GENERAL ECONOMIC EQUILIBRIUM: PURPOSE, ANALYTIC TECHNIQUES, COLLECTIVE CHOICE

Kenneth J. Arrow

Harvard University

Prepared for:

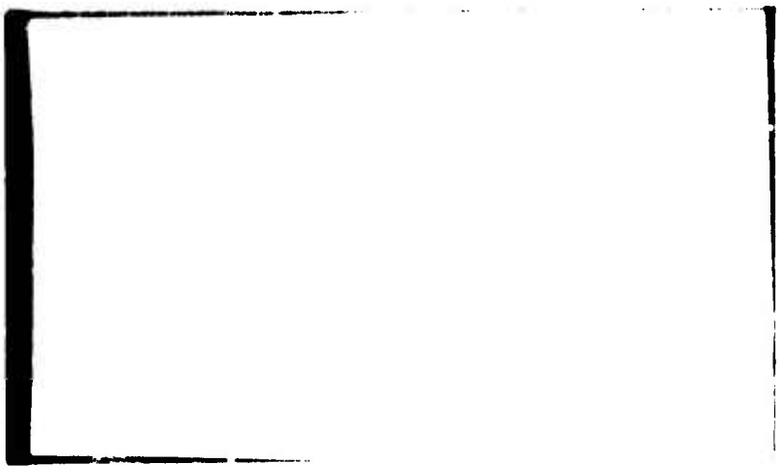
Office of Naval Research

January 1973

PROJECT ON EFFICIENCY OF DECISION MAKING

IN

ECONOMIC SYSTEMS



HARVARD UNIVERSITY

GENERAL ECONOMIC EQUILIBRIUM: PURPOSE, ANALYTIC TECHNIQUES, COLLECTIVE CHOICE

Kenneth J. Arrow

Technical Report No. 12

Prepared under Contract No. N00014-67-A-0298-0019
Project No. NR 047-004
For Office of Naval Research

Harvard University
1737 Cambridge Street, Room 401
Cambridge, Massachusetts 02138

January, 1973

DOCUMENT CONTROL DATA - R & D

Project on Efficiency of Decision Making in Economic Systems, 1737 Cambridge St., #401, Harvard University, Cambridge, Mass. 02138

Unclassified

General Economic Equilibrium: Purpose, Analytic Techniques, Collective Choice

Technical Report No. 12

Kenneth J. Arrow

January, 1973

TOTAL NO. OF PAGES: 50 53

No. OF REFS: 39

N00014-67-A-0298-0019

NR 047-004

Technical Report No. 12

Logistics & Mathematical Statistics Branch, Dept. of the Navy, Office of Naval Research, Washington, D.C. 20360

This lecture is designed to survey the modern development of the theory of the general equilibrium of a competitive economy and its role in the allocation of resources. A survey is given of the state of the theory as developed by J.R. Hicks, P.A. Samuelson, and their predecessors as of about 1945, followed by an evaluation of the needs for further development. The differing general equilibrium tradition in the German-language literature is then summarized. There follow presentations of the more modern developments in (a) the interrelations between Pareto efficiency and competitive equilibrium and the importance of the convexity assumptions; (b) the conditions which insure the existence of competitive equilibrium and the mathematical tools needed to demonstrate their sufficiency; (c) the introduction of uncertainty into the general equilibrium theory; and (d) the theory of social choice conceived of as an evaulation of resource allocations.

DD FORM 1473 (BACK)

Unclassified
Security Classification

GENERAL ECONOMIC EQUILIBRIUM: PURPOSE, ANALYTIC TECHNIQUES, COLLECTIVE CHOICE*

Kenneth J. Arrow

1. The Coordination and Efficiency of the Economic System.

From the time of Adam Smith's Wealth of Nations in 1776, one recurrent theme of economic analysis has been the remarkable degree of coherence among the vast numbers of individual and seemingly separate decisions about the buying and selling of commodities. In everyday, normal experience, there is something of a balance between the amounts of goods and services that some individuals want to supply and the amounts that other, different individuals want to sell. Would-be buyers ordinarily count correctly on being able to carry out their intentions, and would-be sellers do not ordinarily find themselves producing great amounts of goods that they cannot sell. This experience of balance is indeed so widespread that it raises no intellectual disquiet among laymen; they take it so much for granted that they are not disposed to understand the mechanism by which it occurs. The paradoxical result is that they have no idea of the system's strength and are unwilling to trust it in any considerable departure from normal conditions. This reaction is most conspicuous in wartime situations with radical shifts in demand. It is taken for granted that these can be met only by price control,

* Nobel Prize Lecture, 1972

rationing, and direct allocation of resources. Yet there is no reason to believe that the same forces that work in peacetime would not produce a working system in time of war or other considerable shifts in demand. (There are undesirable consequences of a free market system, but sheer unworkability is not one of them.)

I do not want to overstate the case. The balancing of supply and demand is far from perfect. Most conspicuously, the history of the capitalist system has been marked by recurring periods in which the supply of available labor and of productive equipment available for the production of goods has been in excess of their utilization, sometimes, as in the 1930's, by very considerable magnitudes. Further, the relative balance of overall supply and demand in the postwar period in the United States and Europe is in good measure the result of deliberate governmental policies, not an automatic tend ncy of the market to balance.

Nevertheless, when all due allowances are made, the coherence of individual economic decisions is remarkable. As incomes rise and demands shift, for example, from food to clothing and housing,the labor force and productive facilities follow suit. Similarly, and even more surprising to the layman, there is a mutual interaction between shifts in technology and the allocation of the labor force. As technology improves exogenously, through innovations, the labor made redundant does not become permanently unemployed but finds its

place in the economy. It is truly amazing that the lessons of both theory and over a century of history are still so misunderstood. On the other hand, a growing accumulation of instruments of production raises real wages and in turn induces a rise in the prices of labor-intensive commodities relative to those which use little labor. All these phenomena show that by and large and in the long view of history, the economic system adjusts with a considerable degree of smoothness and indeed of rationality to changes in the fundamental facts within which it operates.

The problematic nature of economic coordination is most obvious in a free enterprise economy but might seem of lesser moment in a socialist or planned society. But a little reflection on the production and consumption decisions of such a society, at least in the modern world of complex production, shows that in the most basic aspects the problem of coordination is not removed by the transition to socialism or to any other form of planning. In the pure model of a free enterprise world, an individual, whether consumer or producer, is the locus both of interests or tastes and of information. Each individual has his own desires, which he is expected to pursue within the constraints imposed by the economic mechanism; but in addition he is supposed to have more information about himself or at least about a particular sphere of productive and consumptive activity than other individuals. It might be that in an ideal socialist

economy, all individuals will act in accord with some agreed ideas of the common good, though I personally find this concept neither realistic nor desirable, in that it denies the fact and value of individual diversity. But not even the most ideal socialist society will obviate the diversity of information about productive methods that must obtain simply because the acquisition of information is costly. Hence, the need for coordination, for some means of seeing that plans of diverse agents have balanced totals, remains.

How this coordination takes place has been a central preoccupation of economic theory since Adam Smith and received a reasonably clear answer in the 1870's with the work of Jevons, Menger, and above all, Léon Walras: it was the fact that all agents in the economy faced the same set of prices that provided the common flow of information needed to coordinate the system. There was, so it was argued, a set of prices, one for each commodity which would equate supply and demand for all commodities; and if supply and demand were unequal anywhere, at least some prices would change, while none would change in the opposite case. Because of the last characteristics, the balancing of supply and demand under these conditions may be referred to as <u>equilibrium</u> in accordance with the usual use of that term in science and mathematics. The adjective, "general," refers to the argument that we cannot legitimately speak of equilibrium with respect to any one commodity; since supply and demand on any one market depends on

the prices of other commodities, the overall equilibrium of the economy cannot be decomposed into separate equilibria for individual commodities.

Now even in the most strictly neoclassical version of price theory, it is not precisely true that prices alone are adequate information to the individual agents for the achievement of equilibrium, a point that will be developed later. One brand of criticism has put more stress on quantities themselves as signals, including no less an authority than the great Keynes [1936]; see especially the interpretation of Keynes by Leijonhufvud [1968, especially Chapter II] More recently the same argument has been advanced by Kornai [1971] from socialist experience. Nevertheless, while the criticisms are, in my judgment, not without some validity, they have not given rise to a genuine alternative model of detailed resource allocation. The fundamental question remains, how does an overall total quantity, say demand, as in the Keynesian model, get transformed into a set of signals and incentives for individual sellers?

If one shifts perspective from description to design of economies it is not so hard to think of non-price coordinating mechanisms; we are in fact all familiar with rationing in one form or another. Here, the discussion of coordination shades off in that of efficiency. There has long been a view that the competitive price equilibrium is efficient or optimal in some sense that rationing is not. This sense

and the exact statement of the optimality theorem were clarified by Pareto [1909, Chapter VI, sections 32-38] and, in the 1930's by my teacher, Hotelling [1938] and by Bergson [1938]. An allocation of resources is Pareto efficient (or Pareto optimal) if there is no other feasible allocation which will make everyone better off (or, as more usually stated, make everyone at least as well off and at least one member better off). Then, by an argument that I shall sketch shortly, it was held that a competitive equilibrium necessarily yielded a Pareto efficient allocation of resources.

It was, of course, recognized, most explicitly perhaps by Bergson, that Pareto efficiency in no way implied distributive justice. An allocation of resources could be efficient in a Pareto sense and yet yield enormous riches to some and dire poverty to others.

2. The Hicks-Samuelson Model of General Equilibrium.

I will state more formally the model of general competitive equilibrium as it had been developed by about 1945, primarily through the detailed developments and syntheses of Hicks [1939] and Samuelson [1947]. Competitive analysis is founded on two basic principles: optimizing behavior on the part of individual agents in the presence of prices taken as given by them and the setting of the prices so that, given this individual behavior, supply equals demand on each market. The outcome of the competitive process is then to be evaluated in terms of Pareto efficiency and additional conditions on the

resulting distribution of goods.

The maximizing behavior of individuals has been well surveyed by Samuelson in his Nobel lecture [1971], and I will not go over that ground here. I just want to remind the listener of a few elementary points. The first is that the consumer's choices are subject to a budget constraint. The consumer starts with the possession of some quantities of economically valuable goods, such as labor of particular types, land, or other possessions. Let us imagine there are n commodities altogether, and let $\bar{x}_{hi}$ be the amount of commodity i owned initially by individual h (this may well be zero for most commodities). If p_i is the price of the i^{th} commodity, then his total income available for expenditure is $\sum_{i=1}^{n} p_i \bar{x}_{hi}$. Hence, he can choose for consumption any bundle of goods, $x_{h1}, \ldots, x_{hn}$, which cost no more than his income,

$$\sum_{i=1}^{n} p_i x_{hi} \leq \sum_{i=1}^{n} p_i \bar{x}_{hi}.$$

Within this budget set of possible consumption bundles, the individual is presumed to choose his most preferred bundle. The most usual interpretation of "most preferred" in this context is that there is a preference ordering over all possible bundles, according to which, for every pair of bundles, one is preferred to the other or else the two are indifferent; and these pairwise judgments have the consistency property known to logicians as "transitivity;" thus,

for example, if bundle A is preferred to bundle B and B to C, then A will be preferred to C. This "ordinalist" view of preferences was originally due to Pareto and to Irving Fisher, about 1900, and represented an evolution from the earlier "cardinalist" position, according to which a measurable satisfaction or "utility" was associated with each bundle, and the consumer chose that bundle which maximized utility within the budget set. Obviously, a cardinal utility implies an ordinal preference but not vice versa; and if the only operational meaning of utility is in the explanation of consumer choice, then clearly two utility functions which defined the same preference ordering are operationally indistinguishable.[1]

The most preferred bundle then is a function, $x_{hi}(p_1,\ldots,p_n)$ of all prices. Notice that, from this viewpoint, all prices clearly enter into the determination of the demand for any one commodity. For one thing, the rise in any one price clearly diminishes the residual income available for all other commodities. More specifically, however, the demands for some commodities are closely interrelated with others; thus, the demand for gasoline is perhaps more

---

1/ The ordinalist view in fact only began to have wide currency in the 1930's, and indeed the treatments of Hicks and Samuelson, along with a paper of Hotelling's [1935], did much to make the ordinalist view standard. Interestingly enough, both Hicks and Samuelson have studied consumer choice by alternative axiom systems even weaker than ordinalism; see Hicks and Allen [1935], Samuelson [1938].

influenced by the use of automobiles and therefore by their price than it is by its own price. The interrelation of all demands is clearly displayed here.

The characterization of consumer choice by optimization can, as we all know, be made more explicit. Let us recall Hicks's definition of the marginal rate of substitution between two commodities for any individual. For any given bundle, $(x_1^o, \ldots, x_n^o)$, consider all bundles indifferent to it, i.e., neither preferred to it nor inferior to it. If we hold all but two commodity quantities constant, say $x_k = x_k^o (k \neq i,j)$ we can consider x_i as a function of x_j on this indifference surface. Then $- dx_i/dx_j$, evaluated at the point $x_i = x_i^o$, all i, is the marginal rate of substitution of commodity j for commodity i; it is, to a first approximation, the amount of commodity i that would be required to compensate for a loss of one unit of commodity j. The optimizing consumer will equate this marginal rate of substitution to the price ratio, p_j/p_i; for if the two were unequal, it would be possible to move along the indifference surface in some direction and reduce spending.

But since the marginal rate of substitution for any pair of commodities is equal to the price ratio for all individuals, it is also true that the marginal rate of substitution for any two commodities is the same for all individuals. This suggests in turn that there is no possibility that two or any number of individuals can

gain by trading with each other after achieving a competitive equilibrium. The equality of the marginal rates of substitution means that a trade which would leave one individual on an indifference surface would do the same to the other. Hence, a competitive equilibrium satisfies the same kinds of conditions that are satisfied by a Pareto optimum.

(It will be observed that the stated conditions for a consumer optimum and for a Pareto optimum are first-order conditions in the differential calculus. Hotelling, Hicks, and Samuelson also developed the second-order conditions which distinguish maxima from minima and showed that these had important implications.)

Evaluation of the performance of an economy with regard to distributive justice was far less studied, not surprsingly, since the deepest philosophical issues are at stake. The Anglo-American tradition had incorporated in it one viewpoint, tacitly accepted though rarely given much prominence, the utilitarian views of Bentham and Sidgwick, given formal expression by Edgeworth. The criterion was the maximization of the sum of all individuals' utilities. This criterion only made sense if utility were regarded as cardinally measurable. With the rise of ordinalist doctrines, the epistemological basis for the sum-of-utilities criterion was eroded. It was to this issue that Bergson's famous paper [1938] was addressed. As already noted, a given preference ordering corresponds to many

different utility functions. For any given set of preference orderings for the members of the economy, choose for each one of the utility functions which imply that preference ordering, and then the social welfare is expressed as some function, $W(U_1,\ldots,U_n)$ of the individual utilities. The function W will change appropriately if the utility indicator for the given preference orderings is changed, so that the entire representation is consistent wit' the ordinalist interpretation. However, the function W is not uniquely prescribed, as in the Edgeworth-Bentham sum of utilities, but is itself an expression of social welfare attitudes which may differ from individual to individual.

So far, I have, for simplicity, spoken as if there were no production, an omission which must be repaired. A productive unit or firm is characterized by a relation between possible outputs and inputs. A firm may have, of course, more than one output. Then firm f may be characterized by its transformation surface, defined by an equation, $T(y_{f1},\ldots,y_{fn}) = 0$, where y_{fi} is taken to be an output if positive and input if negative; the surface is taken to define the efficient possible input-output vectors for the firm, that is, those which yield maximum output of one commodity for given inputs and given outputs of other commodities. The optimizing behavior of the firm is taken to be the maximization of profit among the points on its transformation surface. Because of the sign conventions for

inputs and outputs, the firm is seeking to maximize,

$$\sum_{i=1}^{n} p_i y_{fi}.$$

It is assumed in the treatment by Hicks and by Samuelson in the books referred to that the transformation surface is differentiable, so that the maximum-profit position is defined by suitable marginal equalities, and that the result is a function, $y_{fi}(p_1,\ldots,p_n)$ $(i = 1,\ldots,n)$.

Two remarks should be made at this point: (1) Clearly, if all prices are multiplied by the same positive constant, the budget constraint for households is really unchanged, and hence so are the consumer demands. Similarly, the profits are multiplied by a positive constant, so that the profit-maximizing choice of a firm is unchanged. Hence, the functions $x_{hi}(p_1,\ldots,p_n)$ and $y_{fi}(p_1,\ldots,p_n)$ are homogeneous of degree zero in their arguments. (2) The firms' profits have to be treated as part of the income of the households that own them. This causes a modification of the previous budget constraint for the individual, which I shall not spell out in symbols here but will refer to below.

For any commodity i, there will be some demands and some supplies at any given set of prices. Following Hicks, we will speak of the <u>excess demand</u> for commodity i as the sum over all individuals and firms of demands and supplies, the latter being taken negative. The

demand by individual h is $x_{hi}(p_1,\ldots,p_n)$, so that the total demand by all households is,

$$\sum_h x_{hi}(p_1,\ldots,p_n).$$

The supply by households is the aggregate amount they have to begin with, i.e.,

$$\sum_h \bar{x}_{hi}.$$

Finally, the aggregate supply by firms is,

$$\sum_f y_{fi}(p_1,\ldots,p_n);$$

some firms may be demanders rather than suppliers, but the sign convention assures that the above sum gives the aggregate net supply by firms, i.e., after cancelling out demands by one firm which are supplied by another. Hence, the market excess demand for commodity i is,

$$z_i(p_1,\ldots,p_n) = \sum_h x_{hi}(p_1,\ldots,p_n) - \sum_h \bar{x}_{hi} - \sum_f y_{fi}(p_1,\ldots,p_n).$$

Since each term is homogeneous of degree zero, so is the total, z_i. Further, the satisfaction of the budget constraint for each individual also restricts the excess demand functions. Since for each individual the monetary value of expenditures planned at any set of prices equals the monetary value of his initial endowments plus his share of the profits, we have in the aggregate that the money value of planned expenditures by all households equals the money value of total endowments plus total profits, or,

$$\sum_h \sum_{i=1}^{n} p_i \, x_{hi}(p_1,\ldots,p_n) \equiv \sum_h \sum_{i=1}^{n} p_i \, \bar{x}_{hi} + \sum_f \sum_{i=1}^{n} p_i \, y_{fi}(p_1,\ldots,p_n),$$

or, from the definition of excess demand,

$$\sum_i p_i \, z_i(p_1,\ldots,p_n) \equiv 0,$$

where the identity symbol reminds that this relation, called by Lange [1942] Walras' Law, holds for all values of the prices.

The general equilibrium of the economy is then the set of prices which equate all excess demands to zero,

$$z_i(p_1,\ldots,p_n) = 0 (i = 1,\ldots,n).$$

These appear to be n equations in n unknowns; but there are two offsetting complications in the counting. On the one hand, since the equations are homogeneous, no solution can be unique, since any positive multiple of all prices is also a solution. In effect, the equations really only determine the n-1 price ratios. On the other hand, the equations are not independent; if n-1 are satisfied, then the n^{th} must be by Walras' Law.

3. The Need for Further Development.

There were, however, several directions in which the structure of general equilibrium theory was either incomplete or inconsistent with doctrines which had strong currency in economic theory.

(1) There was no proof offered that the system of equations defining general equilibrium had a solution at all; that is, it was

not known that there existed a set of prices which would make excess demand zero on every market. This was the most serious unresolved problem.

(2) The assumptions on production were not the same as those used in the analysis of production itself. In the latter, a common, though not universal, assumption was that of constant returns to scale; if any production process can be carried out, with given inputs and outputs, then the process can be carried out at any scale. That is, if the inputs are all multiplied by the same positive number, then it is possible to produce the same multiple of all the outputs. But in this case, there cannot be a unique profit-maximizing position for any set of prices. For suppose there were a position which yielded positive profits. Then doubling all inputs and outputs is feasible and yields twice as great profits. Hence, there would be no profit-maximizing position, since any one could be improved upon. On the other hand, zero profits can always be obtained by having no inputs and no outputs. It can be concluded that, if prices are such that there is some profit-maximizing set of inputs and outputs not all zero, the corresponding profits must be zero, and the same profits can be achieved by multiplying all inputs and outputs by any positive number.

Therefore, under constant returns to scale, there is never a single-valued function, $y_{fi}(p_1,\ldots,p_n)$ defining inputs and outputs as a function of prices; rather, for any given set of prices, either

there is no profit-maximizing input-output vector or else there is a whole ray of them. But then the notion of equating supply and demand must be redefined.

Of somewhat lesser importance in this regard is the fact that the transformation surface need not be differentiable in very plausible circumstances. A frequently-held view was that production of a given output required prescribed amounts of each input; in some circumstances, at least, it is impossible to reduce the need for one input by increasing the amount of another. This is the fixed-coefficient technology. In this case, it can easily be seen that though the transformation surface is well defined, it is not differentiable but has kinks in it.

(3) The relation between Pareto-efficient allocations and competitive equilibria was less clearly formulated than might be desired. What had really been shown was that the necessary first-order conditions for Pareto efficiency were the same as the first-order conditions for maximization by firms and individuals when the entire economy is in a competitive equilibrium.

(4) Actually, the condition for individual optimization (equating of marginal rates of substitution to price ratios) required some modification to take care of corner maxima. It is obvious to everyday observation that for each individual there are some (indeed, many) commodities of which he consumes nothing. Similarly, for every firm,

there are some commodities which are neither inputs to nor outputs of it. But then the argument that the marginal rate of substitution must equal the price ratio for each individual breaks down. For consider an individual for whom the marginal rate of substitution of commodity j for commodity i is less than the price ratio, p_j/p_i, but the individual consumes nothing of commodity i. A small increase in the consumption of i with a compensating decrease in j to stay on the same indifference surface would involve an increase in costs. The only way to achieve a decrease in cost without moving to a less preferred position would be to decrease the consumption of i; but this is impossible, since consumption cannot fall below zero. It is true, however, that the marginal rate of substitution of j for i cannot exceed the price ratio.

Similarly, if one individual consumes nothing of commodity i, it is possible to have Pareto efficiency with his marginal rate of substitution of j for i less than that for some other individual. Since marginal rates of substitution do not have to be equated across individuals either for competitive equilibria or for Pareto-efficient allocations, the relation between the two concepts was seen to need further study.

(5) Still another question is whether supply and demand are necessarily equal. Clearly, demand cannot exceed supply, for there would have to be unfilfilled demands. But as we look around us,

we see that there are goods, i.e., flows which we prefer to have, which nevertheless are so abundant that we have no desire for more. Air and sunlight come immediately to mind. Characteristically, such highly abundant goods are free; no price is charged for their use.

This elementary observation has been made a number of times by economists. A distinction was drawn between scarce goods and free goods, the former alone being the proper subject matter of economics. But it is easy to see from a mathematical viewpoint that the classification of goods in this way is not a given but depends on those parameters of the system which govern tastes, technology, and initial supplies. Suppose, for example, that we have two commodities, A and B, which serve as factors of production only. Suppose further it so happens that the two factors are always used together and always in the same proportion, say, one unit of A with two units of B. Finally, suppose that A and B are not themselves produced goods but are natural resources available in equal quantities. Then clearly commodity B is the bottleneck; commodity A is a free good in the usual economic sense, since a small change in the quantity available would have no effect on production. But this classification of the two goods into free and scarce is relative to the technology and to the initial supplies of the two goods. If a technological

innovation reduced the need for B so that one unit of A required the cooperation of less than one unit of B, A would become the free good, and B, the scarce one; and the same would happen if the initial supply of B were reduced, perhaps by some catastrophe, to less than half of that of A.

The conditions for equilibrium then have to be modified. We require now that excess demand be non-positive and that, for any commodity for which it is negative, the price be zero. In symbols,

$$z_i(p_1,\ldots,p_n) \leq 0 (i = 1,\ldots,n),$$

$$\text{if } z_i(p_1,\ldots,p_n) < 0, \text{ then } p_i = 0.$$

The commodities for which the inequality holds are the free goods. Equilibria in which there are free goods are referred to as <u>corner equilibria</u>.

The problem just raised illustrates a general tendency in the evolution of general equilibrium theory for a shift from a local to a global analysis. If we consider small shifts in the parameters which determine tastes, technology, and initial supplies, the classification of goods into free and scarce remains unchanged. Hence, from a local viewpoint, the list of scarce goods could legitimately be taken as given. We need not debate here the relative virtues of local and global analysis: clearly a global analysis is always preferable if it is possible, but a local analysis

will normally produce more specific implications. But it turns out that the first of the problems raised, that of the existence of equilibrium prices, cannot be handled at all except from a global viewpoint; and the realization of the possibility of corner equilibria turned out to be an indispensable step in the development of an existence proof.

To avoid a misinterpretation of this list of the needs for further development, two points should be stressed: (1) the general aims and structure of general equilibrium theory have remained those already set forth by Hicks, and the subsequent development would have been impossible and indeed meaningless except on his foundations; (2) I have summarized here only the most general and foundational aspects of the work of Hicks and Samuelson, since those are most relevant for my present purpose, but the primary interest of both was rather in the laws of working of the general equilibrium system, results not summarized above, than in the questions of existence and the like.

4. The German-Language Literature.

We must turn from the Anglo-American work to a variant strand of neoclassical thought, published primarily in German, and written to a considerable extent by mathematicians rather than economists. The whole literature might be described as an extended commentary

on a formulation of general equilibrium theory by Cassel [1918], a statement rather different in nature from that of Hicks. In particular, maximizing behavior hardly appeared in Cassel's model. With regard to individual consumers, Cassel also assumed that the demand of individual households was a function of prices; he did not, however, seek to derive this demand from a preference or utility maximization. With regard to production, he assumed a fixed-coefficient technology, so that there was in effect no scope for profit maximization by firms; the demands for inputs were completely defined by the outputs, independent of prices. More explicitly, Cassel differentiated commodities into produced goods and primary factors, the two classes being assumed distinct. Individuals owned initially only primary factors, and they demand only produced goods. Produced goods were made by inputs of primary factors; let a_{ij} be the amount of factor j used in the production of one unit of good i. Let P be the set of produced goods, F, that of factors.

At any set of prices, the total demand for produced good i is,

$$\sum_h x_{hi}(p_1,\ldots,p_n) = x_i;$$

the demand for factor j by the industry producing good i is then $a_{ij}\, x_i$, and the total demand for factor j is obtained by summing this demand over all producing industries. On the other hand, the initial supply of factor j is $\sum_h \bar{x}_{hj}$, so that the condition

for equality of supply and demand for factor j is,

$$\sum_h \bar{x}_{hj} = \sum_{i \in P} a_{ij} x_i .$$

As j varies over F, we have a system of linear equations in the x_i's. Now von Stackelberg [1933] observed that this system might easily have no solution, for example, if there are more factors than produced goods.

Cassel completed the system by using the condition that, under constant returns to scale, there must be zero profits. Then, for each produced good, the price must equal the cost of the factors used in making one unit, or

$$p_i = \sum_{j \in F} a_{ij} p_j \quad (i \in P)$$

About contemporaneously with von Stackelberg, Neisser [1932] showed that it could easily be true that the complete Cassel system could be satisfied only if some factor prices were negative.

It was at this point that the Viennese banker and amateur economist, K. Schlesinger [1933-4], decisively affected the subsequent discussion. He observed that the criticisms raised by von Stackelberg and by Neisser could be met by recognizing the possibility of corner equilibria, particularly with regard to primary factors. Some may simply be superfluous and have to be regarded as free goods. Thus, the equality of supply and demand for factors has to be replaced by

the following conditions:

$$\sum_h \bar{x}_{hj} \leq \sum_{i \in P} a_{ij} x_i, \; p_j = 0 \text{ if the strict inequality holds } (j \in P).$$

With this amendment, Schlesinger conjectured, it could be shown that there existed an equilibrium in which all prices are non-negative. He interested the mathematician Abraham Wald in this problem, and the latter showed in a brilliant series of papers [1933-4, 1934-5], summarized in [1936], that equilibrium indeed existed, though rather strong assumptions had to be made and the analysis was confined to variations of the Cassel model. Wald's reasoning was formidably complex, and his work published in a German-language mathematics journal; it was only some ten years later that American mathematical economists began to be aware of it.

Within the same period, the mathematician John von Neumann published a paper [1937] which had in the longer run a deeper impact, though its subject matter was less relevant. This was a development of Cassel's model of steady growth of the economy. The aim was to show the existence of a growth path with maximum proportional expansion in all commodities. From an economic point of view the model was somewhat strange in that there was no consumption at all; the outputs of one period were inputs into activities which generated the outputs of the next period. There were three noteworthy points which had great influence on the development of general equilibrium theory:

(1) The structure of production was characterized in a novel way. It was assumed that there were a fixed set of activities, each being characterized by a vector of possible inputs and outputs and each being technologically capable of operation at any scale. This generalized the fixed-coefficient model, in which there was one activity for each output. The feasible combinations of activities were those for which the total usage of each input did not exceed the amount available from previous production. (2) The maximum growth path could be characterized as a sort of competitive equilibrium, in the sense that it was mathematically possible and meaningful to introduce a new set of variables, which could be regarded as prices. Any activity that was run at all yielded zero profits; other activities yielded zero or negative profits. Hence, the choice of activity levels could be described as profit-maximizing, where the maxima may involve some corners. Further, the price of any commodity for which the demand as input fell short of the amount available had to be zero; hence, the competitive equilibrium could require corners. (3) The method of proof of the existence of prices and relative quantities which yielded a maximum growth rate required the use of a tool from combinatorial topology, a generalization of Brouwer's fixed point theorem. From a mathematical viewpoint, the existence of equilibrium in the von Neumann growth model was a generalization of the minimax theorem for zero-sum two-person games, which von Neumann

had studied a few years earlier. The interest in game theory following the publication of the great book of von Neumann and O Morgenstern [1944] was a strong collateral force in introducing new mathematical techniques, particularly in the theory of convex sets, into general equilibrium theory.

A simplification of von Neumann's fixed point theorem was developed a few years later by S. Kakutani [1941] and has become the standard tool for proving existence theorems. Let us review briefly the fixed point theorems of Brouwer and Kakutani. Recall that a set of points is said to be compact if it is closed and bounded and to be convex if every line segment joining two points of the set lies entirely within the set. Let C be a compact convex set. Let f(x) be a vector function which assigns to every point of C a point of C. Then Brouwer's theorem asserts that if the mapping f(x) is continuous, then there is at least one point, x*, which is mapped into itself, i.e., for which $f(x^*) = x^*$.

In the indicial notation which we have used hitherto, we have n real-valued functions $f_i(x_1,\ldots,x_n)$ of n variables. If these functions are continuous and if the point $(f_1,\ldots,f_n)$ lies in some compact convex set C whenever $(x_1,\ldots,x_n)$ lies in that set, then the system of equations, $f_i(x_1,\ldots,x_n) = x_i$, has at least one solution in C.

The relevance of such a mathematical tool to the problem of existence is obvious. However, we have already noted above that once

we permit constant returns to scale, we have to allow for the possibility that the profit-maximizing choice of production process may be a whole set, all equally profitable, for some given set of prices. Hence, instead of dealing with functions, we need to concern ourselves with the more general notion of a point-to-set mapping, or correspondence, as it is sometimes termed. Kakutani's theorem deals with this more general situation. To every point $x = (x_1,\ldots,x_n)$ in a compact convex set C, we associate a subset of C, say $\Phi(x)$. We say that x^* is a fixed point of this correspondence if the point x^* belongs to the set associated with x^*, i.e., to $\Phi(x^*)$.

Kakutani's theorem tells us that such a fixed point will exist if two conditions are fulfilled: for each x, $\Phi(x)$ is a convex set; and as x varies, $\Phi(x)$ is continuous in a certain sense, more technically, that it has the property known as upper semi-continuity.

5. Pareto Efficiency, Competitive Equilibrium, and Convexity.

My own interest first centered on the relations between Pareto efficiency and competitive equilibrium. In particular, there was considerable discussion among economists in the late 1940's about the inefficiencies resulting from rent control and different proposals for arriving at the efficiency benefits of a free market by one or another transition route. Part of the informal efficiency arguments hinged on the idea that under rent control people were buying the

wrong kind of housing, say, excessively large apartments. It struck me that an individual bought only one kind of housing, not several. The individual optima were at corners, and therefore one could not equate marginal rates of substitution by going over to a free market. Yet diagrammatic analysis of simple cases suggested to me that the traditional identification of competitive equilibrium and Pareto efficiency was correct but could not be proved by the local techniques of the differential calculus.

I soon realized that the theory of convex sets, and, in particular, the separation theorem, was the appropriate tool. Start with a Pareto-efficient allocation, and consider all logically possible allocations which would be preferred to it by everyone. Of course, no such allocation can be feasible; otherwise the allocation we started with would not be Pareto efficient. Each such allocation is a statement of demand or supply of each commodity by each individual or firm. Hence, by adding up over individuals and firms, with appropriate attention to signs, we can define the excess demand for each commodity. Let Z be the set of all excess demand vectors $(z_1,\ldots,z_n)$ generated this way. Since they are all infeasible, it must be true for each one that there is positive excess demand for at least one commodity. In the language of set theory, the set Z is disjoint from the non-positive orthant, i.e., the set of vectors $(z_1,\ldots,z_n)$ such that $z_i \leq 0$ for all i.

The separation theorem for convex sets asserts that if two convex sets are disjoint, there is a hyperplane which separates them, so that one set is on one side and the other set on the other. In symbols, if C_1 and C_2 are disjoint convex sets in n-dimensional space, there exists numbers $p_i (i = 1,\ldots,n)$, not all zero, c, such that $\sum_{i=1}^{n} p_i x_i \leq c$ for all $x = (x_1,\ldots,x_n)$ in C_1, $\sum_{i=1}^{n} p_i x_i \geq c$ for all x in C_2. Let us apply this theorem to the present case. The non-positive orthant is obviously a convex set; let us assume for the moment that Z is convex. Then we can find numbers $p_i (i = 1,\ldots,n)$, not all zero, c such that,

$$\sum_{i=1}^{n} p_i z_i \leq c \text{ for } z = (z_1,\ldots,z_n) \text{ in } Z,$$

$$\sum_{i=1}^{n} p_i z_i \geq c \text{ if } z_i \leq 0 \text{ for all } i.$$

From the second condition, it can easily be seen that we cannot have $p_i < 0$ for any i. Hence, p_i is non-negative for all i and (since there is at least one non-zero p_i) positive for at least one i. This is customarily expressed by saying that the vector $p = (p_1,\ldots,p_n)$ is <u>semi-positive</u>.

It follows that,

$$\sum_{i=1}^{n} p_i z_i \leq 0 \text{ if } z_i \leq 0 \text{ for all } i,$$

and therefore we can assume without loss that $c \leq 0$. On the other hand, if we set $z_i = 0$ for all i, we see that $c \geq 0$. Hence, we can set $c = 0$.

The conditions for a Pareto-efficient allocation then become,

$$\sum_{i=1}^{n} p_i z_i \geq 0 \text{ for } z \text{ in } Z, \text{ p semi-positive.}$$

Let $z^o = (z_1^o, \ldots, z_n^o)$ be the vector of excess demands defined by the Pareto efficient allocation under consideration. It is feasible, so that $z_i^o \leq 0$, all i, and therefore,

$$\sum_{i=1}^{n} p_i z_i^o \leq 0.$$

Now assume, as is usually reasonable, that there are points in Z as close as one wishes to z^o. Then clearly we must have,

$$\sum_{i=1}^{n} p_i z_i^o = 0,$$

and hence, since $z_i^o \leq 0$, all i, $p_i \geq 0$, all i, that,

$$\text{if } z_i^o < 0, \; p_i = 0.$$

We begin to see that a Pareto efficient allocation is an equilibrium of supply and demand in the generalized sense which includes corners.

We also see that,

$$\sum_{i=1}^{n} p_i (z_i - z_i^o) \geq 0 \text{ for } z \text{ in } Z.$$

Let us go back to the definition of excess demand, as a sum of individual and firm demands and supplies.

$$z_i = \sum_h x_{hi} - \sum_h \bar{x}_{hi} - \sum_f y_{fi},$$

where $y_f(y_{f1},\ldots,y_{fn})$ is a technologically possible vector of inputs and outputs for firm f and $x_h = (x_{h1},\ldots,x_{hn})$ is a possible vector of consumptions for individual h. In particular, the excess demands defined by the Pareto efficient allocation can be written in this form,

$$z_i^o = \sum_h x_{hi}^o - \sum_h \bar{x}_{hi} - \sum_f y_{fi}^o,$$

and then, if z belongs to Z, we must have, for each h, that the consumption vector of individual h, $(x_{h1},\ldots,x_{hn})$ is preferred to that under the Pareto efficient allocation $(x_{h1}^o,\ldots,x_{hn}^o)$. Then,

$$\sum_h \left(\sum_{i=1}^n p_i x_{hi} - \sum_{i=1}^n p_i x_{hi}^o\right) - \sum_f \left(\sum_{i=1}^n p_i y_{fi} - \sum_{i=1}^n p_i y_{fi}^o\right) \geq 0$$

if, for each h, x_h is preferred by individual h to x_h^o.

Now the elementary point about this inequality is that the variable vectors x_h, y_f are independent of each other. It is not hard to see that this inequality can hold only if it holds for each individual and each firm separately. For a firm f, this means that,

$$\sum_{i=1}^n p_i y_{fi}^o \geq \sum_{i=1}^n p_i y_{fi} \text{ for all possible } y_f,$$

that is, if we interpret the p_i's as prices, each firm is maximizing its profits. The corresponding interpretation for individuals is

somewhat less simple; it is that the consumption vector prescribed by the given Pareto efficient allocation is the cheapest way of deriving that much satisfaction.

Taken altogether, it has been shown that if Z is a convex set, the Pareto efficient allocation can be achieved as a competitive equilibrium of the market, in the sense that prices and a suitable initial allocation of resources can be found such that each individual is achieving his satisfaction level at minimum cost, each firm is maximizing profits, and the markets are all in equilibrium in the generalized sense which permits corner equilibria.

The need to assume that Z is convex puts in sharper focus the convexity assumptions which had always implicitly underlain neoclassical theory. The convexity of Z could be derived from the following two assumptions: (1) for each individual, the set of consumption vectors preferred to a given vector is convex; (2) for each firm, the set of technologically possible vectors is convex.

The result states that, under suitable convexity conditions, a necessary condition for an allocation to be Pareto efficient is that it be realizable in the market as a competitive equilibrium. A by-product of the investigation was the proof of the converse theorem: a competitive equilibrium is always Pareto efficient, and this theorem is true without any convexity assumption.

These results were embodied in Arrow [1951a]. But the idea that

the theory of convex sets was the appropriate tool was clearly in the air. While I was working at Stanford, Gerard Debreu [1951] obtained very much the same results at the Cowles Commission for Research in Economics at Chicago.

6. The Existence of Competitive Equilibrium.

Again working independently and in ignorance of each other's activities, Debreu and I both started applying Kakutani's fixed point theorem to the problem of existence. In this case, we exchanged manuscripts in sufficient time to realize our common efforts and also to realize the need for relaxing an excessively severe assumption we had both made (Arrow and Debreu [1954]).

An essential precondition for our studies was the basic work of Tjalling Koopmans [1951] on the analysis of production in terms of activity analysis. In this he extended von Neumann's work into a systematic account of the production structure of the economy. He saw it as a set of activities, each of which could be operated at any level but with the overall levels constrained by initial resource limitations. The crucial novelty was the explicit statement of the assumptions which insured that the feasible set of outputs would be bounded for any finite set of initial resources. It turned that this limitation is a "global" property. That is, conditions on the nature of individual activities (for example, that every activity had to have at least one input) were not sufficient to insure the

boundedness of the economy as a whole. It was necessary to require that no combination of activities as a whole permitted production without inputs.

The first question is the definition of equilibrium when the behavior of firms is described by a correspondence rather than a function. For simplicity, I will continue to assume that the decisions of individual consumer h can be represented by single-valued functions of prices, $x_h(p)$. A set of prices defines a competitive equilibrium if supply and demand balance on each market, including the possibility of corners, with some choice of the profit-maximizing input-output vector for each firm. Formally, we will say that a price vector p^*, an input-output vector y_f^* for each firm, and a consumption vector, $x_h^* = x_h(p^*)$, for each individual together constitute a competitive equilibrium if the following conditions hold:

(a) p^* is semi-positive;

(b) for each commodity i, $\sum_h \bar{x}_{hi} + \sum_f y_{fi}^* \leq \sum_h x_{hi}^*$;

(c) for any commodity for which the strict inequality holds in (b), we must have $p_i^* = 0$;

(d) y_f^* is one of the input-output vectors which maximizes profits, $\sum_{i=1}^{n} p_i^* y_{fi}$, among all the input-output vectors technologically possible for firm f.

It is, of course, understood that the demand function for individuals, $x_h(p)$ is defined, as before, as the most preferred consumption

pattern consistent with the budget constraint,

$$\sum_{i=1}^{n} p_i x_{hi} \leq \sum_{i=1}^{n} p_i \bar{x}_{hi};$$

for the present purposes, I will ignore the possibility that individuals' incomes also include profits; this modification can be handled at the cost of some analytic complexity but no true difficulty.

It will be assumed (1) that the set of possible input-output vectors for any firm is convex, and (2) the individual demand functions are continuous; this assumption will be discussed again below.

Since the total production possibilities of the economy are bounded, it can be shown there is no loss of generality in assuming the set of possible input-output vectors for each firm is bounded (actually, we assume the set to be compact). Then for any set of prices there is at least one profit-maximizing input-output vector, but in general there may be a whole set of them, say $Y_f(p)$. However, this set is certainly convex and further, as p varies, the correspondence so defined is upper semi-continuous.

Define an excess demand correspondence as follows: For each f, consider any possible selection of a vector y_f from the profit-maximizing correspondence $Y_f(p)$. For each such selection for each firm, form the excess demand for each commodity for the entire economy,

$$z_i = \sum_h x_{hi}(p) - \sum_f y_{fi} - \sum_h \bar{x}_{hi}.$$

Let $Z(p)$ be the set of all vectors $(z_1,\ldots,z_n)$ which can be formed by all possible selections of the vectors y_f from the profit-maximizing correspondence $Y_f(p)$, the selections for different firms being made independently of each other. It is not hard to show that $Z(p)$ is convex for each p and is an upper semi-continuous correspondence for p as a variable. It is also true and important that Walras' Law holds; that is,

$$\text{if } z \text{ belongs to } Z(p), \text{ then } \sum_{i=1}^{n} p_i z_i = 0.$$

The correspondence $Z(p)$ assigns to each price vector a set of excess demands; an equilibrium price vector p^* would be one such that $Z(p^*)$ has at least one element for which $z_i \leq 0$, all i. We now introduce a mapping from excess demands; very roughly, we want low excess demands to have low or more precisely zero prices. Since the whole system is homogeneous of degree zero in the prices, the general level of the prices can be set arbitrarily with no loss of generality. It will be assumed then that

$$\sum_{i=1}^{n} p_i = 1.$$

Since prices are semi-positive, it is also assumed that $p_i \geq 0$, all i; the set of price vectors satisfying these conditions will be denoted by P, the price simplex. Then we define the following correspondence, assigning to each vector of excess demands, a subset of the price simplex: for any $z = (z_1,\ldots,z_n)$, let $\bar{z}$ be the largest

of the components z_i; then define P(z) to be the set of price vectors in the unit simplex for which $p_i = 0$ for all commodities i for which $z_i < \bar{z}$. In words, total prices must add up to one, but this total is to be distributed only over those commodities with maximum excess demand. This rule is somewhat artificial, but it suffices for the proof.

Consider the set of all pairs (z,p) of vectors, one an excess demand vector and one a price vector. To any such pair we assign a set of pairs, Z(p) x P(z)(for any pair of sets, S, T, the notation S x T means the set of ordered pairs of vectors obtained by taking any vector from S followed by any vector from T). With some further argument, it can be shown that Kakutani's theorem applies. The mapping of pairs has a fixed point, (z^*, p^*) belonging to $Z(p^*) \times P(z^*)$. By definition,

$$z^* \in Z(p^*), \quad p^* \in P(z^*).$$

Let $\bar{z}^*$ be the largest component of $z^* = (z_1^*,\ldots,z_n^*)$. Then $p_i^* = 0$ for $z^* < \bar{z}^*$. Therefore,

$$p_i^* \, z_i^* = p_i^* \, \bar{z}^*.$$

By Walras' Law,

$$0 = \sum_{i=1}^{n} p_i^* \, z_i^* = \sum_{i=1}^{n} p_i^* \, \bar{z}^* = \bar{z}^* \sum_{i=1}^{n} p_i^* = \bar{z}^*;$$

since the largest excess demand is zero, all excess demands are non-positive, and therefore p^* is indeed an equilibrium price vector.

Many variations of this argument are possible and are illuminating in different ways. Independently of my work with Debreu, Lionel McKenzie [1954] proved the existence of equilibrium; he simply assumed the existence of supply and demand functions rather than analyzing them in terms of the underlying production and consumption structures. For systematic presentations of the existence theorems for competitive equilibrium, see Debreu [1959] and Arrow and Hahn [1971, Chapters 2-5].

There is one loose end that should now be picked up. It has been assumed that the demand functions of the individual are continuous. But one of the surprsing discoveries that Debreu and I made in the course of our study was that even under all the usual strong assumptions about the behavior of individuals, this cannot be true everywhere in the price simplex except under very artificial conditions. The trouble is that the individual's income also depends upon prices, and if the prices of those commodities which the individual owns originally fall to zero, his income falls to zero. When some prices and income are zero, however, the demand for the now-free goods may jump discontinuously. To illustrate, suppose an individual owned initially only one good, say, labor. So long as the price of that good was positive, he might retain some for his own use, but in any case could never consume more than he had initially. But when the price fell to zero, he could demand the same labor from others and

in any amount he chooses.

The existence of competitive equilibrium then does depend on assumptions which insure that for each individual there is at least one commodity he owns initially which is bound to have positive value. I will not state these assumptions here; the original set in Arrow-Debreu has been refined through the work of Gale [1957], McKenzie [1959, 1961], and Arrow and Hahn [1971, Chapter 5, section 4].

7. General Equilibrium and Uncertainty.

Once the broad approach to the analysis of existence was set, it could be applied in many different directions. One was the analysis of models which represented in one way or another imperfections in the competitive system. The requirement of proving an existence theorem in each case leads to the need for a rigorous spelling out of assumptions, a requirement which seems to be proving very fruitful. Much of this work is now going on, in such areas as the analysis of futures markets, expectations, and monetary theory, but time does not permit comment on what is in any case a rapidly changing field.

Another approach is to retain the competitive assumptions but interpret them in new contexts. One example of this is the extension of general equilibrium theory to uncertain outcomes (Arrow [1953]; Debreu [1959, Chapter 7). Suppose there is some uncertainty in production due, for example, to the weather. One type of weather

will benefit one kind of producer and injure another, while another type will do the opposite. If we assume that individuals are averse to risk, there is room for a mutually profitable trade in insurance. Even apart from risk aversion, individuals and firms in planning for an uncertain future may want to make sure that their demands and outputs are mutually compatible. Thus, if there is uncertainty about the supply of grain, a miller may prefer to make future contracts for labor contingent on that uncertainty.

We take from the theory of probability the concept of a <u>state of the world</u>, which is a description of the world so precise that it completely defines all initial holdings of goods and all technological possibilities. Uncertainty is not knowing which state will in fact hold. The initial holdings of commodity i by individual h if state s should hold can be designated by $\bar{x}_{his}$. Similarly, the set of possible input-output vectors for a firm may depend upon the state s; let $y_{fs} = (y_{fls},\ldots,y_{fns})$ be a possible input-output vector for firm f if s is the state.

The feasibility of any allocation will then depend upon the state s, and therefore commitments to consumption and production must vary similarly. Hence the decision by any individual must be a separate vector $x_{hs} = (x_{hls},\ldots,x_{hns})$ for each state s. But clearly it is optimal for all concerned to make all these decisions simultaneously, in advance of knowing which state of the world will in fact prevail;

it is this advance decision which permits the possible gains from insurance, from the reduction in risk-bearing. Hence, we should really think of the vector x_h, which, for fixed h, contains components x_{his} where i and s range over commodities and states of the world, respectively.

What we are led to is considering the same physical commodity in different states of the world as economically different commodities. The procedure is exactly analogous to Hicks's analysis of present and future goods [1939]; the same physical commodity at different points of time define different commodities.

The whole previous analysis can then be applied, with a suitable reinterpretation. Commodities in the ordinary sense are replaced by contingent commodities, promises to buy or sell a given commodity if and only if a certain state of the world occurs. The market will then determine contingent prices. Clearing of the markets means clearing of the contingent markets; the commitments made are sufficiently flexible so that they can always be satisfied.

It should be noted that preference orderings over vectors of contingent commodities contain elements of judgment about the likelihoods of different states of the world as well as elements of taste in the ordinary sense. Other things being equal, one will invest less heavily in a demand contingent upon a state deemed unlikely.

One can work out the implications of this model. Clearly, the

contingent commodities called for do not exist to the extent required, but the variety of securities available on modern markets serves as a partial substitute. In my own thinking, the model of general equilibrium under uncertainty is as much a normative ideal as an empirical description. It is the way the actual world differs from the criteria of the model which suggests social policy to improve the efficiency with which risk-bearing is allocated.

In fact, it is not a mere empirical accident that not all the contingent markets needed for efficiency exist, but a necessary fact with deep implications for the workings and structure of economic institutions. Roughly speaking, information about particular events, even after they have occurred, is not spread evenly thoughout the population. Two people cannot enter into a contract contingent on the occurrence of a certain event or state if only one of them in fact will know that the event has occurred. A particular example of this is sometimes known as "moral hazard" in the insurance and economic literature. The very existence of insurance will change individual behavior in the direction of less care in avoiding risks. The insurance policy that would be called for by an optimal allocation of risk bearing would only cover unavoidable risks and would distinguish their effects from those due to behavior of the individual. But in fact all the insurer can observe is a result, for example, a fire or

the success or failure of a business, and cannot decompose it into exogenous and endogenous components. Contingent contracts, to speak generally, can be written only on mutually observed events, not on aspects of the state of the world which may be known to one but not both of the parties.

Although I cannot argue the point here, I would hold that the allocational difficulties arising from the inequality in information are of importance in such diverse fields as medical care and racial discrimination (see Arrow [1963a, 1972]). The difficulty of achieving optimal allocation of risk-bearing because of differences in information was first stated in a general form by Radner [1968].

8. The Theory of Social Choice.

General competitive equilibrium above all teaches the extent to which a social allocation of resources can be achieved by independent private decisions coordinated through the market. We are assured indeed that not only can an allocation be achieved, but the result will be Pareto efficient. But, as has been stressed,there is nothing in the process which guarantees that the distribution be just. Indeed, the theory teaches us that the final allocation will depend on the distribution of initial supplies and of ownership of firms. If we want to rely on the virtues of the market but also to achieve a more just distribution, the theory suggests the strategy of changing the

initial distribution rather than interfering with the allocation process at some later stage.

Thus even under the assumptions most favorable to decentralization of decision-making, there is an irreducible need for a social or collective choice on distribution. In point of fact, there are a great many other situations in which the replacement of market by collective decision-making is necessary or at least desirable. In their different ways, both political scientists and economists have discussed the necessary role of the state. Among economists, these discussions have revolved around the concepts of externalities, increasing returns, and market failure; the clarification and application of these ideas have been among the major achievements of modern economic thought, but I have time now merely to recall them to you as helping to create the need for normative and descriptive analysis of collective decision-making.

In the context of social choice, each individual may be assumed to have a preference ordering over all possible social states. This ordering expresses not only his desire for his own consumption but also social attitudes, his views on justice in distribution or on benefits to others from collective decisions. The ordinalist viewpoint forbids us from ascribing a definite quantitative expression to this preference, at least a quantitative expression which would have any interpersonal validity.

Classical utilitarianism specifies that choices among alternative social states be judged in terms of their consequences for the members of the society; in the present terminology, this means in terms of the individual preference scales for social choices. This is obviously not a sufficient basis for choice in view of the diversity of individual preferences. It is implicit in classical utilitarianism and explicit in Bergson's work that there is a second level at the individual judgments are aggregated into what might be termed a welfare judgment.

Thus the formation of welfare judgments is logically equivalent to what I will call a constitution. Specifically, a constitution is a rule which associates to each possible set of individual preference orderings a social choice rule. A social choice rule, in turn, is a rule for selecting a socially preferred action out of any set of alternatives which may be feasible.

So far, I would hold that the description of a constitution is a tautology, at least if we start from the view that social choice has to be based on the individual preference orderings. The real question is what conditions are to be imposed on the constitution.

One condition, which is already contained in Bergson's work, is that for any given set of individual preferences, the social choice rule defined by them shall satisfy the technical conditions of an ordering, that is, that all possible alternative social states should

be capable of being ranked and then the social choice from any particular set of alternatives should be the most preferred alternative, according to the ordering, in the available set. This is sometimes called the condition of Collective Rationality.

A second condition, again in agreement with Bergson, is the Pareto principle; the social choice process shall never yield an outcome if there is another feasible alternative which everyone prefers according to his preference ordering.

A third hardly controversial condition is that of Non-Dictatorship; the constitution shall not be such that there is an individual whose preferences automatically become those of society regardless of anyone else's preferences.

The fourth condition which I have suggested, that of the Independence of Irrelevant Alternatives, is more disputable, though I would argue that it has strong pragmatic justification: the social choice made from any set of alternatives will depend on only the orderings of individuals among alternatives in that set. To see what is at stake, suppose that a society has to make a choice among some alternatives and does so. After the decision is made, an alternative which has not previously been thought of is mentioned as a logical possibility, although it is still not feasible. The individuals can expand their preference orderings to place this new alternative in its place on their ranking; but should this preference information

about an alternative which could not be chosen in any case affect the previous decision?

Any form of voting certainly satisfies the condition of Independence of Irrelevant Alternatives; the preferences of voters as between candidates and non-candidates or as between non-candidates are, of course, never asked for or taken into account.

It turns out (Arrow [1951b, 1963b]) that these four reasonable-sounding requirements are contradictory. That is, if we devise any constitution, then it is always possible to find a set of individual orderings which will cause the constitution to violate one of these conditions. In one special form, this paradox is old. The method of majority voting is an appealing method of social choice. Like any other voting method, it satisfies Independence of Irrelevant Alternatives and certainly the Pareto principle and the condition of Non-Dictatorship. But as Condorcet pointed out as far back as [1785], majority voting may not lead to an ordering. More specifically, intransitivity is possible. Consider the following example. There are three alternatives x, y, and z, among which choice is to be made. One-third of the voters have the ranking x, y, z; one-third, the ranking y, z, x; and one-third, the ranking z, x, y. Then a majority of the voters prefer x to y, a majority prefer y to z, and a majority prefer z to x. Unfortunately, this result is not due to a removable imperfection in the method of majority voting. The four conditions

on social choice are mutually contradictory.

The philosophical and distributive implications of the paradox of social choice are still not clear. Certainly, there is no simple way out. I hope that others will take this paradox as a challenge rather than as a discouraging barrier.

REFERENCES

Arrow, K.J. 1951a An extension of the basic theorems of classical welfare economics. In J. Neyman (ed.) Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. Berkeley and Los Angeles: University of California Press, pp. 507-532.

Arrow, K.J. 1951b Social Choice and Individual Values. New York: Wiley.

Arrow, K.J. 1953 Le rôle des valeurs boursières pour la repartition la meilleure des risques. Econometrie, Colloques Internationaux du Centre National de la Recherehe Scientifique, Vol. XI, pp. 41-47.

Arrow, K.J. 1963a Uncertainty and the welfare economics of medical care. American Economic Review 53: 941-973.

Arrow, K.J. 1963b Social Choice and Individual Values, second edition, New York: Wiley, and New Haven: Yale University Press.

Arrow, K.J. 1972 Models of racial discrimination. In A.H. Pascal (ed.) Racial Discrimination in Economic Life. Lexington, Mass.: D.C. Heath, Chapter 2.

Arrow, K.J., and G. Debreu 1954 Existence of equilibrium for a competitive economy. Econometrica 22:265-290.

Arrow, K.J., and F.H. Hahn 1971 General Competitive Analysis. San Francisco: Holden-Day, and Edinburgh: Oliver & Boyd.

Bergson, A. 1938 A reformulation of certain aspects of welfare economics. Quarterly Journal of Economics 52:310-334.

Cassel, K.G. 1918 Theoretische Sozialökonomie. Erlangen: W. Scholl.

Condorcet, Marquis de 1785 Essai sur l'application de l'analyse à la probabilité des decisions rendues à la pluralité des voix. Paris.

Debreu, G. 1951 The coefficient of resource utilization. Econometrica 19:273-292.

Debreu, G. 1959 Theory of Value. New York: Wiley.

Gale, D. 1957 General equilibrium of linear models. Unpublished.

Hicks, J.R. 1939 Value and Capital. Oxford: Clarendon Press.

Hicks, J.R., and R.G.D. Allen 1935 A reconsideration of the theory of value. Economica N.S. 1:52-76, 196-219.

Hotelling, H. 1935 Demand functions with limited budgets. Econometrica 3:66-78.

Hotelling, H. 1938 The general welfare in relation to problems of taxation and railway and utility rates. Econometrica 6:242-269.

Kakutani, S. 1941 A gereralization of Brouwer's fixed-point theorm. Duke Mathematical Journal 8:451-459.

Keynes, J.M. 1936 The General Theory of Employment Interest and Money. New York: Harcourt Brace.

Koopmans, T.C. 1951 Analysis of production as an efficient combination of activities. In T.C. Koopmans (ed.) Activity Analysis of Allocation and Production. New York: Wiley, Chapter III.

Lange, O. 1942 Say's Law: a restatement and criticism. In O. Lange, F. McIntyre, and T. Yntema (eds.) Studies in Mathematical Economics and Econometrics. Chicago: University of Chicago Press, pp. 49-68.

Leijonhufvud, A. 1968 On Keynesian Economics and the Economics of Keynes: A Study of Monetary Theory. New York: Oxford University Press.

McKenzie, L. 1954 On equilibium in Graham's model of world trade and other competitive systems. Econometrica 22:147-161.

McKenzie, L. 1959 On the existence of general equilibrium for a competitive market. Econometrica 27:54-71.

McKenzie, L. 1961 On the existence of general equilibrium: some corrections. Econometrica 29:247-248.

Neisser, H. 1932 Lohnhöhe und Beschäftigungsgrad im Marktgleichgewicht. Weltwirtschaftliches Archiv 36:415-455.

Neumann, J. von 1937 Über ein ökonomisches Gleichungssystem

und eine Verallgemeinerung der Brouwerschen Fixpunktsatzes. Ergebnisse eines Mathematischen Kolloquiums 8: 73-83.

Neumann, J. von, and O Morgenstern 1944 Theory of Games and Economic Behavior. Princeton: Princeton University Press.

Pareto, V. 1909 Manuel de l'économie politique. Paris: Droz.

Radner, R. 1968 Competitive equilibrium under uncertainty. Econometrica 36:31-58.

Samuelson, P.A. 1938 A note on the pure theory of consumer behavior. Economica N.S. 4:61-71.

Samuelson, P.A. 1947 Foundations of Economic Analysis. Cambridge, Mass.: Harvard University Press.

Samuelson, P.A. 1971 Maximum principles in analytical economics. In Les Prix Nobel en 1970. Stockholm: P.A. Norstedt & Söner, pp. 273-288.

Schlesinger, K. 1933-4 Über die Produktionsgleichungen der ökonomische Wertlehre. Ergebnisse eines Mathematischen Kolloquiums 6:10-11.

Stackelberg, H. von 1933 Zwei Kritische bemerkungen zur preistheorie Gustav Cassels. Zeitschrift für Nationalökonomie 4:456-472.

Wald, A. 1933-4 Über die eindeutige positive Lösbarkeit der neuen Produktionsgleichungen. Ergebnisse eines Mathematischen Kolloquiums 6:12-20.

Wald, A. 1934-5 Über die Produktionsgleichungen der ökonomische Wertlehre. Ergebnisse eines Mathematischen Kolloquiums 7:1-6.

Wald, A. 1936 Über einige Gleichungssysteme der mathematischen Ökonomie. Zeitschrift für Nationalökonomie 7:637-670.