

SCOPE AND METHOD OF ECONOMETRICS
ILLUSTRATED BY APPLICATIONS TO AMERICAN
AGRICULTURE.¹

By GERHARD TINTNER.

Department of Applied Economics, Cambridge University.

(Read before the Society on Monday, 21st March, 1949).

(a) *Economics and Econometrics*

Econometrics is the application of a specific method in the general field of economic science in an effort to achieve numerical results and to verify economic theorems² It consists in the application of mathematical economic theory and statistical procedures to economic data in order to establish numerical results in the field of economics.

Econometrics has to be distinguished from mathematical economics and from statistical economics. It is, however, closely related to both, and utilizes results achieved in these fields

Mathematical economics formulates economic theory in mathematical terms and uses the methods of mathematics to derive economic relationships from certain basic assumptions or axioms, e g, from certain ideas about maximization³ In order to do this we have to construct economic models which involve structural economic relationships Such relationships are derived from economic theory

We may, for instance, construct the following model A consumer maximizes his satisfaction, under the condition that his income and the prices of all goods and services are given independently of his actions Then we can derive from this model theoretically the demand functions for all commodities and services. These demand functions have certain properties, e g, concerning the price and income elasticities of demand With the help of econometric methods we may try to establish numerical

¹A paper given before the Statistical and Social Inquiry Society of Ireland

²J Schumpeter, "The Common Sense of Econometrics," *Econometrica* vol 1, 1933, pp 1 ff

R Frisch, "The Responsibility of the Econometrician," *ibid*, vol 14, 1946, pp 1 ff

J Tinbergen, *Econometrie*, Groningen, 1941

H T Davis, "Theory of Econometrics," *Bloomington Ind* 1941

W Leontieff, "Econometrics," in *A Survey of Contemporary Economics*, Philadelphia, 1948

³F Kaufmann, *Methodology of the Social Sciences*, London, 1944, pp. 141 ff

values for these elasticities and also test hypotheses regarding them, *e.g.*, relating to the relative magnitude of the price and the income elasticities of demand. This, is of course, only possible if suitable data are available, *e.g.*, from family budgets.

The above model is static, since it does not involve time explicitly. If the results of an econometric investigation should not be satisfactory, we may try a dynamic model. In such a model satisfaction may, for instance, depend upon the present income and upon past income, etc. By comparing the residuals from a fit of the dynamic and the static model we may form an opinion about the relative merits of the two models.

In a similar way we may, for instance, investigate the actions of a firm under free competition, monopoly, or other forms of market organisation. We assume that the firm maximizes its profit.

A more complicated system which deals with the interdependence of economic units and uses the theory of games of strategy has been developed by von Neumann and Morgenstern⁴.

Such models are micro-economic, *ie*, they deal with the economic behaviour of individual units, *e.g.*, consumers or firms. Macro-economic models are concerned with the behaviour of the total economy. They are probably more useful in practice than the other kind. But the construction of macro-economic models involves the problem of aggregation or the general index number problem, which still presents great theoretical and practical difficulties.

Economic theory is not merely descriptive but aims to establish a system of economic laws. These laws may be of a statistical nature. They are supposed to describe regularities of economic behaviour. They ought to enable us to make valid predictions about empirically observable economic actions.

There is in my opinion no fundamental difference between mathematical economics and economic theory which utilizes non-mathematical methods. Many economic theorems have first been formulated in a literary way, and later restated in mathematical terms. The best example of such a procedure is perhaps the Keynesian⁵ theory, which was first stated by Keynes in a non-mathematical way. Later it was reformulated by many mathematical economists in terms of mathematical equations⁶. This reformulation has certain advantages compared with the original theory. It has brought out the basic hypotheses and has enabled us to see more clearly the difference between Keynesian and what has come to be called "classical" economics.

The relation between logic and mathematics is much disputed. But I for one believe that Bertrand Russell⁷ has shown that mathematics

⁴J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour*, Princeton, 1944.

⁵J. M. Keynes, *The General Theory of Employment, Interest and Money*, London, 1936.

⁶See *e.g.*, J. R. Hicks, "Mr Keynes and the Classics," *Econometrica*, vol 5, 1937, pp 147 ff. O. Lange, "The Rate of Interest and the Optimum Propensity to Consume," *Economica*, vol 5, new series, 1938, pp 12 ff. F. Modigliani, "Liquidity Preference and the Theory of Money," *Econometrica* vol 12, 1944, pp 45 ff. L. R. Klein, "The Keynesian Revolution," New York, 1947, technical appendix, pp 189 ff.

⁷B. Russell, *Introduction to Mathematical Philosophy*, New York, 1919.

may be considered as a branch of logic. Nobody who is familiar with the vast literature of logistics can deny that it is possible to state logical relationships symbolically, *i.e.* in mathematical terms. On the other hand, I believe that mathematics can be stated and formulated in purely logical terms, *i.e.*, without the use of symbols. Mathematics may be based upon logic, though logic of a non-Aristotelian kind. Mathematics and logic are very closely related, if not identical.

All economic laws are conditional. The propositions of welfare economics⁸ deal with the way in which ideal use of resources can be obtained, given a certain social objective. This objective is normative, and welfare propositions are conditioned upon it.

It may be said, for instance, that the lowering of tariffs will under certain conditions increase the national product. National product must, of course, be suitably defined. This proposition may or may not be true, but is certainly verifiable. It can be tested for instance by studying cases in which tariffs have actually been lowered. Econometric methods may be quite useful in performing such a test.

But economics does not tell us anything about the question *whether* tariffs should be lowered. This will depend upon the social goals pursued in policy. These goals will in most cases be multiple and not unique. For instance, the lowering of tariffs may be considered desirable because, on the whole, it makes people better off. Against this has to be balanced a potential loss of self-sufficiency, which may be serious for the national existence of a country in times of war. Which policy is actually chosen will not depend upon economic considerations. These can only supply the (true or false) statement: if tariffs are lowered, the national product will increase. Econometrics may even be able to tell us by how much it is likely to increase. But the choice of policy is a matter of politics, ethics and similar considerations.

But why is mathematical economics not enough? The work of Marshall, Walras, Pareto and their modern followers is certainly an imposing intellectual achievement. But, as Marshall himself has pointed out,⁹ something has still to be added. We desire numerical results in economics in order to be able to make quantitative predictions. This can be done with the help of statistics. Statistics is necessary if we want to proceed from the abstract formulations of mathematical economics to numerical results. These results may enable us to verify the economic theorems involved.

It is one thing to develop the concept of elasticity of demand, as it has been done by Cournot, Marshall and their followers. But it is much more interesting to try to evaluate the price and the income elasticity of demand for a given commodity by econometric methods, and to try to form an idea about the reliability of the results. We may also test various hypotheses about the relative magnitude of the elasticities, etc.

The importance of these and other results of econometric research

⁸M. Reder, *Studies in the Theory of Welfare Economics*, New York, 1947, p. A. Samuelson, *Foundations of Economic Analysis*, Cambridge Mass. 1947, pp. 203 ff.

⁹A. Marshall "On the Graphic Method of Statistics," Jubilee Volume, *Royal Statistical Society*, London, 1885, p. 260.

is obvious We shall try to illustrate this point in the last section of this essay

Econometrics has also to be distinguished from statistical economics. Statistical economics frequently declines the use of economic theory and presents statistical summary of the economic data themselves. This point of view underlies for instance much of the work of the National Bureau of Economic Research¹⁰ This has recently been called rather aptly "Measurement without theory"¹¹ It may be doubted if such a procedure is fruitful and even possible Some kind of fundamental conceptions underlie even the work of the most institutionalist-minded statistical economist The selection of the data, their organisation, etc, imply already some kind of underlying general theoretical conception, even if it is not specifically stated

Econometrics represents an intermediary position between the extreme non-theoretical empiricism of the statistical economists and the non-empirical theoretizing of some "pure" economists¹²

The aversion of many statistical economists to the use of theory and econometric methods is based upon the idea that there are no laws in the social sciences, or that these laws are very ephemeral and unstable¹³ There is a grain of truth in these objections and the econometricists should be careful not to overstate their case

The same objection against the existence of laws can also be made in the natural sciences Certainly, the case against the existence of stable laws is a better one in the social than in the natural sciences But ultimately only the existence of a large body of valid numerical economic relationships, which is the goal of econometrics, can disprove the contentions of the radical institutionalists and followers of the historical school

It is not impossible that the case for econometrics has sometimes been overstated by enthusiastic econometricians Econometrics is a useful method of economic research, but certainly not the only one suitable for the verification of economic theorems It can for instance throw very little light on the problems of economic development, e.g., the origin and evolution of the capitalist system Here historical research is much more fruitful, if only because the data are too scarce to allow us the successful application of econometric methods The study of economic institutions, especially of the legal framework of economic activity, is also a very useful procedure in economic research Econometrics cannot claim a monopoly as a method of economic research.

Economics is a social science¹⁴ It deals with a special aspect of society; e.g., the administration of scarce resources in order to satisfy

¹⁰See e.g., A. F. Burns and W. C. Mitchell, *Measuring Business Cycles* New York, 1946

¹¹T. C. Koopmans, "Measurement without Theory," *Review of Economic Statistics*, vol. 29, 1947, pp. 161 ff

¹²L. von Mises, *Grundprobleme der Nationalökonomie*, Jena 1933, see also L. Robbins, *An Essay on the Nature and Significance of Economic Science*, London, 1932

¹³L. Robbins, *op cit*

¹⁴O. Lange, "The Scope and Method of Economics," *Review of Economic Studies*, vol. 13, 1945, pp. 19 ff

wants It is an empirical science like physics and biology and not an *a priori* science like logic or mathematics Economic theorems can, at least in principle, be verified by empirical data ; in effect, econometrics is perhaps the outstanding method for such a verification.

But economics has one characteristic in common with other social sciences like, *e.g.*, psychology. Apart from observation of external economic events like, *e.g.*, the price formation on a market, we have also an additional source of information about economic events this is introspection.¹⁵ Introspection is particularly helpful in connection with problems in the field of consumption This source of information is perhaps not entirely reliable, but should by no means be neglected Results of introspection should be carefully checked and compared with other economic observations.

The pure economist may for instance by introspection find that the more chocolate he consumes the less is the enjoyment of an additional piece of chocolate But before generalizing this into a universal law of diminishing marginal utility (or, in more modern terms, of diminishing rate of substitution) he should carefully check his results with economic observations performed on other subjects than himself Market studies of the demand for certain commodities may be of great help Budget studies will also have some importance Econometric research may eventually give some numerical estimates of the phenomenon in question.

(b) *Econometrics and Statistics.*

Statistics, and especially modern statistical theory, is of paramount importance for econometrics Statistical methods of sampling are already useful for collecting the data which are the raw material of econometric studies.¹⁶ It has the additional advantage of giving us an idea about the reliability of the data But the econometrician is in general not concerned with the collection of data and similar questions, which is the job of the professional statistician It might, however, be advantageous if more econometricians could gain influence upon the way in which statistical data are collected and presented It is particularly unfortunate that so much of our basic data is still the by-product of administrative processes

Modern methods of statistical analysis are indispensable for the econometrician The econometric relationships are always of a statistical nature, as already indicated above

Modern statistical methods have been developed, especially by R. A. Fisher¹⁷ and his school, mostly for practical use in the biological sciences, where they have been so successful The same methodology can also frequently be used in the social sciences, especially in economics. But

¹⁵F Kaufmann, *op cit*, pp 143 ff

¹⁶See *e.g.*, A J King and R J Jessen, "The Master Sample of Agriculture," *Journal of the American Statistical Association*, vol 40, 1945, pp 38 ff

¹⁷R A Fisher *Statistical Methods for Research Workers*, 10th Ed Edinburgh 1946

there is a difference between biological and social sciences. In economics, as in astronomy and meteorology, no experimentation is possible¹⁸. The two examples of astronomy and meteorology indicate that this is not a peculiarity of the social sciences. The example of astronomy, one of the oldest and surely one of the most successful of the natural sciences, also shows that this difficulty should not prevent the emergence of a vast number of empirical laws, which are well confirmed by experience. The blind application of statistical methods which have proved useful in biology and agricultural experimentation is not possible in economics.

This has been recognized by many econometricians for a long time, but formerly they have been using older and not very reliable methods because no better procedures were available. But thanks to the effort of econometricians who saw clearly the fundamental problem, we have now more promising methods especially designed for dealing with non-experimental data¹⁹. The methods developed at the Cowles Commission (University of Chicago) are based on the assumption that there are errors in the equations but none in the variables. It should be emphasised that these methods use the same fundamental ideas of modern statistical theory as biological statistics, *e.g.*, maximum likelihood, tests of hypotheses, fiducial or confidence limits, etc. But they apply these methods no longer blindly and by mere analogy to statistical procedures which are useful and successful in other fields. We can now make use of these methods in such a way that the special difficulties of the lack of experimentation are overcome. One new problem which arises is the problem of identification.

Modern statistics is based upon the idea of probability²⁰. There are a number of conflicting ideas about this concept, which is fundamental for all scientific methodology. Some authors hold that probability statements refer to propositions and are hence logical and not empirical. Two outstanding theoreticians who hold this view are Keynes²¹ and Jeffreys²². Another school thinks that probability refers to the outcome of frequently repeated experiments, as the number of trials increases²³.

Carnap²⁴ has recently shown that both these concepts are legitimate and useful. The econometrician ought to use the first concept when talking about the probability of a theory or hypothesis. *E.g.*, we might discuss the question if, on the basis of empirical evidence, the Keynesian or the "classical" theory is more "probable". A theory of this probability concept, based upon the fundamental ideas of modern mathematical logic, has been developed by Carnap. He calls it the degree

¹⁸T. Haavelmo, "The Probability Approach in Econometrics," *Econometrica*, vol. 12, 1944 supplement.

¹⁹T. Koopmans, "Statistical Estimation of Simultaneous Economic Relations," *Journal of the American Statistical Association*, vol. 40, 1945, pp. 448 ff.

²⁰E. Nagel, "Principles of the Theory of Probability," *International Encyclopedia of Unified Science*, vol. 1, No. 6, Chicago, 1939.

²¹J. M. Keynes, *A Treatise on Probability*, London, 1921.

²²H. Jeffreys, *Theory of Probability*, Cambridge, 1939.

²³R. von Mises, *Probability, Statistics and Truth*, London, 1939.

²⁴R. Carnap, "On Inductive Logic," *Philosophy of Science*, vol. 12, 1945, pp. 72 ff.; "The Two Concepts of Probability," *Philosophy and Phenomenological Research*, vol. 5, 1945, pp. 513 ff.

of confirmation. It must be confessed, however, that these ideas are not yet applicable to any except the simplest problems in the field of statistics. The use of these ideas in econometrics has to await further developments of the theory.

Another and entirely different probability concept refers to the limit of the relative frequency of an event, as the number of trials increases indefinitely. We may for instance throw a coin frequently and note the relative frequency of heads in each series of throws. As we get more and more experiments, the relative frequency of heads among the total number of throws will tend to a limit. This limit will, under certain conditions, be the probability of obtaining a head with a throw of this particular coin. This concept of probability related to relative frequency (defined *e g*, in the manner of von Mises²⁵) has to be distinguished from the first concept of probability.

The econometrician may for instance consider the relative frequency of business failures, *e g*, the percentages of businesses which fail each year. If he takes a larger and larger sample of business enterprises he may talk about the probability of a business failure as the limit of the relative frequency of failures in a given sample, as the sample becomes larger and larger. It is evident that this second concept of probability has to be sharply distinguished from the first.

Since the first probability concept is not yet useful for any except that simplest problems of statistical inference, we will follow statistical practice and use, for the time being, only the second concept. But we should bear in mind that the results reached in this way are not very satisfactory from a philosophical point of view.

The foundations of probability theory are disputed and will remain so for some time to come. There is also not very much agreement about the methods to be used in statistical inference. These methods are of great importance for econometrics.

Statistical inference deals with the way in which conclusions are drawn from the sample about the population. Economic relationships can always be regarded as samples from an unknown population of all possible economic relationships. Econometricians use statistical methods in order to obtain numerical results or estimates. These estimates may consist of one single figure (point estimates) or of limits computed with a given probability (interval estimates). They also make use of statistical methods in order to test certain hypotheses about the (unknown) population. This is useful in the testing and verification of economic laws.

Modern statistical inference is based upon the idea of the random variable. This is a variable which can assume certain values with definite probabilities (probability is here understood in the second sense defined above). A good example is the outcome of a throw with a die. This random variable can assume the values one to six with definite probabilities. With a true die, these probabilities will all be one sixth.

The first problem in statistical inference is point estimation. In this case we obtain a single figure for an estimate of the unknown quantity.

²⁵*Op cit*

The statistician uses here a number of methods, the two most important of which are the method of maximum likelihood²⁶ and the method of least squares²⁷. The method of maximum likelihood chooses as the estimate the particular value which maximizes the probability. The method of least squares chooses the value which minimizes the sum of the squares of the deviations from the chosen value. Both methods have desirable properties and lead frequently to the same estimates. There are also other methods available, but they are not in general use.

Sometimes we desire more information than a single value for our estimate. The statistician computes what is called fiducial or confidence limits²⁸. The theory of fiducial²⁹ and of confidence limits is not identical³⁰ and the computed limits do not always coincide. But we will neglect these difficulties for the moment. They do not arise in the simple cases we are going to discuss here.

Fiducial or confidence limits are computed in such a fashion that the probability that the true value or population value (which is unknown) will fall between them is a preassigned number, *e.g.*, 95 per cent. This has to be interpreted in the following sense. If a statistician computes a great many confidence or fiducial limits on the 95 per cent. probability basis, then in the long run on the average these limits will enclose in 95 per cent. of the cases the true population value, whereas in 5 per cent. of the cases the population value will fall outside the limits.

Another problem arising in statistical inference is the testing of statistical hypotheses³¹. A statistical hypothesis is of course not derived from the data. It is given independently of the statistical investigation *e.g.*, from considerations arising in economic theory. Here we follow closely the Neyman-Pearson theory. They distinguish two types of errors in testing a hypothesis. Type one error occurs if we reject a true hypothesis, type two error, if we do not reject a false hypothesis.

Tests of hypotheses should be designed in this way. For a given probability of type one error (called level of significance) we use the test which at the same time minimizes the type two error.

Having constructed a test in this manner, we proceed as follows. We will choose a level of significance, *e.g.*, 5 per cent., then we will reject all hypotheses which have a probability of less than 5 per cent. and not reject those which have a higher probability. If this procedure is carried on for many tests, then on the average we shall reject a true hypothesis in about 5 per cent. of the cases. The test is also constructed

²⁶M. G. Kendall, *The Advanced Theory of Statistics*, vol. 2, London, 1946, pp. 1 ff, H. Cramer, *Mathematical Methods of Statistics*, Princeton, 1946, pp. 498 ff.

²⁷F. N. David and J. Neyman, "Extension of the Markoff Theorem on Least Squares," *Statistical Research Memoirs*, vol. 2, London, 1938, pp. 105 ff.

²⁸M. G. Kendall, *op. cit.*, pp. 62 ff.

²⁹M. G. Kendall, *op. cit.*, pp. 85 ff.

³⁰M. G. Kendall, *op. cit.*, pp. 269 ff, H. Cramer, *op. cit.*, pp. 525 ff.

³¹*Ibid*

in such a fashion that it rejects more false hypotheses than any other test with the same level of significance

A special and very important case of tests of hypotheses are tests of significance³² These are constructed for the testing of a special hypothesis, called the null-hypothesis, which says that the magnitude in question is zero. Again we choose arbitrarily a level of significance, say 5 per cent. We shall reject the null-hypothesis if its probability is less than 5 per cent, and not reject it otherwise. If we reject the null-hypothesis, we will say the quantity tested is significant, *i.e.* significantly different from zero.

The meaning of the level of significance (5 per cent) is as follows. If we carry out many tests of significance, we may expect to reject in the long run on the average a null-hypothesis in 5 per cent of the cases, if it is true.

It should perhaps be mentioned that A. Wald³³ has developed a most ingenious theory of statistical inference, which is a generalisation of the Neyman-Pearson theory of testing hypotheses but includes also estimation and other problems. One of these problems which is probably of great importance for econometrics is multiple choice, *i.e.* the choice between several hypotheses. It is based upon the assumption of the existence of a risk-function which describes the consequences of committing errors of various types. Since there is rarely agreement on the goals of social policy in economic matters, such an assumption does not seem to be useful in our field. But it is a most useful approach, *e.g.* in industrial applications of statistics, where the consequences of the various types of error can easily be expressed in monetary terms.

Carnap has developed certain ideas which seem more adequate for the construction of a pure, *i.e.* non-pragmatic, theory of statistical inference. This theory is, however, not yet complete, and most of the problems of practical interest to the economic statistician and econometrician are still out of its reach³⁴.

(c) *Illustrations from a Study of American Agriculture*

In this section we shall try to illustrate econometric methods by an example taken from a study of demand and supply in American agriculture. We shall indicate the theoretical basis and the statistical procedures, without going into details of numerical computations. We shall also indicate tentatively the possibility of utilizing the results in economic policy.

In principle, it would have been advantageous to construct a model of the total economy, which includes also the market of agricultural products. But a verification of such a model was beyond the possibilities of this study. Hence we consider the market of agricultural products

³²G. W. Snedecor, *Statistical Methods*, 4th Ed., Ames, Iowa, 1946, pp. 43 ff.

³³A. Wald, *On the Principles of Statistical Inference*, Notre Dame, Ind., 1942.

³⁴G. Tintner, *Foundations of Probability and Statistical Inference*, unpublished essay to be published in the *Journal of the Royal Statistical Society*.

in isolation. This has the disadvantage that we cannot deal with the interactions between the agricultural and other markets. These interactions may be quite important as far as applications to policy are concerned.

We have indicated above the problem of identification. For useful applications in policy we must have estimates of the structural relationships which can be identified as the demand function and the supply function of agricultural products. But a simple linear regression of the quantity of agricultural produce on the price of these products, or the regression of the price of agricultural products on the quantity bought and sold cannot give us these structural relationships. In order to achieve identification we have to include other variables.

Furthermore, in order to simplify our computations we must make the assumption that the demand and the supply function of agricultural products are linear. These linear functions can only be considered as very rough approximations to the "real" demand and supply function which are of course not linear. But in this way we may achieve some estimates of the structural parameters, *i.e.* the price elasticity and the income elasticity of the demand for agricultural products, etc.

Denote by M_1 agricultural prices, by M_2 income, by M_3 the quantity of agricultural products, by M_4 time and by M_5 a cost factor. All these quantities are supposed to be the "true" values, unaffected by errors of observations, etc. Denote further by w_1 the effect of other factors not included in the demand equation on the demand for agricultural commodities, and by w_2 the effect of other factors not included in the supply equation on the supply of these products. Then we have from our assumption of linear relationships

$$(1) M_3 = k_{11}M_1 + k_{12}M_2 + k_{14}M_4 + k_{10} + w_1$$

$$(2) M_3 = k_{21}M_1 + k_{24}M_4 + k_{25}M_5 + k_{20} + w_2$$

The first equation is the linear approximation to the demand function for agricultural products. We note that we assume that demand for agricultural products (M_3) depends upon the prices of these products (M_1), upon the income (M_2), has a time trend (M_4) and also depends upon other factors which are not included in the equation but are represented by w_1 . Such factors are prices of industrial commodities, state of business, etc.

The second equation is a linear approximation to the supply function of agricultural commodities. The supply of these products (M_3) depends upon their prices (M_1), has a time trend (M_4) and depends upon a cost factor for agricultural products (M_5). It depends also on other factors not explicitly included in this equation, but represented by w_2 . Such factors are the state of agricultural technology, interest rates, agricultural wages, etc.

We have mentioned above that the variables M_1 , M_2 , M_3 and M_5 are not actually the observed quantities but the "true" values of the

³G. Tintner, "Multiple Regression for Systems of Equations," *Econometrica*, vo 14, 1946, pp 5 ff

variables in question. They are actually affected by errors which are in the nature of errors of observations.

Our data for the analysis are the variables $X_1 = M_1 + y_1$, $X_2 = M_2 + y_2$, $X_3 + M_3 + y_3$, $X_5 = M_5 + y_5$. The quantities y_1, y_2, y_3, y_5 are the errors similar to errors of observations. We will call them errors in the variables, whereas the variables w_1 and w_2 are errors in the equations.

There are two methods available for estimating the constants $k_{11}, k_{12}, k_{14}, k_{10}, k_{21}, k_{24}, k_{25}, k_{20}$. (1) We can assume that the errors in the variables are absent, then we use the method of weighted regression. (2) Or we can assume that the errors in the equations are absent, then we use the methods developed by the Cowles Commission.³⁶ Neither assumption is of course completely justified. We have certainly errors of observations in our empirical data, and it is also obvious that not all pertinent variables are included in our equations. But a method which deals simultaneously with errors in the equations and with errors in the variables has not yet been developed.

In our study we have neglected errors in the equations and deal only with errors in the variables. We must hope that the effort of the variables which have not been included in our equations is not serious enough to vitiate our analysis.

It should be noted that our equations are identified. This has been achieved by making the assumption that the variable M_2 (income) enters into the first equation, but not into the second. And we have also assumed that the variable M_5 (cost factor) enters into the second equation, but not into the first. These assumptions are only approximately justified. Without these or similar assumptions it would however not be possible to identify our equations.

Since we deal only with errors in the variables and not with errors in the equations we use the method of weighted regression. With this method we do not minimize the sum of squares of the deviations of the dependent variable from the fitted values, as with ordinary multiple regression. We minimize a weighted sum of squares. The weights are determined by the error variances and covariances of all our variables.

For simplicity's sake we make the assumption that the errors in our variables are not correlated. The variances of the errors in the variables are estimated by the Variate Difference method. This procedure is justified by the fact that the systematic parts of our variables (M_1, M_2 , etc.) are "smooth" functions of time.

The data are 24 annual observations from the period 1920-1943. X_1 is prices received by farmers, index numbers, 1910-14=100; X_2 , national income, in billions of dollars, X_3 an index of agricultural production, 1935-39=100, X_4 time, origin between 1931 and 1932, X_5 prices paid by farmers, index numbers, 1910-14=100. The arithmetic means of the variables are given in the following table.

³⁶T. Haavelmo, "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, vol. 11, 1943, pp. 1 ff., "The Probability Approach in Econometrics," *Econometrica*, vol. 12, supplement, 1944. T. Koopmans, "Statistical Estimation of Simultaneous Economic Relations," *Journal of the American Statistical Association*, vol. 40, 1945, pp. 448 ff.

Variable	Symbol	Arithmetic Mean
Prices received by farmers	X_1	127 417
National income	X_2	72 375
Agricultural production	X_3	100 625
Time	X_4	0 000
Prices paid by farmers	X_5	136 460

The error variances of the variables are estimated by the Variate Difference Method. This procedure is justified, if we assume that the systematic parts of the variables, *i.e.*, M_1, M_2, M_3, M_5 are "smooth" functions of time.³⁷ If this is the case, these systematic parts can be eliminated, or approximately eliminated, by taking finite differences. The variances of the difference series which results after the elimination of the systematic parts may be taken as the estimate of the error variances, *i.e.*, the estimates of the variances of y_1, y_2, y_3, y_5 , which are the errors in the variables. The variable X_4 (time) is not affected by errors.

Methods have been developed which also make it possible to form an idea about the number of linear relationships probably existing among a set of variables.³⁸ An application of these procedures shows that there are probably two linear independent relations between the systematic parts of the five variables. These two relations are the demand and the supply function. There is also in all probability one linear relationship between M_1, M_2, M_3, M_4 , and one between M_1, M_3, M_4, M_5 . The first linear equation can be identified with the demand function, the second with the supply function.

On the basis of these results we achieve an estimate of the constants in equations (1) and (2). These estimates are derived by weighted regression methods. They are

$$(3) M_3 = -0.097 M_1 + 0.424 M_2 + 0.313 M_4 + 82.297$$

$$(4) M_3 = 1.721 M_1 + 0.809 M_4 - 3.611 M_5 + 374.097$$

Tests of significance show that all the regression coefficients in equation (3) (demand function) are significant at the 5 per cent level of significance. But none of the regression coefficients in equation (4) (supply function) are significant at the same level of significance.

This indicates that we have been reasonably successful in estimating a demand function for agricultural products in the United States. But we have not succeeded in determining a supply function for these products.

³⁷G. Tintner, *The Variate Difference Method*, Bloomington, Ind. 1940.

³⁸G. Tintner, "A Note on Rank, Multicollinearity and Multiple Regression," *Annals of Mathematical Statistics*, vol. 16, 1945, pp. 304 ff. See also R. C. Geary, "Studies in the Relations between Economic Time Series," *Journal of the Royal Statistical Society, Series B*, vol. 10, 1948; T. W. Anderson, *The Asymptotic Distribution of the Roots of certain Determinantal Equations*, *ibid.*

The reason for this is probably the inadequacy of our model. We have left out at least one important factor which enters into the supply function of agricultural products—namely, weather conditions. It is also probable that a dynamic supply function would be more adequate than the static one in our model. In this case we would have to introduce the prices of agricultural products with a lag.

We will now discuss the demand function (3) in more detail. We note first that the coefficient of M_1 (agricultural prices) is negative, and the coefficient of M_2 (income), is positive. These signs are consistent with theoretical expectations: if agricultural prices rise, the demand for these products can be expected to decline, and if income increases, the demand for agricultural products will increase. The coefficient of M_4 (time) is also positive. This indicates a rising trend in demand for agricultural products in the United States during the period. This result is also in agreement with other observations.

From equation (3) we can derive an estimate of the price elasticity of demand for agricultural products. This estimate is computed at the average of the variables during the period. We estimate the price elasticity as -0.123 . This should be interpreted in the following way: if conditions are the same as during the period analysed and if the condition of *ceteris paribus* is fulfilled, then an increase in agricultural prices by 1 per cent will be followed by a decline of the demand for agricultural products of not quite $\frac{1}{8}$ of 1 per cent.

We can use this result tentatively in economic policy. Suppose the American government decides to raise agricultural prices by 10 per cent, for instance by price fixing. Then the consumption of agricultural products will decline by about 1 per cent, if conditions are similar to the ones during the period analysed, and if other things remain equal. Thus last condition may easily not be fulfilled, since, for instance, income may be affected by the government price policy.

We can also derive an approximate standard error of the price elasticity of demand for agricultural products. This standard error is 0.034. We may use this standard error to test statistically various hypotheses about the price elasticity.

Assume, for instance, that an economist claims that the "true" elasticity of demand for agricultural products is -1 . This is to say, the hypothesis is: to a one per cent. increase in agricultural prices there corresponds also, *ceteris paribus*, a one per cent. decline in agricultural consumption.

To test this hypothesis we use the *t*-test. The level of significance is 5 per cent, we will reject a hypothesis whose probability is less than 5 per cent. We divide the difference between the empirical elasticity (-0.123) and the hypothetical one by its standard error. The resulting *t* is -25.65 . But at the 5 per cent level for 20 degrees of freedom we have a permissible (positive or negative) *t* of only 2.086. Hence the probability of the divergence between our hypothesis (elasticity -1), and the empirical result (elasticity -0.123) is much less than 5 per cent. The hypothesis must be rejected.

On the other hand we may test the significance of our price elasticity. We test then the null-hypothesis that in the population the (unknown) "true" price elasticity is zero. This, by the way, is also a situation which may be postulated theoretically by some economists. It corresponds to a completely inelastic demand for agricultural products. Under

these conditions the prices of agricultural produce would exert no influence upon the quantity sold

We again use the t -test, as an approximation. We divide the empirical elasticity, (-0.123) by its standard error. The level of significance is again 5 per cent. This is to say, we will reject the null-hypothesis if it could not have happened by chance in at least 5 per cent of all cases. Then we will say that our elasticity is significant.

We divide our empirical elasticity (-0.123) by its standard error (0.034), the resulting t is -3.6 . But at the 5 per cent level of significance we should have a t which is not larger than 2.086 for 20 degrees of freedom. Because of the symmetry of the t -distribution the t may be positive or negative. Our empirical t is much larger than the one permitted. Hence we have to reject the null-hypothesis. The elasticity is significant.

Finally, we can also use the t -distribution to compute approximate fiducial or confidence limits for our price elasticity of the demand for agricultural products. We use the 95 per cent confidence coefficient. This is to say the limits are computed in such a way, that they will include the (unknown) "true" elasticity in the population in 95 per cent of all cases, in the long run, on the average.

The limits, computed in this fashion, are -0.052 and -0.195 . They have to be interpreted in the following way: if the conditions are approximately the same as in the period analysed, if also the condition of *ceteris paribus* is fulfilled, then we can state with a probability of 95 per cent after an increase in agricultural prices of 1 per cent the demand for agricultural products will decrease by not less than about 1/20 and not more than about 1/5 of one per cent.

Apart from the price elasticity we can also derive the income elasticity of the demand for agricultural products. We compute this elasticity again for the average of all the variables in the period considered. This income elasticity of the demand for agricultural products in the United States is estimated as 0.307 . This result has to be interpreted in the following way: if conditions are on the whole the same as during the period considered, and also if other things remain equal, then an increase of 1 per cent in income will be followed by an increase of about 3/10 of one per cent in the demand for agricultural products.

We give again a tentative example for the application of this result in economic policy. Suppose the American government decides to raise income by 10 per cent. This may, for instance, be accomplished by subsidies, by public works, by an appropriate wage policy, etc. Then, if other things remain equal and if conditions are approximately the same as during the period analysed, this increase in income will be followed by an increase in the demand for agricultural products by about 3 per cent. If the increase in income is, for instance, a part of a more comprehensive economic plan, then the government must take the increased demand for agricultural products into account when it is planning for agricultural production.

The income elasticity (0.307) appears to be much larger than the price elasticity (-0.123) for agricultural products. Is this difference significant? A statistical test reveals that the difference between the two elasticities has to be considered significant at the 5 per cent level of significance. The probability that this difference could have arisen by pure chance is much less than 5 per cent.

This is also a result which is of some importance for economic policy. It has to be taken into account when considering, for instance, the effects

of a policy of price fixing compared with a policy of full employment which would increase and stabilize incomes

We consider finally another possible and tentative application of our results to economic policy. This is the problem of taxation of agricultural commodities or of subsidies for agricultural commodities. A tax of amount t will shift the supply function upwards by t units. A subsidy of t units can be considered as a negative tax and will shift the supply function downwards by t units.

Consider now our equations (3) and (4). Assume that M_2, M_4, M_5 remain fixed at the averages indicated in the table. Then after the imposition of a tax of t units for each unit of agricultural products sold or a subsidy of $-t$ units for each unit of agricultural products produced, the price of agricultural products will be

$$(5) M_1 = 127.417 + 0.55t$$

and the quantity sold will be

$$(6) M_3 = 100.625 - 0.053t$$

If $t=0$ then we obtain the average values of price and quantity 127.417 and 100.625.

Assume now that a tax of 10 units is imposed on each unit of agricultural products sold. This may, for instance, be accomplished by a sales tax for agricultural products. Then the equilibrium price will be 132.915, and the quantity sold will be 99.095.

Next assume a subsidy of 10 units for each unit of agricultural commodities produced. Then the price will be 121.915, and the quantity sold will be 101.155.

These tentative applications to policy are certainly very suggestive. They should, however, not be taken too seriously, because of the uncertainty of our results. They may point the way to more useful applications based upon more extensive and detailed models. I have only tried to indicate the potential power and usefulness of econometric methods in economic research and applications to policy.

DISCUSSION.

THE PRESIDENT (DR GEARY). To signalise in a special way our sense of the importance of this occasion, I shall adopt the unusual course, for the President, of proposing the vote of thanks to our distinguished lecturer. In order to discuss the paper we have been at the disadvantage that we necessarily had not the text of the paper beforehand, though it will, of course, be published in full in the Journal. I had intended taking as a text some sentences out of the synopsis circulated to members and had drafted some notes of the relations between mathematics, mathematical statistics and economics. Having heard Professor Tintner's most interesting lecture, however, with its strong emphasis on the application of mathematical statistics to an important problem in economics, I cannot resist the call of the mathematical wild and I shall abandon my notes forthwith.

In order to establish the relationships between time series, it seems desirable to extend these series by using data relating to quarters, months or even weeks. This would increase the number of series over a short term of years and lessen the sampling error of estimates of the coefficients. At the present time yearly data was used almost univer-

sally in these studies. A disadvantage in using yearly data was that the time period of a year was too long and that the yearly data very often smoothed away very significant causes or consequences. I would like to have Professor Tintner's view on the possibility of using shorter time intervals (even of using continuous time series). In these studies also when lagged terms were used, the lag was invariably regarded as constant whereas in actual practice the lag might vary—the lag itself might be a stochastic variable. I would like also to have Professor Tintner's opinion on the possibilities of using varying lags applied to continuous series.

I confess myself to be a disciple of Professor Tintner's in his general approach to the ascertainment of linear relationships between economic variables, in that in our researches we deem known the variables entering into the relationships, though these variables are deemed to have errors of observation. I must admit that I am somewhat sceptical of the usefulness of the so-called error in equation approach because the estimate of the coefficients of the variables appearing explicitly must surely depend on the number of variables in the equation. If, for instance, seven variables appear explicitly and we compute the coefficient of the independent variable X_2 , what guarantee have we that this coefficient will remain unchanged, even if the number of sets of observations is indefinitely large, when we retain (say) five variables in the equation? And if the coefficient has a value which depends on the number of variables retained, what significance can be attached to it?

After all, the number of possible factors is strictly limited, and it would not appear to be an insuperable task to investigate relationships on empirical grounds including all conceivable variables. Despite Professor Tintner's depreciation, I must confess myself to be something of an empiricist. On the latter point I would like to have Professor Tintner's opinion of a point which seems to me of fundamental importance in determining the relations between theoretical and applied economics. The question is: To what extent is economic theory useful in determining statistical relations between economic variables? It is a strong point in favour of the theoretical economic approach that Professor Tintner has been able to show that the number of relations between the variables which he discusses in his paper is probably two and from the signs (+ or -) of the coefficients of price, it would appear that these two equations are identifiable as the demand equation and the supply equation and, therefore, concordant with the theoretical economic approach. On the other hand it is a curious circumstance that satisfactory demand relations are far easier to determine than are supply relations. In fact, I am unaware of any satisfactory determination of a supply relation.

Professor Tintner has shown how to determine linear relations between economic variables subject to errors of observation, the variances of which are known in advance and, perhaps more important, has discovered a technique for ascertaining the probable number of linear relations, i.e., whether none, one, two, three, etc. It is fascinating to know that by Professor Tintner's methods applied to the agricultural supply and demand factors, two relations have been found clearly identifiable as the supply and demand equations. May I say that my personal debt to Professor Tintner's work is great since a large part of my Cambridge research consisted in an extension of his results. I suggested, in fact, that, in dealing with time series instead

of proceeding with new data, one should try if the individual series could be smoothed by algebraic curves. If they could, there were tests for number of relations and the estimates of the coefficients were more efficient in the statistical sense than were the estimates from the unsmoothed data.

MR THORNTON said that the work of the econometrician was essentially based on the contributions of the economist and the statistician, seeking, by statistical methods, to evaluate the effects of various economic policies.

He suggested that the United States Bureau of Labour Statistics had adopted very similar methods to those of Professor Tinbergen in deriving full employment patterns in the United States for 1950, on the basis of Professor Wassily Leontief's monumental work on American inter-industry relationships. Mr Thornton was sure that the Bureau had derived supply functions for this purpose, thus answering the President's question on this subject.

While confessing to an admiration for Professor Tinbergen's paper, Mr Thornton thought that, apart from the question of the choice between an auto-regressive scheme and the "Geary" scheme for eliminating the random element, there were two important limitations to the system of multiple regression analysis of economic time-series. Firstly, the error variances of the variables, in practice, are usually not known a priori and the variances computed from the "sample" time-series must be used, thus reducing the number of degrees of freedom and so diminishing the utility of any significance tests subsequently applied. Secondly, the applicability of numerical evaluations based on the past to policy making in the future was of some doubt, particularly when the time-series was of short duration.

MR HONOHAN said he was disappointed that econometrics appeared not to offer much hope of evading "ceteris paribus". The isolation of particular features of a problem, however valuable as a study, did not yield results of sufficient moment to capture the attention of politicians and public administrators. He also felt that, while the analytical methods of econometrics might be suitable in application to a closely-knit economy, they would scarcely be appropriate in a loose and scattered economy such as we had in this country.

He thought, however, that it was important not merely to indicate in what direction effect would follow cause but what the measure of such effect was. This was surely a legitimate and necessary field of inquiry and seemed to bear some resemblance to that of the actuary. As the actuary's profession was essentially practical, however, he could not usually fall back on "ceteris paribus", but had to come down boldly on whatever conclusions his knowledge and experience indicated, having regard to all the circumstances bearing on the matter.