

## Mathematical Models in the Social Sciences\*

KENNETH J. ARROW

### I. THE USEFULNESS OF MATHEMATICAL REASONING

IT IS a commonplace remark among many social scientists that mathematics, however useful it may have proved in the physical sciences, can play no essential role in the development of the social sciences because the phenomena studied are somehow different—"human beings are not amenable to mathematical law." The social scientist who thinks a little more about the matter will perhaps add that mathematical analysis is quantitative, while his field calls for qualitative analysis. Doubtless he will concede that certain elementary facts of a numerical nature can be tabulated (e.g., distribution of income or population); and he will usually admit that for certain purposes, one might be permitted to add up a column of the table. Nevertheless, it is held that the judgment and intuition of the skilled investigator are fundamentally more useful in the social sciences than mathematical formulas based on quantitative observations.

To the mathematician or the individual trained in the spirit of modern mathematics, the views just presented seem to be based on nothing more profound than a misunderstanding. "Mathematics," said the American physicist Gibbs, "is a language." If this be true, any meaningful proposition can be expressed in a suitable mathematical form, and any generalizations about social behavior can be formulated mathematically. Mathematics,

\* This chapter will be reprinted as Cowles Commission Paper, New Series, No. 48. The author wishes to thank J. Marschak, of the Cowles Commission and the University of Chicago, for valuable comments.

in this view, is distinguished from the other languages habitually used by the social scientist chiefly by its superior clarity and consistency.<sup>1</sup> Furthermore, it is simply not true that mathematics is useful only in quantitative analysis. Doubtless many branches of mathematics—especially those most familiar to the average individual, such as algebra and the calculus—are quantitative in nature. But the whole field of mathematical or symbolic logic is purely qualitative. We can frame such questions as the following: Does the occurrence of one event always imply the occurrence of another? Is it impossible that two events should both occur? The events here may be of a purely qualitative nature, such as the presence or absence of traits in a culture complex. It must further be observed that there may very well be a quantitative aspect to the study of even the most definitely qualitative phenomena; if we realize that we will rarely be able to assert universally valid laws about the relation between different traits, we will be willing to ask in what proportion of the observed phenomena will two traits be found together, i.e., what is the probability of their coexistence. The Mendelian theory of inheritance is the prototype of the transformation of qualitative into quantitative analysis via the probability calculus.<sup>2</sup> It is, of course,

<sup>1</sup> The doctrine that mathematics is a superior language goes back to Leibnitz. The inclusion of all forms of logical reasoning in mathematics was begun by George Boole and continued by Charles S. Peirce, Gottfried W. Frege, and Bertrand Russell. See F. C. S. Schmitt's *Introduction to Mathematical Philosophy* (1st ed., 1920). The best elementary presentation of the present state of mathematical logic is found in Alfred Tarski, *Introduction to Logic* (1941).

<sup>2</sup> It is of interest to observe that the modern theory of statistical inference fundamental to any

Reprinted from

THE POLICY SCIENCES

Edited by Daniel Lerner and Harold D. Lasswell

Stanford University Press, Stanford, California

also clear that quantitative phenomena enjoy equal claim with qualitative ones in the study of social forces. An understanding of a community doubtless requires knowledge of its religious and social beliefs, but it also involves knowledge of the distribution of income by size and by occupation, the total population and its distribution by occupation and social class, and the proportion of resources devoted to various social and individual needs.

Finally, the argument that only trained intuition can yield worth-while social analysis is rejected as meaningless by the mathematically trained. If the intuition of the investigator is reliable, it will yield the same judgments every time it is confronted with the same set of facts. But any such unique correspondence can always be represented by a mathematical relation of sufficiently complicated form. Hence, any intuitive knowledge can always be reduced to mathematical terms. Apart from this, there is the general presupposition that scientific knowledge should be interpersonally valid and transmittable and hence expressible in an objective, consistent language.

These arguments seem logically irrefutable, and yet, outside the realm of economics, very little use has been made of mathematical and symbolic methods. Even in economics, only a small minority of the theorists use anything more complicated than a very elementary form of calculus. How can we explain this failure on the part of social scientists to accept in their practice the theoretically superior language of mathematics? The simple fact that they would like to shun such a difficult subject as mathematics can be regarded at best as only a partial explanation. There have always been enough mathe-

quantitative analysis in the social sciences owes its inception to the interest of Francis Galton and Karl Pearson in the problems of inheritance and evolution. A further discussion of the significance of probability laws in the social sciences is given later in this chapter.

maticians interested in the social sciences to have made the superiority of mathematics manifest if it were a clearly better tool.

There must be some flaw in the arguments advanced above for the use of mathematical methods in the social sciences; the most important one is concealed in the statement that every proposition "can" be expressed in mathematical form. The statement is doubtless true if we mean that there exists in some Platonic realm of being a mathematical expression of every given proposition, but it is not true if we mean that the mathematical expression in question can be given within the realms of mathematical theory now existing. Every mathematician realizes what a small part of all the potentially available mathematical knowledge is actually grasped at the present time.<sup>3</sup> The usual reaction of the "literary" social scientist when confronted with a mathematical system designed as a model of reality is to assert that it is "oversimplified," that it "does not represent all the complexities of reality." In effect, he is saying that the symbolic language in which the mathematical model is expressed is too poor to convey all the nuances of meaning which he can carry in his mind. What happens is that the very ambiguity and confusion of ordinary speech give rise to a richness of meaning which surpasses for the social scientist the limited resources of mathematics, in which each symbol has only one meaning. It is not surprising that there should be a difference between the social and the natural sciences in this regard. Language is itself a social phenomenon, and the multiple meanings of its symbols are very likely to be much better adapted to the con-

<sup>3</sup> Consider the following apparently simple statement: Every even number is the sum of two odd prime numbers (a prime number is a number divisible only by itself and by 1). There is no known exception to the statement, yet it has never been proved in general. This is known as Goldbach's problem and dates from the 18th century.

veying of social concepts than to those of the inanimate world. Furthermore, the empirical experience on which one's understanding of the social world is based consists to a large extent of symbolic expressions of other individuals; one can apprehend these expressions directly because one is himself part of the social world he observes.<sup>4</sup> Such apprehension must inevitably take place on a largely unconscious level unamenable to mathematical expression (which is surely the acme of consciousness). It is precisely in the field of economics, where the individuals studied are engaged in relatively highly conscious calculating operations, that mathematical methods have been most successful.<sup>5</sup>

It is true, then, that there are certain limitations of mathematical methods in the social sciences. Nevertheless, it must be insisted that the advantages are equally apparent and may frequently be worth a certain loss of realism. In the *first* place, clarity of thought is still a pearl of great price. In particular, the multiplicity of values of verbal symbols may be a great disadvantage when it comes to drawing the logical consequences of a proposition. Consider, for example, the following verbal arguments: (a) If prices are high, people will tend to buy less; when people buy

less, manufacturers produce less, since they tend to produce only what they can sell; therefore, high prices are associated with low production. (b) If prices are high, manufacturers will produce more, since it is more profitable for them to do so; therefore, high prices are associated with high production. At a verbal level, both arguments are convincing; yet obviously they cannot both be valid. Let us try to isolate the postulates of the two arguments and express them symbolically. Let  $x_1$  denote the amount which people will buy,  $x_2$  the amount manufacturers will produce, and  $p$  the price. The first argument says that  $x_1$  is a decreasing function of price, i.e., that  $x_1 = f(p)$ ; and that  $x_1 = x_2$ . The second argument says that  $x_2$  is an increasing function of price,  $x_2 = g(p)$ . Then we have three equations in three unknowns; thus, in general there is no inconsistency, all three relations simultaneously determining  $x_1$ ,  $x_2$ , and  $p$ . These relations express the behavior of consumers, the market, and producers respectively. So long as no change occurs in anyone's behavior pattern, the values of  $x_1$ ,  $x_2$ , and  $p$  will remain constant. There will be no question whether supply varies directly or inversely with price, since neither moves at all. Suppose, however, that there were a shift in the tastes of consumers, so that  $f(p)$  changed to another function, say  $f_1(p)$ . To obtain the values of  $x_1$ ,  $x_2$ , and  $p$ , we would solve the new system of simultaneous equations,  $x_1 = f_1(p)$ ,  $x_1 = x_2$ ,  $x_2 = g(p)$ . Note now that the solutions  $x_1'$ ,  $x_2'$ ,  $p'$  to this equation system satisfy the conditions that  $x_2' = g(p')$ . Hence, if  $p'$  is greater than  $p$ ,  $x_2'$  will be greater than  $x_2$ . If therefore we have a sequence of observations in which consumers' tastes are varying but production conditions, as expressed by  $g(p)$ , remain constant, prices and production will move together. On the other hand, if consumers' tastes are constant but production conditions are variable, high

<sup>4</sup> This point has been stressed by Frank H. Knight in "The Limitations of Scientific Method in Economics," in *The Ethics of Competition and Other Essays* (1935), pp. 105-47.

<sup>5</sup> Credit for the first significant use of mathematics in economics is due the great French economist Augustin Cournot, who published in 1838 his *Recherches sur les principes mathématiques de la théorie des richesses*. His work was largely neglected. It was the contribution of W. Stanley Jevons in *The Theory of Political Economy* (1871) and especially the contribution of L. Walras in *Éléments d'économie politique pure* (1874), which brought the power of the mathematical methods to the attention of economists. The present status of mathematical economics is best summarized in Paul A. Samuelson, *Foundations of Economic Analysis* (1947); see also J. R. Hicks, *La théorie mathématique de la valeur en régime de libre concurrence* (1937), and *Value and Capital* (1939), particularly the Mathematical Appendix, for a more intensive survey of more limited areas than Samuelson's work. These works do not cover, however, a good deal of significant work published in the various economic journals, and particularly *Econometrica*.

prices will be associated with low production. Thus, mathematical symbolism resolves the apparent contradiction between the two arguments and shows when each is valid.

In addition to the problem of clarity in logical deductions, there is another methodological question related to the formulation of theoretical models: the problem of inductive inference. It is by now a platitude of the scientific method that if theory without empirical evidence is unreliable, empirical inquiry without theoretical background is unfruitful. The theory of statistical inference, as it has been developed in recent years by Jerzy Neyman, Egon S. Pearson, and Abraham Wald on the basis of the earlier work of R. A. Fisher and Karl Pearson, is precisely the mathematical expression of the logic of induction. It has shown clearly how the optimum statistical methods depend critically on the theoretical model assumed.<sup>6</sup> Now the derivation of the statistical methods appropriate for making inferences within a given theoretical model is usually a matter of considerable mathematical difficulty.<sup>7</sup> To proceed with the derivation at all, the underlying theoretical presumptions must be set forth in explicit symbolic form. Hence, a *second* argument, besides that of greater clarity of thought, for the explicit formulation of theories in mathematical terms is the resultant opportunity to tap the great resources of modern theoretical statistics as an aid in empirical verification.

<sup>6</sup> The concept of "optimum" is, of course, always relative to a value system. As will be seen below, the statement in the text is not meant to imply that all the philosophical difficulties in the theory of induction have been completely resolved. We can, however, be reasonably sure that the optimum methods, however defined, for making a choice among alternative theories on the basis of given empirical evidence will depend critically upon the theoretical background of the inquiry, i.e., upon the range of theories which are believed to be sufficiently plausible that they are admitted to the set among which the choice is to be made.

<sup>7</sup> The level of mathematics needed in modern statistical theory can be seen in such a work as Harald Cramér, *Mathematical Methods of Statistics* (1946).

The observations made above as to the limitations of our present knowledge of mathematics are applicable here also. It is unfortunately true that it is very easy to formulate theoretical models in which the determination of the optimum statistical methods leads to mathematical problems which have not been solved; in other cases, the resultant problems can be solved in principle, but the computations needed to find the solution in any given case take an impractical amount of time.<sup>8</sup> Here again we must resort to simplification. The customary procedure is to substitute a mathematically practicable theory, as similar as possible to the desired one, and use that as the basis for deriving statistical methods. For example, the assumption is frequently made in theoretical models in biology and economics that there exists some linear combination of the variables that is normally distributed. The assumptions of normality and linearity are introduced primarily because of the relative mathematical simplicity of deriving optimum methods of estimation and tests of significance.<sup>9</sup>

In Section II, some alternative possibilities in the types of mathematical models which are being proposed are discussed in a classificatory way. In Section III, extended examination is given to one proposed model, the "game" theory of social behavior developed by John von Neumann and Oskar Morgenstern. Section IV will take up more briefly some other pro-

<sup>8</sup> An attempt, for example, to estimate statistically the economic laws governing a large modern country on the basis of a relatively simple model (say, a few hundred equations to be fitted) could easily occupy the best computing machines now available for the next five hundred to a thousand years.

<sup>9</sup> Of course there are other justifications for the widespread use of the normal distribution, in particular the Central Limit Theorem, which asserts that under certain general conditions the sum of a large number of independent variates, each contributing a small part to the whole, will be normally distributed. See, for instance, Harald Cramér, *Mathematical Methods of Statistics* (1946).

posed mathematical theories of the social sciences. Finally, certain very important recent advances in the methodology of empirical work as guided by theoretical models will be taken up in Section V. While the literature of current economic thought is frequently referred to, the emphasis is on those parts which have wide applicability in the social sciences.

## II. SOME CLASSIFICATIONS OF MODELS

### 1. *Individualistic versus collective basis*

In most mathematical and, generally, in most deductive studies in the social sciences, the starting point is the behavior of the individual. Each individual is conceived of as acting in a way determined partly by his psychology and his physical surroundings and partly by the actions of others. If there are  $n$  individuals, we may denote the actions of individual  $i$  by  $A_i$ , and the nonsocial determinants of his behavior by  $P_i$ . Then the actions of the first individual may be described by a symbolic equation,

$$A_1 = f(P_1, A_2, \dots, A_n). \quad (1)$$

There is one such equation for each individual. Together, they constitute  $n$  equations in the  $n$  variables  $A_1, \dots, A_n$ . In general, these may then be solved to express the actions of all individuals in terms of the data  $P_1, \dots, P_n$ .<sup>10</sup> Therefore, given the reaction of each individual to his total (social and other) environment, as expressed in relations of type (1), and given the nonsocial environmental factors, which we may term *exogenous*, we can determine the behavior of society in the sense that we can determine the behavior of any individual in society.

This individualistic viewpoint, as we

<sup>10</sup> The fact that the actions  $A_i$  may not be representable by single numbers does not alter the validity of the principle that  $n$  equations suffice, in general, to determine  $n$  unknowns.

may term it, is explicit in the main tradition of economic thought and is completely accepted in the von Neumann-Morgenstern game theory. It seems also to be accepted by the other theorists whose work is discussed below, though George Zipf seems at times to be referring to the laws of behavior of society in some total sense. The individualistic viewpoint has been challenged recently by Rutledge Vining in the course of a methodological controversy with Tjalling C. Koopmans.<sup>11</sup> Vining has stated, "I think that in a positive sense the aggregate has an existence over and above the existence of Koopmans' individual units and behavior characteristics that may not be deducible from the behavior of these component parts."<sup>12</sup> Taken literally, this position seems indefensible. As Koopmans points out, a full characterization of each individual's behavior logically implies a knowledge of group behavior; there is nothing left out.<sup>13</sup> The rejection of the organism approach to social problems has been a fairly complete, and to my mind salutary, rejection of mysticism. But as usual in these problems, there is something to be said for at least the possibility of a collective basis for social theorizing, if not taken too literally. Much of our casual observation and much of our better statistical information relate to groups rather than individuals. We may therefore be led to generalize and form a theory whose subject is the total

<sup>11</sup> Rutledge Vining, "Koopmans on the Choice of Variables to Be Studied and of Methods of Measurement," *Review of Economics and Statistics*, XXXI (1949), 77-86. See also Tjalling C. Koopmans, "A Reply," *Review of Economics and Statistics*, XXXI (1949), 86-91; Rutledge Vining, "A Rejoinder," *Review of Economics and Statistics*, XXXI (1949), 91-94. Vining's first paper was a criticism of an earlier paper of Koopmans entitled "Measurement Without Theory," *Review of Economics and Statistics*, XXIX (1947), 161-72. The Koopmans paper, in turn, was a review article of a book by Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (1946).

<sup>12</sup> Vining, "Koopmans on the Choice of Variables to Be Studied and of Methods of Measurement," p. 81.

<sup>13</sup> Koopmans, "A Reply," pp. 86-87.

behavior of a group. So long as it is understood that such a theory is really a resultant of certain as yet unanalyzed laws of individual behavior, no harm will be done, and the greater convenience of empirical analysis on groups may be highly beneficial.

In fact, even in economics, the unit of the theory of production is not really the individual but the firm, which is an operating organization of individuals. Similarly, the unit of the theory of consumption is really the household, not the individual consumer. In empirical economics, the investigator is usually forced to use a collectively based model by the nature of the data. The Keynesian theory postulates that the consumption of a community is an increasing function of its total income. As an empirical equation, this is a statement about group behavior, not individual behavior.<sup>14</sup> Similar remarks apply to all macroeconomic models designed as a basis for empirical fitting, in which the variables that enter are obtained by aggregating the behavior of many individuals.<sup>15</sup> They apply even to many

models constructed for the purposes of theoretical analysis. The various representations of the Keynesian theory in mathematical form all involve functional relations among magnitudes which cannot be identified with the behavior of any individual.<sup>16</sup>

The usual feeling among economists is that these macroeconomic models could be justified on the basis of an individualistic theory if suitable definitions of the aggregate magnitudes in terms of those pertaining to individuals were given. The problem of finding such definitions, generally known as the *aggregation problem*, has received a certain amount of attention in recent years. Because of the nature of the discussion, at once technical and unresolved, it is not summarized here. But one methodological principle emerges clearly: in order to have a useful theory of relations among aggregates, it is necessary that they be defined in a manner derived from the theory of individual behavior. In other words, even the definition of such magnitudes as national income cannot be undertaken without a previous theoretical understanding of the underlying individual phenomena. It also seems evident that the aggregated model must include among its variables some which characterize the distribution of various magnitudes among the individuals of the society; e.g., in the consumption function some measure of income inequality should be introduced as an additional variable. Doubtless, the same remarks apply to aggregation in other social realms.<sup>17</sup>

ment; the laws relating them are laws of behavior of the society as a whole and only by implication are they laws of individual behavior.

<sup>14</sup> Oscar Lange, "The Rate of Interest and the Optimum Propensity to Consume," *Economica*, n.s., V (1938), 12-32; John R. Hicks, "Mr Keynes and the 'Classics': A Suggested Interpretation," *Econometrica*, V (1937), 147-59. The anti-Keynesian work of A. C. Pigou, *Employment and Equilibrium* (1941) follows the same methodological pattern on this point.

<sup>17</sup> See Francis W. Dresch, "Index Numbers and the General Equilibrium," *Bulletin of the American Mathematical Society*, XLIV (1938), 139-41,

<sup>14</sup> John Maynard Keynes, *The General Theory of Employment, Interest, and Money* (1936), pp. 89-131. There are a large number of variant empirically derived consumption functions; see, for example, Arthur Smithies, "Forecasting Postwar Demand: I," *Econometrica*, XIII, No. 1. (1945), 1-14; Jacob L. Mosak, "Forecasting Postwar Demand: III," *Econometrica*, XIII, No. 1 (1945), 25-53; and Trygve Haavelmo, "Methods of Measuring the Marginal Propensity to Consume," *Journal of the American Statistical Association*, XLII (1947), 105-22.

<sup>15</sup> Jan Tinbergen, *Statistical Testing of Business-Cycle Theories* (1939); Lawrence R. Klein, "The Use of Econometric Models as a Guide to Policy," *Econometrica*, XV (1947), 111-51, and *Cyclical Fluctuations in the United States, 1929-1941* (Cowles Commission for Research in Economics, Monograph No. 11 [1950]); Mordecai Ezekiel, "Savings, Consumption, and Investment," *American Economic Review*, XXXII (1942), 22-49, 272-77; and Colin Clark, "A System of Equations Explaining the United States Trade Cycle, 1921 to 1941," *Econometrica*, XVII (1949), 93-124. These are all attempts to find a comprehensive system of equations which will explain cyclical fluctuations in business. Mention should also be made of the empirical studies on production made by Paul H. Douglas and his associates, summarized in Douglas's, "Are There Laws of Production?" *American Economic Review*, XXXVIII (1948), 1-41. All these studies employ magnitudes such as total consumption, total output, and aggregate invest-

## 2: The principle of rationality

A postulate frequently encountered in theoretical economics and elsewhere in social theory is that the behavior of the individual or group can be described by saying that the individual or group is seeking to maximize some quantity. Thus, in the theory of the firm, the economist postulates that the individual seeks to choose that mode and scale of operation which will yield more profit than any other possible choice. In the theory of consumption, it is assumed that among all the combinations of commodities an individual can afford, he chooses that combination which maximizes his utility or satisfaction. Behavior of this type is frequently referred to as rational.

The basis for the assumption of rationality is the following seemingly quite general formulation of individual behavior in a social situation: Each individual at a given moment of time is free to choose among several possible courses of action; he decides among them on the basis of their consequences. The range of actions open to him and the consequences of these actions are determined by the contemporary actions of others, by the past actions of himself and of others, and by the exoge-

nous factors. To put it briefly, we may say that the individual can choose at any instant among a limited range of consequences according to his tastes. It is natural to suppose that as the range varies, his choices from different ranges bear some sort of consistent relation to each other. In particular, we will suppose that he makes the same choice each time he is confronted with the same set of alternatives; further, if, when confronted with the two alternatives *A* and *B*, he chooses *A*, and when confronted with the two alternatives *B* and *C*, he chooses *B*, then it is reasonable to suppose that when confronted with a choice between *A* and *C*, he chooses *A*. Under these circumstances, the process of decision may be described as follows: we can imagine the individual as listing, once and for all, all conceivable consequences of his actions in order of his preference for them; then, on being informed of the choices actually available to him, he selects that one which stands highest on his list. The list represents his tastes, while the alternatives actually available to him are restricted by obstacles.<sup>18</sup> So far there has been no quantitative element at all in the formulation. In certain problems, of course, the desirability of the various courses of action is associated naturally with a quantitative variable; e.g., profits in the case of a firm deciding upon its productive operations, probability of winning in the case of a country deciding whether or not to go to war. Even if there is no natural quantitative variable associated with the problem, it is sometimes useful for analysis to introduce one artificially, by assigning numbers to each consequence of an individual's actions in such a way that, of two consequences, the higher number will be assigned to the preferred one. Such an assignment of numbers is

and "Continuous Index Numbers and Quantitative Study of the General Economy," in Jerzy Neyman (ed.), *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability* (1949), pp. 2-11; Lawrence R. Klein, "Macroeconomics and the Theory of Rational Behavior," *Econometrica*, XIV (1946), 93-108, and "Remarks on the Theory of Aggregation," *Econometrica*, XIV (1946), pp. 303-12; Kenneth May, "The Aggregation Problem for a One-Industry Model," *Econometrica*, XIV (1946), 285-98, and "Technological Change and Aggregation," *Econometrica*, XV (1947), 51-63; Shou Shan Pu, "A Note on Macroeconomics," *Econometrica*, XIV (1946), 299-302; Kenneth J. Arrow, "Summarizing a Population of Behavior Patterns," *Econometrica*, XVI (1948), 203; and André Nataf, "Sur la possibilité de construction de certains macromodèles," *Econometrica*, XVI (1948), 22-44. The works of Dresch and May, in particular, make great use of the index numbers introduced by François Divisia in his *Économique rationnelle* (1928), pp. 260-73. For the aggregation problems in the consumption function, see also Jacob Marschak, "Personal and Collective Budget Functions," *Journal of Economics and Statistics*, XXI, No. 4 (1943), 161-70.

<sup>18</sup>This formulation is essentially due to Vilfredo Pareto, *Manuel d'économie politique* (2d ed., 1927), pp. 150-79.

known as a *utility index*. Let  $x$  stand for the consequences of some action, and let  $U_1(x)$  be the utility assigned to  $x$  under one assignment of the type described. Then  $U_1(x_1) > U_1(x_2)$  if and only if the individual prefers  $x_1$  to  $x_2$ . There is, clearly, nothing unique about the assignment of a utility index, at least within the conditions laid down thus far. Let  $F(u)$  be any strictly increasing function of a real variable  $u$ , i.e., if  $u_1 > u_2$ , then  $F(u_1) > F(u_2)$ . Then, if we define  $U_2(x)$  as  $F[U_1(x)]$ , it is clear that  $U_2(x)$  will also serve as a utility index.<sup>19</sup>

Under the hypothesis of rationality, then, the individual's behavior depends on his tastes—as expressed, say, by a utility index—and upon the obstacles, which are determined by exogenous factors, by the actions of others, and possibly by past actions of the individual and of others. We may say that the individual maximizes his utility, subject to the obstacles. We are thus led to an expression of form (1) above. If we postulate nothing more than that there exists some utility function in terms of which the individual's behavior can be described in the above way, we have imposed some restrictions on the possible forms of (1), since the principle of rationality is not a pure tautology; i.e., it is logically possible that an individual may choose  $A$  over  $B$ ,  $B$  over  $C$ , and yet choose  $C$  over  $A$ .<sup>20</sup> However, the degree of restrictiveness thus obtained is not great. It is customary to supplement the general principle of rationality in any particular in-

<sup>19</sup> It should be clearly understood that the variable  $x$  need not be numerical. The consequences of an action may be a power or prestige situation or a state of religious ecstasy, as well as a bundle of commodities for consumption or a sum of money. All that is required is that given two well-defined situations resultant from his actions, an individual should be able to say that he prefers one to the other.

<sup>20</sup> For a discussion of the refutable, and therefore empirically meaningful, consequences of the principle of rationality in the field of consumers' demand, see Paul A. Samuelson, *Foundations of Economic Analysis* (1947), chap. v.

stance by further assumptions as to the nature of the preferences involved. For example, in consumer's demand theory, it is ordinarily understood that more of a commodity is preferred to less, all other things being equal.

A number of objections have been raised to the usefulness of the principle of rationality: (a) If the complicated nature of the range of choices possible in an actual social situation is even approximately taken into account in the theoretical model, the mathematical problems to be solved in the maximization of utility will become extremely complex, and it will be hard to derive results which have any simple meaning. This objection has been frequently raised, for example, against the Walrasian scheme of general equilibrium in economics. (b) There is no real reason to suppose that individual behavior does conform to the principle of rationality. This argument is partly related to the previous one; it is argued that if the rational choice is too difficult for the trained mathematician to find, it is certainly unreasonable to suppose that the untrained, unreflecting, average individual will be able to locate it.<sup>21</sup> (c) The utility function itself, even if it plays the role assigned to it, is highly unstable over time; hence, for an understanding of social processes, more interest attaches to the determinants of the variation of tastes than to the line of causation from the utility index to the actual decision made.<sup>22</sup> (d) There is a fundamental ambiguity in the concept of rationality in a social situation. An individual will soon realize that his

<sup>21</sup> Objections of this type and others were raised by Richard A. Lester in his "Shortcomings of Marginal Analysis for Wage-Employment Problems," *American Economic Review*, XXXVI (1946), 63-82. For a reply see Fritz Machlup, "Marginal Analysis and Empirical Research," *American Economic Review*, XXXVI (1946), 519-54.

<sup>22</sup> Vining attributes this view to Thorstein Veblen. See his article, "Koopmans on the Choice of Variables to Be Studied and Methods of Measurement," *Review of Economics and Statistics*, XXXI (1949), 77-86, esp. 82-83.

actions, in addition to their other consequences, will alter the obstacles faced by others, thereby affecting their actions and in turn altering the obstacles controlling his choices. Hence, his actions will be partly controlled by his realization of their repercussions on the actions of others. But the same statement is true of each other individual; thus, each will be concerned with the effect of his action on the others, and no determinate solution will be possible.<sup>23</sup>

There is no single sweeping principle which has been erected as a rival to that of rationality. To the extent that formal theoretical structures in the social sciences have not been based on the hypothesis of rational behavior, their postulates have been developed in a manner which we may term *ad hoc*. Such propositions are usually drawn from introspection or casual observation; sometimes they are of the nature of empirical regularities. They depend, of course, on the investigator's intuition and common sense.

An example of this approach is Lewis F. Richardson's approach to international relations.<sup>24</sup> What will make a government increase its armaments? Clearly, it will increase them more if the armaments of a potential enemy are greater; but it will be deterred from increasing them by the expense. Let  $x$  represent the armaments of this country,  $y$  those of a potential enemy, and  $t$  time; and let  $k$  and  $a$  be fixed coefficients. Then, the above theory may be expressed in the equation

$$dx/dt = ky - ax. \quad (2)$$

A similar equation holds for the other

country. Taken together, they form a complete system of differential equations which, with the initial conditions and the values of the coefficients, define the course of  $x$  and  $y$  over time. Richardson then deals with the implications of these results and those of more general models.

Because of the first and fourth difficulties mentioned above, theoretical economics has generally had an admixture of *ad hoc* assumptions which limit the scope of the principle of rationality. The hypothesis of a perfect competition removes the fourth difficulty because each individual supposes his effect on other individuals is so small that their actions are not influenced by him. Under conditions of monopoly it is assumed that while one individual can affect other individuals, none of them is strong enough to affect him. The monopolist's behavior incorporates his realization of this situation, thereby again avoiding indeterminacy. In those cases where neither hypothesis can be maintained (e.g., bargaining between a labor union and a monopolistic employer, or the behavior of an industry in which there are a few large firms) still more complicated *ad hoc* assumptions have been made about the way individual firms took into account the anticipated reactions to their actions in order eventually to reduce the problem to a simple maximization.<sup>25</sup>

One somewhat digressive remark on the principle of rationality may be in order: a rational theory always has a dual interpretation. On the one hand it may be taken as a description of reality to the extent that individuals really are consistent in the sense assumed. On the other hand it may be taken rather as a normative theory, which prescribes what individuals ought to do. Thus,

<sup>23</sup> This fourth objection to the principle of rationality has been recognized in a general way in economics in the theories of oligopoly and bilateral monopoly. It was given a definitive statement by Oskar Morgenstern in his *Wirtschaftsprognose* (1928), p. 98.

<sup>24</sup> Lewis F. Richardson, *Generalized Foreign Politics* (*British Journal of Psychology*, Monograph Supplement No. 23 [1939]).

<sup>25</sup> The most extended treatment of various possible assumptions of this type is found in Ragnar Frisch, "Monopole, polypole: La Notion de la force dans l'équilibre économique," *Nationaløkonomisk Tidsskrift* (1934).

theoretical economics has been used to analyze what the optimum state of economic welfare would be and how to attain it.<sup>26</sup> This subject of "welfare economics" is not new, being indeed as old as theoretical economics itself, but clarification of its basic principles has been a slow process. Even today there are a number of profound unresolved difficulties, principally revolving about the problem of comparing the welfares of different individuals in arriving at a concept of a social optimum.<sup>27</sup> Since we are here concerned with the problems of a descriptive social science, we shall not pursue this matter any further.

Statistical inference may be viewed broadly as the behavior of an individual under a certain set of conditions; namely, he does not know completely the consequences of his actions. We may therefore speak of a rational theory of statistical inference. More precisely, statistical inference may be described as follows: The true state of the universe under investigation is known, a priori, to belong to one of a class of states. A sample of elements of the universe is drawn; the probability distribution of the sample depends on the true state of the universe. The statistician must then take some action (e.g., estimate some state to be the true one, assert that the state lies in some subclass of the one given a priori; or, in industrial applications, accept or reject a lot of goods), the consequences of which are a function of the action and of the true state, being favorable if the action is really appropriate to the true state and unfavorable otherwise. For simplicity, let us consider the case where the true state is known to be one of a finite number, the possible samples to be drawn are finite

in number (e.g., suppose we draw a sample of three observations, each of which is either "yes" or "no," as in a questionnaire or in a quality inspection; then one possible sample is "yes, yes, no," another is "no, yes, no," and so on; there are altogether eight possible samples), and the number of actions the statistician can take is finite. Let  $i$  stand for an action,  $j$  for a true state,  $k$  a sample,  $p_{kj}$  for the probability of observing sample  $k$  when the true state is  $j$ , and  $r_{ij}$  the loss to the statistician if he takes action  $i$  when the true state is  $j$  ( $r_{ij}$  is small or negative if  $i$  is appropriate to  $j$ , large otherwise). The statistician's problem is to choose a decision function  $i(k)$ , which tells him what action to take for each possible sample  $k$ . The function  $i(k)$  is to be so chosen as to minimize in some sense the probable loss.

This formulation includes all the classical problems of statistics and more. For example, in the above illustration the problem of estimation is the case where an action consists of naming a true state. The range of possible actions is then the same as the range of possible states of nature. Another case would be that of two possible actions, one affirming and one denying that  $j = j_0$ , where  $j_0$  is a fixed possible true state. This is the classical problem of testing a hypothesis.<sup>28</sup>

We have here clearly a problem of

<sup>26</sup> This general formulation of the problem of statistical decision is due to A. Wald, "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, XV (1947), pp. 279-313, especially pp. 279-83; and his "Statistical Decision Functions," *Annals of Mathematical Statistics*, XX (1949), 165-205, esp. 172-73. For a more elementary presentation of the problem in a somewhat less general form, see Wald's "Contributions to the Theory of Statistical Estimation and Testing Hypotheses," *Annals of Mathematical Statistics*, X (1945), 299-326, and his "Statistical Decision Functions Which Minimize the Maximum Risk," *Annals of Mathematics*, XLVI (1945), 265-80. For the currently standard methods of estimation and testing hypotheses, see Harald Cramér, *Mathematical Methods of Statistics* (1946), chaps. xxx-xxxvii. An excellent introduction is to be found in Wald, *On the Principles of Statistical Inference* (Notre Dame Mathematical Lectures, No. 1 [1942]).

<sup>26</sup> Recent systematic expositions of this subject will be found in Paul A. Samuelson, *Foundations of Economic Analysis* (1947), chap. viii, and Melvin W. Reder, *Studies in the Theory of Welfare Economics* (1947), Parts I and III.

<sup>27</sup> Kenneth J. Arrow, *Social Choice and Individual Values* (1951).

behavior; the statistician must choose among the various possible functions  $i(k)$ , and the consequences are given, though not with certainty, by the conditions of the problem, the  $p_{kj}$ 's and  $r_{ij}$ 's. The foundations of statistical inference, from the normative point of view of proper scientific method, are then nothing but the application of a suitable principle of rationality to the problem just described. Indeed, it will be seen that Wald's important contributions to this field are closely related to an important development in social theory, the von Neumann-Morgenstern theory of games.

This relation between the theoretical analysis of behavior and the foundations of statistical inference can, of course, be applied in reverse. To the extent that actual behavior under conditions of uncertainty as to the consequences of any action is governed by the principle of rationality, the theory of statistical inference may also be interpreted as a descriptive theory. This point of view has especially been stressed in economics by Jacob Marschak.<sup>29</sup>

### III. THE THEORY OF GAMES<sup>30</sup>

The von Neumann-Morgenstern theory of games is an attempt to provide a theory of social (primarily economic) interaction by analogy with ordinary games of strategy (such as chess or card games) as they would be played by thoroughly rational individuals. It goes well beyond any other systematic social

theorizing in the complexity of its structure and the rigorous nature of its formal logic. Yet it is interesting to note that virtually no mathematics more difficult than algebra is employed although the chains of reasoning are frequently long and complicated.

#### 1. *Rational behavior in situations involving risk*

As a preliminary to the theory of games, though not strictly part of it, von Neumann and Morgenstern formulate the principle of rationality applied to a situation in which the consequences of the different actions open to an individual are expressed as probability distributions rather than as certainties. A simple example is that of an individual choosing between two lottery tickets, one of which pays \$1,000 with probability 0.01 and nothing otherwise, while the other pays \$100 with probability 0.10 and nothing otherwise. More generally, the choice is among a number of "lottery tickets," each of which is a promise to pay amounts  $m_1, m_2, \dots$  with probabilities  $p_1, p_2, \dots$ , the quantities  $m_1, m_2, \dots$  and  $p_1, p_2, \dots$  varying from ticket to ticket. Also, of course, in general the "payments" need not be sums of money but consequences of any sort which matter to the individual. A "lottery ticket" is a paradigm for any choice made by a human being in which the outcome of his action is not known with certainty but in which, on the basis of experience or intuition, he believes he knows the probabilities of the different possible outcomes.

In accordance with the principle of rationality, the individual may order all

<sup>29</sup> For a specific application, see Jacob Marschak, "Role of Liquidity Under Complete and Incomplete Information," *American Economic Review*, XXXIX, No. 3 (1949), esp. 95, esp. 192-95; see also Franco Modigliani, "Liquidity and Uncertainty: Discussion," *American Economic Review*, XXXIX, No. 3 (1949), 201-8, esp. 203-8. For earlier work, see Marschak's "Money and the Theory of Assets," *Econometrica*, VI (1938), 311-25, and Leonid Hurwicz, "Theory of the Firm and Investment," *Econometrica*, XIV (1946), 109-36, esp. 133-36. The relation between the problems of statistical inference and of behavior under certainty is clearly indicated in Frank H. Knight, *Risk, Uncertainty, and Profit* (1921), Part III, esp. chap. vii.

<sup>30</sup> John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior* (2d ed., 1947). Some of the review articles give excellent expositions of the theory; see especially Leonid Hurwicz, "The Theory of Economic Behavior," *American Economic Review*, XXXV (1945), 901-25; and Jacob Marschak, "Neumann's and Morgenstern's New Approach to Static Economics," *Journal of Political Economy*, LIV (1946), 97-115. For a more critical appraisal of the theory in its application to economics, see Carl Kaysen, "A Revolution in Economic Theory?" *Review of Economic Studies*, XIV (1946-47), 1-15.

possible probability distributions of outcomes; he may then choose a utility index having the properties described in the second part of Section II of this chapter and act so as to maximize utility. The concept of rationality as applied to this situation has been formalized by von Neumann and Morgenstern in a series of axioms.<sup>31</sup> The principal condition laid down, in addition to the general requirement that the probability distributions be ordered in accordance with the individual's preferences, may be expressed loosely as follows: If the individual prefers the outcome  $x_1$  to the outcome  $x_2$ , then he will prefer the certainty of  $x_1$  to an even chance of getting either  $x_1$  or  $x_2$ . With the aid of this highly reasonable condition, and some other more technical ones, it is shown that among the many utility indexes which can be assigned to the possible probability distributions, there is one with the property that the utility attached to a probability distribution is the mathematical expectation of the utilities of the possible outcomes. I.e., if a given choice has possible outcomes  $x_1 \dots x_n$  with probabilities  $p_1 \dots p_n$ , respectively, and if  $U(x)$  is the utility (in the index just described) of the outcome  $x$ , then the utility attached to the given choice is  $\sum_{i=1}^n p_i U(x_i)$ . (Note that this utility index is not unique; if  $U(x)$  is one such index and  $a$  is a positive number, then  $aU(x) + b$  is another such index.) Rational behavior in the choice of risky alternatives can thus be simplified to the statement that the individual seeks to maximize the expected value of his utility.<sup>32</sup> Milton Friedman

<sup>31</sup> John von Neumann and Oskar Morgenstern, *op. cit.*, chap. 1, sec 3; Appendix.

<sup>32</sup> The explanation of choice among risky alternatives in terms of maximizing the expected value of utility dates back to Daniel Bernoulli, "Specimen theoriae novae de mensura sortis" (1738); German translation, *Die Grundlage der modernen Wertlehre* (1896). Bernoulli was interested in showing, in connection with the famous St. Petersburg paradox, that an individual might not take a gamble even if the expected money

and L. J. Savage have developed an interesting hypothesis about the shape of the von Neumann-Morgenstern utility curve for money which will explain some phenomena observed in the behavior of individuals in connection with gambling and insurance.<sup>33</sup>

Since probability distributions are basic in the theory of games, the above construction enables the authors to speak of a numerical utility, having the convenient property that the utility attached to a chance event is precisely the mathematical expectation of the utilities of the various possible outcomes. However, they need further the idea of a transferable utility, i.e., some sort of measure the units of which mean the same, in some sense, to both players. In effect, they then restrict the realm of their theory greatly by assuming that the outcome of each game is expressible in some common units, which we may take to be money, and that the utility of a given sum of money, for each player, is simply a linear function of the sum, so that each player is seeking to maximize the expected amount of money he will receive. This assumption is to be understood as an intentional simplification to make the problem mathematically manageable. (Compare the remarks in the first section above.)

## 2. *The general concept of a game*<sup>34</sup>

In general, a game will have a certain number of players, say  $n$ . The game is composed of moves, which are of two

reward were infinite; his argument was that the individual was seeking to maximize the expected utility of money, which was given, he suggested, by the logarithm of the amount of money received. With the introduction of the marginal utility theory of value, it was natural to discuss choices among risky alternatives in these terms; see Alfred Marshall, *Principles of Economics* (8th ed., 1920), pp. 1, 843. However, Von Neumann and Morgenstern have been the first to show the assumptions about behavior which are implied in this theory.

<sup>33</sup> Milton Friedman and L. J. Savage, "The Utility Analysis of Choices Involving Risk," *Journal of Political Economy*, XVI (1948), 279-304.

<sup>34</sup> See von Neumann and Morgenstern, *op. cit.*, chap. ii.

types; personal, made by one of the players, and chance, in which one of several possible outcomes is selected by a chance device acting in accordance with certain probability laws. Thus, in a game of cards the distribution of the hands is made by a random device, while the players still have certain choices as to how to play their hands. The rules of the game prescribe the following: (a) whether the first move shall be personal or chance; if personal, who shall make it, and what range of choice he shall have; if chance, what the possible outcomes are and what the probability is of each; (b) after  $k$  moves have been made, whether the  $(k+1)$ st move shall be personal or chance; if personal, who shall make it, what range of choices he shall have, and what information he shall have about the outcomes of the previous moves; if chance, what the range of possible outcomes is and what probability is attached to each (note well that the rules may make each of these prescriptions dependent on the outcome of the preceding moves, it being required only that at each stage of the game the prescriptions in question must be unambiguously defined); (c) when the game shall stop and what amount shall be paid to each of the participants depending on the outcomes of the various moves. Thus, in chess, the rules specify that all moves are personal, with the two players taking alternate moves; that at each move each player knows the outcome of every previous move; and that in any given move, the range of permissible actions depends upon the present configuration of the pieces, which in turn depends upon the outcomes of all past moves. In card games, on the other hand, the dealing of the hands is a chance move, the outcome of which is not revealed to the other players.

It is easy to see that many economic and social situations can be described in these terms. For example, the opera-

tions of a market may be described as a game, with the buyers and sellers as players, and bids, offers, and agreements to conclude transactions as moves. The initial distribution of goods and money may be taken as part of the rules of the game; the final distribution of goods and money describes the pay-off to the various players attendant upon the particular set of moves which they made. Similarly, conflicts between nations or between social and economic groups can be brought within the same rubric.

Now suppose, for the purpose of simplified description, that each individual does not play the game by waiting each time until his move comes and then deciding among the various alternatives presented to him, but rather that he prepares beforehand a *strategy* which specifies for each possible situation with which he may be confronted in the course of the play what choice he shall make. Indeed, something of the sort must be done by skillful players, since they realize that the effects of their choices at various moves are interrelated; a good move at one point may improve the range of choices available at another. A good choice of any one move may therefore involve consideration of the alternative possible developments of the game following each possible action. Suppose each individual were to write down his strategy and hand it to an umpire. It would then be unnecessary actually to play the game; the umpire would need only to follow the instructions of the appropriate individual's strategy at each personal move and make the indicated chance moves. Therefore, every game can be reduced to the following type: each player makes a single move, namely the choice of a strategy; then the pay-off to each player is a random variable whose probability distribution is determined by the strategies of the various players. Since, according to the assumption made at the end of Section I, the individual is in-

interested only in the mathematical expectation of the pay-off, we can consider the pay-off to be a single number, instead of a probability distribution. Thus every game, no matter how complicated, can be reduced to the following *normal form*: each player makes one move, knowing nothing about the moves of the others, from a range of choices fixed in advance; the pay-off to each player is then a function of the moves of all the players. This reduction permits a consideration of the basic nature of games undisturbed by the complications of particular rules. From now on, we will consider all games to be in normal form.<sup>35</sup>

To simplify the analysis further, von Neumann and Morgenstern assume that the game will be terminated in a finite number of moves and that at each move there are only a finite number of possible choices. Then it can be seen that the number of possible strategies is finite. In the normal form, therefore, each individual has the choice among a finite number of alternatives for his move. It should be noted, though, that the finite number of strategies may be enormous; merely to enumerate all possible situations in chess is a tremendous task, and the number of possible ways of specifying responses to these situations staggers the imagination. If the restriction to a finite number of choices at each move is not made, then some extremely subtle mathematical difficulties arise. Some work has been done on this problem; it will be cited in footnotes where appropriate.

### 3. The zero-sum two-person game<sup>36</sup>

Consider now the special case of a game with two players in which the sum of the pay-offs to the two players is zero for all choices of strategies; e.g., if the outcome of a play of the game yields 5 to player *a*, then it necessarily constitutes a loss of 5 to player *b*. This is the model of a situation of pure conflict,

since the gain of one player is precisely the loss of the other. Since there are only two players, there is nothing whatever to be gained by agreement between them. Assume that the game is in normal form; let player *A* have the choice of strategies 1, . . . , *m*, player *B* the choice of strategies 1, . . . , *n*. If *A* chooses strategy *i* and *B* strategy *j*, let the pay-off to player *A* be  $a_{ij}$ ; then the pay-off to player *B* is  $-a_{ij}$ . Clearly, *A* wishes to maximize  $a_{ij}$ , while *B* seeks to minimize the same magnitude. Suppose *A* is contemplating the use of a particular strategy *i*. It is natural for him to think of the worst that *B* could do to him; the latter will try to choose that *j* which will make  $a_{ij}$  as small as possible for the given strategy *i*. Hence, if player *A* chooses *i*, all he can be sure of is  $\min_j a_{ij}$  (read, "minimum with respect to *j* of  $a_{ij}$ "). He can evaluate this magnitude for each value of *i* and choose that value of *i* which maximizes it; by this choice, he insures that his return is at least  $\max_i \min_j a_{ij}$ . Similarly, by suitable choice of *j*, player *B* can guarantee that player *A* will not receive more than  $\min_j \max_i a_{ij}$ . Suppose the pay-off matrix  $a_{ij}$  is such that,

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij}. \quad (3)$$

If the common value of the two sides of (3) is *v*, then *A* can, by choosing a suitable strategy *i'*, guarantee himself at least *v*, while *B*, by choosing a suitable strategy *j'*, can guarantee that *A* will not get more than *v*. It is clear that in this case, *A* will choose *i'*, *B* will choose *j'*, and the outcome of the game (termed its

<sup>35</sup> The concept of the economic agent as choosing a strategy—a plan for meeting contingencies—rather than merely making specific decisions from time to time, was clearly expressed, in a somewhat different context, by Albert G. Hart in his "Risk, Uncertainty, and the Unprofitability of Compounding Probabilities," in *Studies in Mathematical Economics and Econometrics*, edited by Oscar Lange, Francis McIntyre, and T. O. Yntema (1942), pp. 110–18.

<sup>36</sup> See von Neumann and Morgenstern, *op. cit.*, chaps. iii–iv.

value) will be that *B* pays an amount *v* to *A*. (The quantity *v* could be negative, in which case the payment would be from *A* to *B*.) Equation (3) can be valid, as shown in the pay-off matrix given in Table I in which each player has the choice of two strategies.

TABLE I

Strategies of Player <i>A</i>	Strategies of Player <i>B</i>	
.....	1	2
1	2	1
2	3	0

If player *A* chose strategy 1 and player *B* chose strategy 2, *A* would receive 1 unit from *B*. If *A* chose strategy 2, he would run the risk of receiving nothing; if *B* chose strategy 1, he might have to pay 3. Hence, *i* = 1, *j* = 2, and *v* = 1 represent a stable solution; even if one player knew in advance the other one's strategy, he would have no incentive to change his own.

Unfortunately, equation (3) need not hold. Consider the pay-off matrix in Table II.

TABLE II

Strategies of Player <i>A</i>	Strategies of Player <i>B</i>	
.....	1	2
1	1	-2
2	0	2

In this case,  $\max_i \min_j a_{ij} = 0$ ,  $\min_j \max_i a_{ij} = 1$ . If player *B* played so as to minimize  $\max_i a_{ij}$ , he would choose strategy 1; but if player *A* knew this, he would choose strategy 1, rather than strategy 2, which he would choose if he were maximizing  $\min_j a_{ij}$ . If, then, *B* realized that *A* would reason in this way, he would choose strategy 2, and so forth, leading to a seemingly endless regress. Indeed, this is precisely a chief

objection to the principle of rationality, as noted in the second part of Section II, above.

To arrive at a determinate stable solution, the following procedure is suggested: let player *A* have a random device which will choose strategy 1 with probability 0.4 and strategy 2 with probability 0.6. Then, if player *B* selects strategy 1, player *A* will have an expected return of  $0.4(1) + 0.6(0) = 0.4$ ; if *B* chooses strategy 2, the expected return is  $0.4(-2) + 0.6(2) = 0.4$ . Hence, *A* can guarantee himself an expected return of 0.4. Similarly, *B*, by playing strategy 1 with probability 0.8 and strategy 2 with probability 0.2, can hold the expected pay-off to *A* down to 0.4. This solution, in terms of randomized or mixed strategies, has the stability found in the first case; even if one player found out the probabilities used by the other, he could not gain from this information.

In general, the players can be considered not as choosing a strategy but as choosing a probability for each strategy (possibly zero). Let player *A* play strategies 1, . . . , *m* with probabilities  $x_1, \dots, x_m$ , respectively, while player *B* plays his strategies with probabilities  $y_1, \dots, y_n$ . Let  $\max_x$  mean

the maximum with respect to permissible variations in  $x_1, \dots, x_m$  and let  $\min_y$  have a corresponding meaning.

The expected pay-off to *A* if he chooses  $x_1, \dots, x_m$  and *B* chooses  $y_1, \dots, y_n$

is  $\sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j$ . Then the analogue to (3) is

$$\max_x \min_y \sum_{k=1}^m \sum_{j=1}^n a_{kj} x_k y_j = \min_y \max_x \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j. \quad (4)$$

If (4) holds, then we have a stable solution in terms of mixed strategies. It turns out, rather remarkably, that (4)

always holds.<sup>87</sup> The theory of rational behavior in zero-sum two-person games can therefore be regarded as definitely solved, at least within the limitations of a transferable utility. However, the objection that the principle of rationality cannot be realistic because it imposes too great burdens on the ability of the individual is not refuted; the solution of even relatively simple social games is frequently very difficult.<sup>88</sup>

Before leaving the discussion of zero-sum two-person games, consider the special case of a game in which, before reduction to normal form, the rules provide that at every stage of the game each individual knows the outcome of all previous moves, i.e., a game of *perfect information*. Chess and backgammon are familiar examples. The more enduring features of many social conflicts are of this type; there is no possibility of concealing the existence of a major strike from the management, though certain preliminary tactical plans may be concealed for a while. It is shown that for such games, relation (3) always holds. Mixed strategies become unnecessary

<sup>87</sup> See von Neumann and Morgenstern, *op cit.*, pp. 134-43, p. 55. The theorem was first proved, with the aid of very deep mathematical developments, by von Neumann in his "Zur Theorie der *n*-Personen-Spiele," *Mathematische Annalen*, C (1928), 205-21. The first elementary proof was given by Jean Ville, "Sur la théorie générale des jeux où intervient l'habileté des joueurs," in Émile Borel et al., *Traité du calcul des probabilités et de ses applications* (1938), Vol. IV, pp. 105-13. The simplest proof of (4) has been given in a slightly more general form by Lynn H. Loomis, "On a Theorem of von Neumann," *Proceedings of the National Academy of Sciences*, XXXII, No. 8 (1946), 213-15.

The case where the number of strategies may be infinite is considerably more difficult mathematically, and the generalization of (4) does not always hold. See Loomis, *op. cit.*; Wald, "Generalization of a Theorem by von Neumann Concerning Zero-Sum Two-Person Games," *Annals of Mathematics*, XLVI (1945), 281-86; Wald, "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, XV (1947), 279-313, esp. 285-96; and Wald, "Statistical Decision Functions," *Annals of Mathematical Statistics*, XX (1949), 165-205, esp. 166-71. A recent volume, edited by H. W. Kuhn and A. W. Tucker, *Contributions to the Theory of Games* (*Annals of Mathematics*, Study No. 24 [1950]), contains some papers on this and other developments in the theory of zero-sum two-person games.

<sup>88</sup> See the detailed discussion needed for even a highly simplified form of poker in von Neumann and Morgenstern, *op. cit.*, chap. iv, sec. 19.

for a determinate solution. This suggests that mixed strategies may be regarded essentially as a device for withholding information from the opponent, for bluffing.

#### 4. The zero-sum *n*-person game<sup>89</sup>

Consider first the following simple game: There are three players; each is to write down the name of one of the others. If two players write down each other's name, they each get  $\frac{1}{2}$ , and the third player pays 1. If no two players write down each other's name, then no one gets anything. The rational method of playing seems fairly obvious; two of the players will agree in advance to write down each other's name and then collect. In what sense is this really a determinate solution? In particular, it is to be noted that the solution is not unique; there are three possible pairs of players who can form coalitions. Von Neumann and Morgenstern argue as follows: Consider all possible distributions of payments to the players. These will include transfers not provided for by the rules of the game, since we permit one player to bribe another into following an acceptable course of action. Since the game is zero-sum, the sum of the payments must be zero; since each player can get  $-1$  without any coalition, no distribution can be enforced which gives any player less than  $-1$ . Let  $(a_1, a_2, a_3)$  denote a distribution which gives  $a_1$  to the first player,  $a_2$  to the second player, and  $a_3$  to the third player. Then, in symbols,

$$a_1 + a_2 + a_3 = 0, \quad (5)$$

$$a_i \geq -1 \text{ for each } i. \quad (6)$$

Any set of numbers satisfying (5) and (6) will be termed *imputation*. The particular imputations which seem intuitively to be the rational ones are, then,  $(\frac{1}{2}, \frac{1}{2}, -1)$ ,  $(\frac{1}{2}, -1, \frac{1}{2})$ , and  $(-1, \frac{1}{2}, \frac{1}{2})$ . Let  $V$  be the set con-

<sup>89</sup> *Ibid.*, chaps. v and vi.

sisting of these imputations. Then the essential properties of  $V$  are (1) for every imputation not in  $V$ , there are two players who would both prefer a particular imputation in  $V$  to the one not in  $V$ ; (2) for any two imputations in  $V$ , there are not two players who would prefer one to the other. These are stability properties; no imputation not in  $V$  can be maintained, while there is no drive to change from one imputation in  $V$  to another, since in our simple game the behavior of pairs of players is decisive.

The decisive character of a pair of players who choose to agree consists in the following: any pair of players can, by agreement, get the amount 1 between them, while no imputation can, from (5) and (6), give them more than 1. Therefore, they are in a position to choose any imputation whatever, since no imputation gives them more than they can get by the rules of the game. These concepts all lead to natural generalization. Consider the following zero-sum  $n$ -person game: Let  $S$  be any set of players; let the rules specify a *characteristic function*  $v(S)$  which states the total amount any set  $S$  will get if they form a coalition. Because of the zero-sum character of the game,  $v(-S) = -v(S)$ , where  $-S$  is the set of all players not in  $S$ . In particular, let  $(i)$  be the set consisting of the player  $i$  alone, so that  $v([i])$  is the amount player  $i$  will get (presumably negative) if he does not enter into a coalition with anyone. Then an imputation is a set of payments  $(a_1, \dots, a_n)$  to the respective players having the properties,

$$a_1 + \dots + a_n = 0, \quad (7)$$

$$a_i \geq v([i]) \text{ for each } i. \quad (8)$$

A set of players  $S$  is said to be *effective* for an imputation  $(a_1, \dots, a_n)$  if  $\sum_{i \in S} a_i \leq v(S)$ , where " $i \in S$ " means " $i$  belongs to  $S$ ," i.e., if the imputation does not give the players in  $S$  more than they

could get by forming a coalition. The imputation  $(a_1, \dots, a_n)$  is said to *dominate* the imputation  $(b_1, \dots, b_n)$  if there is a set  $S$  which is effective for  $(a_1, \dots, a_n)$  and such that  $a_i > b_i$  for every player  $i$  in  $S$ ; under these circumstances we would certainly expect the players in  $S$ , by agreement, to change the imputation from  $(b_1, \dots, b_n)$  to  $(a_1, \dots, a_n)$ . A set of imputations  $V$  is said to be a *solution* if (1) every imputation not in  $V$  dominates another imputation in  $V$ ; (2) no imputation in  $V$  dominates another imputation in  $V$ . A solution thus defines the outcome of the above game when the players behave rationally.

It might seem that the above discussion relates only to a very special type of game, in which coalitions enter through the explicit statement of the rules. It is argued, however, that every  $n$ -person zero-sum game can be expressed in this manner. Suppose a coalition  $S$  forms, which seeks to maximize the sum of the pay-offs to its members. The coalition will naturally suppose that the players not in  $S$  will form a countercoalition; in view of the zero-sum character of the game, the coalition  $-S$  will seek to minimize the sum of the pay-offs to members of  $S$ . This situation is precisely a zero-sum two-person game, which has a determinate value, as has been seen. Therefore, we may define  $v(S)$  to be the value to the coalition  $S$  of the zero-sum two-person game just described. This can be done for each possible set of players  $S$ , so that the game is reduced to the form discussed in the previous paragraph.

Games are classified as *essential* or *inessential*. An inessential game is one in which a coalition can get no more than the sum of the amounts that each of its members could get by playing separately, i.e.,  $v(S) = \sum_{i \in S} v([i])$ . In

such a game, there is no incentive to form coalitions and hence no specifi-

cally social element. In an essential game, on the other hand, such as the three-person game which introduced this discussion, there is a positive incentive to form coalitions. We have already observed that a solution may contain more than one imputation. This is a general phenomenon: it can be shown that in an essential game, a solution must contain at least two imputations. There seems to be some sort of fundamental indeterminacy in the purely rational theory of social behavior; the true significance of this indeterminacy must be regarded as obscure.

The definitions as framed do not exclude the possibility that there is more than one solution to a given game, and indeed this can happen, although there are also games with only one solution. In the three-person game just considered, let  $c$  be any number less than  $\frac{1}{2}$ . Consider the class of all imputations in which  $a_3 = c$ . It can easily be verified that this class is in fact a solution, within the meaning of the previous definitions. For each value of  $c$ , there is a corresponding solution, so that the number of solutions is infinite. Unlike the first solution, these solutions do not have any symmetry. They may be termed *discriminatory* solutions; the ruling coalition (1, 2) agrees to give the excluded player a fixed amount  $c$ , and then divide up the remainder between them. It is remarkable that discriminatory practices turn out to have certain stability properties even when the basic rules of the game are completely symmetric as between players. Even within the first solution found, there was a certain element of unfairness; each given imputation within the solution was unfair to one player, but there was an element of symmetry in the solution as a whole. So it can be seen that even a symmetrical game is not necessarily fair, in that it does not guarantee that each individual will be as well off as every other. Much more

than equality of opportunity is needed to insure equality of outcome.

There is thus a hierarchy of successively more restrictive conditions on the state of society: the rules of the game, which define the class of possible imputations, the solution, and the imputation which actually prevails at any given instant of time. It is suggested that the first corresponds to the basic limitations of the society—physical, biological, and possibly psychological; the second to a socially stable standard of behavior, which, however, does not uniquely determine the actual distribution of values among the members of the society. More than one standard of behavior may be compatible with the rules, but only those found by the theory will be stable enough to survive.

##### 5. *The general n-person game*<sup>40</sup>

The games thus far considered have lacked generality in one important respect. Society as a whole could not gain because of the zero-sum restriction. In the study of many social conflicts, this may not be an important limitation, but it does serve to eliminate all economic aspects. The crucial feature of economic behavior is precisely the possibility that by co-operation all individuals can be made better off.

The general  $n$ -person game is handled by a reduction to a zero-sum  $(n + 1)$ -person game. Introduce a fictitious player,  $n + 1$ , who can make no moves but who receives the negative of the sum of the amounts paid to the other players. This enlarged game can be described in terms of the theory in the preceding section;  $v(S)$  again represents the amount that each coalition can guarantee for itself. The concept of domination goes through as before. However, not all the solutions of the extended game can be accepted. We may take it for granted that the genuine players will get as much for themselves

<sup>40</sup> See von Neumann and Morgenstern, *op. cit.*, chap. xi.

collectively as they can, since this increases the amount of wealth to be divided among them. In effect, the genuine players will discriminate against the fictitious player. The only solutions that are therefore admissible are those in which every imputation gives the fictitious player the minimum possible.

In the case of a general two-person game, it turns out that there is just one solution, namely, the set of all imputations which maximize the sum of the returns to the two players. Suppose *A* and *B* are bargaining over a house which *A* owns but *B* values more highly. Then the house will be transferred to *B*, who will pay a price at least equal to the value *A* places on it but not more than the value of *B*. This conclusion extends readily to all cases of bilateral monopoly, even when there is not a unique indivisible object but a continuously divisible commodity: the amount of goods which changes hands is determinate, being that amount which maximizes the sum of the profits of the two bargainers, but the buyer will pay the seller at least enough so that he will not be worse off than if he sold nothing and not so much that the buyer will be worse off than if he bought nothing.<sup>41</sup>

The discussion of the two-person game yields a purely common-sense answer. Much less obvious results follow in the general three-person game, e.g., one seller and two buyers. In the case of a single indivisible object, again, it will be transferred to the individual who values it most highly, but the possible accompanying patterns of payments may be very complicated indeed. The game theory fully takes into ac-

count such possibilities as having one buyer bribe another to stay out of the market. The determination of the solutions to this game cannot be presented here in detail.<sup>42</sup>

#### 6. Statistical inference as a game

In the discussion of statistical inference as a problem in rational behavior (in the second part of Section II, above), it has been shown that the statistician must choose a function  $i(k)$  which determines the action  $i$  he will take for any possible sample  $k$ . Suppose he chooses such a function, while the true state of nature is  $j$ . If the sample  $k$  should come up, he will take action  $i(k)$  and therefore lose  $r_{i(k),j}$ . But sample  $k$  will come up with probability  $p_{kj}$ . Therefore, if the statistician chooses a function  $i(k)$ , his expected loss when the true state is  $j$  is given by

$$\sum_k r_{i(k),j} p_{kj}, \quad (9)$$

which depends on the statistician's choice of  $i(k)$  and nature's "choice,"  $j$ . Wald has suggested interpreting statistical inference as a zero-sum two-person game, in which nature chooses a state  $j$ , the statistician chooses a decision function  $i(k)$ , and the pay-off to nature is given by (9). The optimal behavior of the statistician is to choose that decision function (possibly mixed) which minimizes the maximum expected loss.

### IV. SOME OTHER MATHEMATICAL MODELS OF SOCIAL BEHAVIOR

#### 1. Rashevsky's theory of human relations<sup>43</sup>

Nicholas Rashevsky of the University of Chicago, who is known primarily for his work in the application of mathematical methods to biology, has

<sup>41</sup> The earlier discussions of this problem in economic literature had given smaller ranges of indeterminacy; the contribution of game theory is to show that the earlier answers were artificial from the viewpoint of rational behavior (though not necessarily from that of actual behavior). However, Gerhard Tintner has already given the solution indicated above. See his "Note on the Problem of Bilateral Monopoly," *Journal of Political Economy*, XLVII (1939), 263-70.

<sup>42</sup> See von Neumann and Morgenstern, *op. cit.*, chap. xi, sec. 62; an excellent elementary discussion is contained in Jacob Marschak, "Neumann's and Morgenstern's New Approach to Static Economics," *Journal of Political Economy*, LIV (1946), 97-115.

<sup>43</sup> Nicholas Rashevsky, *Mathematical Theory of Human Relations* (1947).

suggested that analogous methods could be applied to the study of human relations. His approach is of a generally *ad hoc* character, partly collective, partly individualistic. The bulk of his work is devoted to the interaction of social classes. A great variety of different hypotheses as to the nature of this interaction are proposed and formulated mathematically, and some consequences drawn. An example of his analysis will suffice to give the general flavor. It is assumed that there are two classes, the active and the passive. Active individuals do not imitate anyone, while passive individuals are influenced by active individuals in proportion to the number with whom they come into contact and also, to a lesser extent, by other passive individuals. Suppose that there are two possible activities, *A* and *B*, that  $x_0$  and  $y_0$  of the active individuals engage in *A* and *B* respectively, and that at any instant of time  $x$  and  $y$  passive individuals engage in *A* and *B* respectively. The number of passive individuals engaged in *A* will tend to be increased by the number of active individuals engaged in *A* and, to a lesser extent, by the number of passive individuals engaged in *A*, and to be decreased by the number of active and passive individuals engaged in *B*. Symbolically,

$$dx/dt = a_0x_0 + ax - c_0y_0 - ay, \quad (10)$$

with a similar equation for  $dy/dt$ . These equations can be solved to reveal the variation in  $x$  and  $y$  over time as functions of the coefficients and to find the points of stable equilibrium to which the system will tend.<sup>44</sup>

It is to be noted that the situation in question might be interpreted as a conflict between the active individuals engaged in *A* and those engaged in *B* for control over the passive individuals. It could easily therefore be brought within the general framework of game theory.

<sup>44</sup> Rashevsky, *op. cit.*, chap. iii.

It is noteworthy that Professor Rashevsky's model is broadly similar to models of conflict in other spheres, such as Richardson's work on international relations—compare (10) with (2) above—and Lanchester's theoretical treatment of military strategy.<sup>45</sup> There seem to be emerging two general analyses of social conflict situations employing mathematical methods: rationalistic game theory, and *ad hoc* dynamic analysis by means of systems of linear differential equations. There may be some complementarity in the two approaches; in particular, note that game theory, at least in its present state, deals only with equilibrium situations. If additional assumptions could be made as to how each player learns about the others from experience, the theory would be a dynamic one.

Rashevsky presents the above case as highly oversimplified and suggests more complicated variations of the same scheme. Some of these involve economic considerations in that the interaction of classes may be productive of wealth.<sup>46</sup> In the economic analyses, the *ad hoc* laws which relate, e.g., to certain aspects of the distribution of income, are not sufficient to define the situation uniquely; within the framework of these postulates, the individuals are assumed to act rationally in some appropriate sense, i.e., to maximize satisfaction or profits. One simple case is that of two classes of individuals, one of which is composed of organizers under whose supervision the members of the second class produce more than they would otherwise. It is assumed that the members of the first class decide on a fixed fraction of the total output to be divided among the members of the second class in proportion to the amount of work each does, while each

<sup>45</sup> Lewis F. Richardson, *Generalized Foreign Politics* (*British Journal of Psychology*, Monograph Supplement No. 23 [1939]); and F. W. Lanchester, *Aircraft in Warfare* (1916), chaps. v and vi.

<sup>46</sup> Rashevsky, *op. cit.*, cbaps. v, vi, xix, xx.

member of the second class decides how much work to do on the basis of the given fraction. The first class then chooses the fraction in question so as to maximize their total profits. There are several *ad hoc* elements here; the fraction to be given to the second class need not be independent of the amount of work they do, it need not be divided among the members of that class in proportion to the amount of work each does, and each member of the first class may try to maximize his own profits, possibly by special bargains with some members of the second class, rather than maximize the total class profit. Some of these assumptions are similar to the assumption of perfect competition in classical economics.

The same techniques of analysis are used throughout Rashevsky's work, though there are many variant sets of assumptions made. Applications are made to the urban-rural distribution of population, the distribution of city sizes, international relations, historical change, and war. Rashevsky emphasizes the tentative and simplified character of his approach and suggests that its chief function is to make available a number of alternative models which might be useful in further theoretical and empirical work.

It might be remarked that the standards of mathematical rigor are high. The methods used are drawn from the calculus and the theory of ordinary linear differential equations, with a few tentative steps toward the use of integral equations.

## 2. Zipf's principle of least effort and Stewart's social physics

The work of George K. Zipf<sup>47</sup> is an extraordinarily comprehensive effort to subsume the major part of human behavior, both individual and collective, under a single principle of least effort:

<sup>47</sup> George K. Zipf, *Human Behavior and the Principle of Least Effort* (1949).

individuals and societies act so as to minimize the expected average rate of work. We shall here be concerned only with his discussion of social behavior in Part Two of his book. In terms of the classification in Section II of this chapter, the model is certainly rational; since it seems to be assumed that both individuals and society seek to minimize effort, the model has both collective and individualist elements.

It must be stated, however, that Dr. Zipf's work does not constitute a properly developed mathematical model. The fundamental postulates are nowhere stated explicitly; though mathematical symbols and formulas are sprinkled rather freely through a long work, the derivations involved are chiefly figures of speech and analogies, rather than true mathematical deductions; in some cases, they are simply wrong. Thus, as an attempt at a systematic social theory, Zipf's work can only be regarded as a failure.

However, two empirical regularities do emerge which are highly suggestive and may prove promising for further research. It is held, with considerable empirical support, that if the cities of a nation are ranked in order of population, then the rank of a city is inversely proportional to its population; e.g., the largest city, with rank 1, has twice the population of the next largest, which has rank 2. This relation, it is suggested, is an equilibrium condition in that a nation in which it does not hold will either expand or break up until the relation does obtain. Some, though inconclusive, empirical evidence is adduced for this last conclusion, which is used to generalize about international relations. Similar relations hold for other types of frequency data, such as the number of stores of each kind.

An attempt is made to analyze a hierarchy of social classes with the aid of analogous rank-frequency relations, it being assumed that each class ex-

exploits the one beneath it and is exploited by the one above it. The strength of a class, and therefore its potential for rebellion, is measured by its total income, while the incentive to remain *in statu quo* is proportional to the income of an individual. To have equilibrium, the two magnitudes must be brought into a suitable relation. Dr. Zipf suggests that, if the classes are ranked upward from the bottom, the income of each individual should be proportional to his class rank while the number of individuals in any one class should be inversely proportional to the square of its rank. From this analysis, certain implications are drawn for the distribution of income. The postulates of this theory of social classes are again never clearly stated, beyond some generalizations that each individual would prefer exploiting others to working as a source of income. It would be interesting to define the problem more precisely and then compare the game-theoretical analysis with the suggestions of Dr. Zipf.

Another empirical regularity found is that the interaction between two cities is inversely proportional to the distance between them. This applies particularly to the flow of traffic or information between them. The study of this relation had been begun earlier by John Q. Stewart of Princeton University, an astronomer.<sup>48</sup> Stewart has stressed the formal analogies of this relation to the law of gravitation. Let  $P_i$  be the population of a place  $i$ , and  $D_i$  be the distance of place  $i$  from a given place  $A$ ; then the *demographic potential* at  $A$  is defined to be  $\sum (P_i/D_i)$ , the sum being taken over all populated places  $i$ . Under the above law, the demographic potential should represent the total amount of transactions per unit population at  $A$  and therefore should correlate with other economic magnitudes. Some evi-

<sup>48</sup> John Q. Stewart, "Demographic Gravitation: Evidence and Applications," *Sociometry*, XI (1948), 31-58, and earlier work cited therein.

dence has been found to support this assertion, but it can hardly be described as proved.<sup>49</sup>

#### V. THE USE OF THEORETICAL MODELS IN INDUCTIVE INFERENCE

The familiar proposition in the methodology of science that empirical study without previous theoretical development is fruitless has recently been developed in detail with particular reference to the field of economic analysis by a group of individuals associated with the Cowles Commission for Research in Economics, at the University of Chicago. The general viewpoint was expressed at length by Trygve Haavelmo.<sup>50</sup> The associated problems in mathematical statistics have been studied by Henry B. Mann and A. Wald;<sup>51</sup> Tjalling C. Koopmans, Herman Rubin, and Roy B. Leipnik;<sup>52</sup> and Theodore W. Anderson and H. Rubin.<sup>53</sup> The most elementary exposition is that of J. Marschak;<sup>54</sup> slightly more technical are

<sup>49</sup> There must be mentioned, if only in a footnote, certain other recent works relevant to the mathematical formulation of social theory. Stuart C. Dodd and Eliot D. Chapple have devoted much attention to the proper symbolization of social problems as a basis for the construction of models and for empirical research. See Dodd's *Dimensions of Society* (1947) and "A Systematics for Sociometry and for All Science," *Sociometry*, XI (1948), 113-30; and Chapple's *Measuring Human Relations: An Introduction to the Study of the Interaction of Individuals* (*Genetic Psychology Monographs*, Vol. XXII, No. 1 [1940]). Their work has been principally methodological rather than substantive. For other work of this type, see the journal *Sociometry*.

Norbert Wiener, in his *Cybernetics* (1948), has initiated the systematic study of communication as a feature common to physiology, modern servomechanisms, and human society; however, he has made little specific application to social study and, indeed, has expressed pessimism over the possibilities of any sort of social science.

<sup>50</sup> "The Probability Approach in Econometrics," *Econometrica*, XII (1944), Supplement.

<sup>51</sup> "On the Statistical Treatment of Linear Stochastic Difference Equations," *Econometrica*, XI (1943), 220.

<sup>52</sup> "Measuring the Equation Systems of Dynamic Economics," in Tjalling C. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models* (1950).

<sup>53</sup> "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, XX (1947), 46-63.

<sup>54</sup> "Economic Structure, Path, Policy, and Prediction," *American Economic Review*, XXXI, No. 2 (1947), 81-84.

those of Koopmans,<sup>55</sup> Marschak,<sup>56</sup> and Gershon Cooper.<sup>57</sup> Though the work done is strongly influenced by the special problems of economics, there is unquestionably a great portion applicable to the social sciences in general.

### 1. *Exact versus stochastic relations*

In economics and in other social sciences, we may certainly expect that no exact relation will hold between the variables we measure for at least two reasons: (1) Not all the variables which are relevant are included in the analysis, and we always omit a host of unimportant factors which are too difficult to measure. (2) The variables we do observe are not measured precisely. In the statement of a relation, then, we must include not only the explicitly enumerated variables but an additional unmeasurable variable, known as a *disturbance*. Such a relation is said to be *stochastic*. Thus, we may say that, for an individual, consumption ( $c$ ) is related to income ( $y$ ) by a relation of the type  $c = f(y) + u$ , where  $u$  is a disturbance which stands for all the omitted variables which influence consumption. Merely to say that a relation holds with a disturbance in it is a tautology. For any function  $f(y)$ , we can make the above relation true by defining  $u$  to be  $c - f(y)$ . To give a relation empirical content, we must assert some regularities in the behavior of the disturbances.

We say that the disturbance in a given relation is a random variable with a probability distribution which is the same for each time the variables are observed. It is usually also assumed that the disturbances at different times are independent. Thus, in

<sup>55</sup> "Statistical Estimation of Simultaneous Economic Relations," *Journal of the American Statistical Association*, XL (1945), 448-66.

<sup>56</sup> "Statistical Inference in Economics: An Introduction," in Koopmans (ed.), *Statistical Inference in Dynamic Economic Models* ( . . .

<sup>57</sup> "The Role of Economic Theory in Econometric Models," *Journal of Farm Economics*, Vol. XXX (1948).

our consumption example, if the subscript  $t$  denotes time, the disturbances  $u_t = c_t - f(y_t)$  for different values of  $t$  can be regarded as forming a random sample from a fixed probability distribution. These concepts extend themselves naturally to all types of social laws. In general, then, we may say that the formulation of a generalization in social science is equivalent to an assertion about the probability distributions of certain disturbances.<sup>58</sup>

### 2. *Simultaneous relations and the concept of a structure*

If we take the individual viewpoint as developed in the first part of Section II of this chapter, we find that the actions of all individuals are simultaneously determined by all the equations of form (1). There is no unique line of causality; the actions of all individuals enter symmetrically. The situation is not altered if we assume that the equations (1) are stochastic. The variables  $P_i$ , on the other hand, play a different role. If they refer to exogenous factors, they will in general be determined by processes which are independent of the social context in which equations (1) are stochastic.

With respect to actions  $A_1, \dots, A_n$ , the exogenous factors have a strictly unidirectional causal influence. The same is true, with some qualifications, if the variables  $P_1, \dots, P_n$  contain past actions of individuals as well as truly exogenous factors, for at any given instant of time, the past has also a unidirectional causal influence on the present choice of actions. The causal variables  $P_i$  are referred to as *predetermined*; the mutually interdependent variables  $A_i$  are referred to as *endogenous* or *jointly dependent*. At any given instant of time, therefore, the endogenous variables are simultaneously dependent

<sup>58</sup> For a more extended discussion of the ideas of this paragraph, see Haavelmo, "The Probability Approach in Econometrics," *Econometrica*, Vol. XII (1944). Supplement, chap. iii

terminated by a set of stochastic equations,

$$A_i = f_i(A_1, \dots, A_{i-1}, A_{i+1}, \dots, A_n, P_i, u_i). \quad (11)$$

Here again, equations (11) are tautologous unless the distribution of the disturbances  $u_i$  is specified. The equations and the distribution together are known as the *structure* of the social system. The predetermined character of the  $P_i$ 's is expressed in the assertion that they are statistically independent of the  $u_i$ 's.

Suppose the system (11) to be solved for the endogenous variables:

$$A_i = g_i(P_1, \dots, P_n, u_1, \dots, u_n). \quad (12)$$

Equations (12) are known as the *reduced form* of the structure. The reduced form is itself a structure, and it might seem at first that a knowledge of it conveys as much information about the endogenous variables as does (11). This is not so if we realize that one or more of equations (11) may change; for example, one of the equations may refer to a government policy, or it may refer to an industry whose behavior will be affected by technological change. If one of equations (11) changes, then all of equations (12) change, but there is no necessity that the other equations (11) will be altered. The equations (11) have a higher degree of *autonomy*<sup>59</sup> than (12), in the sense that each equation of (11) is invariant under a wider class of structural changes than the equations of (12).

If the laws under study were truly immutable, autonomy would be an irrelevant concept, and the reduced form of a structure would be as useful a description of reality as any other form. The need for an autonomous structure is especially critical when the results of social analysis are to be used for policy purposes. Here, the question is that of choosing among several possible forms for the equation expressing government

behavior. What we wish to do is to consider the structure for each possible form of the government equation and predict the behavior of the endogenous variables under each; that form is then chosen which yields the most satisfactory behavior. This procedure can be carried out, however, only if the equations other than the government equation are invariant with respect to changes in it. For prediction under changed structure, and in particular for policy decisions, it is therefore important that the structure be expressed in as autonomous a form as possible. The stress on models based on individual behavior arises from the argument that they are more apt to be autonomous than collective models.

### 3. *Models, statistical inference, and the problem of identification*

Suppose now that we have a sample of observations on a system of endogenous and predetermined variables and we are faced with the problem of inferring the structure which generated them. This is a problem in statistical inference. Here, the states of nature are the different possible structures. As in the general formulation of the problem given earlier (see the second part of Sec. II and the sixth part of Sec. III, above), we assume it known that the possible structures belong to a certain class. That is, certain statements can be made about the structure on an a priori basis, so that we need not consider structures not compatible with those assertions. The class of admissible structures is known as the *model*. The determination of the optimum statistical methods for selecting one structure out of the model is then a technical problem which has been solved, at least in an approximate sense.<sup>60</sup>

<sup>59</sup> The fundamental statistical papers are Koopmans, Rubin, and Leipnik, "Measuring the Equation Systems of Dynamic Economics," in Koopmans (ed.), *Statistical Inference in Dynamic Economic Models* (1950); and Anderson and Rubin, "Estimation of the Parameters of a Single

<sup>60</sup> This term is Ragnar Frisch's.

There is a certain difficulty in the statistical analysis which does not seem to arise in the natural sciences. Take first a very simple nonstochastic problem. Assume that prices are determined by the equating of supply and demand, where each is a function of price. If, during the period of observation, neither supply nor demand shifted, the price and the quantity exchanged would not alter, and we would have only one point from which to infer the supply and demand curves. This is obviously impossible, since there are an infinite number of possible pairs of curves passing through the observed point. This same problem arises in stochastic structures; consider again the supply-and-demand problem:

$$q = a_1 p + a_2 + u \text{ (demand), (13)}$$

$$q = b_1 p + b_2 + v \text{ (supply), (14)}$$

where  $q$  is quantity exchanged,  $p$  is price, and  $u$  and  $v$  are disturbances with a given probability distribution. Equations (13) and (14) can be solved for  $q$  and  $p$  in terms of  $u$  and  $v$ , and we can then find the distribution of  $q$  and  $p$  from that of  $u$  and  $v$ . For example, suppose that  $a_1 = -1$ ,  $a_2 = 2$ ,  $b_1 = 1$ ,  $b_2 = 1$ , and  $u$  and  $v$  are normally and independently distributed with means zero and variances 1. Then it can be shown that  $p$  and  $q$  are normally and independently distributed with means  $1/2$  and  $3/2$  respectively, and variances  $1/2$ . Now suppose that  $a_1 = -1/2$ ,  $a_2 = 7/4$ ,  $b_1 = 2$ ,  $b_2 = 1/2$ , and  $u$  and  $v$  are normally and independently distributed with means zero and variances

Equation in a Concrete System of Stochastic Equations," *Annals of Mathematical Statistics*, XX (1949), 46-63. In practice, chiefly the methods of the latter paper have been used. For more elementary expositions, see Koopmans, "Statistical Estimation of Simultaneous Economic Relations," *Journal of the American Statistical Association*, XL (1945), 448-66; and M. A. Girshick and T. Haavelmo, "Statistical Analysis of the Demand for Food: Examples of Simultaneous Estimation of Structural Equations," *Econometrica*, XV (1947), 79-110. It is important to observe that equations (11) should not be fitted by taking the least squares regression of  $A_1$  on the other variables.

$5/8$  and  $5/2$ , respectively. Then  $q$  and  $p$  have the same distribution as in the previous case. Since observations on a sample of  $p$  and  $q$  values can only yield information as to the distribution of  $p$  and  $q$ , it is clear that no such sample, no matter how large, could enable a choice to be made between the two structures. A model containing both structures is said to be *unidentified*; in general, a model in which no two structures generate the same probability distribution of the endogenous variables is said to be *identified*. In order to make useful statistical inferences, the range of possible structures must be sufficiently restricted by a priori considerations so that the model is identified.

The conditions for identification have been discussed for models in which the structures are systems of linear equations. The concept of identification has also been generalized in various ways.<sup>61</sup>

#### 4. Model-building and scientific tactics

The method of scientific investigation indicated in the preceding paragraphs calls then for intensive a priori thinking to formulate a model, followed by the selection of a best-fitting structure from that model by appropriate statistical techniques. It is the virtue of the Cowles Commission approach to have set forth this process clearly and to have resolved many of the statistical difficulties in the way of its fulfillment. It is clear that the crucial step is the choice of a model. If we can say very little on purely a priori grounds about the nature of the process under investigation, then the resulting model is unidentified, and further progress is stopped.

<sup>61</sup> Koopmans, Rubin, and Leipnik, *op. cit.*; Leonid Hurwicz, "Generalization of the Concept of Identification," in Koopmans (ed.) *Statistical Inference in Dynamic Economic Models* (1950). For an excellent exposition of the identification concept, see Koopmans, "Identification Problems in Econometric Model Construction," *Econometrica*, XVII (1949), 125-44.

An alternative procedure employed by a number of economists<sup>62</sup> is to start with a very wide and vaguely stated model and investigate empirical data which seem to be relevant. By examination of these data, more definite models will be suggested which will, in turn, provide the basis for further empirical research, and so forth. It might be asked how, if the original vague model is really unidentified, this procedure can lead to any results. Two related answers seem to be implied in the discussion: (1) The observations will select out of the original model a collection of structures compatible with the observations; among these, the "simplest" in some sense is selected.<sup>63</sup> (2) There is really an identified model in the minds of the investigators, which, however, has not been expressed formally because of the limitations of mathematics

or of language in general. Perhaps we may interpret the "simple" structures as those found in this unconsciously maintained model. The choice between the alternative scientific tactics indicated depends on the stage of formalization of the underlying theory. No dogmatism is possible; a certain amount of oversimplification is tolerable (and necessary) to gain the advantages of formalization and the use of optimum statistical methods, but there are limits; since the statistical methods are best only on the assumption that the model is correct, a serious error in formulating the model may invalidate all further empirical work based on it.

<sup>62</sup> See, for example, the works named in footnote 11.

<sup>63</sup> See Milton Friedman, "Lange on Price Flexibility and Employment," *American Economic Review*, XXXVI (1946), 613-31. See pages 616-22 on the use of Occam's razor.