat{M}^{I-1}$;

b) it is <u>efficient</u>: for all I-tuples $\langle(\hat{m}_i,\hat{x}_i(\cdot))\rangle_{i=1}^{I}$ satisfying a), the resulting decisions $(\hat{x}_o(\hat{m}),\langle\hat{x}_i(\hat{x}_o(\hat{m}))\rangle_i)$ solve the organizational decision problem D; and

c) it is <u>non-capricious</u>: if $\langle(\hat{m}_i,\hat{x}_i(\cdot))\rangle_i$ and $\langle(\hat{m}_i',\hat{x}_i'(\cdot))\rangle_i$ both satisfy a), then

$$\hat{x}_o(\hat{m}) = \hat{x}_o(\hat{m}')$$

and $\hat{\omega}_i[\hat{x}_i(\hat{x}_o(\hat{m})),\hat{m}] = \hat{\omega}_i[\hat{x}_i'(\hat{x}_o(\hat{m})),\hat{m}']$ for all i.

The organizational control problem may be then simply stated as the problem of finding an optimal control mechanism given the decision problem D.

In order to have a meaningful problem we assume henceforth that problem D satisfies the following regularity conditions that are sufficient to guarantee the problem has a solution.

<u>Regularity Conditions</u>

A.1: The functions f_j, $j = 0,1,\ldots,I$ are upper semi-continuous functions and the functions g_j, $j = 0,1,\ldots,I$ are continuous functions;

A.2: The sets $X_o \equiv \{x_o | g_o(x_o) \leq 0\}$, $X_i \equiv \{x_i | g_i(x_i,x_o) \leq 0$ for some $x_o \in X_o\}$ for all i, and $X_i(x_o) \equiv \{x_i | g_i(x_i,x_o) \leq 0\}$ for every x_o are compact;

A.3: $X_o^i \equiv \{x_o | g_i(x_i,x_o) \leq 0$ for some $x_i \in \mathbb{R}^{N_i}\}$ is closed for all i;

A.4: $\bigcap_{i=1}^{I} X_o^i \cap X_o \neq \emptyset$.

Under (A.1) - (A.4) problem D has a solution:

<u>Proposition 1</u>: Under (A.1) - (A.4)

$$\pi_i(x_o) \equiv \text{Max}\{f_i(x_i,x_o) | x_i \in X_i(x_o)\}$$

is an upper semi-continuous function on $X_o^i \cap X_o$.

<u>Proof</u>: It is straightforward to verify that $X_i(\cdot)$ is a non-empty upper semi-continuous mapping on $X_o^i \cap X_o$. Then, since $f_i(\cdot)$ is upper semi-continuous, $\pi_i(\cdot)$ is also, by Berge's theorem [3 , Theorem 2, p. 116].

<u>Proposition 2</u>: Under (A.1) - (A.4), problem D has a solution.

<u>Proof</u>: By Proposition 1, $\Sigma_{i=1}^{I} \pi_i(x_o) + f_o(x_o)$ is an upper semi-continuous function on $\bigcap_i X_o^i \cap X_o$ which is a compact set. Thus, $\Sigma_i \pi_i(\cdot) + f_o(\cdot)$ attains a maximum, say $\hat{x}_o$, on $\bigcap_i X_o^i \cap X_o$. Furthermore, by definition of π_i, $f_i(x_i,\hat{x}_o)$ attains its maximum on $X_i(\hat{x}_o)$ at, say, $\hat{x}_i$. Note that $\langle \hat{x}_j \rangle$ satisfies the constraints of problem D.

Let $\langle x_j \rangle$ be any other decisions satisfying the constraints. Thus, since $g_i(x_i,x_o) \leq 0$ for all i, $x_o \in X_o^i$ for all i. Also, since $g_o(x_o) \leq 0$, $x_o \in X_o$. Thus $x_o \in \bigcap_i X_o^i \cap X_o$. In addition, since $g_i(x_i,x_o) \leq 0$, $x_i \in X_i(x_o)$. But, by the definition of $\pi_i(\cdot)$, $\pi_i(x_o) \geq f_i(x_i,x_o)$.

Thus

$$\Sigma\pi_i(x_o) + f_o(x_o) \geq \Sigma_i f_i(x_i,x_o) + f_o(x_o).$$

Also, by the definition of x_o and x_i

$$\Sigma_i f_i(\hat{x}_i,\hat{x}_o) + f_o(\hat{x}_o) = \Sigma_i \pi_i(\hat{X}_o) + f_o(\hat{x}_o) \geq \Sigma_i \pi_i(X_o) + f_o(x_o).$$

Hence

$$\Sigma_i f_i(\hat{x}_i,\hat{x}_o) + f_o(\hat{x}_o) \geq \Sigma_i f_i(x_i,x_o) + f_o(x_o)$$

for all $\langle x_j \rangle$ satisfying the constraints. Thus $\langle \hat{x}_j \rangle$ solves problem D.

IV. A General Solution of the Control Problem

IV.1 A Class of Optimal Control Mechanisms

In this section a particular control mechanism is defined which is then proved to be optimal. Further, we prove following Green and Laffont [10] that given the mechanism's language and decision rule for the Center the evaluation measures are, in a sense, unique. Finally some extensions and limitations of these results are discussed.

To begin, the language M^* of the control mechanism is defined as:

$$M^* = \{m_i : \tilde{X}_o^i \rightarrow \mathbb{R} \mid \bar{x}_o \in \tilde{X}_o^i \text{, a closed subset of } \mathbb{R}^{N_o}, m_i \text{ is an upper semi-continuous function on } \tilde{X}_o^i\} \quad (4.1a)$$

where $\bar{x}_o \in \bigcap_{i=1}^{I} X_o^i \cap X_o$ (see Regularity condition A.4).

Thus to specify the language it is necessary to know a priori any feasible decision vector $\bar{x}_o$ for the Center. This seems a weak requirement. A message $m_i \in M^*$ is interpreted as a reported maximal divisional profit function, i.e. $m_i(x_o)$ is the amount of divisional profit i reports he will contribute to the organization if the Center takes the decisions x_o

and i takes his maximal decisions x_i given x_o. The domain $\tilde{X}_o^i$ is interpreted as the set of Center's decisions x_o that admit a feasible and maximizing decision x_i.

If we define a division's "true" maximal divisional profit function π_i by:

$$\pi_i(x_o) \equiv \text{Max}\{f_i(x_i,x_o) | x_i \in X_i(x_o)\} ,$$

Proposition 1 ensures that π_i is defined (at least) over $X_o^i \cap X_o$ and is also u.s.c. on this set. Furthermore $\bar{x}_o \in X_o^i \cap X_o$. Thus, interpreting $\tilde{X}_o^i$ as the largest closed domain over which π_i is u.s.c., π_i is thus an element of M^*.

Given this language, the Center's decision rule $x_o^*(\cdot)$ is defined by:

$$x_o^*(m) = x_o^* \quad \text{maximizes } \Sigma_{i=1}^I m_i(x_o) + f_o(x_o) + \text{constant} \tag{4.1b}$$

$$\text{subject to } x_o \in X_o \equiv \{x_o | g_o(x_o) \le 0\} ,$$

for any constant.

Since $\bar{x}_o \in \bigcap_i \tilde{X}_o^i \cap X_o$, $\tilde{X}_o^i$ is closed, and X_o is compact, $\bigcap_i \tilde{X}_o^i \cap X_o$ is a non-empty compact set. Further, since m_i, all i, and f_o are u.s.c. functions, $\Sigma_i m_i(x_o) + f_o(x_o)$ attains a maximum on $\bigcap_i \tilde{X}_o^i \cap X_o$. Although this maximum may not be unique, the rule $x_o^*(m)$ picks a specific maximizer, and furthermore picks the same maximizer for all m' such that $m_j' = m_j$ + constant. The decision rule $x_o^*(\cdot)$ is obviously interpreted as the rule maximizing the total organization's reported profits.

Suppose, in particular, that the divisions all send their "true" maximal profit functions π_i and the Center takes the decisions $\hat{x}_o \equiv x_o^*(\pi)$. Since $\hat{x}_o \in X_o^i \cap X_o$ for every i, $f_i(x_i,\hat{x}_o)$ attains its maximum at some point (not necessarily unique), say, $\hat{x}_i$ in $X_i(\hat{x}_o)$; that is $\pi_i(\hat{x}_o) = f_i(\hat{x}_i,\hat{x}_o)$. The proof of Proposition 2 then establishes immediately that the decisions $\langle \hat{x}_j \rangle_{j=o}^{I}$ are a solution to problem D.

Thus, if the divisions send their "true" profit functions π_i and then, given the Center's decisions $\hat{x}_o \equiv \hat{x}_o^*(\pi)$, maximize their own profits $f_i(x_i,\hat{x}_o)$ (subject to feasibility), the resulting decisions solve problem D. It is the role of the evaluation measures E_i^* to induce this behavior, i.e. give the division managers an incentive to send these messages and take these decisions. Note again that because of the sequencing of decisions, manager i will know the Center's decisions $x_o = x_o^*(m)$ at the time he must make his decisions x_i and thus is able to maximize $f_i(x_i,\hat{x}_o)$ subject to $x_i \in X_i(\hat{x}_o)$ if $X_i(\hat{x}_o)$ is non-empty.

Finally, the evaluation measures $E_i^*(\cdot)$ are defined in terms of:

$$E_i^o(f_i,m) = f_i + \Sigma_{j\neq i}[m_j(x_o^*(m)) - m_j(\bar{x}_o)] + [f_o(x_o^*(m)) - f_o(\bar{x}_o)], \text{ all } i \quad (4.1c)$$

The measure E_i^o thus evaluates the i^{th} manager on the basis of his division's realized profits plus the sum of the deviations of expected reported profits from reported profits at $\bar{x}_o$ of all the other divisions and the deviation of the expected profits of the Center from profits at $\bar{x}_o$ when the Center takes the decision $x_o^*(m)$. However, as is shown below, this particular evaluation measure is only one of many optimal evaluation measures. Specifically we define the class $\mathcal{E}^*$ of evaluation measure $\langle E_i^*(\cdot)\rangle$ by:

$$\mathcal{E}^* \equiv \{\langle E_i^*(\cdot)\rangle_i \,|\, E_i^*(f_i,m) = \alpha_i(m\setminus m_i)E_i^o(f_i,m) + \beta_i(m\setminus m_i)\} \quad (4.1d)$$

where $\alpha_i(m\setminus m_i)$ is any strictly positive and β_i is any arbitrary function of all divisions' messages except the i^{th} that is constant on the sets $\{m'\setminus m_i' \,|\, m_j' = m_j + \text{constant}\}$

where $m \setminus m_i \equiv (m_1, \ldots, m_{i-1}, m_{i+1}, \ldots, m_I)$. An interesting particular member of the class $\mathcal{E}^*$ is given by setting $\alpha_i(m \setminus m_i) = 1$ for all $m \setminus m_i$, and

$$\beta_i(m \setminus m_i) = \Sigma_{j \neq i}(m_j(\bar{x}_o) - m_j(x_o^i)) + f_o(\bar{x}_o) - f_o(x_o^i)$$

where $x_o^i = x_o^i(m \setminus m_i)$ maximizes $\Sigma_{j \neq i} m_j(x_o) + f_o(x_o)$

subject to $g_o(x_o) \leq 0$.

Then

$$\bar{E}_i(f_i, m) = f_i - \{\Sigma_{j \neq i} m_j(x_o^i) + f_o(x_o^i) - [\Sigma_{j \neq i} m_j(x_o^*(m)) + f_o(x_o^*(m))] \quad . \tag{4.2}$$

and it is easy to see that $\bar{E}_i$ is a member of $\mathcal{E}^*$ since x_o^i maximizes $\Sigma_{j \neq i} m_j(x_o) + f_o(x_o)$ + constant, subject to $g_o(x_o) \leq 0$ for any constant. The measures $\bar{E}_i(f, m)$ thus evaluate the i^{th} division on the basis of his division's realized profits less the total expected impact on all the other divisions' reported and the Center's realized profits attributable to the i^{th} division's message. In the special case when division i's message does not affect the Center's decisions, they will be $x_o^i(m \setminus m_i)$ and division i will be evaluated solely on its own realized profits. Otherwise it is assessed for the expected reported impact on the rest of the organization by causing the Center to choose $x_o^*(m)$ instead of $x_o^i(m \setminus m_i)$.

Given a control mechanism $C^* = \{M^*, x_o^*(\cdot), \langle E_i^*(\cdot) \rangle_i\}$ defined by (4.1a-d) where $\langle E_i^*(\cdot) \rangle$ is in the class $\mathcal{E}^*$ of evaluation measures, the decision problem confronting the i^{th} devision manager is:

D_i^*: Choose a decision rule $x_i^*(\cdot)$ as a function of the Center's decisions x_o and a message m_i^* such that

$$\omega_i^*[x_i(x_o^*(m/m_i)), m/m_i] \equiv E_i^*[f_i(x_i(x_o^*(m/m_i))\ x_o^*(m/m_i))$$

is maximized subject to $g_i[x_i(x_o^*(m/m_i)),\ x_o^*(m/m_i)] \leq 0$, i.e. $x_i(x_o^*(m/m_i)) \in X_i(x_o^*(m/m_i))$, for every $m \setminus m_i \in M^{*I-1}$.

As shown above, if $\hat{m}_i = \pi_i$ and $\hat{x}_i(x_o)$ maximizes $f_i(x_o,x_o)$ subject to $x_i \in X_i(x_o)$ for all $x_o \in X_o^i$, then the decisions $(x_o^*(\hat{m}), \langle\hat{x}_i(x_o^*(\hat{m}))\rangle_i)$ solve problem D. Thus if $(\hat{m}_i,\hat{x}_i(\cdot))$ is the only solution of the i[th] manager's decision problem, then $\{M^*, x_o^*(\cdot), \langle E_i^*(\cdot)\rangle_i\}$ would be optimal. However $(\hat{m}_i,\hat{x}_i(\cdot))$ as defined above are not the unique solutions of problem D_i^*. So instead we first characterize all solutions of D_i^* (Theorem 1) and then show that all such solutions yield decisions solving problem D (Theorem 2). These results will then establish the optimality of any C^* in $\mathcal{C}^* = \{\{M^*, x_o^*(\cdot), \langle E_i^*(\cdot)\rangle_i\} \mid \langle E_i^*(\cdot)\rangle_i \in \mathcal{E}^*\}$ (Theorem 3):

The following theorem characterizes all solutions of D_i^*:

<u>Theorem 1</u>:[16] The pair $(m_i^*, x_i^*(\cdot))$ maximizes

$$\omega_i^*[x_i(x_o^*(m)),m] \text{ subject to } x_i(x_o^*(m)) \in X_i(x_o^*(m))$$

for every $m\backslash m_i \in M^{*I-1}$ if and only if

(a) $x_i^*(x_o)$ maximizes $f_i(x_i,x_o)$ on $X_i(x_o)$ for every $x_o \in X_o^i \cap X_o$

and

(b) $m_i^*(x_o) = \pi_i(x_o)$ + constant for every $x_o \in X_o^i \cap X_o$ and is undefined on $(X_o^i)^c \cap X_o$, where $(X_o^i)^c$ is the complement of the set X_o^i.

To prove Theorem 1 we first establish two Lemmata:

<u>Lemma 1</u>: The decision rule $x_i^*(\cdot)$ maximizes

$$\omega_i^*[x_i(x_o^*(m)),m] \text{ subject to } x_i(x_o^*(m)) \in X_i(x_o^*(m))$$

for every m such that $X_i(x_o(m))$ is non-empty, i.e. $x_i^*(\cdot) = \hat{x}_i(\cdot)$ if $x_i^*(x_o)$ maximizes $f_i(x_i,x_o)$ subject to $x_i \in X_i(x_o)$ for every $x_o \in X_o^i \cap X_o$.

Furthermore, if $x_i^*(\hat{x}_o)$ does not maximize $f_i(x_i,\hat{x}_o)$ subject to $x_i \in X_i(\hat{x}_o)$ for some $\hat{x}_o \in X_o^i \cap X_o$, then there exists some $\hat{m} \in M^{*I}$ such that $X_i(x_o^*(\hat{m}))$ is not empty and

$$\omega_i^*[\hat{x}_i(x_o^*(\hat{m}),\ \hat{m}] > \omega_i^*[x_i^*(x_o^*(\hat{m}/m_i),\ \hat{m}/m_i]$$

for all $m_i \in M^*$ such that $X_i(x_o^*(\hat{m}/m_i))$ is not empty.

<u>Proof of Lemma 1</u>: (first part - if) Let m be such that $X_i(x_o^*(m))$ is not empty, i.e. $x_o^*(m) \in X_o^i \cap X_o$. Then

$$\omega_i^*[x_i^*(x_o^*(m)),m] - \omega_i^*[x_i(x_o^*(m)),m]$$

$$= \alpha_i(m\backslash m_i)[f_i(x_i^*(x_o^*(m)),x_o^*(m)) - f_i(x_i(x_o^*(m)),x_o^*(m))]$$

$$= \alpha_i(m\backslash m_i)[\max_{x_i \in X_i(x_o^*(m))} f_i(x_i,x_o^*(m)) - f_i(x_i(x_o^*(m)),x_o^*(m))] \geq 0$$

for every $x_i(\cdot)$ such that $x_i(x_o^*(m)) \in X_i(x_o^*(m))$.

(second part) Let $\hat{m}_i = \pi_i$ and let $\tilde{x}_o$ maximize $\pi_i(x_o) + f_o(x_o)$ on $X_o^i \cap X_o$. Define $A \equiv \pi_i(\tilde{x}_o) + f_o(\tilde{x}_o)$ and let $(\hat{m}\backslash\hat{m}_i)$ be any (I-1)-tuple in M^{*I-1} such that

$$\Sigma_{j\neq i}\ \hat{m}_j(x_o) = \begin{cases} \pi_i(\hat{x}_o) + f_o(\hat{x}_o) + \epsilon & \text{where } \epsilon > 0 \text{ for } x_o = \hat{x}_o \\ -A & \text{for all } x_o \in X_o^i \cap X_o,\ x_o \neq \hat{x}_o \end{cases}$$

(Clearly there exists such an $(\hat{m}\backslash\hat{m}_i)$ in M^{*I-1}).

It follows simply that $\hat{x}_o$ maximizes $\Sigma_j\hat{m}_j(x_o) + f_o(x_o)$ on X_o <u>uniquely</u>. Hence $x_o^*(\hat{m}) = \hat{x}_o$. Let $\hat{x}_i(\hat{x}_o) = \hat{x}_i$.

Now, for any $m_i \in M^*$ such that $X_i(x_o^*(\hat{m}/m_i))$ is not empty,

$$\omega_i^*[\hat{x}_i(x_o^*(\hat{m})),\hat{m}] - \omega_i^*[x_i^*(x_o^*(\hat{m}/m_i)),\hat{m}/m_i]$$

$$= \alpha_i(\hat{m}\backslash\hat{m}_i)\{f_i(\hat{x}_i,\hat{x}_o) + \Sigma_{j\neq i}\ \hat{m}_j(\hat{x}_o) + f_o(\hat{x}_o)$$

$$- f_i[x_i^*(x_o^*(\hat{m}/m_i)),x_o^*(\hat{m}/m_i)] - \Sigma_{j\neq i}\ \hat{m}_j[x_o^*(\hat{m}/m_i)] - f_o[x_o^*(\hat{m}/m_i)]\}$$

There are two cases to consider: Either (1) $x_o^*(\hat{m}/m_i) \equiv x_o^*$ maximizes $\pi_i(x_o) + \Sigma_{j\neq i} \hat{m}_j(x_o) + f_o(x_o)$ on X_o, or (2) it does not.

Case 1: In this case $x_o^* = \hat{x}_o$ as $\hat{x}_o$ is the unique maximizer. But then

$$\omega_i^*[\hat{x}_i(\hat{x}_o),\hat{m}] - \omega_i^*[x_i^*(x_o^*),\hat{m}/m_i] =$$

$$= \alpha_i(\hat{m}\backslash\hat{m}_i)[f_i(\hat{x}_i,\hat{x}_o) - f_i(x_i(\hat{x}_o),\hat{x}_o)] > 0$$

since $x_i(\hat{x}_o)$ does not maximize $f_i(x_i,\hat{x}_o)$ over $X_i(\hat{x}_o)$ but $\hat{x}_i$ does.

Case 2: In this case

$$\omega_i^*[\hat{x}_i(\hat{x}_o),\hat{m}] - \omega_i^*[x_i^*(x_o^*),\hat{m}/m_i]$$

$$= \alpha_i(\hat{m}\backslash m_i)[\pi_i(\hat{x}_o) + \Sigma_{j\neq i}\hat{m}_j(\hat{x}_o) + f_o(\hat{x}_o) - f_i(x_i(x_o^*),x_o^*) - \Sigma_{j\neq i}\hat{m}_j(x_o^*) - f_o(x_o^*)]$$

$$> \alpha_i(\hat{m}\backslash m_i)[\pi_i(x_o^*) + \Sigma_{j\neq i}\hat{m}_j(x_o^*) + f_o(x_o^*) - f_i(x_i(x_o^*),x_o^*) - \Sigma_{j\neq i}\hat{m}_j(x_o^*) - f_o(x_o^*)]$$

$$= \alpha_i(\hat{m}\backslash m_i)[\max_{x_i \in X_i(x_o^*)} f_i(x_i,x_o^*) - f_i(x_i(x_o^*),x_o^*)] \geq 0.$$

Thus, in either case $\omega_i^*[\hat{x}_i(\hat{x}_o),\hat{m}] > \omega_i^*[x_i^*(x_o^*),\hat{m}/m_i]$. |

Lemma 2: The message m_i^* maximizes

$$\omega_i^*[\hat{x}_i(x_o^*(m)),m] \text{ subject to } \hat{x}_i(x_o^*(m)) \in X_i(x_o^*(m))$$

for every $m\backslash m_i \in M^{*I-1}$ if and only if

$$m_i^*(x_o) = \pi_i(x_o) + \text{constant for every } x_o \in X_o^i \cap X_o$$

and is undefined on $(X_o^i)^c \cap X_o$.

Proof of Lemma 2: (if)

$$\omega_i^*[\hat{x}_i(x_o^*(m/m_i^*)),m/m_i^*] = \alpha_i(m\backslash m_i)[\pi_i(x_o^*(m/m_i^*)) + \Sigma_{j\neq i}m_j(x_o^*(m/m_i^*))$$

$$+ f_o(x_o^*(m/m_i^*))] + \beta_i(m\backslash m_i).$$

But since $m_i^*(x_o) = \pi_i(x_o)$ + constant, $x_o^*(m/m_i^*)$ maximizes $\pi_i(x_o) + \Sigma_{j\neq i} m_j(x_o) + f_o(x_o)$ on $X_o^i \cap X_o$. Thus

$$\omega_i^*[\hat{x}_i(x_o^*(m/m_i^*)),m/m_i^*] = \alpha_i(m\backslash m_i)[\operatorname*{Max}_{x_o}(\pi_i(x_o) + \Sigma_{j\neq i} m_j(x_o) + f_o(x_o))] + \beta_i(m\backslash m_i)$$

$$\geq \alpha_i(m\backslash m_i)[\pi_i(x_o^*(m/m_i)) + \Sigma_{j\neq i} m_j(x_o^*(m/m_i)) + f_o(x_o^*(m/m_i))] + \beta_i(m\backslash m_i)$$

$$= \omega_i^*[\hat{x}_i(x_o^*(m/m_i)),m/m_i]$$ for any m_i such that $X_i(x_o^*(m))$ is not empty.

(only if): Suppose $m_i^* \neq \pi_i$ + constant on X_o. Then there exist x_o^1 and x_o^2 in X_o such that

$$m_i^*(x_o^1) - m_i^*(x_o^2) > \pi_i(x_o^1) - \pi_i(x_o^2).$$

Now let A be any number satisfying

$$m_i^*(x_o^1) - m_i^*(x_o^2) > A > \pi_i(x_o^1) - \pi_i(x_o^2)$$

and B be defined by:

$$B = \operatorname{Min}\{\pi_i(x_o^1) - \operatorname*{Max}_{x_o \in X_o^i \cap X_o} [\pi_i(x_o) + f_o(x_o)],$$
$$m_i^*(x_o^2) + A - \operatorname*{Max}_{x_o \in \tilde{X}_o^i \cap X_o} [m_i^*(x_o) + f_o(x_o)]\}$$

Let $(\hat{m}\backslash\hat{m}_i)$ be any (I-1)-tuple in M^{*I-1} such that

$$\Sigma_{j\neq i}\, \hat{m}_j(x_o) = \begin{cases} -f_o(x_o^1) & \text{if } x_o = x_o^1 \\ -f_o(x_o^2) + A & \text{if } x_o = x_o^2 \\ B & \text{otherwise for all } x_o \in X_o \end{cases}$$

[Clearly such an $(\hat{m}\backslash\hat{m}_i)$ exists in M^{*I-1}].

It is easy to verify that $x_o^*(\hat{m}/\pi_i) = x_o^2$ and $x_o^*(\hat{m}/m_i^*) = x_o^1$.

Hence,

$$\omega_i^*[x_i^*(x_o^*(\hat{m}/\pi_i)),\hat{m}/\pi_i] - \omega_i^*[x_i^*(x_o^*(\hat{m}/m_i^*)),\hat{m}/m_i^*]$$

$$= \alpha_i(\hat{m}\backslash\hat{m}_i)[\pi_i(x_o^2) + \Sigma_{j\neq i}\hat{m}_j(x_o^2) + f_o(x_o^2) - \pi_i(x_o^1) - \Sigma_{j\neq i}\hat{m}_j(x_o^1) - f_o(x_o^1)]$$

$= \alpha_i(\hat{m}\backslash\hat{m}_i)[\pi_i(x_o^2) + A - \pi_i(x_o^1)] > 0$ which contradicts the assumption that m_i^* maximizes $\omega_i^*[x_i^*(x_o^*(m)),m]$ subject to $\hat{x}_i(x_o^*(m)) \in X_i(x_o^*(m))$ for every $m\backslash m_i \in M^{*I-1}$.

Finally, if $m_i^*(x_o)$ is defined for some $\tilde{x}_o \in (X_o^i)^c \cap X_o$, then for $(\hat{m}\backslash\hat{m}_i)$ defined as above with $\tilde{x}_o$ substituted for x_o^2 (in definition of B also), $x_o^*(\hat{m}/m_i^*) = \tilde{x}_o$ and $X_i(\tilde{x}_o) = X_i(x_o^*(\hat{m}/m_i^*))$ is empty. |

We can now prove Theorem 1:

Proof of Theorem 1: (if) Follows immediately from the if parts of Lemmata 1 and 2.

(only if). Suppose not

(a) If $x_i^*(x_o)$ does not maximize $f_i(x_i,x_o)$ on $X_i(x_o)$ for every $x_o \in X_o^i \cap X_o$, the second part of Lemma 1 establishes that $(m_i,x_i^*(\cdot))$ cannot be a solution for any $m_i \in M^*$.

(b) If $x_i^*(x_o)$ does maximize $f_i(x_i,x_o)$ on $X_i(x_o)$ for every $x_o \in X_o^i \cap X_o$, but $m_i^*(x_o) \neq \pi_i(x_o)$ + constant for every $x_o \in X_o^i \cap X_o$ then by the second part of Lemma 2, there exists some $\hat{m}\backslash\hat{m}_i \in M^{*I-1}$ such that

$$\omega_i^*[x_i^*(x_o^*(\hat{m}/\pi_i)),\hat{m}/\pi_i] > \omega_i^*[x_i^*(x_o^*(\hat{m}/m_i^*)), \hat{m}/m_i^*].$$

Also, by Lemma 2, $m_i^*(x_o)$ cannot be defined for any $x_o \in (X_o^i)^c \cap X_o$ or otherwise there exists an $\hat{m}\backslash\hat{m}_i$ such that $X_i(x_o^*(\hat{m}))$ is undefined. QED Theorem 1.

Next we show that any I-tuple of pairs $(m_i^*, x_i^*(\cdot))$ characterized in Theorem 1 yield decisions solving the organization decision problem D.

<u>Theorem 2</u>: If for every i, $(m_i^*, x_i^*(\cdot))$ satisfy a) and b) of Theorem 1, then $(x_o^*(m^*), \langle x_i^*(x_o^*(m^*))\rangle_i$ is a solution of problem D.

<u>Proof</u>: Let $\langle x_j^*\rangle_{j=o}^I \equiv (x_o^*(m^*), \langle x_i^*(x_o^*(m^*))\rangle_i)$.

Since $m_i^*(x_o)$ is defined on $X_o^i \cap X_o$ and undefined on $(X_o^i)^c \cap X_o$ $x_o^* \equiv x_o^*(m^*)$ is defined and satisfies the constraint $g_o(x_o) \leq 0$, i.e. $x_o^* \in X_o$. Also, since $X_i(x_o^*)$ is not empty, $x_i^* = x_i^*(x_o^*)$ is defined and satisfies the constraint $g_i(x_i, x_o^*) \leq 0$, for all i.

Let $\langle \tilde{x}_j\rangle_{j=o}^I$ be any other decisions also satisfying the constraints. Then

$$\Sigma_i f_i(x_i^*, x_o^*) + f_o(x_o^*) - \Sigma_i f_i(\tilde{x}_i, \tilde{x}_o) - f_o(\tilde{x}_o)$$

$$\geq \Sigma_i \pi_i(x_o^*) + f_o(x_o^*) - \Sigma_i \max_{\tilde{x}_i \in X_i(\tilde{x}_o)} f_i(\tilde{x}_i, \tilde{x}_o) - f_o(\tilde{x}_o)$$

$$= \Sigma_i \pi_i(x_o^*) + f_o(x_o^*) - \Sigma_i \pi_i(\tilde{x}_o) - f_o(\tilde{x}_o)$$

$$= \max_{x_o \in X_o} [\Sigma_i \pi_i(x_o) + f_o(x_o)] - \Sigma_i \pi_i(\tilde{x}_o) - f_o(\tilde{x}_o) \geq 0. \quad \text{QED of Theorem 2.}$$

<u>Theorem 3</u>: Any control mechanism C^* in the class $\mathcal{C}^*$ is optimal. where by (4.1a-d)

$$\mathcal{C}^* = \{\{M^*, x_o^*(\cdot), \langle E_i^*(\cdot)\rangle_i\} \mid \langle E_i^*(\cdot)\rangle_i \in \mathcal{E}^*\} \tag{4.3}$$

<u>Proof</u>: Theorem 1 establishes the decisiveness of C^* (a) and Theorem 2 establishes the efficiency of C^*(b). Thus, we need only show its non-capriciousness (c).

Let $\langle(\hat{m}_i, \hat{x}_i(\cdot))\rangle_i$ and $\langle(\hat{m}_i', \hat{x}_i'(\cdot))\rangle_i$ satisfy a) of the definition of optimality. Then, by Theorem 1,

$$\hat{m}_i = \hat{m}_i' + a_i \quad \text{on } X_o^i \cap X_o \text{ for some } a_i = \text{constant}$$

and

$$\hat{x}_i(x_o) \text{ and } \hat{x}_i'(x_o) \text{ maximize } f_i(x_i,x_o) \text{ on } X_i(x_o) \text{ for every } x_o \in X_o^i \cap X_o.$$

Thus, by the definition of $x_o^*(\cdot)$, $x_o^*(\hat{m}) = x_o^*(\hat{m}') = x_o^*$, and

$$\omega_i^*[\hat{x}_i(x_o^*(\hat{m})),m^*] - \omega_i^*[\hat{x}_i'(x_o^*(\hat{m}')),\hat{m}']$$

$$= \alpha_i(\hat{m}\backslash\hat{m}_i)[f_i(\hat{x}_i(x_o^*),x_o^*) + \Sigma_{j\neq i}[\hat{m}_j(x_o^*) - \hat{m}_j(\bar{x}_o)] + f_o(x_o^*) - f_o(\bar{x}_o)]$$

$$- \alpha_i(\hat{m}'\backslash\hat{m}_i')[f_i(\hat{x}_i'(x_o^*),x_o^*) + \Sigma_{j\neq i}[\hat{m}_j(x_o^*) + a_i - \hat{m}_j(\bar{x}_o) - a_i] + f_o(x_o^*) - f_o(\bar{x}_o)] = 0$$

since $\alpha_i(m\backslash m_i)$ is constant on $\{m\backslash m_i | m_j = m_j' + \text{constant}\}$ and $f_i(\hat{x}_i(x_o),x_o) = f_i(\hat{x}_i'(x_o),x_o)$ for every $x_o \in X_o^i \cap X_o$ and $x_o^* \in X_o^i \cap X_o$. Q.E.D.

IV.2 A Partial Characterization of Optimal Control Mechanisms

Although any control mechanism C^* in the class $\mathcal{C}^*$ is optimal by Theorem 3, one might wonder if $\mathcal{C}^*$ contains all optimal control mechanisms. A partial answer to this question can be given based on some recent work. of Green and Laffont [10].[17] Denote by $\tilde{\mathcal{C}}$ the class of all control mechanisms $\tilde{C} = \{M, x_o(\cdot), \langle\tilde{E}_i(\cdot)\rangle_i\}$ such that

$$\tilde{E}_i(f_i,m) = \gamma_i(m\backslash m_i)[f_i + T_i(m)] \tag{4.4}$$

where $T_i : M^I \to \mathbb{R}$ is an arbitrary function and $\gamma_i : M^{I-1} \to \mathbb{R}_{++}$ is an arbitrary strictly positive function. It is easy to see that $\mathcal{C}^*$ is a subclass of mechanisms in $\tilde{\mathcal{C}}$. Now, one can show that if $\tilde{C}$ in $\tilde{\mathcal{C}}$ is optimal, then there exists a control mechanism C^* in $\mathcal{C}^*$ that is equivalent to $\tilde{C}$ in the sense that C^* leads to the same optimal decisions x_o for the Center and the same evaluations of the divisions as $\tilde{C}$. Thus, the class $\mathcal{C}^*$ spans the equivalence classes of optimal control mechanisms of the form $\tilde{C}$.

This result is given as Theorem 4 below. However to establish it, several characteristics and properties of control mechanisms are defined and several preliminary propositions established.

Definition: Given any control mechanism C in which the language M is the space M^* [c.f. (4.1a)], the mechanism is called truth-inducing if and only if the pairs $(m_i^*, x_i^*(\cdot))$ [c.f. Theorem 1] are the unique message-decision rule pairs maximizing $\omega_i[x_i(x_o(m)),m]$ subject to $x_i(x_o(m)) \in X_i(x_o(m))$ for every $m\backslash m_i \in M^{*I-1}$.

Theorem 1 establishes that every control mechanism C^* in $\mathcal{C}^*$ is truth-inducing.

Another property of mechanisms is defined for a subset $\hat{\mathcal{C}}$ of mechanisms in $\tilde{\mathcal{C}}$ where

$$\hat{\mathcal{C}} = \{\hat{C} \in \tilde{\mathcal{C}} | \text{a) } M = M^*, \text{ b) } x_o(\cdot) = x_o^*(\cdot), \text{ and}$$
$$\text{c) } \gamma_i(m\backslash m_i) \text{ and } T_i(m) \text{ are constant on the sets } \{m' \in M^{*I} | m_j' = m_j + \text{constant for all } j\}\}$$

The property is defined as:

Definition: A control mechanism $\hat{C}$ in $\hat{\mathcal{C}}$ satisfies property 0 if and only if, for all i

a) $T_i(m)$ is independent of m_i at $x_o^*(m)$; i.e. if, for $(m/m_i) \in M^{*I}$ and $\hat{m}_i \in M^*$, $x_o^*(m/m_i) = x_o^*(m/\hat{m}_i)$, then $T_i(m/m_i) = T_i(m/\hat{m}_i)$.

b) $$T_i(m/m_i) - T_i(m/\hat{m}_i) = \Sigma_{j\neq i} m_j[x_o^*(m/m_i)] + f_o[x_o^*(m/m_i)] - \Sigma_{j\neq i} m_j[x_o^*(m/\hat{m}_i)] - f_o[x_o^*(m/\hat{m}_i)]$$

for all $m/m_i \in M^{*I}$, $\hat{m}_i \in M^*$.

Lemma 3: A control mechanism $\hat{C}$ in $\hat{\mathcal{C}}$ is in $\mathcal{C}^*$ if and only if it satisfies Property 0.

Proof: This is straightforward to verify.

Using the characterization of $\mathcal{C}^*$ provided by property 0, it can be shown that any mechanism $\hat{C}$ in $\hat{\mathcal{C}}$ is truth-inducing if and only if it is in $\mathcal{C}^*$.

Proposition 3: A control mechanism $\hat{C}$ in $\hat{\mathcal{C}}$ is truth-inducing if and only if $\hat{C} \in \mathcal{C}^*$.

Proof: (if). This is just Theorem 1.

(only if). By Lemma 3 if $\hat{C}$ satisfies property 0, then it is in $\mathcal{C}^*$. Thus, we consider the negation of both parts of the definition of property 0.

1) Suppose for some i that $T_i(m)$ is not independent of m_i at $x_o^*(m)$; then there exist $(m/m_i) \in M^{*I}$, $\hat{m}_i \in M^*$ with $m_i \neq \hat{m}_i$ such that $x_o^* = x_o^*(m/m_i) = x_o^*(m/\hat{m}_i)$ but $T_i(m/m_i) > T_i(m/\hat{m}_i)$. Let $\pi_i = \hat{m}_i$. (Clearly there exists some problem D satisfying A.1-A.4 such that $\pi_i = \hat{m}_i$.) But then

$$\omega_i[x_i^*(x_o^*(m)),m] - \omega_i[x_i^*(x_o^*(m/\pi_i),m/\pi_i]$$

$$= \gamma_i(m\backslash m_i)[f_i(x_i^*(x_o^*),x_o^*) + T_i(m/m_i) - f_i(x_i^*(x_o^*),x_o^*) - T_i(m/\pi_i)]$$

$$= \gamma_i(m\backslash m_i)[T_i(m/m_i) - T_i(m/\pi_i)] > 0,$$

which contradicts the fact that $\hat{C}$ is truth-inducing.

2) Suppose for some i that $\hat{C}$ does not satisfy b); then there exist $(m/m_i) \in M^{*I}$, $\hat{m}_i \in M^*$ such that

$$T_i(m) - T_i(m/\hat{m}_i) = \Sigma_{j\neq i}[m_j(x_o^*(m)) - m_j(x_o^*(m/\hat{m}_i))]$$

$$+ f_o(x_o^*(m)) - f_o(x_o^*(m/\hat{m}_i)) + \epsilon \quad \text{for } \epsilon > 0.$$

Now suppose

$$\pi_i(x_o) = \begin{cases} -\Sigma_{j\neq i} m_j(x_o^*(m)) - f_o(x_o^*(m)) & \text{for } x_o = x_o^*(m) \\ -\Sigma_{j\neq i} m_j(x_o^*(m/\hat{m}_i)) - f_o(x_o^*(m/\hat{m}_i)) + \delta & \text{for } x_o = x_o^*(m/\hat{m}_i) \\ -\underset{x_o\in X_o}{\text{Max}} [\Sigma_{j\neq i} m_j(x_o) - f_o(x_o)] - \epsilon & \text{for } x_o \neq x_o^*(m), x_o \neq x_o^*(m/\hat{m}_i) \end{cases}$$

where $0 < \delta < \epsilon$. Since $\pi_i(\cdot)$ is upper-semi continuous, $\pi_i \in M^*$. (Also, clearly there exists a problem D satisfying A.1-A.4 such that this holds.)

It is easy to verify that $x_o^*(m/\pi_i) = \hat{x}_o \equiv x_o^*(m/\hat{m}_i)$ and thus, by part 1 above, $T_i(m/\pi_i) = T_i(m/\hat{m}_i)$. Thus, letting $x_o^* \equiv x_o^*(m)$,

$$\omega_i[x_i^*(x_o^*),m] - \omega_i[x_i^*(\hat{x}_o^*),m/\pi_i] = \gamma_i(m\backslash m_i)[\pi_i(x_o^*) - \pi_i(\hat{x}_o) + T_i(m) - T_i(m/\hat{m}_i)]$$

$$= \gamma_i(m\backslash m_i)[-\Sigma_{j\neq i} m_j(x_o^*) - f_o(x_o^*) + \Sigma_{j\neq i} m_j(\hat{x}_o) + f_o(\hat{x}_o) - \delta$$

$$+ \Sigma_{j\neq i} m_j(x_o^*) + f_o(x_o^*) - \Sigma_{j\neq i} m_j(\hat{x}_o) - f_o(\hat{x}_o) + \epsilon] = \gamma_i(m\backslash m_i)(\epsilon-\delta) > 0$$

which again contradicts the fact that $\hat{C}$ is truth-inducing. ▌

Proposition 4: Given any optimal control mechanism $\tilde{C}$ in $\tilde{\mathcal{C}}$ such that the language $M = M^*$, it is truth-inducing if and only if it is also in $\mathcal{C}^*$.

Proof: (if). By Theorem 3, any mechanism in $\mathcal{C}^*$ is optimal and by Theorem 1 it is truth-inducing.

(only if). We need only show that $x_o(\cdot) = x_o^*(\cdot)$, i.e. satisfies (4.1b), and that $\gamma_i(m\backslash m_i)$ and $T_i(m)$ are constant on the sets $\{m' \in M^{*I} | m_j' = m_j + \text{constant for all } j\}$, since then $\tilde{C}$ is in $\hat{\mathcal{C}}$ and thus by Proposition 3 is in $\mathcal{C}^*$ since it is truth-inducing.

Since $\tilde{C}$ is truth-inducing, $m_i^* = \pi_i$ + constant are i's best messages. And, since the mechanism is optimal

$x_o(\pi)$ maximizes $\Sigma_i \pi_i(x_o) + f_o(x_o)$ over X_o for every $\pi \in M^{*I}$ (i.e. for every problem D). Also, since the mechanism is non-capricious, $x_o(m^*) = x_o(m^{*\prime})$ for all best messages m^* and $m^{*\prime}$ i.e. for $m_i^* = \pi_i + a_i$ and $m_i^{*\prime} = \pi_i + a_i'$. Thus $x_o(\cdot) = x_o^*(\cdot)$.

That $\gamma_i(m\backslash m_i)$ and $T_i(m)$ are constant on the sets $\{m' \in M^{*I} | m_j' = m_j + \text{constant for all } j\}$ follows from the non-capriciousness of an optimal mechanism and that these sets are the sets of the division's best messages. ∎

Theorem 4: If $C = \{M, x(\cdot), \langle E_i(\cdot)\rangle_i\}$ is any optimal control mechanism in the class $\tilde{\mathcal{C}}$, then there exist functions $\Psi_i : M \to M^*$ and a mechanism C^* in $\mathcal{C}^*$ such that:

$$x_o(m) = x_o^*(\Psi(m)) \qquad \text{for all } m \in \underset{i=1}{\overset{I}{X}} \tilde{M}_i$$
$$E_i(f_i, m) = E_i^*(f_i, \Psi(m))$$

where $\tilde{M}_i = \{\hat{m}_i \in M | (\hat{m}_i, \hat{x}_i(\cdot))$ solves i's decision problem for some organizational problem D with $(f(\cdot), g(\cdot))$ such that $\pi_i(\cdot) \in M^*\}$ and $\Psi(m) \equiv [\Psi_1(m_1), \ldots, \Psi_I(m_I)]$.

Proof: Note first of all that since $C \in \tilde{\mathcal{C}}$, $E_i(f_i, m) = \gamma_i(m\backslash m_i)[f_i + T_i(m)]$ so that $(\hat{m}_i, \hat{x}_i(\cdot))$ solves i's decision problem only if $\hat{x}_i(x_o)$ maximizes $f_i(x_i, x_o)$ on $X_i(x_o)$.

Now, for every $\pi_i \in M^*$, let $\phi_i(\pi_i)$ be defined as the set of i's best messages under C. That is, since C is optimal, and $\hat{m}_i \in \phi_i(\pi_i)$, then $(\hat{m}_i, \hat{x}_i(\cdot))$ solves i's decision problem when i is characterized by the functions $(f_i(\cdot), g_i(\cdot))$ such that $f_i(\hat{x}_i(x_o), x_o) = \pi_i(x_o)$ on X_o^i.

Let $\phi(m^*) \equiv (\phi_1(m_1^*),\ldots,\phi_I(m_I^*))$ and define

$$\left.\begin{aligned} x_o^*(m^*) &= x_o(\phi(m^*)) \\ T_i^*(m^*) &= T_i(\phi(m^*)) \\ \gamma_i^*(m^*\backslash m_i^*) &= \gamma_i(\phi(m^*)\ \phi_i(m^*)) \end{aligned}\right\} \quad \text{for every } m^* \in M^{*I} .$$

Note that $x_o^*(\cdot)$ is well defined since C is optimal and thus non-capricious which implies that if $\hat{m}_i$ and $\hat{m}'$ are in $\phi_i(m^*)$ for all i, then $x_o(\hat{m}) = x_o(\hat{m}') \equiv \hat{x}_o$. Also, since C is non-capricious, if $\hat{m}_i$ and $\hat{m}_i' \in \phi_i(m^*)$ for all i $\omega_i[\hat{x}_i(x_o(\hat{m})),\hat{m}] - \omega_i[\hat{x}_i(x_o(\hat{m}')),\hat{m}'] = 0$ which implies that:

$$\gamma_i(\hat{m}\backslash\hat{m}_i)[\pi_i(\hat{x}_o) + T_i(\hat{m})] = \gamma_i(\hat{m}'\backslash\hat{m}_i')[\pi_i(\hat{x}_o) + T_i(\hat{m}')]$$

But note that since both $\hat{m}_i$ and $\hat{m}_i'$ maximize $\omega[\hat{x}_i(x_o(m)),m]$ $= \gamma_i(m\backslash m_i)[\pi_i(x_o(m)) + T_i(m)]$, they also maximize $\gamma_i(m\backslash m_i)[\pi_i(x_o(m)) + a_i + T_i(m)]$ for any constant a_i. Thus non-capriciousness of C implies that for any constant a_i

$$\gamma_i(\hat{m}\backslash\hat{m}_i)[\pi_i(\hat{x}_o) + a_i + T_i(\hat{m})] = \gamma_i(\hat{m}'\backslash\hat{m}_i')[\pi_i(\hat{x}_o) + a_i + T_i(\hat{m}')]$$

Thus, $\gamma_i(\hat{m}\backslash\hat{m}_i) = \gamma_i(\hat{m}\backslash\hat{m}_i')$ and $T_i(\hat{m}) = T_i(\hat{m}')$. Hence $T_i^*(\cdot)$ and $\gamma_i^*(\cdot)$ are also well-defined.

Now let $\Psi_i(m_i)$ be any element of $\phi_i^{-1}(m_i)$ for every $m_i \in \tilde{M}_i$. Then, since both m_i and any $m_i' \in \phi_i \cdot \Psi_i(m_i)$ are best messages for $\pi_i = \Psi_i(m)$, i.e. m_i and $m_i' \in \phi_i(\pi_i)$,

$$x_o^*(\Psi(m)) = x_o[\phi\cdot\Psi(m))] = x_o(m),\ T_i^*[\Psi(m)] = T_i(m)$$

and $\gamma_i^*[\Psi(m)\backslash\Psi_i(m_i)] = \gamma_i(m\backslash m_i)$. Thus

$$E_i(f_i,m) = \gamma_i(m\backslash m_i)[f_i + T_i(m)] = \gamma_i^*[\Psi(m)\backslash\Psi_i(m_i)][f_i + T_i(\Psi(m))] \equiv E_i^*[f_i,\Psi(m)]$$

<u>Claim</u>: $C^* = \{M^*, x_o^*(\cdot), \langle E_i^*(\cdot)\rangle_i\}$ is in $\mathcal{C}^*$

<u>Proof</u>: If we show C^* is both truth-inducing and optimal, then by Proposition 3 it will be in $\mathcal{C}^*$.

Note first of all that $(\hat{m}_i, \hat{x}_i(\cdot))$ solves i's decision problem under C^* only if $\hat{x}_i(x_o)$ maximizes $f_i(x_i, x_o)$ on $X_i(x_o)$. Thus, if C^* is not truth-inducing, then for some i, there exists an $\hat{m}_i \neq \pi_i$ + constant $\in M^*$ and $m^*\backslash m_i^* \in M^{*I-1}$ such that

$$\omega_i^*[\hat{x}_i(x_o^*(m^*/\hat{m}_i)), m^*/\hat{m}_i] > \omega_i^*[\hat{x}_i(x_o^*(m^*/\pi_i)), m^*/\pi_i], \quad \text{or}$$

$$E_i^*[\pi_i(x_o^*(m^*/\hat{m}_i)), m^*/\hat{m}_i] > E_i^*[\pi_i(x_o^*(m^*/\pi_i)), m^*/\pi_i] \quad \text{or}$$

$$E_i[\pi_i(x_o(\phi(m^*)/\phi_i(\hat{m}_i))), \phi(m^*)/\phi(\hat{m}_i)] > E_i[\pi_i(x_o(\phi(m^*)/\phi(\pi_i))), \phi_i(\pi_i))), \phi(m^*)/\phi_i(\pi_i)]$$

which contradicts the fact that any $(\hat{m}_i', \hat{x}_i(\cdot))$ where $\hat{m}_i' \in \phi_i(\pi_i)$ is a best message-decision rule pair for i under C. Thus, C^* is truth-inducing.

Since it is truth-inducing C^* is decisive. Also C^* is efficient since $x_o^*[\Psi(\hat{m})] = x_o(\hat{m})$ maximizes $\Sigma_i \pi_i(x_o) + f_o(x_o)$ over X_o since C is efficient and $\hat{x}_i(x_o)$ maximizes $f_i(x_i, x_o)$ on $X_i(x_o)$ for all $x_o \in X_o^i$. C^* is finally non-capricious since C is. This establishes the claim and hence Theorem 4.

Although by Proposition 3 a control mechanism $\hat{C} \in \hat{\mathcal{C}}$ is truth-inducing if and only if $\hat{C} \in \mathcal{C}^*$, it is not true that if $\hat{C}$ is merely optimal that it is also in $\mathcal{C}^*$. Since all mechanisms in $\mathcal{C}^*$ are both optimal and truth-inducing, this means that there exist optimal mechanisms $\hat{C}$ in $\hat{\mathcal{C}}$ that are not truth-inducing even though the language of any $\hat{C}$ is the set of professed profit functions and the Center's decision rule is to maximize the sum of reported (plus its own) profits (subject to feasibility).

A particularly simple optimal, yet non-truth-inducing, member of $\hat{\mathcal{C}}$ may be defined by:

$$M^1 = M^*,\ x_o^1(\cdot) = x_o^*(\cdot)$$

and

$$E_i^1(f_i,m) = f_i + h_i(x_o^1(m)) + \Sigma_{j\neq i}[m_j(x_o^1(m)) - m_j(\bar{x}_o)] + f_o(x_o^1(m)) - f_o(\bar{x}_o)$$

where $h_i: X_o \to \mathbb{R}$ is any arbitrary function such that $\Sigma_{i=1}^I h_i(x_o) \equiv 0$ for all x_o. It is easy to show that $C^1 = \{M^1, x_o^1(\cdot), \langle E_i^1(\cdot)\rangle\}$ is optimal and that i's best messages are all of the form $\hat{m}_i(\cdot) = [\pi_i(\cdot) + h_i(\cdot)]$ + constant. Thus C^1 is not truth-inducing.

Of course, for this example, it is obvious that knowing any $\hat{m}_i(\cdot)$ one can determine $\pi_i(\cdot)$ up to a constant by subtracting $h_i(\cdot)$ from $\hat{m}_i(\cdot)$. An interpretation of Theorem 4 is that this is generally true. Given any optimal mechanism in $\mathcal{C}$ (not just $\hat{\mathcal{C}}$), by applying the function Ψ_i to any $\hat{m}_i$ one can determine the true $\pi_i(\cdot)$ up to a constant. Thus the mechanism might as well be based on asking divisions to report their true profit functions π_i, maximizing reported profits, and then using some evaluation measures in $\mathcal{E}^*$ to evaluate the divisions.

IV.3 Total Profit Allocating Mechanisms

A noteworthy feature of any (optimal) control mechanism C^* in $\mathcal{C}^*$ is that in general the sum of the divisions' evaluations, even when the divisions are reporting "honestly" (as they have an incentive to do under C^*), will <u>not</u> equal the organization's total realized profits: $\Sigma_{i=1}^I f_i(\hat{x}_i,\hat{x}_o) + f_o(\hat{x}_o)$. That is, in general

$$\Sigma_i E_i(\hat{f}_i,\pi) = \Sigma_i \gamma_i(\pi\backslash\pi_i)[f_i(\hat{x}_i,\hat{x}_o) + T_i(\pi)] \tag{4.5}$$

$\neq \Sigma_i f_i(\hat{x}_i,\hat{x}_o) + f_o(\hat{x}_o)$ or for that matter to $\Sigma_i f_i(\hat{x}_i,\hat{x}_o)$,

where $\hat{x}_o = x_o^*(\pi)$, $\hat{x}_i = x_i^*(\hat{x}_o)$, $\gamma_i(\pi\backslash\pi_i) = \alpha_i(\pi\backslash\pi_i)$, and

$T_i(\pi) = \Sigma_{j\neq i}[\pi_j(\hat{x}_o) - \pi_j(\bar{x}_o)] + [f_o(\hat{x}_o) - f_o(\bar{x}_o)] + \beta_i(\pi\backslash\pi_i)/\alpha_i(\pi\backslash\pi_i)$

[c.f. (4.1 c-d) and (4.4)]. In particular, in those cases in which $\gamma_i(m\backslash m_i) = \alpha_i(m\backslash m_i) = 1$ for all $m\backslash m_i$, this means that

$\Sigma_i T_i(\pi) = (n-1)[\Sigma_j[\pi_j(\hat{x}_o) - \pi_j(\bar{x}_o)] + [f_o(\hat{x}_o) - f_o(\bar{x}_o)] + \Sigma_i \beta_i(\pi\backslash\pi_i) \neq f_o(\hat{x}_o)$ or 0.

Whether or not it is possible to find arbitrary functions $\beta_i(m\backslash m_i)$ that are constant on the sets $\{m' \in M^* | m_j' = m_j + \text{constant}\}$ such that equality in (4.5) holds for all $m \in M^{*I}$ is an open question.[18]

This question is an interesting one since many standard divisional accounting procedures and new SEC disclosure regulations requiring line-of-business reporting require that total organizational profits be distributed or allocated or attributed to the various divisions for reporting purposes. If, as seems most likely, it is not possible to find any C^* with the total profit allocating property (i.e. such that $\Sigma_i E_i(\hat{f}_i,\pi) + f_o(\hat{x}) = \Sigma_i f_i(\hat{x}_i,\hat{x}_o) + f_o(\hat{x}_o)$), then, by Theorem 4, whatever accounting measures are used to report the separate contribution of each division, these measures should perhaps not be used as evaluation measures since they will not be part of any optimal control mechanism.

It seems that the interest in having measures for allocating total profits among all divisions (and perhaps the Center too) originates in the view that total organizational achievement must be just the sum of the achievements of the organization's constituent parts. This view may be contrasted with the view that total achievement may be greater than (or less than) the sum of the individual achievements.

Such considerations, in any event, raise the question of the intrinsic meaning of the evaluation measure for an optimal control mechanism. That is, what significance, if any, does the magnitude of the evaluation measure in any particular instance have and can one meaningfully compare the values of the measures for two different divisions? This question is important since, presumably, in order to motivate a divisional manager to maximize his division's evaluation measure some type of compensation scheme that is increasing in the evaluation measure must be used. Furthermore, it would seem desirable for equity considerations if there were some objective basis for distinguishing among the performance of different divisions on the basis of their evaluation measures.

While a full exploration of this issue is beyond the scope of this paper and has not yet been completely developed (c.f. however, Groves and Loeb [16]), it is interesting to note that one of the optimal mechanisms in $\mathcal{C}^*$ may yield measures with some intrinsic meaning. Specifically, consider the mechanism $\bar{C} = \{M^*, x_o^*(\cdot), \langle \bar{E}_i(\cdot) \rangle_i\}$ where $\bar{E}_i$ is defined in (4.2) as:

$$\bar{E}_i(f_i, m) = f_i - \{\Sigma_{j \neq i} m_j(x_o^i) + f_o(x_o^i) - [\Sigma_{j \neq i} m_j(x_o^*(m)) + f_o(x_o^*(m))]\} \quad (4.6)$$

where $x_o^i = x_o^i(m \backslash m_i)$ maximizes $\Sigma_{j \neq i} m_j(x_o) + f_o(x_o)$ subject to $g_o(x_o) \leq 0$. Now, when all division manager's are responding to the incentives so that $m_j(\cdot) = \pi_j(\cdot)$ + constant for all $j \neq i$, the measure $\bar{E}_i(f_i, m)$ measures the opportunity cost to the organization of having the i^{th} division. In other words, $\bar{E}_i(\hat{f}_i,\pi/\hat{m}_i)$ where $f_i = f_i(\hat{x}_i, x_o^*(\pi/\hat{m}_i))$ is the profit the organization would lose if the i^{th} division could be and were abandoned.[19] Thus the measures $\bar{E}_i(f_i,m)$ have an intrinsic opportunity cost meaning even though (except when no division can affect the Center's decision),

$$\Sigma_i\bar{E}_i[f_i(\hat{x}_i,\hat{x}_o)] \leq \Sigma_i f_i(\hat{x}_i,\hat{x}_o) \quad \text{for all } \hat{m} \text{ including } \pi \text{ and } \hat{x}_i,$$

where $\hat{x}_o = x_o^*(\hat{m})$ and $\hat{x}_i$ are the realized decisions taken by the agents.

Another approach to the issue of total profit allocating mechanisms has been developed by Hurwicz [20, 21] and Groves and Ledyard [13] in a quite different context. For the organizational control problem the thrust of their results may be expressed as follows:

Consider the class $\mathcal{C}^+$ of all control mechanisms

$$C^+ = \{M,x_o(\cdot),\langle E_i(\cdot)\rangle_i\} \text{ in which } E_i(f_i,m) = f_i + T_i(m)$$

and such that

$$\Sigma E_i(f_i,m) = \Sigma_i f_i + f_o(x_o(m))$$

or $\qquad$ for all $m \in M^I$

$$\Sigma T_i(m) = f_o(x_o(m))$$

i.e. the evaluation measures are total profit allocating.[20] Now, since it does not appear that there exists any mechanism in $\mathcal{C}^+$ that is optimal, in general we look instead for a mechanism in $\mathcal{C}^+$ that we shall call satisfactory:

Definition: Any control mechanism C is satisfactory if and only if

a) it is stable: there exist an I-tuple $\langle(\hat{m}_i,\hat{x}_i(\cdot))\rangle_i$ of message-decision rule pairs such that for each i

$\omega_i[\hat{x}_i(x_o(\hat{m})),\hat{m}] \equiv E[f_i(\hat{x}_i(x_o(\hat{m})),x_o(\hat{m})),\hat{m}]$

for all $(m_i,x_i(\cdot))$ such that $m_i \in M$ and $x_i(x_o(\hat{m}/m_i)) \in X_i(x_o(\hat{m}/m_i))$

b) it is efficient: for all I-tuples $\langle(\hat{m}_i,\hat{x}_i(\cdot))\rangle$ satisfying a), the resulting decisions $(\hat{x}_o(\hat{m}),\langle(\hat{x}_i(\hat{x}_o(\hat{m}))\rangle_i)$ solve the organizational decision problem D; and

c) it is non-capricious; if $\langle(\hat{m}_i,\hat{x}_i(\cdot))\rangle_i$ and $\langle(\hat{m}_i',\hat{x}_i'(\cdot))\rangle_i$ both satisfy a), then

$$\hat{x}_o(\hat{m}) = \hat{x}_o(\hat{m}')$$

and $\hat{\omega}_i[\hat{x}_i(\hat{x}_o(m)),\hat{m}] = \hat{\omega}_i[\hat{x}_i'(\hat{x}_o(\hat{m}'),\hat{m}']$ for all i.

In essence, a satisfactory mechanism is one for which a Nash (or non-cooperative) equilibrium of divisional strategies (message-decision rule pairs) exists and that is efficient and non-capricious. In contrast, an optimal mechanism is one for which a dominant strategy equilibrium exists and that is efficient and non-capricious. It is clear that an optimal mechanism is satisfactory but not necessarily vice versa.

Whether or not satisfactory mechanisms exist in $\mathcal{C}^+$ (that is, satisfactory mechanisms that are total profit allocating for problem D) in general is an open question. However, for some special classes of problems one can find such mechanisms. To illustrate such a mechanism, consider the following problem:

D^o: $\underset{x_o}{\text{Max}}\ \Sigma_{i=1}^{I}\ \pi_i(x_o) + f_o(x_o)$

subject to $g_o(x_o) \le 0$

Problem D^o is clearly a "reduced form" of problem D arising when the division managers take the decision $x_i^*(x_o)$ for every x_o in X_o^i.

Now consider the following restrictions on problem D^o:

Restrictions on D^o:

B.1: The functions π_i are continuous concave functions defined on $\mathbb{R}_+^{N_o}$, the non-negative orthant of $\mathbb{R}^{N_o}$.

B.2: The function $f_o(x_o)$ is a linear function:

$f_o(x_o) = a_o - b_o \cdot x_o$ where $b_o \in \mathbb{R}_{++}^{N_o}$ is strictly positive.

B.3: The set $X_o = \{x_o \in \mathbb{R}^{N_o} | g_o(x_o) \le 0\} = \mathbb{R}_+^{N_o}$, the non-negative orthant of $\mathbb{R}^{N_o}$.

Under these restrictions, problem D^o may be written as:

D^1: $\underset{x_o \in \mathbb{R}_+^{N_o}}{\text{Max}}\ \Sigma\pi_i(x_o) - b_o \cdot x_o$.

The restrictions on D^o are, of course, very severe, especially with reference to the Center's functions $f_o(\cdot)$ and $g_o(\cdot)$. However, these restrictions can be relaxed.[21]

Now for the problem D^1 we can exhibit a satisfactory mechanism C^+ that is total profit allocating. The language M^+ is defined as:

$$M^+ \equiv \mathbb{R}^{N_o} \tag{4.7a}$$

The Center's decision rule $x_o^+(\cdot)$ is defined as:

$$x_o^+(m) = \Sigma_{i=1}^{I} m_i \tag{4.7b}$$

Given the language and decision rule, a division's message may be interpreted as the increments (or decrements since m_i may be negative in any component) in the levels of the decisions x_o that division i wishes the Center to add to (subtract from) the aggregate levels requested by the other divisions. Note that at a Nash equilibrium of messages $\langle \hat{m}_i \rangle_{i=1}^{I}$, each division must be implicitly requesting the same aggregate level of the decisions $x_o^+(\hat{m}) = \Sigma \hat{m}_i$ or, in other words, they must all agree on the level. It is the role of the evaluation measures to ensure that this will be possible.

The evaluation measures $E_i^+(\cdot)$ are defined by:

$$E_i^+(\pi_i, m) = \pi_i - T_i^+(m) \qquad \text{(as required since } C^+ \in \mathcal{C}^+) \tag{4.7c}$$

where

$$T_i^+(m) = \alpha_i b_o \cdot \Sigma_j m_j + \frac{\gamma}{2} \left[\frac{I-1}{I}(m_i - \frac{1}{I-1}\Sigma_{j \neq i} m_j)^2 - \frac{1}{2(I-1)(I-2)} \Sigma_{j \neq i} \Sigma_{j' \neq i} (m_j - m_j')^2 \right]$$

$$= \alpha_i b_o \cdot \Sigma_j m_j + \frac{\gamma}{2} \left\{ \frac{I-1}{I} [m_i - \mu(m \backslash m_i)]^2 - \sigma(m \backslash m_i)^2 \right\}$$

with $\gamma > 0$, $\Sigma_i \alpha_i = 1$ and

$$\mu_i \equiv \mu(m \backslash m_i) \equiv \frac{1}{I-1} \Sigma_{j \neq i} m_j$$

$$\sigma_i^2 \equiv \sigma(m \backslash m_i)^2 \equiv \frac{1}{2(I-1)(I-2)} \Sigma_{j \neq i} \Sigma_{j' \neq i} (m_j - m_j')^2 = \frac{1}{I-2} \Sigma_{j \neq i} (m_j - \mu(m \backslash m_i))^2$$

An interpretation of the evaluation measures $E_i^+(\cdot)$ is as follows: Let us call $b_o \cdot x_o^+(m)$ the "cost" of decisions $x_o^+(m)$ and $T_i^+(m)$ the i^{th} division's "cost" share. The evaluation measure thus subtracts from the realized "profits" of each division, π_i, a "cost" share consisting of a proportional share of the total "cost" plus a positive multiple $(\frac{\gamma}{2})$ of the difference between the squared deviation of the division's message from the mean of the other divisions' messages (corrected for small sample bias) and the squared standard error of the mean of the others' messages (for small samples). Thus, given the aggregate level of the decisions requested by all divisions, $\Sigma_j m_j$, division i's cost is larger as the amount it requests deviates from the average of the others' request and smaller the greater the squared standard error of the mean of the others' messages.

An alternative interpretation of the control mechanism C^+ follows from the alternative specification: For each message $m_i \in M^+$ define a function $h_i : \mathbb{R}_+^{N_o} \to \mathbb{R}$ by:

$$h_i(x_o;m_i) \equiv (\gamma m_i + \alpha_i b_o)\cdot x_o - \frac{\gamma}{2I} x_o \cdot x_o \ . \tag{4.8}$$

Since each message m_i defines such a function, a division's message m_i can be interpreted as communicating the function $h_i(\cdot;m_i)$. Thus, an alternative specification of the language M^+ of C^+ is:

$$H^+ = \{h_i(\cdot;m_i) | h_i(x_o;m_i) \text{ satisfies } (4.8)\}$$

We call any message $h_i(\cdot)$ in H^+ a (quadratic) reported profit function.

Next, note that given the messages $m \in M^{+I}$, the decision $x_o^+(m) = \Sigma_i m_i$ maximizes $\Sigma_i h_i(x_o;m_i) - b_o x_o$. Finally, in terms of the functions $h_i(\cdot;m_i)$, the evaluation measures $E_i^+(\pi_i,m)$ may be expressed as:

$$E_i^+(\pi_i,m) = \pi_i - T_i^+(m)$$

where

$$T_i^+(m) \equiv \alpha_i b_o \cdot x_o^+(m) + \operatorname*{Max}_{x_o} \{\Sigma_{j\neq i}[h_j(x_o;m_j) - \alpha_j b_o \cdot x_o]\} \\ - \Sigma_{j\neq i}[h_j(x_o^+(m);m_j) - \alpha_j b_o \cdot x_o^+(m)] - \frac{\gamma}{2}\sigma_i^2 \ . \tag{4.9}$$

With this specification, the control mechanism C^+ may be viewed as a parametric representation of a mechanism in the class $\mathcal{C}^*$ [see (4.3) and especially (4.1d)]. Since C^+ can be shown to be a satisfactory mechanism, a division's best message $\hat{m}_i$ given the messages of the other divisions can be shown to be the parameters of a quadratic approximation to the division's true profit function π_i at the level $x_o^+(m) = \Sigma_j m_j$ of the Center's decisions.

The cost share $T_i^+(m)$ in the form (4.9) can be interpreted as assessing the i^{th} division (i) its proportional cost share, plus (ii) the total reduction in the aggregate net reported profits (gross reported profits $h_j(\cdot;m_j)$ less proportional cost shares $\alpha_j b_o \cdot x_o$) caused, in effect, by division i's request for a level of decisions different from the average of the other's requests, less a positive multiple ($\gamma/2$) of the squared standard error of the mean of the others' messages. [Compare this with the mechanism $\bar{C}$, c.f. (4.6).]

We state, without proof,[22] the result that C^+ is a satisfactory control mechanism and is total profit allocating:

<u>Theorem 5</u>: Under the restrictions B.1-B.3, the control mechanism C^+ defined by (4.7a-c) is a satisfactory mechanism and is total profit allocating as well, i.e.

$$\Sigma_{i=1}^I E_i[\pi_i(x_o^+(m)),m] = \Sigma_{i=1}^I \pi_i(x_o^+(m)) + f_o(x_o^+(m))$$

for every $m \in M^I$ such that $x_o^+(m) \geq 0$.

It is important to recognize that under a satisfactory (but not optimal) control mechanism such as C^+ a division's best message m_i will depend on the messages of the other divisions. Thus to implement such a control mechanism some type of iterative adjustment process seems to be called for. However, by appending an iterative adjustment process to a satisfactory control mechanism opens up the question of what self-interested behavior of division managers would be. Essentially, under an iterative adjustment process, a division manager's strategy is a response rule--that is, a function describing the message to be sent by the manager at each stage of the iterative procedure, given the information acquired by the manager up to that stage. Under suitable conditions adjustment processes can be found for problems such as problem D^1 such that, if the division managers follow Cournot behavior (i.e. send their best message assuming the others messages remain fixed), a Nash equilibrium will be arrived at, in the limit at least.

However, Cournot behavior has been frequently criticized as unrealistic, especially in games with relatively few players (as we might suppose to be the case for even large organizations such as a divisionalized firm). Thus, what constitutes reasonable behavior remains to be defined and furthermore, given a behavioral assumption other than Cournot behavior, whether or not a control mechanism with an iterative adjustment process can be found that leads to solutions solving the underlying optimization problem such as D^1 remains an interesting problem.

Two results in this area have been obtained by Dreze and de la Vallee Poussin [8] and Hurwicz [20, 21] . Roughly speaking Dreze and de la Vallee Poussin were able to exhibit an iterative adjustment process and total profit allocating control mechanism such that the prescribed behavior leading to an optimal solution of the optimization problem is a maximin strategy for a player. However, they did not establish that every set of maximin strategies leads to an optimum and it appears in the case they analyze that other maximin strategies exist that dominate strategically the prescribed strategy.[23]

Hurwicz's results are largely negative. Viewing the problem as an n-person game where the strategies are response functions (not messages) he in effect, asks if a Nash equilibrium of response functions leads to optimal decisions where the decision rule is given by some control mechanism (as defined in this paper) and the argument of the decision rule is a fixed point of the Nash equilibrium of response functions. His results show that except for very special cases, if one confines the search to total profit allocating control mechanisms then no non-manipulable mechanism exists--that is, one such that a Nash equilibrium of response rules leads to optimal decisions.

However, it should be noted that Hurwicz's results depends crucially on the requirement that the control mechanism be total profit allocating. Dropping this requirement completely changes the results since for any optimal control mechanism such as any C^* in $\mathcal{C}^*$, a division manager's best message is independent of the messages of the others and thus the constant response function giving this message dominates any other response function.

Footnotes

1. We are concerned in this paper with the problem of designing the organization for decision making rather than decision implementation. That is, even for problems of type P that are small enough for one person to solve, then implementation of the optimal decisions may require many persons. In its broadest scope the organizational design problem would be concerned with the implementation problem as well.

2. See Marschak and Radner [26] and the references cited therein.

3. See, for example, Arrow and Hurwicz [2], Koopmans [23], and Malinvaud [25], and Reiter [27], a highly arbitrary sample.

4. See, for example, Dantzig [4], Dantzig and Wolfe [5 , 6], Geoffrion [9], and Jennergren [22], also an arbitrary sample.

5. Arrow, in his 1964 paper [1], mentions only Goode and McCarthy (complete references are not provided). The accounting literature frequently refers to incentive problems as does the economics literature in discussing both the competitive market system and the "free rider" problem in public goods and externality models. However, as far as I am aware, only Arrow [1] and Hurwicz [19, 19a, 20] have discussed this issue in a general form with the orientation taken in this paper.

5a. The relevant papers are Groves [11, 12], Groves and Ledyard [13, 14], Groves and Loeb [15], Green and Laffont [10] and Loeb [24].

6. This definition of an organizational form is to be distinguished from that of team theory, c.f. Marschak and Radner [26, p. 124].

7. See Jennergren [22] for a detailed study of this model. Jennergren also discusses the incentive difficulties with the Dantzig-Wolfe decomposition algorithm which may be applied to a special form of problem A in which the functions $F_i(\cdot)$ are linear and the constraints $\Sigma_i G_i(x_i) \leq K$ are represented by the linear inequalities:

$$A_i x_i \leq b_i$$

$$\Sigma_i A_{oi} x_i \leq b_o$$

8. See Hurwicz [18]; a seminal paper formalizing concretely communication processes. Team theory and, more generally, statistical decision theory also formulate models with explicit, though abstract, informational processes.

9. As required by various decomposition algorithms and decentralized planning procedures; see reference at note 4 supra. Of course, for an iterative process there is another side--namely messages from the Center to the division managers. It is unnecessary for this paper to formalize this aspect. See however Groves [11].

10. In Groves [11] a similar model is discussed within the team theoretic framework which allows for simultaneous decision making under uncertainty. The effect of the decision sequencing assumption made here is to allow sufficient information to be exchanged to solve the decision problem D. If decisions were to be made simultaneously or more generally if communication is restricted it may be impossible to solve D under any conditions. In such a situation the decision problem may be viewed as a problem under uncertainty and an objective of expected payoff maximization adopted. This is the approach of Groves [11].

11. See, for example, Horgren's text [17] or Demski [7].

12. At least in such a way as to motivate optimal decisions.

13. See Groves [11] and Loeb [24] for a discussion of "profit-sharing." See also discussion at note 14 below.

14. Under profit-sharing, a managers best message depends on the messages and decisions of the other divisions. Thus, profit-sharing is not compatible with an optimal control mechanism as defined here.

15. Although the title of this paper refers to "incentive compatible" control, the term "optimal" is used here since we <u>assume</u> the division managers maximize their evaluations.

16. This theorem is a combination and generalization of the results in Groves [12].

17. All the new results of this section are essentially contained in Green and Laffont [10]. However, since they do not require non-capriciousness of an optimal mechanism their version of Theorem 4 requires that there exist a <u>unique</u> dominant strategy equilibrium. In cases of non-unique dominant strategy equilibria, such as arise here for any optimal control mechanism C^* in $\mathcal{C}^*$, they define an "extended" mechanism in which the evaluation measures are not functions but correspondences. They also only prove their theorems for the very special case in which X_o contains only two points. The modifications and generalization contained here are, however, only rather trivial extensions of their basic ideas.

18. It would appear, however, that Hurwicz's results see below, answer this question in the negative. But a formal proof has yet to be constructed

19. Assuming no fixed costs and also that by abandoning the i^{th} division the constraint $g_i(x_i,x_o) \leq 0$ could be avoided.

20. Only slight modifications would be involved if $\Sigma_i T_i(m) \equiv 0$ for all $m \in M^I$ were required.

21. A rather complicated way of relaxing restrictions B.2 and B.3 is to separate the Center's function into two parts--one "communicating" like the divisions and the other choosing x_o and computing the evaluations E_i. Since this approach would require elaborate respecifications of the model and new definitions of a control mechanism and the relevant properties, it will not be persued here.

22. Proofs are contained in Groves and Ledyard [13, 14].

23. I am indebted to John Ledyard for this observation.

BIBLIOGRAPHY

[1] Arrow, K., "Control in Large Organizations," Management Science, Vol. 10, No. 3 (1964).

[2] ________ and L. Hurwicz, "Decentralization and Computation in Re-Source Allocation," in Pfouts, R. (ed.), Essays in Economics and Econometrics, University of North Carolina Press, Chapel Hill, North Carolina (1960), pp. 34-104.

[3] Berge, C., Topological Spaces, Macmillan Co., New York (1963).

[4] Dantzig, G., Linear Programming and Extensions, Princeton University Press, Princeton, New Jersey (1962), Chapter 23.

[5] ________ and P. Wolfe, "The Decomposition Algorithm for Linear Programs," Econometrica, Vol. 29, No. 4 (October 1961), pp. 767-778.

[6] ________________________, "Decomposition Principle for Linear Programs," Operations Research, Vol. 8, No. 1 (1960.

[7] Demski, J., "Evaluation Based on Controllable Performance," unpublished manuscript (1975).

[8] Dreze, J. and D. de la Vallee Poussin, "A Tatonnement Process for Public Goods," Review of Economic Studies, Vol. 38, No. 2 (April 1971), pp. 133-150.

[9] Geoffrion, A., "Elements of Large-Scale Mathematical Programming: Parts I and II," Management Science, Vol. 16, No. 11 (July 1970), pp. 652-701.

[10] Green, J. and J. J. Laffont, "Characterization of Strongly Individually Incentive Compatible Mechanism for the Revelation of Preferences for Public Goods," Discussion Paper No. 412, Harvard Institute of Economic Research, (May, 1975).

[11] Groves, T., "Incentives in Teams," Econometrica, Vol. 41, No. 4 (July, 1973), pp. 617-631.

[12] _________, "Information, Incentives, and the Internalization of Production Externalities," in Lin, S. (ed.), Theory and Measurement of Economic Externalities, Academic Press, New York (1975 forthcoming).

[13] _________ and J. Ledyard, "Optimal Allocation of Public Goods: A Solution to the Free Rider Problem," Working Paper No. 144, Center for Mathematical Studies in Economics and Management Science, Northwestern University (May, 1975).

[14] _________, "An Existence Theorem for an Optimal Mechanism for Allocating Public Goods," (1975, in preparation).

[15] _________ and M. Loeb, "Incentives and Public Inputs," Journal of Public Economics, Vol. 4 , No. 3, (August, 1975), pp. 211-226.

[16] _________, "Goal Congruent Evaluation of Divisional Management" (in preparation).

[17] Horngren, C., Cost Accounting: A Managerial Emphasis, 3rd Ed., Prentice-Hall, Englewood Cliffs, New Jersey (1972).

[18] Hurwicz, L., "Optimality and Informational Efficiency in Resource Allocation Processes," in Arrow, K., Karlin, S., and Suppes, P. (eds.), Mathematical Methods in the Social Sciences, Stanford University Press, Stanford (1960), pp. 27-46.

[19] _________, "The Design of Mechanism for Resource Allocation," American Economic Review, Vol. 63, No. 2 (May, 1973), pp. 1-30.

[19a] _________, "Organizational Structures for Joint Decision Making: A Designer's Point of View," in Tuite, M. Chisholm, R., and Radnor, M. (eds.), Interorganizational Decision Making, Aldine, Chicago (1972), pp. 37-44.