THE COST OF COMMUNICATION
IN
ECONOMIC ORGANIZATION

Hajime Oniki
I. Introduction and Background

The purpose of the present paper is to propose a method for measuring the cost of communication in alternative economic organizations allocating scarce resources. An economic organization, particularly if large and complex, needs internal communication in order to find desirable allocations, since economic knowledge relevant to resource allocation is initially dispersed among its members (individual agents). Information transmitted for this purpose may be of various kinds and differs among different organizations. It is possible, however, to measure the amount of information by using a unit called bit, and to express the cost of communication in terms of that unit. The foundation for this is provided by communication theory. In this paper we first introduce a method for approximating economic data in order to apply communication theory to our problem. We then compute the costs of communication for organizations using a price mechanism and for those under centralized control, limiting our attention to a simple problem of resource allocation.

Communication cost is one element of transaction or organizational costs (costs of running an economic system), as distinct from "real" or substantive costs (costs of production, transportation, etc.). The need for a theory of transaction costs arises from the fact that the overall efficiency of an economic organization depends not only on its performance in allocating resources but also on the amount of resources devoted to its operation. For example, it may not be desirable to introduce governmental
control on pollution if the administrative cost of control, an organizational cost, is so high as to outweigh the benefits from decreasing pollution.

The existence of transaction costs has long been recognized, if only from limited studies relating to particular problems. In a recent discussion of issues pertinent to the choice of market versus nonmarket allocation, Arrow pointed out that transaction costs must exist with any mode of allocation. The universal existence of transaction costs was also stressed by Hurwicz in his formulation of the problem of choosing desirable economic institutions.

A subclass of transaction costs to which economists have paid close attention is the informational costs of decision-making. Generally information useful for current decision-making may be obtained from the results of an organization's substantive activities in the past. If, however, the organization is large and complex, or if its environment is highly uncertain, it may pay for the organization to engage in informational activities in order to improve its decision. Informational costs are represented by the resources, human or nonhuman, devoted to such informational activities.

A comprehensive survey of informational activities in economic organizations was given by J. Marschak. He pointed out that informational activities could be regarded as a sequence of transformations of one type of information into another; he then identified three major "transformers": inquiring (data collection), communicating (message transmission), and deciding (computation). The existence of other transformers (encoding, decoding, storing, retrieval, etc.) was also acknowledged. Informational costs may be attached to each of these transformers.
In the present paper we limit our attention to communication costs only. The reason for doing so is that we can compute communication costs by applying communication theory (also called information theory), while no such theory is yet available for computing other informational costs. Perhaps this is because communication is one of the simplest transformers of information in the sense that its primary function is to reproduce input (message sent) as output (message received) as precisely as possible; unlike others its function is not to "transform" information. 8

We deal with a simple problem of resource allocation, to be solved by different organizations. There are two firms, each possessing technology for producing an output commodity from an input commodity, and the total supply of the input commodity is fixed. The problem is to allocate the input commodity between the two firms efficiently, i.e., to maximize the total output. This, of course, is done by equating marginal products between the firms. Suppose, however, that each firm knows its own technology but nothing of the other firm's technology nor of the total amount of input. Suppose further that there is a central agent who knows the total amount of the input but not the technology of the firms. 9 The problem is then how the firms and the central agent coordinate themselves to find an efficient allocation. An economic organization may simply be regarded as a set of behavior rules for the three agents such that by following the rules they may find an efficient allocation. 10 We study two organizations. The first is a centralized system in which each firm informs the central agent what its technology is, once and for all, and lets the latter do all the necessary calculations. The second is the ordinary market mechanism which follows the law of supply and demand; the central
agent is an auctioneer who adjusts prices until an equilibrium is reached. In each organization **economic data** are transmitted between the central agent and the firms in order to find an efficient allocation. The data may consist of prices, quantities of a commodity, or knowledge about production technology. Communication costs are attached to each transmission of economic data.

A fundamental property of communication (and, indeed, of any informational activity) is that its cost depends on its accuracy. Of course, if the accuracy is higher, or equivalently, if errors are smaller, the cost will be greater. (It is noted that the performance of an organization also depends on the accuracy of communication.) Two kinds of accuracy may be distinguished: the accuracy of a **channel** transmitting information and the accuracy of the information being transmitted. As for the former, communication theory has solved the problem of relating the accuracy of a channel to its transmission rate in such a unified way that the **amount of information** to be transmitted may be defined without specifying a channel, in particular, without specifying the accuracy of a channel. We make use of this fact, and do not explicitly consider the accuracy of a channel. On the other hand, the accuracy of information to be transmitted will be fully taken into account. In fact, economic data cannot be transmitted in full accuracy by a channel with limited capacity so that one cannot avoid introducing some measure of accuracy. Several methods for expressing the accuracy of, or equivalently the uncertainty of, economic data are possible. What we shall introduce in this study is a very simple method close to those employed in the practice of information processing.
Our results suggest that the relative magnitude of communication costs for the two organizations depends on the accuracy with which an efficient allocation is obtained. If high accuracy is required, then the communication cost for the market mechanism will be much lower than that for the centralized system, as expected. On the other hand, if one is satisfied with rather low accuracy, then the reverse may be the case.

Since the scope of this paper is admittedly limited, we do not intend to provide any overall evaluation of the organizations; the conclusions must be considered partial and tentative. However, we do intend to present methods which might be useful for studying the informational costs of decision-making. For instance, the way we approximate economic data, to be introduced in the following section, may be applicable to analyzing information transformers other than communication.

The problem of economic organization and information is so complicated that it has been dealt with from a number of different aspects. Arrow surveyed several points central to theoretical considerations. He has emphasized that the existence of uncertainty is a key factor in understanding the functioning of large organizations, and distinguished two sources of uncertainty: uncertainty about the external world (external uncertainty) and uncertainty within an organization (internal uncertainty). In Marschak's terminology, inquiring is relevant to external uncertainty, communicating to internal uncertainty. It has been suggested by Arrow that the presence of external uncertainty is one force for creating centralized organizations within a large organization (like firms in a whole economy) because of the economy of scale in inquiring. On the other hand, the problem of controlling large organizations
arises chiefly from the presence of internal uncertainty. In regard to this problem, he stressed the possibility of conflicts of interests between the organization and its members. Furthermore, he pointed out that the organization's behavior could be characterized by dynamic interactions among such phenomena as changes in internal uncertainty through communication among the members, learning from outcomes, and modifications of the organization's operating rules and the members' behavior patterns.

A team is an organization whose members possess a common goal so that no conflicts of interests exist among the members. The theory of teams was constructed by J. Marschak and Radner. The theory stresses the presence of both external and internal uncertainty, and considers observing the state of the external world and communicating it within the team. The team's (substantive) action is chosen so as to maximize the expected utility. The organizations analyzed here are special types of team. This is so particularly in two respects: First, we assume away external uncertainty (in our model the external state is expressed by the technology of each firm and the supply of input). Second, we do not explicitly introduce a utility function; instead we assume that the organizations maximize the total output. These simplifications together with the introduction of approximation have made it possible for us to compute communication costs in terms of Shannon's measure and compare them among the organizations. Marschak and Radner computed communication costs for certain teams in a conventional method (without using Shannon's measure). Marschak, however, has suggested a method for computing communication costs which is a primitive form of what we use in this study.

One of the central subjects of the theory of economic organization has been the choice between centralized and decentralized modes of allocation.
It was a basic issue in the well-known controversy among von Mises, Hayek, Taylor, Lange, Lerner, and others on the possibility of a rational allocation in a socialist state. The controversy resulted in an almost unanimous preference for decentralized decision-making by individual agents coordinated through a price mechanism. The reason for the preference was that the costs of communication and computation would be prohibitively high if the economy's decision-making were largely centralized. This is partly because the number of the agents is large and partly because information relevant to decision-making is difficult to transmit. This rather intuitive reasoning was accepted without further investigation then. Regarding the performance of the price mechanism, it is now well known that, roughly speaking, the competitive allocation is (Pareto) efficient and vice versa, provided that the environment of the economy is classical (divisible commodities, convex isoquants, the absence of externalities, etc.).

Two questions then arise concerning the preference for the decentralized price mechanism. First, is it at all possible to show that the costs of communication and computation are in fact much greater in the centralized allocation than in the decentralized allocation? Second, can we say anything if the environment is nonclassical?

To answer the first question, it is necessary to spell out the process of adjustments in centralized and decentralized modes of allocation. This task was done by a joint work of Arrow and Hurwicz for the decentralized mechanism of price adjustments. They formulated the adjustment process into a system of differential equations, and proved that, under certain assumptions, the solution converges to the efficient allocation. Also, Dantzig-Wolfe's decomposition principle for large linear programs is an algorithm which simulates the process of price adjustments, and has been so interpreted by the authors.
A systematic attempt to answer the first question was made by T. Marschak, utilizing the Arrow-Hurwicz adjustment process. He formulated a general model of resource allocation, and considered communication and computation (information handling) within an organization explicitly. He then assumed that the production technology was expressed by a set of linear activities, and computed these costs in both centralized and decentralized schemes. The unit of communication cost was the transmission of a real number, and the unit of computation cost was an arithmetic operation. It was found that the general preference for one of the centralized and the decentralized allocations over the other could not be established without further substantial restrictions on the model.

The purpose of our study is very close to that of T. Marschak: to compute informational costs of decision-making in centralized and decentralized allocation schemes. Aside from inessentials, however, we may point out the following differences: (1) Marschak considered both communication and computation costs, while we measure communication costs only. (2) He assumed that the production technology was expressed by linear activities, while we do not. (3) We use Shannon's measure of the amount of information to express communication cost, while he measured it by the number of real numbers transmitted. (4) Both Marschak and we agree that efficient allocations cannot be attained with perfect accuracy, and we both explicitly consider the errors (inaccuracy) of the allocations. The two formulations, however, differ significantly. We introduce approximation of economic data at the outset, thus making the choice of the accuracy level entirely exogenous to the analysis. Marschak assumed that, while the adjustment was proceeding, the previous (nonoptimal) allocation was in effect,
so that higher accuracy could be obtained by delaying the time to adopt
the optimal allocation. The desired accuracy may be chosen so as to
maximize the overall utility function.

The problem of comparing different organizations under the condition
that the economy's environment is nonclassical has been investigated by
Hurwicz.\footnote{22} He has formulated models of resource allocation and adjustments
so general that centralized and decentralized organizations are but their
special cases, thus providing a groundwork for comparing different organi-
zations. Various properties of environment and organizations are defined
to clarify concepts relevant to the problem. He stressed the informational
aspect of decentralization, and pointed out two conditions of informational
decentralization: initial dispersion of information and costliness of
channels transmitting information. A channel is consistent with informa-
tional decentralization if its capacity is such that economic data which
are finite dimensional vectors (like prices, commodities) can be transmitted
but those which are infinite dimensional vectors (like production functions
and utility functions) cannot. This distinguishes organizations whose
communication requirement is within that of the decentralized price mechan-
ism from those whose requirement exceeds it. The distinction is important
in evaluating models of economic planning, particularly those permitting the
presence of externalities.\footnote{23} He has further shown that informational decen-
tralization in the sense defined above and certain desirable properties of
an organization are likely to be inconsistent in the presence of externali-
ties.\footnote{24}

Our model is a special case of those presented by Hurwicz in several
respects. The main difference lies in that we do not adopt Hurwicz's
definition of informational decentralization but measure the cost of communication by introducing approximation of economic data. His definition of informational decentralization corresponds to one of our results that the communication cost for centralized systems is much higher than that for decentralized price mechanisms if the required level of accuracy is high.

II. Preliminaries: communication and the amount of information

As stated in the previous section, we shall use certain results obtained in communication theory in order to measure the cost of communication. This section is devoted to summarizing the part of communication theory that is needed for the present work.\(^\text{25}\)

For the present purpose, communication is defined as transmitting a message which states that a particular object has been selected from the set of possible objects.\(^\text{26}\) We assume that the set of possible objects is known to the sender and the receiver of messages, and that the set contains a finite number (=n) of objects. Denote the set by

\[
A = \{a_i\} = \{a_1, a_2, \ldots, a_n\},
\]

where \(a_i\) is a possible object.\(^\text{27}\) The message stating that the choice has been \(a_i\) will be called message \(a_i\). We assume that the probability for message \(a_i\) to occur is \(p_i\):

\[
P = \{p_i\} = \{p_1, p_2, \ldots, p_n\},
\]

where \(p_i > 0\), \(\sum p_i = 1\).

A message is transmitted through a channel in a coded form. The means of coding could be letters, numerals, words, sounds, voices, electric currents or pulses, etc. The power of communication theory lies in its independ-
ence of the choice of a particular channel and coding. Consider binary coding: a coding such that a coded word (or code) which represents a message is a sequence of binary digits (bits), 0 and 1. The length of a code is the number of bits composing it. 28

The cost of transmitting message \( a_1 \) depends on the length of its code. The length of each code, of course, is determined by how each message is coded. It is advantageous to give short codes to messages transmitted frequently, and to give long codes to messages transmitted less frequently, since this minimizes the average length of a code.

It has been shown by Shannon that, if each of the successive messages occurs independently, then the greatest lower bound of the average length of a code is equal to

\[
H(P) = - \sum p_i \log p_i (\geq 0),
\]

where the base of the logarithm is taken to be 2. 29 The right-side of (3) denotes the amount of information transmitted by a message designating an element of \( A \). It is also called the entropy of the system \( (A, P) \), which measures the "degree of uncertainty" of the system. The greater is the degree of uncertainty, the more is the amount of information which resolves it.

If, in particular, each message occurs with equal probability so that \( p_i = 1/n \) for all \( i \), then

\[
H(P) = H(n) = \log n,
\]

as can easily be derived from (3). 30, 31 It is noted that (4) is the maximum of (3) in \( P \). In other words, the amount of information given by a message is maximized when one knows nothing about which message is likely to occur.
The capacity of a channel is defined to be the maximum number of bits which can be transmitted per unit of time. Therefore, if the capacity of a channel is denoted by $C$, then the greatest lower bound of time needed to transmit information in the amount $H(P)$ is equal to $H(P)/C$. A remarkable result which Shannon has obtained is that the above relation can be extended to noisy channels. A noisy channel transmits bits with possible errors: 0 as 1 or 1 as 0. If the probability law governing errors in transmission is known, then it is possible to define the capacity from that probability law only, so that the relation among the capacity, the amount of information, and the transmission time is precisely the same as in noiseless channels.

The cost of communication includes the cost of constructing and operating a channel transmitting messages. In considering the cost, we may take two viewpoints, short-run and long-run. From the short-run viewpoint, we suppose that a channel, noisy or noiseless, of some capacity has already been given. Then the amount of information determines the transmission time, which together with the operating cost of the channel constitutes the (variable) cost of communication. From the long-run viewpoint, we suppose that a channel is yet to be designed and constructed. The choice of an optimum channel depends upon various engineering and economic factors. Since, however, the amount of information is defined independently of channels, we can consider it as the determinant of the long-run cost of communication.

III. The model and the approximation of economic data

In the present section, we formulate a simple problem of resource allocation into a model with respect to which the costs of communication are computed. First, we state the model in the ordinary terminology of economics.
Next, we introduce a method for approximating the economic data which appear in the model formulated, and restate the model in terms of the approximated data. The revised model thus obtained is suitable for computing communication costs.

(1) The original model

Let us assume that there exist two firms, to be denoted by the index \( i = 1,2 \). Each firm possesses technology represented by a production function

\[
y_i = f_i(x_i) \quad (i=1,2),
\]

where \( x_i \) is the input commodity and \( y_i \) the output commodity. It will be assumed that the production function \( f_i \) is continuous, nondecreasing, and concave for all nonnegative \( x_i \)'s.

We next assume that the total supply of the input commodity is fixed at \( \bar{x} > 0 \). The objective of the firms is to find an allocation \((x_1, x_2)\) which maximizes the total output

\[
y = y_1 + y_2 = f_1(x_1) + f_2(x_2),
\]

subject to the input constraints

\[
x_1 + x_2 \leq \bar{x}, \quad x_i \geq 0 \quad (i=1,2).
\]

If the production functions are differentiable, then an interior optimum satisfies

\[
f'_1(x_1) = f'_2(x_2).
\]

(2) The approximation of the economic data

The formulation stated in the previous paragraph is not suitable for the present purpose: to consider informational costs of decision-making.
in general and to compute communication costs in particular. This is because almost any economic decision is made within a limited range and with limited accuracy, upon which the costs of decision-making depend intrinsically. For instance, the knowledge about the production technology may not be so precise as implied by the definition of a function but may be a fuzzy object. In realities, the optimum condition (8) may be satisfied only to some approximate degree. The range of a commodity interesting to a firm is not properly represented by the nonnegative half-line, since a firm never considers producing an infinite (or a very large) amount of goods. Further, price data may not be transmitted with a perfect precision, for this would imply that an extremely large number of digits need to be transmitted; the number of digits in economic data significant to decision-making seldom exceeds three or four. In daily economic activities we allow the existence of errors, uncertainty, or fuzzyness in all economic data we are concerned with. Below, we introduce a simple method for expressing this fact.

Let us first consider the input commodity, its quantity being denoted by \( x \). We assume that there is an upper bound of \( x \) such that the quantity of \( x \) beyond this bound is never considered. For simplicity, we take the total supply \( \bar{x} \) to be this upper bound. (Note that it is safe to do so for the maximization problem (6) and (7).) We may then limit our attention to the interval

\[
X = \{ x; \ 0 \leq x \leq \bar{x} \}. \tag{9}
\]

Furthermore, we assume that it is not interesting to know a precise quantity of \( x \), but that it is sufficient to have an approximate value only. To do
this, let us adjust the unit of \( x \) so that \( \bar{x} = m \), where \( m \) is a positive integer given exogenously. We then assume that only integer values of \( x \) are considered. Define the sets of approximate values of \( x \) by

\[
M = \{1, 2, \ldots, m\},
\]

\[
M_0 = M \cup \{0\}. \tag{10}
\]

The choice of \( m \) determines the degree of accuracy in this approximation. \(^3\)

Let \( p \) be the price of the input commodity in terms of the output commodity. We assume that \( p \) is approximated in the same way as \( x \) is:

\[
P = \{p; 0 \leq p \leq \bar{p}\},
\]

\[
N = \{1, 2, \ldots, n\},
\]

\[
N_0 = N \cup \{0\}. \tag{11}
\]

where \( \bar{p} > 0 \) is the upper bound of \( p \), \( n \) is a positive integer chosen exogenously, and the unit of the output commodity is adjusted so that \( \bar{p} = n \). The output commodity \( y \) is approximated likewise:

\[
Y = \{y; 0 < y \leq \bar{y} = \bar{p} \bar{x} + 1\},
\]

\[
L = \{1, 2, \ldots, mn + 1\}. \tag{12}
\]

The upper bound \( \bar{y} \) is chosen for analytical convenience. \(^4\)

Let us next consider approximating production functions. First of all, we assume that the class of production functions which possibly arise in the allocation problem can be expressed by

\[
F = \{f: X \rightarrow Y; f \text{ is continuous, nondecreasing, and concave}; \\
0 < f(0) \leq 1, \quad 0 \leq f'(0) \leq \bar{p}\}, \tag{13}
\]

where \( f'(0) \) denotes the right-directional derivative of \( f \) at 0. Figure 1 illustrates the range of the graphs of production functions belonging to \( F \).
Among the conditions which characterize the set $F$, the first three are standard in economic theory, calling for no further comments. The last two conditions, stated in terms of inequalities, are imposed chiefly for analytical convenience. 35

In order to take into account the fact that information about production technology is never perfectly accurate, we "approximate" the set $F$ with a finite set, as we approximated previously the sets $X$, $P$, and $Y$ with the sets $M$, $N$, and $L$, respectively, all being finite sets. We do this by considering the values of a production function $f \in F$ on the set $M$, and then approximating those values with points in the set $L$. First of all, let us choose $g(j) = g_f(j) \in L$ such that

$$g(j)-1 < f(j) \leq g(j), \quad j \in M; \quad (14)$$

i.e., $g(j)$ is an integer approximating $f(j)$. The subscript $f$ attached to $g$ reminds us that $g$ is obtained from $f$; the subscript may be omitted if no ambiguity arises. Define
We put \( g(0) = 1 \). We then immediately have
\[
g(j) = g(0) + \sum_{j' = 1}^{j} h(j'), \quad j \in \mathbb{M}. \tag{16}
\]

Inequality (14) shows that the function \( g = g_f \) approximates the original production function \( f \) on \( \mathbb{M} \). We call \( g_f \) the approximate production function of \( f \). From definition (15), it is expected that the function \( h \) approximates the slopes of \( f \). In fact it can be shown that
\[
0 \leq h(j) \leq n, \text{ i.e., } h(j) \in \mathbb{N}_0, \quad j \in \mathbb{M}; \tag{17}
\]
\[
h(j) - 1 < f(j) - f(j-1) < h(j) + 1, \quad j \in \mathbb{M}. \tag{18}
\]
(Proof is given in Appendix I.) As might be suggested by (18), it is not true that the function \( h(j) \), unlike the slope \( f(j) - f(j-1) \), is nonincreasing in \( j \).

We write, for simplicity,
\[
g = g_f = (g(1), \ldots, g(m)) \in \mathbb{L}^m,
\]
\[
h = h_f = h_g = (h(1), \ldots, h(m)) \in \mathbb{N}_0^m. \tag{19}
\]

Define the set of approximate production functions by
\[
G = G_f = \{g_f; f \in \mathbb{F}\} \subset \mathbb{L}^m. \tag{20}
\]

The fact that \( g = g_f \) approximates \( f \) can be shown more precisely. To do this, we extend \( g \), which has been defined only on \( \mathbb{M} \), to the set \( \mathbb{X} \):
\[
g(x) = g([x]) + h([x] + 1)(x-[x]), \quad x \in \mathbb{X}, \tag{21}
\]
where \([x] \in \mathbb{M}_0\) is the greatest integer not exceeding \( x \). The function \( g(x) \)
thus extended is a piecewise linear function; its graph is obtained by joining $(m+1)$ points $(j, g(j))$ successively by segments. The remark in the preceding paragraph implies that $g(x)$ may not be concave. It can be shown (see Appendix II for proof) that

$$|g(x) - f(x)| < \max[n/4, 1], \quad x \in X. \quad (22)$$

That is to say, if $n$ is large, the error of the approximation does not exceed $n/4 = (y-1)/4m$. This implies that the error can be made arbitrarily small if a sufficiently large $m$ is chosen. The approximation of $f$ by $g = g_f$ is illustrated in Figure 2.

**Figure 2**

In the preceding paragraphs, we introduced approximate economic data for the input and the output commodities, the price of the input commodity, and the production technology. We now turn to stating the demand function for the input commodity $x$ in terms of the approximate data. Let us first define
\[ A = A(\pi) = A(\pi, h) \]

\[ = \{ j \in M; h(j') \geq \pi \text{ for all } j' \in M \text{ such that } j' \leq j \}, \]

\[ \overline{A} = \overline{A}(\pi) = \overline{A}(\pi, h) \]

\[ = \{ j \in M; h(j') \leq \pi \text{ for all } j' \in M \text{ such that } j' \geq j \}; \tag{23} \]

\[ \underline{\xi} = \underline{\xi}(\pi) = \underline{\xi}(\pi, h) = \max A, \]

\[ \overline{\xi} = \overline{\xi}(\pi) = \overline{\xi}(\pi, h) = \min \overline{A}, \tag{24} \]

where \( \pi \in N_0 \), \( h = h_f \), and \( f \in F \). It is noted that \( A \) and \( \overline{A} \) are never empty

(For \( 0 \in A \) and \( m \in \overline{A} \) for any \( \pi \) and \( h \)), so that the functions \( \underline{\xi} \) and \( \overline{\xi} \) are well defined by (24).

We show that the functions \( \underline{\xi}(\pi) \) and \( \overline{\xi}(\pi) \) may be interpreted to be the minimum and the maximum, respectively, of the "demand correspondence" at \( \pi \) for the input commodity approximated. In other words, these functions in a sense characterize the "inverse" of the function \( h \); i.e., for any \( \pi \in N_0 \),

\[ 0 < j \leq \underline{\xi}(\pi) \text{ implies } \pi + 1 \leq h(j); \]

\[ \underline{\xi}(\pi) < j \leq \overline{\xi}(\pi) \text{ implies } \pi - 1 \leq h(j) \leq \pi + 1; \]

\[ \overline{\xi}(\pi) < j \leq m \text{ implies } h(j) \leq \pi - 1. \tag{25} \]

(For proof see Appendix IV.)\(^{39}\) For simplicity, we call \( \underline{\xi}(\pi) \) and \( \overline{\xi}(\pi) \) the approximate demand functions for the input commodity. It can also be shown that the demand correspondence formed by \( \underline{\xi} \) and \( \overline{\xi} \) is "continuous" and "non-decreasing" (proved in Appendix III):

\[ \overline{\xi}(\pi-1) \geq \overline{\xi}(\pi) \geq \underline{\xi}(\pi-1) \geq \underline{\xi}(\pi), \quad \pi \in N, \tag{26} \]

\[ \overline{\xi}(0) = m \text{ and } \underline{\xi}(0) = 0. \tag{27} \]

Figure (3) illustrates (26).
(3) The model in terms of the approximate data

The allocation problem formulated previously as (6) and (7) can now be restated in terms of the approximate data which we have just introduced. Suppose that information relevant to the allocation problem is dispersed among the three economic agents: $f_i$ is known to firm $i$ only, and $x$ is known to the central agent only. Suppose further that informational activities for finding an efficient allocation can be carried out by means of the approximate data only. The objective of an organization is then to find a set of allocations, say $A \times > 0 (i=1,2)$, and

$$x_1 + x_2 \leq \bar{x}, \quad x_i \geq 0 \quad (i=1,2), \quad \text{and}$$

$$|f_1(x_1) + f_2(x_2) - y^*| < d,$$

for all $(x_1, x_2) \in A$, where $y^*$ is the maximum output in the original allocation problem in which no restriction is imposed on informational activities, and $d$ is a number denoting the bound of allocation errors.

It is clear that within our framework an organization is characterized by an algorithm for finding a set $A$. Informational activities of each agent
and, hence, informational costs will precisely be determined once an algorithm is specified. Allocation errors are also determined by an algorithm. A general approach to the problem of comparing alternative organizations would be to consider the tradeoff between informational costs and allocation errors. In the present paper, however, we do not adopt such a general approach but limit our attention to answering certain special problems only: For a centralized organization, we compute communication costs. For a decentralized market mechanism, we compute communication costs and estimate allocation errors. All of these tasks will be done in the following section.

IV. The Organizations and the Communication Costs

The present section is devoted to computing communication costs for the two organizations. The discussion of each organization is composed of two parts: (a) to describe the organization's informational activities for solving the (approximate) allocation problem and, especially, to describe the process of communication among the three agents; (b) to compute the costs of communication by using Shannon's measure.

(1) Centralized system

(a) The centralized system to be considered here is like that discussed at the early stage of the socialist-planning controversy. The central agent plays the role of the State Government or the Central Planning Board. It first collects from each firm all the information necessary to compute an efficient allocation. After computation, the central agent informs the efficient allocation to each firm. In terms of the approximate data introduced in the preceding section, the internal communication can be stated as follows: Each firm sends its approximate production function \( g_f = g_{fi} \) to the central agent. The central agent computes a set \( A \) of approximately efficient allocations by
using the information given by \( g_1, g_2, \) and \( m (= x) \). It then chooses from \( A \) an approximately efficient allocation, say \((j_1, j_2) \in M^2\), and sends it back to the firms.

(b) Communication costs for this system arise from two sources: transmission of \( g_1, g_2 \), and transmission of \((j_1, j_2)\). It was the cost of transmitting the \( g_1 \)'s that was considered prohibitively high by most of the participants to the planning controversy at its later stage; centralized decision making for a whole society was considered not feasible. 37

We compute these costs in terms of Shannon's measure. To do this, we have to consider the sets of possible objects of transmission; they are \( G \) and \( M \) for the present case. If we compute \( |G| \) (where \( |G| \) denotes the number of elements in \( G \)), then \( \log |G| \) would determine the communication cost of transmitting \( g_1 \). As summarized in Section II, \( \log |G| \) approximates the minimum number of bits needed to transmit \( g_1 \), if one does not know the probability with which \( g \in G \) arises so that a code of equal length is assigned to each element of \( G \). Likewise, the communication cost of transmitting \( j_1 \) would be determined by \( \log |M| \).

It is difficult to obtain a simple formula with which \( |G| \) can be computed; hence, we consider estimating \( |G| \) indirectly. Define

\[
G = \{g \in M; \ h_g(j-1) \geq h_g(j), \text{ for } j=2, \ldots, m, \text{ and } h_g \in N_0^m \},
\]

\[
G^* = \{g \in M; \ h_g(j-1) + 1 \geq h_g(j), \text{ for } j=2, \ldots, m, \text{ and } h_g \in N_0^m \}. \quad (29)
\]

An element of \( G \), if extended to \( X \) according to formula (21), is a piecewise linear nonincreasing concave function from \( X \) into \( Y \). Also, an element of \( G^* \), if extended to \( X \) likewise, is a piecewise linear continuous function from \( X \)
into \( Y \) with slopes either nonincreasing or increasing at most by one. It is shown in appendix V that

\[
G_* \subseteq G \subseteq G^*.
\]  

(30)

It turns out that

\[
|G_*| = \binom{n + m}{m},
\]

\[
|G^*| = \binom{n + 2m - 1}{m}.
\]  

(31)

(See Appendix VI for proof.)

We then obtain, making use of the formulas \( \binom{a}{b} = \frac{a!}{b!(a-b)!} \) and \( a! = a^ae^{-a\sqrt{2\pi}} \) (Sterling's formula), the following equations:

\[
\log |G_*| = \log \frac{(m+n)(m+n+1/2)}{(m+1/2)(n+1/2)} + r,
\]  

(32)

\[
\log |G^*| = \log \frac{(2m+n-1)(2m+n-3/2)}{(m+1/2)(m+n-1)(m+n-3/2)} + r,
\]  

(33)

where \( r = -\log \sqrt{2\pi} \). If, in particular, \( m=n \), then (as proved in Appendix VII)

\[
\log |G_*| = m(\log 4 + o(m)) + r,
\]

\[
\log |G^*| = m(\log 4 + o(m)) + r,
\]  

(34)

where \( \lim_{k \to \infty} o(k) = 0 \). Since (30) implies \( |G_*| \leq |G| \leq |G^*| \), we can assert that the amount of information \( \log |G| \) increases approximately linearly if \( m = n \) and \( m \) is large.

The cost of transmitting an optimal allocation from the central agent to a firm is easily calculated; it is equal to \( \log |M| = \log m \).
The total amount of information transmitted within this centralized organization is (note that there are two firms)

\[ C_1 = 2\log |G| + 2\log |M|. \quad (35) \]

In view of (34), we obtain (if \( m = n \))

\[ 2\left\{ m(\log 4 + o(m)) + r + \log m \right\} \leq C_1 \leq 2\left\{ m(\log 27/4 + o(m)) + r + \log m \right\}. \quad (36) \]

(2) The competitive market mechanism

(a) This allocative mechanism reflects the working of a competitive market which follows the law of supply and demand. The central agent is the auctioneer of the market. It may represent the Central Planning Board in a socialist state adjusting prices according to Lange-Taylor's scheme, or it may be the invisible hand of a capitalistic market in which price adjustment follows the tâtonnement process.

As is well known, the adjustment in this system proceeds as follows: Given a current market price, each firm computes the demand for each commodity and then reveals the demand at the market. This gives the central agent the aggregate excess demand for each commodity. The central agent then raises (lowers) the price if the excess demand is positive (negative). The process is continued until an equilibrium is reached.

For the allocation problem we are dealing with, this adjustment process may be stated in terms of the approximate data in the following way: First, given a current approximate price \( \pi \in \mathbb{N}_0 \), firm 1 responds to the market with the approximate demand functions \( \bar{f}_1(\pi) = \bar{f}_1(\pi, h_1) \) and \( \bar{f}_2(\pi) = \bar{f}_2(\pi, h_2) \), where \( h_i = h_i \) (\( i = 1, 2 \)). Then, the central agent computes the approximate aggregate demand. If the equilibrium condition

\[ \bar{f}_1(\pi) + \bar{f}_2(\pi) \leq m \leq \bar{f}_1(\pi) + \bar{f}_2(\pi) \]

(37)
is satisfied, the process is terminated. Otherwise, the central agent adjusts the current price $\pi$ following the rule of supply and demand, and announces a new price chosen from $N_0$.

We first consider the equilibrium condition (37). Let $\pi^*$ denote an approximate price satisfying (37). From (26) and (27), it can easily be derived that such $\pi^*$ always exists. Furthermore, we can estimate a bound of allocation errors at $\pi^*$ (for proof see Appendix VIII):

$$|f_1(x_1) + f_2(x_2) - y^*| < \max[n,4] + 2m = d, \quad \text{(38)}$$

for all $(x_1, x_2)$ such that $\xi_1(\pi^*) \leq x_1 \leq \xi_2(\pi^*)$ (i=1,2) and $x_1 + x_2 = \bar{x}$. This implies that the allocation error can be made arbitrarily small by choosing sufficiently large $m$ and $n$, since $d/\bar{y} \to 0$ as $m, n \to \infty$ ($d/\bar{y} = 3/m$, if $m = n \geq 4$).

Let us next consider the scheme for price adjustment in more detail. To do this, we need to specify (i) a rule for choosing the initial price, and (ii) a rule for choosing a new price at each step of adjustments. (The law of supply and demand does not fully determine the choice of a new price; it determines the direction of the adjustment, but not its magnitude.)

We assume that the rules are chosen in such a way that the largest possible number of adjustment steps be minimized. In other words, we assume that the initial price is equal (or close) to the midpoint of the segment $X(= m/2)$, and that the subsequent choice of new prices follows the bisectioning rule approximately. If $m = 2^k$ for some positive integer $k$, then it is possible to follow the bisectioning rule precisely:

$$\pi(1) = 2^{k-1},$$
$$\pi(t) = \pi(t-1) \pm 2^{k-t}, \quad \text{(39)}$$

where $\pi(1)$ is the initial price and $\pi(t)$ is the price chosen at the $t$-th
step of adjustments. It is clear that, for this special case, the largest possible number of steps is equal to \( k = \log_2 m \). If \( m \) is not a power of 2, the bisectioning rule may be followed approximately, and the largest possible number of steps is approximately equal to \( \log_2 m \).

The actual number of steps needed to reach an equilibrium depends on the exogenously given data \((g_1 \text{ and } g_2)\) and the choice of the initial price. It is possible, if not likely, that the initial price hits the equilibrium; no adjustment is necessary then. A sophisticated treatment of this problem would be to compute the expected number of steps needed, assuming that some probability distribution on the set \( C \) is given. In this study, however, we consider the upper bound of the number of steps only.

(b) Given these assumptions, we can easily compute the communication cost. The amount of information sent and received by each firm in one step is \( \log n \) (for transmitting price \( n \)) plus \( 2\log m \) (for transmitting the demand correspondence \( (\overline{x}_1, \overline{x}_2) \)). Hence, we may assert that

\[
C_2 \leq 2\log_2 m (\log n + 2\log m),
\]

where \( C_2 \) is the total communication cost for this system. If, in particular, \( m = n \), then we have,

\[
C_2 \leq 6(\log m)^2 / \log 2.
\]

(3) **Comparison of the two organizations**

In Table 1, we compute the lower and the upper bounds of \( C_1 \) and the upper bounds of \( C_2 \) for some values of \( m \) (we assume \( m = n \) throughout).
Table 1

<table>
<thead>
<tr>
<th>m</th>
<th>Centralized system</th>
<th>The market mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2(mlog4 + logm)</td>
<td>2(mlog(\frac{27}{4}) + logm)</td>
</tr>
<tr>
<td>4</td>
<td>4.89</td>
<td>6.24</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>16.8</td>
</tr>
<tr>
<td>100</td>
<td>124</td>
<td>170</td>
</tr>
<tr>
<td>1000</td>
<td>1210</td>
<td>1665</td>
</tr>
<tr>
<td>10000</td>
<td>12049</td>
<td>16594</td>
</tr>
</tbody>
</table>

Notes to Table 1:

1. The base of logarithm is 10.
2. Figures with * are computed directly from (30), (31), and (35).

We can assert from this table that if m is set greater than 100, that is, if it is required that the relative allocation errors be less than 3% (= 3/100; see (38)), then the communication cost for the market mechanism is definitely lower than that for the centralized system. If, however, we allow somewhat higher relative allocation errors, then it is possible that the centralized decision making is cheaper with respect to communication costs than the market allocation.

Since we are entirely disregarding computation costs of resource allocation, the two organizations can be compared only partially; in general one cannot obtain a definite preference for one over the other from the
results which have so far been obtained. If, however, we can assume that (i) in the centralized system the central agent computes efficient allocations by simulating the adjustment process for the market mechanism, and that (ii) the resources devoted to computation are transferable freely between the central agent and the individual firms (this may be so in the long run), then the difference in communication costs may indicate the difference in the overall efficiency of the organizations.

V. Conclusions and some further remarks

In the preceding sections, we considered a simple problem of allocating resources between two firms. We introduced a method for approximating economic data with finite and discrete sets and computed the costs of internal communication in two organizations seeking efficient allocations. The comparison of communication costs confirms that the decentralized market mechanism is more economical if the required accuracy of allocation is high, but it also suggests that the centralization of information may pay if the required accuracy is relatively low.

We are interested in comparing different economic organizations, since in the real world the environment is often nonclassical (e.g., externalities like pollution) and the market mechanism is likely to fail in nonclassical environment. Further, we are interested in comparing different organizations with respect to communication costs, or more generally with respect to transaction or organizational costs, since these costs affect the organizations' overall efficiency (if one takes the normative viewpoint) or their viability (if one takes the descriptive viewpoint).

The main purpose of this paper, however, is not to derive a general conclusion about the preference on the centralized system and the market
mechanism, nor to consider the allocation problems in the presence of externalities, but to propose certain analytical tools which might turn out to be useful for further studying these problems. Partly for this reason, and partly for the reason of mathematical difficulty, we kept our model within a very simple and limited framework. Below, we consider several possible extensions briefly: (i) "The command system," an adjustment mechanism which is like the market mechanism except that the roles of price and quantity are interchanged: Marglin once pointed out that the market mechanism and the command system are informationally equivalent. I confirmed this (with respect to communication costs) by using a method similar to that used in this paper for analysing the market mechanism. (ii) The case of many firms: All the conclusions obtained in this paper will be carried through, since communication costs are linear in the number of firms. If, however, more than two firms are present, then the possibility of organizations other than those considered in this paper arises (e.g., partially centralized organizations). (iii) The case of many input (and output) commodities: It is mathematically difficult to evaluate $|G|$, if the number of commodities is more than two. Also, it is difficult to extend the "bisection rule" to this case. (iv) The case of increasing returns or nonconcave production functions: It is possible to evaluate $|G|$ at least for the single-input single-output case, so that the centralized system can be dealt with easily. The difficulty seems to be in formulating a decentralized adjustment process in a suitable way. (v) Sequential communication, computation, and decision making, allowing the possibility of variable degrees of accuracy: This is an interesting but seemingly difficult case, on which we have made no investigation.
The discrete approximation of economic data is central to our method. It was introduced into our model based on the observation that any data used for information processing is fuzzy and contains errors. The need for considering the accuracy of data explicitly was demonstrated by the sensitiveness to it of the relative communication costs in the two organizations. It is noted, however, that our method—discrete approximation—is not the only way to do this. For example, to economic data may be added error terms distributing probabilistically. Our method was useful but it also brought inconveniences. First of all, we note that if information is processed exclusively by digital devices, then discretization is natural, since this is the form in which a digital device handles information. But in the real world, of course, information is commonly processed in non-digital forms—typically by human beings. It seems to me that discretization is also a useful way to measure the costs of non-digital information processing, but this is not established. Apart from these points, the usefulness of discrete approximation lies in that (i) it makes everything finite and elementary so that Shannon's measure can easily be used, (ii) and that it makes the adjustment rule for the price mechanism very simple. One of its shortcomings may be expressed by the fact that the Appendices are complicated and lengthy; discrete approximation leaves irregularly distributed residuals, which cannot be dealt with in a simple setting. Whether there is any method for expressing the fuzziness of data which is better than discrete approximation seems to be an open question.
Appendix I (Proof of (17) and (18)): Since $f$ belongs to $F$, (13) implies that $0 \leq f(j) - f(j-1) \leq \frac{p}{2}$. We write (14) as
\[
\begin{align*}
g(j-1) - 1 &< f(j-1) \leq g(j-1), \\
g(j) - 1 &< f(j) \leq g(j).
\end{align*}
\] (a1)

From the above inequalities, we derive
\[
\begin{align*}
0 \leq f(j) - f(j-1) &< g(j) - g(j-1) + 1, \\
g(j) - 1 - g(j-1) &< f(j) - f(j-1) \leq \frac{p}{2} = n.
\end{align*}
\] (a2)

Since $g(j)$'s are integers, we obtain $0 \leq g(j) - g(j-1) \leq n$, which is (17) (See (15)). Also, (18) follows from (a2) and (15).

Appendix II (Proof of (22)): Figure 4a depicts the graph of $g(x)$ and the range of the graphs of $f(x)$ on the interval $(j-1, j)$. The graph of $g(x)$ is segment AC. The graphs of $f(x)$ intersect segments AA' and CC'. Since $f(x)$ is nondecreasing and concave, and its slope is not greater than $n$, its graph is contained in ABCC'A', where the slope of AB is n and BC is horizontal. Therefore, $|g(x) - f(x)|$ does not exceed the maximum of BD and AA' ($=1$), where BD is vertical. It then suffices to show that $BD \leq n/4$.

Let $\eta = h(j)$, which is the slope of AC. Figure 4b reproduces a part of Figure 4a. Let AF be horizontal, and BE and CF be vertical. Then, $BD = BE - DE = \eta - \frac{\eta^2}{n}$. Hence BD is maximized when $\eta = n/2$, and max BD = $n/4$. 
Appendix III (Proof of (26) and (27)): By the definition of $A$ and $\overline{A}$ (see (23)),

$$A(n-1) \supseteq A(n), \quad \overline{A}(n-1) \subseteq \overline{A}(n). \tag{a3}$$

Then, (24) implies the first and the third inequalities of (26). To show the second inequality of (26), suppose that it did not hold: $\overline{A}(n) < \overline{A}(n-1)$. Then, there would exist a $j \in M$ such that $\overline{A}(n) < j < \overline{A}(n-1)$. Since $\overline{A} \subseteq A$ and $\overline{A} \supseteq A$, we would have $h(j) < n$ and $h(j) > n-1$, which is a contradiction since $h(j)$ is an integer.

Appendix IV (Proof of (25)): Let $\xi = \overline{A}(n)$ and $\overline{\xi} = \overline{A}(n)$. The first and the third propositions of (25) follow directly from $\overline{A}(n) \subseteq \overline{A}(n)$ and $\overline{A}(n) \supseteq A(n)$, respectively. It remains to show the second proposition of (25). To do this, observe first that the proposition presupposes $\xi < \overline{\xi}$. We then have $\xi < m$ and $0 < \overline{\xi}$. It follows from (23) and (24) that

$$h(\xi + 1) < n, \quad h(\overline{\xi}) > n. \tag{a4}$$
(i) Assume $\xi + 1 \leq j$. Then, from (18) and the concavity of $f$, we get

$$h(j) - 1 < f(j) - f(j-1) \leq f(\xi+1) - f(\xi) < h(\xi+1) + 1.$$ 

In view of (a4) we have $h(j) \leq h(\xi+1) + 1 \leq \pi + 1$. (ii) Assume $j < \xi$.

Then, again from (18) and the concavity of $f$, we get

$$h(\xi) - 1 < f(\xi) - f(\xi-1) \leq f(j) - f(j-1) < h(j) + 1.$$ 

Hence, we have $\pi - 1 \leq h(\xi) - 1 \leq h(j)$. Combining (i) and (ii) above, we obtain the second proposition of (25). (It is noted that, if $\xi + 1 = \xi$, then $h(\xi+1) = h(\xi) = \pi$ from (a4).)

Appendix V (Proof of (30)): (i) Let $g \in G_\star$. Define an extension $f$ of $g$ to $X$ by $f(x) =$ (the right side of (21)). Then, clearly $g = g_f$ and $f \in F$; i.e., $g \in G = G_F$. This proves the first half of (30). (ii) Let $g \in G = G_F$. Then, there exists an $f \in F$ such that $g = g_f$. Inequality (18) together with the concavity of $f$ implies that $h(j) - 1 < f(j) - f(j-1) \leq f(j-1) - f(j-2) < h(j-1) + 1$, $j=2, \ldots, m$, where $h = h_g = h_f$. From this we get $h(j-1) + 2 > h(j)$, i.e., $h(j-1) + 1 \geq h(j)$ (since $h(j)$'s are integers). This implies $g \in G_\star$, proving the second half of (30).

Appendix VI (Proof of (31)): (i) For a given $g \in G_\star$ and the corresponding $h = h_g$, define a function $h'$ by

$$h'(j) = h(j) - (j-1), \quad j \in M.$$  

(a5)

Then, $h'$ is strictly decreasing (since $h$ is nonincreasing), and maps $M$ into a set $N' = \mathbb{N} \cup \{0, -1, -2, \ldots, -(m-1)\}$. Conversely, if $h'$ is a strictly decreasing function which maps $M$ into $N'$, then the function $h$ obtained from this $h'$ by (a5) is nonincreasing and maps $M$ into $N$. Hence, the function $g$
obtained from this $h$ by (16) belongs to $G_\pi$. In other words, the correspondence between $g$ and $h'$ determined by (15) and (a5) is one-one. Therefore, $|G_\pi|$ is equal to the number of strictly decreasing functions $h'$ which map $M$ into $N'$. Such an $h'$ is specified by first choosing $m$ elements, say, $(k_j)$ from $N'$ without repetition, rearranging them into, say, $(k'_j)$ which is in the decreasing order, and letting $h'(j) = k'_j$. Therefore, $|G_\pi| = (\text{the number of combinations of } m \text{ objects taken from the set of } m+n \text{ distinct objects})$, since $|N'| = m+n$. This proves the first equation of (31). (ii) The second equation of (31) may be proved similarly if one replaces (a5) by $h'(j) = h(j) - 2(j-1), j \in M$.

Appendix VII (Proof of (34)): We prove the second equation of (34). (The first equation may be proved analogously.) Put $m = n$. Then, noting that

$$3m-\frac{1}{2} = (m+\frac{1}{2}) + (2m-\frac{1}{2}) + (-\frac{1}{2}),$$

we have (the first term of the right side of (33))

$$\log \left\{ (3m-1)^{3m-\frac{1}{2}} / m^{\frac{1}{2}} (2m-1)^{2m-\frac{1}{2}} \right\} = \log \left\{ (3-1/m)^{3m(2-1/m)} - (2m-\frac{1}{2}) \right\} = mU(m),$$

where $U(m) = \log \left\{ (3-1/m)^{3(2-1/m)} - (2-\frac{1}{2m}) \right\}$. Since $U(m) + \log (3^3/2^2) = \log \frac{27}{4}$ as $m \to \infty$, we have derived the second equation of (34).

Appendix VIII (Proof of (38)): For simplicity, we write throughout this appendix as follows: $\xi_i^*(\pi^*) = \xi_i^*, \xi_i^*(\pi^*) = \xi_i^* (i=1,2)$, where $\pi^*$ satisfies (37);

$x = (x_1, x_2); f(x) = f_1(x_1) + f_2(x_2); g(x) = g_1(x_1) + g_2(x_2);$ 

$x = (x; x_1 + x_2 = \bar{x}, x_1 \geq 0); x^* = (x; x_1 + x_2 = \bar{x}, \xi_1^* \leq x_1 \leq \xi_1^*) \subset X;$

$y^* = \max f(x), x \in X; \quad g^* = \max g(x), x \in X;$

$g^* = \max g(x), x \in X^* (g^* \leq g^*).$
Noting that
\[ |f(x)-y^*| \leq |f(x)-g(x)| + |g(x)-g^*| + |g^*-g| + |g-y^*|, \quad (a6) \]
for all \( x \in \mathbb{X} \), we show (i)-(iv) below:

(i) \[ |f(x)-g(x)| \leq 2\max\{n/4,1\}, \quad x \in \mathbb{X}. \]

For,
\[ |f(x)-g(x)| = |f_1(x_1)+f_2(x_2)-g_1(x_1)-g_2(x_2)| \]
\[ \leq |f_1(x_1)-g_1(x_1)| + |f_2(x_2)-g_2(x_2)| \leq 2\max\{n/4,1\}. \quad \text{(See (22)).} \]

(ii) \[ g^*-g(x) \leq 2m, \text{ for all } x \in \mathbb{X}. \]
To show this, let \( x \in \mathbb{X} \). In view of (21), we may write
\[ a = g(x) - g(\xi_1^*, \xi_2^*) = \sum_{j=1}^{\xi_1^*+1} h_1(j) + \frac{1}{2} \sum_{j=1}^{\xi_2^*+1} h_2(j) + h_2([x_2]+1)(x_2-[x_2]). \quad (a7) \]

(iii) \[ g^*-g = 0. \]
To prove this, let \( x = (x_1,x_2) \in \mathbb{X}, \ 0 \leq x_1 < \xi_1^* = x_1 + \gamma (\gamma > 0), \]
and \( \delta = x_2 - \gamma \). Then, since \( \xi_1^* + \xi_2^* \leq \overline{x} \) and \( x_1 + x_2 = \overline{x} \), we get \( \gamma = \xi_1^* - x_1 \leq x_2 - \xi_2^*, \ x_1 \geq \xi_1^*, \delta \geq \xi_2^*, \xi_1^* + \delta = \xi_1^* + x_2 - \gamma = x_1 + x_2 = \overline{x}, \) and \( \delta \in \mathbb{M}. \)

We show
\[ g(x) \leq g(\xi_1^*, \delta) = \beta. \quad (a8) \]
To prove (a8), we write, in view of (16) and \( [x_1] \leq \xi_1^* \)
\[ \beta = g_1(0) + g_2(0) + \sum_{j=1}^{\xi_1^*} h_1(j) + \sum_{j=1}^{\xi_2^*+1} h_2(j) = g_1(0) + g_2(0). \]
In view of the first and the third propositions of (25), we have $h_1(j) \geq \pi^* + 1$

for $j \leq \xi_1^*$ and $\xi_2^* \leq \delta < j'$. Hence, $\beta - g_1(x_1) - g_2(0) \geq \sum_{j=1}^{\delta} h_2(j) + \sum_{j=\delta+1}^{\lambda} h_2(j) + h_1((x_1)+1)((x_1)+1-x_1)$, where $\lambda = \xi_1^* - [x_1] + \delta - 1$

$= [x_1] - 1$. Note $(x_1-[x_1]) + (x_2-[x_2]) = 0$ if $x_1 \in M_0$ (then $x_2 \in M_0$), and $= 1$

otherwise. From this we get $\beta - g_1(x_1) - g_2(0) \geq \sum_{j=1}^{\xi_2^*} h_2(j) + h_1((x_1)+1)(x_1-\lfloor x_1 \rfloor) + h_1((x_1)+1)-x_1)$. Therefore, we have

$\beta \geq g_1(x_1) + g_2(x_2)$, which is (a8).

We have shown that, if $x \in X$ but $x_1 < \xi_1^*$, then $g(x) \leq g^*$. Similarly, if $x \in X$ but $x_2 < \xi_2^*$, then $g(x) \leq g^*$. By using analogous reasoning, we can show that if $x \in X$ but $\xi_1^* < x_i$ (i=1 or 2), then $g(x) \leq g^*$. This proves (iii).

(iv) $|g^* - y^*| \leq 2\max[n/4,1]$. This can easily be shown from (i).

The inequality (38) is obtained from (a6) and (i)-(iv) above.
FOOTNOTES

* I owe much to Professors K. J. Arrow and L. Hurwicz for general guidance and valuable comments. Mr. P. Petri gave me useful suggestions. The work was supported by the Office of Naval Research under Contract No. N00014-67-A-0298-0019 (Project No. NR 047-004).


2. See K. J. Arrow, "The Organization of Economic Activity: Issues Pertinent to the Choice of Market versus Nonmarket Allocation," in The Analysis and Evaluation of Public Expenditures: The PPB System, U.S. Congress, Joint Economic Committee, printed by the U.S. Government Printing Office, 1969, vol. I, pp. 47-64. Arrow states: "The distinction between transaction costs and production costs is that the former can be varied by a change in the mode of resource allocation, while the latter depend only on the technology and tastes, and would be the same in all economic systems." (Ibid., p. 60.)


6. Hurwicz, op.cit., Section I.

8. This property is shared by transformers which store or retrieve information, since they are means of intertemporal communication.


10. Ibid.

11. This is an implication of Shannon's coding theorems in noisy and noiseless channels. See Shannon, op.cit., and Section II below.


18. In fact, they presented two alternative formulations of the adjustment process: the gradient method and the price-adjustment method. In our study the latter will be considered for the organization representing the market mechanism.


21. He considered another allocation scheme, which is neither centralized nor decentralized.


25. See Shannon, *op. cit.*, Introduction and Sections I, II.

26. For example, consider today's change in a stock price. The set of possible objects may be composed of three elements: UP, DOWN, UNCHANGED.

27. In the above example, \( a_1 = \text{UP}, a_2 = \text{DOWN}, a_3 = \text{UNCHANGED}, \) and \( n = 3 \).

28. In the above example, we might code: UP = 00, DOWN = 01, and UNCHANGED = 1. The length of the first two codes is 2, while that of the last is 1.

29. The choice of the base determines the unit for measuring the amount of information. In binary coding, the base 2 is natural, since then the amount is expressed in terms of the number of bits. (See the examples explained in footnote 28.) If a coding with three symbols (e.g., +, -, 0) is used, then the base 3 will be convenient. The choice of a base, like that of a unit for measurement in general, may be done arbitrarily, since changing the base is the same as multiplying a constant.
30. The average length of the codes in footnote 28 is \((2+2+1)/3 = 1.667\), if each of the three messages occurs with equal probability. Since the greatest lower bound for the average length is \(H(3) = \log_2 3 = 1.585\), it is possible to shorten the average length by improving coding. This can be done by coding a sequence of messages at one time, rather than coding single messages separately. See Shannon, *op. cit.*, Section I, Theorem 9.

31. Equation (4) can easily be verified if \(n\) is a power of 2. The simplest case is \(n = 2\), i.e., \(A = \{a_1, a_2\}\). Then we code: \(a_1 = 0\) and \(a_2 = 1\), say. Since the length of each code is 1, so is the average length; this agrees with (4): \(\log_2 2 = 1\). If \(n = 4\), then the coding may be: \(a_1 = 00\), \(a_2 = 01\), \(a_3 = 10\), \(a_4 = 11\); the average length is 2 (= \(\log_2 4\)).

32. Shannon, *op. cit.*, Section II.

33. For the present purpose, it is not necessary that the distance between two adjacent points in \(M\) is constant; all that is needed is that \(M\) be a finite set.

34. The construction above implies that the output commodity is more finely approximated than the input commodity and the price. This asymmetry may seem to lead to some inefficiency of an organization's informational activities. Nothing definite can be said at this stage of our investigation. The construction, however, may be considered natural, because \(x\) and \(p\) are treated symmetrically so that the approximation in the dual space may be obtained immediately from the present approximation.

35. The condition \(0 < f(0) \leq 1\) could easily be replaced by a more general one, say, \(0 < f(0) \leq y_0\), where \(y_0\) is a positive integer. (We should then
replace \( \bar{y} = \bar{px} + 1 \) by \( \bar{y} = \bar{px} + y_0 \) for the obvious reason.) This formulation will be desirable, if the origin \( x = 0 \) is not the point at which physical quantity of \( x \) is zero but the point which expresses a lower bound of \( x \).

36. The complexity of equations (23) through (25) arises from the fact that the function \( h(j) \), which represents approximate marginal products, is not monotone in \( j \) (although it is either nonincreasing or increasing at most by 1, as shown in appendix V). If \( h \) were nonincreasing, then we would put \( \xi^- = \min \{ j; h(j) = n \} - 1 \) and \( \xi^+ = \max \{ j; h(j) = n \} \). Furthermore, if \( h \) were strictly decreasing, then we would put simply \( \xi^- = \xi^+ = h^{-1} \).

37. Decision making based on the centralization of such information, however, exists in the real world, if not on a large scale as considered in the planning controversy but on much smaller scales as seen frequently in modern business, government, and other organizations. This means that the transmission of information like production technology may pay sometimes, or at least that it is not always prohibitively costly.

38. The result obtained here is at least not inconsistent with the assertion in the preceding footnote. Since the number of bits needed to transmit \( g_i = (\log |G_i|) \) is approximately linear in \( m \), the transmission cost may be expected to be reasonable for \( m \) that is not very large.


40. That (37) is an equilibrium condition in terms of the approximate data may be seen from the following: Choose \( (j_1, j_2) \) in such a way that \( j_1 + j_2 = m \) and \( \bar{\xi}_i \leq j_i \leq \bar{\xi}_i \) (\( i = 1, 2 \)), where \( \bar{\xi}_i = \xi_i(\pi) \), \( \bar{\xi}_i = \xi_i(\pi) \), and \( \pi \) satisfies (37).
Then, it follows from (25) that (i) $h_i(j_i) \geq \pi + 1$ and $h_i(j_i+1) \leq \pi - 1$ if $j_i = \xi_i = \bar{\xi}_i$, and that (ii) $\pi - 1 \leq h_i(j_i) \leq \pi + 1$ if $\xi_i < j_i \leq \bar{\xi}_i$.

That (37) is an approximate equilibrium condition with respect to the originally given data follows from (38). It is noted, however, that the "true" optimum $(x_1^*, x_2^*)$ may not satisfy $\xi_i \leq x_1^* \leq \bar{\xi}_i$.

The paper proposes a method for measuring communication costs in alternative economic organizations allocating scarce resources. A simple allocation problem for two firms with single-input and single-output production functions are considered, and the economic data (commodities and production functions) are approximated with finite and discrete sets. The paper then computes and compares the costs of internal communication in a centralized system and a competitive market mechanism by using Shannon's measure of information. It is found that (i) the market mechanism is definitely more economical with respect to communication costs if the required accuracy of allocation is high, but that (ii) the centralized system may be cheaper if the accuracy is allowed to be relatively low.