COWLES COMMISSION FOR
RESEARCH IN ECONOMICS
Monograph No. 11

COWLES COMMISSION

FOR RESEARCH IN ECONOMICS

The Cowles Commission for Research in Economics is a not-for-profit corporation founded in 1932 for the purpose of conducting and encouraging investigations into economic problems. The results of research by members of the Commission's staff are published in two series: *Cowles Commission Monographs* in book form, and shorter papers, usually reprints of articles from journals, as *Cowles Commission Papers, New Series*. The Commission is affiliated with the Econometric Society, an international society for the advancement of economic theory in its relation to statistics and mathematics.

THE UNIVERSITY OF CHICAGO

COWLES COMMISSION MONOGRAPHS

No. 1. *Dynamic Economics*, by CHARLES F. ROOS. 1934. 275 pages. Out of print.

No. 2. *NRA Economic Planning*, by CHARLES F. ROOS. 1937. 596 pages. Out of print.

No. 3. *Common-Stock Indexes*, by ALFRED COWLES and ASSOCIATES. Second Edition, 1939. 499 pages. Price $6.00.

New monthly indexes of (1) stock prices, (2) stock prices adjusted for the reinvestment of cash dividends, and (3) yield expectations; and annual indexes of (4) yields, (5) dividend payments, (6) earnings-price ratios, and (7) earnings; all for 69 industry groups or combinations of groups, and running from 1871 through 1938.

No. 4. *Silver Money*, by DICKSON H. LEAVENS. 1939. 439 pages. Price $4.00.

A sketch of the history of the monetary use of silver, followed by more detailed consideration of recent developments.

No. 5. *The Variate Difference Method*, by GERHARD TINTNER. 1940. 175 pages. Price $2.50.

A full account of the history and use of this method for the analysis of time series. Certain new devices of treatment are presented, and extensive tables are given to aid calculation.

No. 6. *The Analysis of Economic Time Series*, by HAROLD T. DAVIS. 1941. 620 pages. Price $5.00.

This book reviews the historical development of the subject, describes the methods used, and makes applications to a variety of economic phenomena.

No. 7. *General-Equilibrium Theory in International Trade*, by JACOB L. MOSAK. 1944. 187 pages. Price $2.50.

This study applies the modern theory of economic equilibrium (as expounded by J. R. Hicks and others) to an important field.

No. 8. *Price Flexibility and Employment*, by OSCAR LANGE. 1944. 114 pages. Price $2.00.

The author tries to clarify important concepts that have had much currency in the practical discussion of depressions and wars but have so far remained too vague to allow of useful treatment.

No. 9. *Price Control and Business*, by GEORGE KATONA. 1945. 246 pages. Price $3.00.

A study of the working of price control based on field studies among producers and distributors of consumers' goods in the Chicago area, 1942–44.

No. 10. *Statistical Inference in Dynamic Economic Models*, edited by TJALLING C. KOOPMANS, with Introduction by JACOB MARSCHAK. 1950. 438 pages. Price $6.00.

This monograph contains original contributions from many authors concerning statistical problems encountered in economic model construction.

No. 11. *Economic Fluctuations in the United States, 1921–1941*, by LAWRENCE R. KLEIN. 1950. 174 pages. Price $4.00.

This study applies the methodology of econometric model construction to business cycle analysis with possible implications for prediction and policy making.

No. 12. *Social Choice and Individual Values*, by KENNETH J. ARROW. Forthcoming.

Methods of symbolic logic are applied to the question whether a social valuation of alternatives can be consistently derived from given, partly conflicting, individual valuations.

No. 13. *Activity Analysis of Production and Allocation*, Proceedings of a Conference, edited by TJALLING C. KOOPMANS. Forthcoming.

Contributions from economists and mathematicians on the theory and techniques of efficient allocation of resources and programming of activities.

Orders for Monographs 3–9
should be sent to
THE PRINCIPIA PRESS, INC.
Bloomington, Indiana

Orders for subsequent monographs
should be sent to
JOHN WILEY & SONS, INC.
440 Fourth Ave., New York City

ECONOMIC FLUCTUATIONS IN THE UNITED STATES
1921-1941

By

LAWRENCE R. KLEIN

JOHN WILEY & SONS, INC., NEW YORK
CHAPMAN & HALL, LIMITED, LONDON
1950

PREFACE

It cannot be overemphasized that this book contains only tentative and preliminary steps in econometric model building. It is the result of somewhat less than three years' work undertaken at the Cowles Commission for Research in Economics during the period 1944–1947. It is with all modesty that the research results are now presented in published form, and the main hope is that the work contained here may be of some help to other researchers in the field. If I were to begin today the project of reconstructing an econometric model for the United States, I believe that I would proceed somewhat differently, benefiting from knowledge of the weaknesses of the models in this volume. I would expect others to do the same.

The stated attempt here is to estimate laws of human behavior in economic life. If we reach the goal of discovering autonomous behavior patterns, we should expect them to hold under a variety of circumstances—for example, under both prewar and postwar conditions. Although I believe that the models presented in the ensuing pages give a reasonably good description of the prewar economic process, they do not carry over in all respects to the postwar situation. To some extent, I believe, the estimated behavior relationships do not take into account some variables which have not been quantitatively important in the past, but which have become more important in the postwar situation. On the other hand, some of the equations may even give a wrong picture of the prewar interrelationships among variables that were then important. But, on the whole, I feel that the prewar model is reasonably good; hence the title of this monograph.

The main weaknesses of the models in their applications to postwar phenomena are the inadequacy of relations that serve to determine absolute price levels and household consumption. The postwar price inflation and heavy consumer purchases seem to contradict the prewar relationships. Of course, to take only one problem for which econometric models may be used, postwar forecasting of economic aggregates should not be carried out by the mere extrapolation of models fitted to the prewar observations. For every forecast period the parameters of the model should be recomputed for all observations prior to the forecast period, and then extrapolations should be made from the newly estimated models. This procedure is absolutely necessary because there are sampling errors in all the estimates and because it serves as a "troubleshooter" indicating the points at which the models break down. With a suitably dynamic forecasting technique, I believe that econometric models can be of great value; however, we have no published records of adequate scientific attempts. Then, too, there is a problem in connection with the statistical data of economics; they are constantly being revised and changed with respect to concept. The estimates of this volume are based on data that are already outmoded. It is highly desirable that the entire model be recomputed on the basis of the new national accounts of the U. S. Department of Commerce for prewar and postwar years. This has not yet been done.

Econometric research, with a few notable exceptions, has been based on a relatively small number of annual observations. In most instances annual observations for the interwar period have been the main source of information. There seems to be room for improvement in this respect, and that is another strong reason why I regard this volume as preliminary and tentative. At present some workers are studying time-series observations for 50 to 60 years in the past, and others, including myself, are studying quarterly or monthly observations for the interwar period and cross-section material (spatial rather than time variations) in order to get more refined estimates of fundamental behavior patterns. The end product of all this intensive research should eventually be an eminently more satisfactory econometric model. It is not a simple job, however, because quarterly or monthly data introduce difficult problems of the treatment of seasonal variation and high serial correlation of disturbances, and real difficulties are involved in the reconciliation of time-series and cross-section estimation results. Difficulties like these have not been faced in the present volume because it treats only annual time series of the interwar period; however, these perplexing problems cannot be avoided indefinitely if we are to get satisfactory econometric models.

I shall assume full responsibility for the contents of this volume, yet it was developed under a scheme of great mutual cooperation among economists and mathematical statisticians at the Cowles Commission for Research in Economics during the period 1944–1947. Numerous staff meetings and informal discussions introduced and conditioned many of the ideas herein expressed. Many outsiders in research, academic, and government positions also were of great assistance on many points, but they are too numerous to mention by name. I am especially grateful to Professor Jacob Marschak who introduced me to this research project and who served continuously as a guiding influence and as a moderator

on my more reckless ventures. Other Cowles Commission staff members to whom I owe a large debt of gratitude are T. W. Anderson, Jr., K. J. Arrow, G. Cooper, T. Haavelmo, L. Hurwicz, T. C. Koopmans, K. May, D. Patinkin, H. Rubin, and S. Tekiner. I profited very much by the reading of a set of critical notes on the manuscript solicited from E. D. Domar. S. Reiter gave valuable help in checking the time series used and the source references in the Appendix. R. B. Leipnik and B. A. de Vries prepared and supervised many of the statistical computations in which these time series were used for estimation of parameters and standard errors.

Thanks are due also to Mrs. Emily R. Strand, Publications Secretary in the Cowles Commission, to Mrs. Jane Novick, her successor, and to Miss Jean Curtis and Miss Virginia Atherton for their assistance in preparing the manuscript for publication.

L. R. K.

New York City
March, 1949

TABLE OF CONTENTS

Chapter I

MODEL BUILDING—GENERAL PRINCIPLES

Tinbergen did a great service to the study of economics when he prepared his volumes on the statistical testing and measurement of business-cycle theories.[1] This book is written in the spirit of Tinbergen's investigations and is intended as an improvement and extension of his results. As a consequence of the extensive theoretical discussions since 1936, when Keynes published his *General Theory*, it has become possible to formulate more sharply the structure of the economic system and thereby to gain added simplicity and accuracy not available to Tinbergen at the time of his work. Furthermore, there has been a decided improvement and extension of data in recent years, and we should be eager to take immediate advantage of this good fortune.

The purpose in building econometric models is to describe the way in which the system actually operates. We want to do more than is suggested by the title to Tinbergen's work (though not by the book itself), i.e., more than the mere testing of business-cycle theories. We want also to discover the best possible theory or theories which explain the fluctuations that we observe. If we know the quantitative characteristics of the economic system, we shall be able to forecast with a specified level of probability the course of certain economic magnitudes such as employment, output, or income; and we shall also be able to forecast with a specified level of probability the effect upon the system of various economic policies. In the course of the search for models which are suitable for our purposes of forecasting and making policy recommendations, we shall inevitably have to consider several alternative economic theories as admissible hypotheses. The acceptance or rejection of these hypotheses in the course of our search for truth will be our contribution to the problem of testing business-cycle theories. This is an important problem for which contributions are needed, because too often writers have made bold statements about the operation of

[1] See J. Tinbergen, *Statistical Testing of Business-cycle Theories*, Vol. I, "A Method and Its Application to Investment Activity," Vol. II, "Business Cycles in the United States of America, 1919–1932," League of Nations Economic Intelligence Service, Geneva, 1939.

the system without examining the factual data to determine whether or not their statements are true. As a result we have had continuing economic fluctuations and continuing publication of business-cycle theories.

CONCEPTS

The construction of econometric models is relatively new, and most of the ideas and terminology are not familiar to the majority of economists. At this stage of the book we shall introduce those concepts which are essential to the discussion of the general problems considered. As we progress, additional concepts will be introduced according to the need for them.

We view the economic system as describable by a set of simultaneous equations expressing all the interrelationships among the measurable economic magnitudes which guide economic behavior. The variables in this set of equations are classified into two main types, endogenous and exogenous. The endogenous variables are those variables which are determined within the system of economic forces in a narrow sense, and they include such familiar magnitudes as output, employment, prices, profits, rents, interest. The exogenous variables are those which represent forces outside the confines of the economic system. They are determined by natural, technological, sociological, political, or institutional forces which are assumed here to be non-economic. It is, of course, not satisfactory to separate sociology and politics from economics, but our purpose in pioneering will be served best if we make assumptions which simplify our model as much as possible. Eventually we may hope to develop a complete social theory which leaves in the exogenous category only such variables as weather, earthquakes, and other "acts of God." The only known theory today which embraces politics, sociology, economics, etc., is the Marxian theory, but not enough work has been done on quantifying this theory to render it readily applicable to making a variable like government spending endogenous.

We shall treat the supply of money, government spending, taxes, population growth, and time trends as exogenous variables, although this procedure may be questionable in some cases. We justify our classification of the supply of money as exogenous on the grounds that the banking system, an institutional phenomenon, can regulate the supply of money at will. If the Board of Governors of the Federal Reserve system decides to make drastic changes in discount rates, reserve requirements, or open-market operations, they can create or destroy almost any amount of money that they see fit to create or destroy.[2] Economists have formulated no laws of behavior which the Federal Reserve Board will obey in making its decision as to the supply of money. It is possible that some social theory might explain the behavior patterns of the Federal Reserve Board, but this theory would probably be so complicated that we should have difficulty in making use of it, even if we could develop it. The same remarks apply to government spending and taxing. The legislative branch of our government is the institution which finally decides upon the levels of spending and taxing, but we have no satisfactory theory to explain the behavior of legislators.[3]

Population growth is related to economic variables, but it is also highly dependent upon institutions such as the public health service, birth-control knowledge, and social customs covering family size, etc. At this point we shall be satisfied to call it exogenous.

Economists have, over a period of years, developed theories of economic behavior which are the basis for the determination of our endogenous variables. It has been stated above that we do not know how the Federal Reserve Board reaches its decisions regarding the supply of money, but we economists do claim to know how an entrepreneur reaches a decision determining how much of his product he will supply to the market, given the prevailing conditions of prices, wages, transportation costs, etc. Entrepreneurs, households, speculators are assumed to behave according to some fundamental rational patterns, which can often be written in the form of mathematical equations. For example, we assume that entrepreneurs behave so as to maximize profits, subject to the constraint that they operate according to the technological possibilities expressed by their production functions. The profit equation, the production function, and the profit-maximizing equations are all structural equations of our model. They are called structural equations because they show the basic structure of the economic system. We assume further that households behave so as to maximize their satisfactions or utilities, subject to budgetary constraints; and in this way we obtain the equations of consumer demand. Finally, we introduce equations of the market to show the interactions of households and entrepreneurs in determining the levels of prices.

[2] In our particular institutional setting, the Federal Reserve Board is restricted in some of its above decisions by another institution, namely, Congress. The ultimate decision is made, in any case, by an institution of our society.

[3] For a discussion of the distinction between variables that are autonomously controlled by policy makers and those that are exogenous for statistical purposes, see T. C. Koopmans, "When Is an Equation System Complete?" *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10, John Wiley & Sons, Inc., 1950.

Our system is complete when the number of equations in the entire set of structural equations is just enough to determine all the endogenous variables, given the exogenous variables. We must develop as many structural equations as we have endogenous variables. For the "explanation" of exogenous variables we do not attempt to develop any structural equations.

Let us define $y_{i,t-k}$ as the ith endogenous variable occurring in the $(t-k)$th period, z_i as the ith exogenous variable,[4] and u_{it} as the ith random disturbance of the tth period. The mathematical model of the economic system will be

$$(1.1.1)\quad f_i(y_{1,t}, \cdots, y_{n,t}, y_{1,t-1}, \cdots, y_{n,t-1}, \cdots, y_{1,t-p}, \cdots, y_{n,t-p}, z_1, \cdots, z_m) = u_{it} \qquad i = 1, 2, \cdots, n$$

The f_i-functions define the structural equations, equal in number to the number of endogenous variables in the system. An important econometric problem is to estimate all the parameters (structural parameters) of the f_i-functions. In economic terminology, the f_i-functions are the production functions, marginal productivity equations of profit maximization, the demand equations of consumer behavior, the speculative equations of inventory fluctuations, etc. The parameters of the f_i-functions are the production elasticities, the demand elasticities, the marginal propensities to consume, etc.

The econometrician who is interested only in forecasting may proceed somewhat differently. He is interested in the structural parameters of the system only to the extent that these parameters aid him in making forecasts. It may very well happen that all the individual parameters are not necessary for making forecasts. Instead, certain combinations of the parameters may be sufficient for this purpose. For example, we may solve (1.1.1) for each of the endogenous variables in terms of the lagged endogenous variables and the exogenous variables to get

$$(1.1.2)\quad y_{it} = g_i(y_{1,t-1}, \cdots, y_{n,t-1}, \cdots, y_{1,t-p}, \cdots, y_{n,t-p}, z_1, \cdots, z_m, u_{1t}, \cdots, u_{nt}) \qquad i = 1, 2, \cdots, n$$

The set of equations (1.1.2) is called the reduced form as distinct from the structural equations (1.1.1). If (1.1.1) is a linear system, then

[4] We omit the subscript denoting the time period of the exogenous variables because we follow the convention of defining lagged exogenous variables as separate exogenous variables.

(1.1.2) will also be a linear system, and we may replace the terms $u_{1t}, \cdots, u_{nt}$ in the ith equation with v_{it}, which is a linear combination of the random disturbances u_{it}. If the u_{it} are random time series, the v_{it} are also random time series.

The parameters of (1.1.2) are in a different form than are the parameters of (1.1.1). Elementary methods of statistical estimation applied to system (1.1.2) will not always lead us to the estimates of the parameters in system (1.1.1). For linear systems, our procedure in the case of (1.1.2) is obvious. Treat y_{it} as the dependent variable and all the variables on the right-hand side of (1.1.2) as independent variables. Then apply the familiar least-squares techniques to estimate the coefficients of the linear g_i-functions. Many [5] theorems on properties of least-squares estimates hold in this situation, because all the independent variables are predetermined; [6] they do not vary jointly with the dependent variables.

Once the parameters are thus estimated, the forecasting procedure is straightforward. Insert into equations (1.1.2) the *known* lagged values of the endogenous variables and the a priori given values of the exogenous variables. Except for random error, this procedure determines the value of the endogenous variable y_{it}.

The above procedure will not always be the most efficient method of forecasting because some information is neglected, i.e., that contained in (1.1.1). The f_i-functions provide information about the structural characteristics of the system, and the best forecasts can be made by making use of all this information. In some cases, the information contained in (1.1.2) is equivalent to that contained in (1.1.1). These are the cases in which the system is just identified (see page 11 for a definition of this term). If the system is overidentified (see page 11), (1.1.1) is to be preferred to (1.1.2), although the forecasts obtained from the least-squares estimates of the latter system will still have many desirable statistical properties.

If we are more ambitious and want to estimate the parameters of the f_i-functions, we must proceed quite differently. Assume a specific probability distribution of the disturbances in (1.1.1). There are as many endogenous variables in (1.1.1) as there are disturbances; hence (1.1.1) can be considered as a set of transformation functions which express the disturbances in terms of the endogenous variables, given

[5] Some small sample theorems do not hold if there are lags in the model. See L. Hurwicz, "Least-Squares Bias in Time Series," *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph 10, T. C. Koopmans, editor, John Wiley & Sons, Inc., 1950.

[6] Predetermined variables are either lagged endogenous variables or exogenous variables.

the exogenous variables and the lagged endogenous variables. If the probability distribution of the disturbances and the transformation functions connecting the disturbances and the endogenous variables are known, it is possible to derive the joint probability distribution of the endogenous variables.

Formally, the probability distribution of the disturbances

$$p(u_{11}, \cdots, u_{nT}) \qquad t = 1, 2, \cdots, T \tag{1.1.3}$$

in combination with (1.1.1) leads to

$$p^*(y_{11}, \cdots, y_{nT}) = \left|\frac{\partial\, (u_{11}, \cdots, u_{nT})}{\partial\, (y_{11}, \cdots, y_{nT})}\right| p(u_{11}, \cdots, u_{nT}) \tag{1.1.4}$$

where p^* is the probability distribution of the endogenous variables, and the first term on the right-hand side is the Jacobian of the transformation (1.1.1). The parameters of (1.1.1), such as the marginal productivities and marginal propensities, are now parameters of p^* and can be estimated by well-known statistical methods. In particular, the method of maximum likelihood consists in computing from the observed values of the y's those values of the parameters of (1.1.1) that maximize the probability p^*. In any case, the method used should be one that is well adapted to the specific model and that has certain properties considered to be desirable. We should want to use a method of estimation of the parameters of (1.1.1) that gives consistent estimates.[7]

Suppose that θ is a structural parameter of (1.1.1); then the estimate $\hat{\theta}$ is said to be consistent if it has the property

$$\operatorname*{plim}_{N \to \infty} \hat{\theta} = \theta$$

where "plim" means the probability limit. A statistic $\hat{\theta}$ from a sample of N is said to have the probability limit θ if the probability of the inequality $|\hat{\theta} - \theta| > \epsilon$, being satisfied, tends to zero as $N \to \infty$, for all $\epsilon > 0$. Thus, as the number of observations increases indefinitely, the probability that a consistent estimate differs from the true value of the estimated parameter by less than a given amount (however small) tends toward certainty.

In our model, the maximum likelihood estimates (using all available information regarding the form of the structural equations) have the property of consistency. There are also other computationally cheaper methods which yield consistent estimates. If (1.1.1) is linear in the

[7] The basic development of this procedure is found in T. Haavelmo's "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, Vol. 11, January, 1943, pp. 1–12, and "The Probability Approach in Econometrics," Supplement to *Econometrica*, Vol. 12, July, 1944.

parameters, it is possible to transform the least-squares estimates of the parameters of (1.1.2) into consistent estimates of the parameters of (1.1.1). This has been called the "method of reduced forms" [8] or, equivalently, the "limited information maximum likelihood method." The second name emphasizes the fact that this method can be regarded as maximum likelihood estimation using only available information regarding the form of the structural equation estimated.[9] The statistical results presented later in this volume have been obtained by the method of maximum likelihood and based in some cases on full use of a priori information, in some cases on limited use.

A very simple example will enable us to clarify many of these ideas. In the simplest Keynesian model of economic behavior, consumption is a linear function of income and investment (private and public) is entirely exogenous. The economic justification for this model will be discussed later, but it is often used today, and we shall assume its validity for the moment. Our system of structural equations now becomes

$$C = \alpha + \beta Y + u \tag{1.1.5}$$

$$C + I = Y \tag{1.1.6}$$

$$I = \text{exogenous} \tag{1.1.7}$$

where C = consumption, Y = income, I = investment. Equation (1.1.5) is a structural behavior equation and is derived from some rational principle of household behavior. Explicit derivations will be given at length in Chapter II. There is a disturbance u attached to equation (1.1.5) to show that human behavior is not exact but is subject to random perturbations. Equation (1.1.6) and equation (1.1.7) are definitions and, consequently, are not subject to random disturbances.

The reduced form of the simple model is obtained by solving for either of the endogenous variables, C or Y, in terms of the exogenous variable I. The reduced forms are

$$Y = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} I + \frac{u}{1-\beta} \tag{1.1.8}$$

$$C = \frac{\alpha}{1-\beta} + \frac{\beta}{1-\beta} I + \frac{u}{1-\beta} \tag{1.1.9}$$

[8] This method has been developed by M. A. Girshick, T. W. Anderson, Jr., and H. Rubin. See T. W. Anderson, Jr., and H. Rubin, "Estimation of Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, Vol. XX, March, 1949, pp. 46–63.

[9] Plus a list, not necessarily complete, of predetermined variables occurring in the complete system of equations.

Equations (1.1.8) and (1.1.9) correspond to the system (1.1.2). It is immediately obvious that the parameters of (1.1.8) and (1.1.9) are functions of the structural parameters in (1.1.5). The structural parameters are the marginal propensity to consume, β, and the constant term, α. The parameters in (1.1.8) and (1.1.9) are specific mathematical functions of α and β. In this particularly simple model it happens that the transformation from the parameters of (1.1.8) and (1.1.9) to the parameters of (1.1.5) is trivial, but this triviality does not extend to other models with which we shall deal in this book.

We shall specify further the statistical properties of the simple economic model. Let us assume $E(u) = 0$, $E(uu_{-\tau}) = 0$ for all $\tau > 0$, $E(u^2) = \sigma_u^2$, where E denotes the operation of mathematical expectation, and σ_u^2 is the variance of the disturbance.

In equations (1.1.8) and (1.1.9), all the statistical hypotheses of the Markoff theorem on least squares hold,[10] and we can estimate the parameters of the reduced forms without bias. This is to say that we can obtain consistent estimates of $\alpha/(1 - \beta)$, $1/(1 - \beta)$, and $\beta/(1 - \beta)$. If we are interested only in forecasting C and Y, we can stop at this point. If, however, we are interested in structural estimation, we must go another step. Knowing an estimate of $1/(1 - \beta)$, we can obtain an estimate of β, a structural parameter, and, knowing also an estimate of $\alpha/(1 - \beta)$, we can obtain an estimate of α, the other structural parameter. The method of maximum likelihood will give exactly the same results in this model whether applied directly or through the equivalent (in this case) procedure of the reduced-form method.[11]

It is interesting to observe that in solving this problem we have never had to make a choice between dependent and independent variables in the sense of the familiar regression theory. We did have to classify the variables into endogenous and exogenous categories on economic grounds; however, once the economist has made this decision the statistician may proceed unambiguously with the equation-systems methods of estimation. We have entirely left behind the questions of the early days of econometrics: Do we take the regression of price on quantity or of quantity on price? Do we take the regression of consumption on income or of income on consumption? The older methods always produced two estimates in a two-variable problem, neither one of which had the property of consistency. Now we get a unique estimate (in a confidence interval) that is consistent.

[10] See F. N. David and J. Neyman, "Extension of the Markoff Theorem on Least Squares," *Statistical Research Memoirs*, London, Vol. 2, December, 1938, pp. 105–116.

[11] For a more detailed treatment of this model see T. Haavelmo, "Methods of Measuring the Marginal Propensity to Consume," *Journal of the American Statistical Association*, Vol. 42, March, 1947, pp. 105–122.

In the simple example it was possible to transform uniquely from the estimates of the reduced-form coefficients to the structural parameters as a result of a fundamental concept in econometrics, namely, that of identification. An identifiable system is one in which we can relate unambiguously (identify) our statistical estimates to the structural parameters of the system. For example, we want to be able to tell whether an estimate of a parameter that relates price variation to quantity variation in a specific market is an estimate of the demand elasticity or the supply elasticity or vice versa. In our example we showed how to calculate a parameter, $1/(1 - \beta)$, identify this parameter with the "multiplier," and, in turn, identify this parameter with the marginal propensity to consume.

In the early days of econometrics, the identification problem arose in the estimation of demand and supply functions. The simplest model for this problem is

$$q^D = \alpha_0 + \alpha_1 p + u \tag{1.1.10}$$

$$q^S = \beta_0 + \beta_1 p + v \tag{1.1.11}$$

$$q^D = q^S + w \tag{1.1.12}$$

where q^D = demand, q^S = supply, p = price, u, v, w = random disturbances. If we substitute (1.1.12) into (1.1.10) we have

$$q^S = \alpha_0 + \alpha_1 p + u - w \tag{1.1.10*}$$

$$q^S = \beta_0 + \beta_1 p + v \tag{1.1.11}$$

We can observe, generally, q^S *or* q^D *and* p, but not u, v, or w. From the statistical point of view we have no basis for distinguishing between (1.1.10*) and (1.1.11). Both equations have the same variables, and both have random disturbances. As far as our observations are concerned, they both look alike, and time series data on p and q^S, alone, will not be enough to determine both the supply and demand equations. Unless there is more information we cannot identify any statistical estimates of the parameters as estimates of either α_0, α_1, or β_0, β_1. We may introduce the a priori information that demand and market adjustment are stable, while supply is very unstable, or, to paraphrase this remark, that the variance of the shift parameter in (1.1.10*), σ_{u-w}^2, is small, while the variance of the shift parameter of (1.1.11), σ_v^2, is large. Then we may be able to estimate α_0, α_1.

The derivation of the model (1.1.10), (1.1.11), (1.1.12) from good economic theory may provide more information that will enable us to identify the parameters of the system. Both supply and demand equations may not be linear. The non-linearities may serve to distinguish the statistical estimates of one set of parameters (estimated from price-quantity time series) from the estimates of the other set of parameters (estimated also from the price-quantity time series). Or economic theory may lead us to introduce lags or exogenous variables into one or both of these equations. If each linear equation has a unique, predetermined variable and is not subject to any other restrictions, the system is just identified.[12]

A slight modification of our simple Keynesian model will also lead to a lack of identification. Suppose that we replace equation (1.1.7) with the admissible alternative

$$I = \gamma + \delta Y + v \tag{1.1.7*}$$

in conjunction with

$$C = \alpha + \beta Y + u \tag{1.1.5}$$

$$C + I = Y \tag{1.1.6}$$

This system is not identifiable. Let us substitute (1.1.5) into (1.1.6) to get

$$u + \alpha + \beta Y + I = Y$$

$$I = -\alpha + (1 - \beta)Y - u \tag{1.1.6*}$$

There is now no way of distinguishing between (1.1.7*) and (1.1.6*); they are both linear relations between I and Y, subject to random error. From the time series data on C, I, Y, alone, we cannot estimate α, β, γ, δ for this model. We are in exactly the same position as were the econometricians who wanted to get supply or demand elasticities from price-quantity data alone. The previous economic specification that I is exogenous was just sufficient to make the system identifiable, but we have to be certain that this economic specification is justified. Perhaps the true system is not identifiable, but, if it is, we can hope to estimate the structural parameters. We must first formulate our economic theory and from it the structural equations. Then we must examine these equations for identification properties.

Identifying restrictions, i.e., specifications as to which variables are excluded from which equations, and similar prescriptions based on

[12] The reader must not get the impression that economic theory is called upon at this moment in order to achieve identification. Economic theory is called upon to provide the true structure of the systems of equations. The parameters of the true system may or may not be identifiable. However, if we fail to get an identified system because certain variables have been omitted from the equations or because the equations are not true, we must use economic theory to improve the equations until they do represent the truth. If the truth permits identification of the parameters, we may proceed with statistical estimation.

economic theory may be insufficient in number, or just sufficient, or more than sufficient to achieve identification of a given equation or parameter. In these cases we speak of underidentified, just identified, and overidentified equations or parameters. An identified or overidentified equation or parameter may possess either weak or strong estimability.

The concept of overidentification is important because there is an essential difference in the statistical treatment of systems that are just identified (neither over- nor under-) and systems that are overidentified. If a linear system is just identified as is the case in equations (1.1.5), (1.1.6), (1.1.7), it is always very simple to pass from the estimated parameters of the reduced forms to the estimated parameters of the structural equations. In these cases the statistical procedures are simple and straightforward. Obtain the reduced forms; estimate the parameters of the reduced forms by the familiar method of least squares; then transform from the parameters of the reduced forms to the structural parameters. However, when there are more than enough conditions to insure identifiability (the case of overidentification), simple algebraic transformations from estimated parameters of the reduced forms to the estimated parameters of the structural equations do not exist. These transformations can be made, but only by more elaborate statistical methods.[13]

EQUATION SYSTEMS VERSUS SINGLE EQUATIONS

The approach to econometric problems adopted in this book has been developed only in recent years, especially by members of the Cowles Commission staff. Formerly, econometricians singled out an isolated equation of the economic system and attempted to estimate the structural parameters by the method of least squares or some other simple method whose statistical properties were not usually satisfactory. When the earlier statisticians fitted their equations to the data by the method of least squares they seldom knew in which direction they should minimize the sum of squares; i.e., which should be the "dependent" and the "independent" variables. They were aware of the problem of identification, but they failed to solve it adequately. Now many of these difficulties are eliminated. If we specify both the economic and statistical properties of the model and treat the set of equations as a unit, instead of treating each equation in isolation from the rest of the system, we are not faced with the problems that formerly were so troublesome.

[13] See T. W. Anderson, Jr., and H. Rubin, *op. cit.*

In order to specify the economic properties of the model, we must know the form of all the equations that connect the several variables of the system; we must know which variables are endogenous and which are exogenous. The statistical properties of the model are the assumptions that must be made about the disturbances—their relationship to the economic variables, their distributions, their autocorrelations, their intercorrelations, etc. The model must be specified in advance, and then the appropriate statistical methods can be determined.

The reduced-form (or limited-information) method of estimation has the virtue that all the equations of a model need not be known, entirely, in order to estimate the parameters of any single equation. In the single equation, for which the parameters are to be estimated, it is necessary to know which variables are endogenous and predetermined. It is necessary to know enough predetermined variables in other equations of the system to insure identification of the equation being considered. The efficiency of the estimates increases as the number of predetermined variables from other equations of the system, used in estimating the parameters of the given equation, increases. The information about predetermined variables in other equations can often be obtained even if the full properties of these equations are not known.

There are still many unsolved problems that we inherit from the earlier days and some new ones created by the new methods. For example, multicollinearity [14] was and still is a problem. When several economic variables move together in the same general time patterns, we shall not be able to measure their separate influences in the equations of the system. The new approach does nothing to skirt the dangers of multicollinearity.

An annoying problem that arises with the new methods is the laboriousness and complexity of computation. Very economical techniques of dealing with multiple correlation problems have been perfected, but they can no longer be used except in special cases like the above example where the system is just identified. Unless we develop more economical computational methods or more efficient computing machines, the problems will remain beyond the reach of individual research workers.

[14] The term is used in the sense employed by R. Frisch.

Chapter II

ECONOMIC THEORY

The economic theory which underlies the construction of our model is classical in its methodology. We view the economic system as composed of two groups. One group consists of households and the other of business firms. It is assumed that the individuals in each group follow specific types of behavior patterns which lead to the structural equations of the model. We have, as a matter of fact, the profit- or utility-maximizing equations of the business firms, the utility-maximizing equations of the households, and the interactions of these two groups in the market to determine prices, wages, rents, etc. These equations make up the system.

Aggregation

Most economic theories are "microeconomic"; they are couched in terms of individual behavior patterns. But there are so many individual units in our economic system that it is hopeless to consider estimating a complete set of equations involving the variables of microeconomics. We should have to deal perhaps with several million equations in several million variables, a hopeless maze. As an alternative, we must sacrifice detailed information and develop systems of macroeconomic equations which involve a much smaller number of (aggregative) variables. It is a very difficult problem to pass from the theories of microeconomics to the theories of macroeconomics. The principal vehicles of this transformation are index numbers and other similar aggregates. We shall, in every case, develop a theory of microeconomics and then show exactly how we construct aggregates in order to arrive at a theory of macroeconomics.

There are two approaches open to handling the problem of aggregates. We are given a theory of microeconomics. We can then accept the published or other aggregates such as national income statistics or index numbers and develop whatever macroeconomic system follows from these two given data. Alternatively, we can write down both our micro- and our macroeconomic systems a priori and then determine what aggregates will be consistent with these two systems.[1] The consistent aggregates may be quite different from the published aggregates. In this volume we shall adopt the former approach.

THE THEORY OF THE FIRM

Profit Maximization

It is usually assumed that entrepreneurs behave so as to maximize their profits. This assumption will be adopted at this stage, and a more general assumption will be introduced later. However, prior to the process of profit maximization, the entrepreneur must take into account his technological possibilities of production. Profit maximization is subject to the restraint of his input-output relationship or production function.

We shall denote output by x, the input of labor-hours by n, and the input of capital-hours by d. Other factors of production are taken to be fixed. The production function is

$$x = f(n, d, t, u_1) \tag{2.1.1}$$

where t = time. There may be technological changes and other dynamic elements in production which require the insertion of the time variable. The variable, u_1, is a random disturbance.

We have already performed a certain amount of implicit aggregation at this stage by lumping all output of the firm into x, all labor-hours of the firm into n, and all capital-hours of the firm into d. This lumping is strictly correct only if the firm produces one good and uses one type of each factor, or if the firm maintains certain fixed proportions among the various components of x, of n, and of d. However, we neglect in the interests of simplification the error introduced by aggregation at this stage.

The entrepreneur has much choice in the way he is to acquire the services of capital equipment in the form of machine-hours or plant-hours. He has capital of various ages with different productivities corresponding to each age group. The newer capital will almost always be more productive than the old. The decision to use a certain number of capital-hours in the most efficient way will require a choice among the utilization of various types of capital goods according to the equation

$$d = g(v, v_{-1}, v_{-2}, v_{-3}, \cdots, u_2) \tag{2.1.2}$$

[1] Simple examples of this procedure are given in the author's paper "Macroeconomics and the Theory of Rational Behavior," *Econometrica*, Vol. 14, April, 1946, pp. 93–108.

where v = gross investment and the subscripts refer to the time period of the investment. As before, u_2 is the disturbance.

We are now ready to construct the profit function which is to be maximized subject to (2.1.1). The profit function is an anticipated profit function. Economic operations involving a time-consuming production process must always be planned on the basis of anticipations referring to the period when output will be ready for market. We shall assume that our entrepreneurs make plans at the beginning of each time period, and that these plans extend over a future horizon based on anticipated variables during that horizon. Profits are thus given by

(2.1.3) $\text{an}\ \pi =$

$$\int_0^T \left[\text{an}\ p(x - \dot{h}) - \text{an}\ wn - \text{an}\ qd + \frac{d(\text{an}\ ph)}{dt} - \delta(h, u_3)\right]e^{-\rho t}\, dt$$

The explanation of this equation follows. The sign "an" represents magnitudes anticipated at time 0; π = profits; p = price of output; x = output; h = stock of inventories; $\dot{h}$ = rate of change of inventories; $x - \dot{h}$ = sales; w = wage rate; n = labor-hours; q = price of capital services; d = capital-hours; $[d(\text{an}\ ph)]/dt$ = the anticipated change in value of inventories; $\delta(h, u_3)$ = storage cost of inventories; u_3 = random disturbances; T = length of horizon; ρ = discount rate = interest + risk.

The anticipated profits (2.1.3) depend upon three variables which are entirely at the disposal of the entrepreneur, namely, n, d, h. The first two are usually taken to be his factors of production, and we can look upon h, the working capital, in much the same way. The entrepreneur is free to choose any values for n, d, h as long as he does not violate (2.1.1). We can most easily account for (2.1.1) in the maximization process by substituting directly (2.1.1) for x in (2.1.3). He will be assumed to make his choice so that an π is as large as possible; hence we shall obtain the three maximizing equations:

(2.1.4) $$\frac{\partial\ \text{an}\ \pi}{\partial n} = 0 \qquad \text{(demand for labor)}$$

(2.1.5) $$\frac{\partial\ \text{an}\ \pi}{\partial d} = 0 \qquad \text{(demand for fixed capital)}$$

(2.1.6) $$\frac{\partial\ \text{an}\ \pi}{\partial h} = 0 \qquad \text{(demand for inventories)}$$

Let us first consider the simple cases where p, w, q, ρ are independent of the decisions of the single entrepreneur; i.e., perfect competition.

We get for this situation, identically in time t over the period of the horizon,

$$\frac{\partial x}{\partial n} = \text{an}\,\frac{w}{p} \tag{2.1.4}$$

$$\frac{\partial x}{\partial d} = \text{an}\,\frac{q}{p} \tag{2.1.5}$$

$$\frac{\partial[\delta(h, u_3)]}{\partial h} = \text{an}\,\dot{p} \tag{2.1.6}$$

The first two equations state the familiar proposition that the marginal productivity of labor (capital) equals the anticipated real price of labor (capital), in equilibrium. The third equation states that the marginal storage cost of inventories must equal the anticipated rate of change of price. Equations (2.1.4)–(2.1.6) are only the first-order conditions for profit maximization. We shall not go into all the second-order conditions in detail, except to note that we must also have

$$\frac{\partial^2 x}{\partial n^2} < 0 \qquad \frac{\partial^2 x}{\partial d^2} < 0 \qquad \frac{\partial^2[\delta(h, u_3)]}{\partial h^2} > 0$$

There must be diminishing marginal productivity and increasing marginal storage costs under perfect competition.

In order to test the model (2.1.1)–(2.1.6), we must specify more precisely the functions involved. Furthermore, anticipated values are never observed; they exist only in a subjective sense in the entrepreneur's mind. We must show how the anticipations are formed. The anticipated values of economic magnitudes vary, of course, with the psychology of the individual who makes the anticipations, but some objective elements other than individual psychology must play a role. The simplest set of objective data which can form a basis for anticipated values is given by the most recently realized values of the variable in question. For example, the immediate past level, rate of change, acceleration, etc., of prices would be a likely set of data on which to form expectations of future prices. The same is true of wages and the costs of capital goods. In this study we shall adopt the method of expressing anticipations by functions of lagged variables plus random perturbations which account for the subjective variations.

We must also specify the forms of the unknown functions $f(n, d, t, u_1)$, $g(v, v_{-1}, v_{-2}, \cdots, u_2)$, $\delta(h, u_3)$. We know certain general properties of these functions a priori, such as: (1) The marginal productivities of f must be non-negative. (2) There must be diminishing marginal productivity at the profit-maximizing point. (3) The marginal storage cost must be non-negative, etc. As a final choice, however, we can do little more than choose those functions which are consistent with the data and which do not unnecessarily complicate our model. We shall make the following choices:

(2.1.1a) $$x = A\, n^{\alpha_1} d^{\alpha_2} e^{\alpha_3 t + \alpha_4 t^2} \cdot u_1$$

(2.1.2a) [2] $$d = \beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2$$

(2.1.7a) $$\delta(h, u_3) = \gamma_1 + \gamma_2 h + \gamma_3 h^2 + u_3$$

where k_{-1} = the stock of fixed capital existing at the end of the period just prior to the laying of plans. If we define i as *net* investment, we have

$$k_t = \int_{-\infty}^{t} i(\theta)\, d\theta$$

Hence k depends upon all the past values of v in a particular way. Instead of writing (2.1.2) as a function of each of the past values of v, we lump all the past values into k_{-1} and carry the current value, v, along as a separate variable.

We may substitute (2.1.2a) into (2.1.1a) to get

$$x = A\, n^{\alpha_1}(\beta_1 + \beta_2 k_{-1} + \beta_3 v + u_3)^{\alpha_2} e^{\alpha_3 t + \alpha_4 t^2} \cdot u_1$$

from which it follows that

$$\frac{\partial x}{\partial k_{-1}} = \alpha_2 \beta_2 A\, n^{\alpha_1}(\beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2)^{\alpha_2 - 1} e^{\alpha_3 t + \alpha_4 t^2} \cdot u_1$$

$$\frac{\partial x}{\partial v} = \alpha_2 \beta_3 A\, n^{\alpha_1}(\beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2)^{\alpha_2 - 1} e^{\alpha_3 t + \alpha_4 t^2} \cdot u_1$$

$$\frac{\dfrac{\partial x}{\partial k_{-1}}}{\dfrac{\partial x}{\partial v}} = \frac{\beta_2}{\beta_3}$$

Thus the ratio of parameters of (2.1.2a), β_2/β_3, is the ratio of the marginal productivity of old capital to that of new capital. Since we think that

[2] We might expect $\beta_1 = 0$ since there cannot be any services derived from capital if $k_{-1} = 0$ and $v = 0$. However, (2.1.2a) is merely a linear approximation of a more complicated function. The linear approximation may become much closer if we allow ourselves the freedom of one more parameter such as β_1.

new capital is technologically superior to old capital, we should expect $\beta_2/\beta_3 < 1$.

Let us substitute (2.1.1a), (2.1.2a), and (2.1.7a) into (2.1.4), (2.1.5), and (2.1.6). The result is

$$\frac{\partial x}{\partial n} = \alpha_1 \frac{x}{n} = \text{an}\, \frac{w}{p} \tag{2.1.4a}$$

$$\frac{\partial x}{\partial d} = \alpha_2 \frac{x}{\beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2} = \text{an}\, \frac{q}{p} \tag{2.1.5a}$$

$$\frac{\partial[\delta(h, u_3)]}{\partial h} = \gamma_2 + 2\gamma_3 h = \text{an}\, \dot{p} \tag{2.1.6a}$$

These equations can be rewritten as

$$\text{an}\, wn = \alpha_1(\text{an}\, px) \tag{2.1.4a}$$

$$\beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2 = \alpha_2 \left(\text{an}\, \frac{px}{q}\right) \tag{2.1.5a}$$

$$h = -\frac{\gamma_2}{2\gamma_3} + \frac{1}{2\gamma_3}(\text{an}\, \dot{p}) \tag{2.1.6a}$$

We must now introduce our conventions with regard to the transformation from anticipated values to observable values. We shall treat wages differently from other prices. We shall take anticipated wages equal to prevailing wages. For the combined variables, an px, and an (px/q) we shall form a linear combination of current and past values.

$$\text{an}\, px = \alpha_1' + \alpha_2' px + \alpha_3'(px)_{-1} + u_4 \tag{2.1.8}$$

$$\text{an}\, \frac{px}{q} = \alpha_1'' + \alpha_2'' \frac{px}{q} + \alpha_3'' \left(\frac{px}{q}\right)_{-1} + u_5 \tag{2.1.9}$$

The anticipated rate of change of prices, an $\dot{p}$, will be treated separately in the form [3]

$$\text{an}\, \dot{p} = \alpha_1''' + \alpha_2'''\Delta p + \alpha_3'''(\Delta p)_{-1} + u_6 \tag{2.1.10}$$

Our basic equations are now

$$wn = \alpha_1\alpha_1' + \alpha_1\alpha_2'\, px + \alpha_1\alpha_3'(px)_{-1} + \alpha_1 u_4 \tag{2.1.4b}$$

$$\begin{aligned} \beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2 \\ = \alpha_2\alpha_1'' + \alpha_2\alpha_2'' \left(\frac{px}{q}\right) + \alpha_2\alpha_3'' \left(\frac{px}{q}\right)_{-1} + \alpha_2 u_5 \end{aligned} \tag{2.1.5b}$$

[3] Throughout this volume, the convention, $\Delta x = x - x_{-1}$, for any variable x, will be adopted.

$$(2.1.6b) \quad h = -\frac{\gamma_2}{2\gamma_3} + \frac{1}{2\gamma_3}\alpha_1''' + \frac{1}{2\gamma_3}\alpha_2'''\Delta p + \frac{1}{2\gamma_3}\alpha_3'''(\Delta p)_{-1} + \frac{1}{2\gamma_3}u_6$$

All the variables of (2.1.4b)–(2.1.6b), except the disturbances, are measurable.

It must be remembered that the profit-maximizing unit is the individual firm, and that equations (2.1.4b)–(2.1.6b) must be applied only to the individual firm. Our goal, however, is to obtain equations which describe behavior patterns in the economy as a whole; therefore we must aggregate the equations for the individual firm. It will be noticed that the above equations are all linear; this condition greatly simplifies the aggregation problem. It is possible to establish a general rule for aggregating linear equations. Suppose that a relation for the ith individual or firm [4] is

$$(2.1.11) \qquad a_{1i}x_{1i} + a_{2i}x_{2i} + \cdots + a_{ni}x_{ni} + a_{0i} = 0 \quad i = 1, 2, \cdots, m$$

One type of aggregation is simple summation, such as

$$\sum_{i=1}^{m} (a_{1i}x_{1i} + a_{2i}x_{2i} + \cdots + a_{ni}x_{ni} + a_{0i}) = 0$$

The variable x_{ji} can be of two distinct types. (1) The x_{ji} for all i may have the same dimensions, such as dollars. (2) The x_{ji} may have different dimensions for different i, such as physical output. For case (1) we define

$$(2.1.12) \qquad \bar{a}_j = \frac{\sum_{i=1}^{m} a_{ji}x_{ji}}{\sum_{i=1}^{m} x_{ji}} \qquad \sum_{i=1}^{m} x_{ji} = X_j$$

Hence the term $\sum_{i=1}^{m} a_{ji}x_{ji}$ in the summation may be replaced by $\bar{a}_jX_j$.

The $\bar{a}_j$ are weighted arithmetic means of the individual a_{ji}, the weights being the individual x_{ji}. The parameters $\bar{a}_j$ are constants only if the distributions of the x_{ji} do not change; but, if the distributions change only slightly or gradually, it may be assumed with little cause for error

[4] In this example x_{ji} refers to any economic variable and not to output in particular.

that the $\bar{a}_j$ are constants. For case (2), the above procedure must be modified. We now have

$$\sum_{i=1}^{m} a_{ki}x_{ki} = \bar{a}_k X_k \tag{2.1.13}$$

where $X_k \neq \sum_{i=1}^{m} x_{ki}$, in general. The variable X_k is an aggregate, say a quantity index; but it is not the sum of the individual quantities for the reason that this sum is not meaningful when each of the components is reckoned in different units. In practice, we choose X_k equal to the published index or some other aggregate which is specified a priori.[5] The parameter $\bar{a}_k$ is the proportionality factor between the a priori aggregate and the aggregate obtained as a weighted sum of individual quantities with a_{ki} as the weights. Some examples computed from actual data have shown the $\bar{a}_k$ to be fairly stable over twenty-year time periods.[6]

Case (2) is by no means arbitrary, although it may seem to be so at first glance. Equation (2.1.13) states that a weighted sum of individual quantities is proportional to the customary indexes. Most of the published indexes are also weighted sums of the same quantities. The weights for the published indexes are not computed according to any rules based on the structural parameters of the system, but are computed according to certain intuitive ideas about what the indexes ought to show. The components in both aggregates include the same quantities; the only difference is in the weights. Even if the weights are not identical, the two aggregates may come very close to being proportional, as, in fact, they are in certain tested examples.[7] What is true for quantity indexes is true in the same way for other aggregates whose elements are not couched in identical units.

Applying these general principles of aggregation in linear systems to (2.1.4b)–(2.1.6b), we obtain

[5] For example, if x_{ki} = real output of the ith firm, X_k is the sum of money values of individual outputs deflated by the appropriate price indexes.

[6] See the author's "A Post-Mortem on Transition Predictions of National Product," *Journal of Political Economy*, Vol. 54, August, 1946, pp. 289–308.

[7] A. L. Bowley has shown why variations in weights often do not have much influence on averages. See A. L. Bowley, *Elements of Statistics*, P. S. King and Son, Ltd., London, 1926, pp. 86–94. S. S. Wilks has shown that if we form two linear aggregates of a set of correlated variables, using different weights, the correlation coefficient between the two aggregates is distributed with mean approaching unity as the number of variables increases, provided that the weights have non-zero means and are independent of the correlation coefficients between the variables. See S. S. Wilks, "Weighting Systems for Linear Functions of Correlated Variables when There Is No Dependent Variable," *Psychometrika*, Vol. 3, No. 1, March, 1938, p. 27.

(2.1.4c) $$\mathsf{W} = \epsilon_0 + \epsilon_1 pX + \epsilon_2 (pX)_{-1} + u_7$$

(2.1.5c) $$V = \zeta_0 + \zeta_1 \frac{pX}{q} + \zeta_2 \left(\frac{pX}{q}\right)_{-1} + \zeta_3 K_{-1} + u_8$$

(2.1.6c) $$H = \eta_0 + \eta_1 \Delta p + \eta_2 (\Delta p)_{-1} + u_9$$

where W = aggregate wage bill, pX = aggregate value of output, V = aggregate gross real investment, K = aggregate stock of real capital, H = aggregate stock of real inventories, q = price index of capital goods, p = price index of output. All these aggregates can be measured from the published tables on national income for the entire economy, or for some appropriate subsection, and from the published price indexes.

Equation (2.1.5c) is finally derived with V = gross real investment as the relevant variable, but it would have been equally valid to have derived it with I = net real investment as the dependent variable. To show this point, let us reconsider (2.1.2a):

(2.1.2a) $$d = \beta_1 + \beta_2 k_{-1} + \beta_3 v + u_2$$

The variable d is called the input of capital-hours into the production process. This variable can also be interpreted as the amount of fixed capital used up in the production process; i.e., real depreciation. In fact, d is measured from actual accounting records of depreciation for the purposes of empirical work in this book. If d represents the real value of depreciation, we may write the fundamental identity

(2.1.14) $$i + d = v$$

where i = net real investment for the individual firm. This leads, by substitution in (2.1.2a), to

(2.1.2a*) $$d = \beta_1 + \beta_2 k_{-1} + \beta_3 (i + d) + u_2$$

$$d = \frac{1}{1 - \beta_3} + \frac{\beta_2}{1 - \beta_3} k_{-1} + \frac{\beta_3}{1 - \beta_3} i + \frac{u_2}{1 - \beta_3}$$

We can now go through the same steps as above, replacing (2.1.2a) everywhere by (2.1.2a*). The final result will be

(2.1.5c*) $$I = \zeta_0^* + \zeta_1^* \frac{pX}{q} + \zeta_2^* \left(\frac{pX}{q}\right)_{-1} + \zeta_3^* K_{-1} + u_8$$

It should be remarked that the aggregation was taken over the set of individual firms, but this was not necessary. We could have performed double summations over both firms and products or factors simultaneously. The model was simplified by writing one product, one type of labor, and one type of capital for each firm, but this simplification was not essential. The same equations hold in the case in which the single firm produces more than one product and uses more than one type of labor and capital.

Another observation on the system developed thus far is that equations (2.1.4c)–(2.1.6c) are partially reduced forms which contain several substitutions. The fundamental structural equations of the system are (2.1.1a), (2.1.2a), (2.1.3), (2.1.4), (2.1.5), (2.1.6), (2.1.7a), (2.1.8), (2.1.9), (2.1.10). It is perfectly legitimate to eliminate (2.1.3), the profit function, because it is merely a definitional equation which defines a new variable, an π. Unless we are particularly interested in measuring this variable, we can reduce the dimensionality of the system by one degree. We eliminate one equation and one variable. Equations (2.1.4)–(2.1.6), as stated, are in terms of anticipated values which are not measurable. We have substituted (2.1.8)–(2.1.10) into (2.1.4)–(2.1.6) in order to transform the non-measurable variables into measurable variables. This step is necessary if we are to carry out statistical calculations. Finally, for a technical reason, (2.1.1a) and (2.1.2a) have been eliminated by substitution. All the equations of the system except (2.1.1a) can be written as linear functions of their parameters when the statistical variables are given in arithmetic form, while (2.1.1a) can be written as a linear function of its parameters only when the statistical variables are written in logarithmic form. The method of maximum likelihood estimation in equation systems becomes extremely complicated if variables of the joint probability distribution occur in some equations in logarithmic units and in others in natural units. We have simplified the system by substitution of (2.1.1a) and (2.1.2a) into (2.1.4)–(2.1.6). Equation (2.1.2a) has also been substituted along with (2.1.1a) because the variable d occurs in (2.1.1a) and (2.1.2a) in such a way that this substitution is very simple and because the parameters of (2.1.2a) are not of particularly great importance by themselves. Similarly (2.1.7a) has also been substituted. We have no data which would enable us to measure δ; hence we have eliminated this variable from the system. All these substitutions combined lead to (2.1.4c)–(2.1.6c).

There are obviously limits to the extent to which structural equations should be eliminated from the system. If the objectives of the statistical analysis are clearly defined before estimation begins, it is possible to set limits to the amount of elimination that is permissible. Elimination of structural equations reduces the amount of information that can be obtained from the statistical models. Elimination should not be carried beyond the point which leaves just enough information in the model to realize the objectives of the investigation.

The above model was developed under certain restrictive assumptions, but we can now begin to generalize and observe the effects on the final structural equations. In the first place, the variables w, q, p were taken to be independent of the entrepreneurial maximizing decisions. It may not be erroneous to assume that, on the whole, w and q, the prices paid for the factors of production, are independent of the maximizing decisions of the individual firm. But the same cannot be said of p, the price of output. The basic assumption of the theory of imperfect competition is that the firm is faced by an entire demand schedule rather than by a single market price. The demand schedule facing the firm is a relationship between price and output, along with other variables, perhaps. We may write

$$\text{an}\, p = p(x, p_{-1}, p_{-2}, \cdots, u_{10}) \tag{2.1.15}$$

where the lagged prices show the development of p up to the period when plans are laid. First, neglect the possibility of making inventory gains. Then, when we develop the equilibrium conditions for profit maximization, we must replace

$$\frac{\partial x}{\partial n} = \text{an}\, \frac{w}{p} \tag{2.1.4}$$

$$\frac{\partial x}{\partial d} = \text{an}\, \frac{q}{p} \tag{2.1.5}$$

by

$$\frac{\partial x}{\partial n} = \text{an}\, \frac{w}{p\left(1 - \frac{1}{\eta}\right)} \tag{2.1.16}$$

$$\frac{\partial x}{\partial d} = \text{an}\, \frac{q}{p\left(1 - \frac{1}{\eta}\right)} \tag{2.1.17}$$

where η = elasticity of demand.[8] If the relationship between an p and x is such that η = constant, no serious problem is involved, because the constant $(1 - 1/\eta)$ will then be absorbed with the other constant parameters, and equations (2.1.4c), (2.1.5c) are not changed in form, although the parameters of these equations, ϵ_i and ζ_i, will have differ-

[8] $\eta = -\dfrac{\text{an}\, p}{x} \dfrac{\partial x}{\partial\, \text{an}\, p}$.

ent meanings. In the case of the production function we did not know the exact form of the function on the basis of a priori reasoning; hence we selected a function which simplified the results and which was reasonable in terms of the observable data. We must do the same thing in this case, where we do not know, in advance, the specific form of the function (2.1.15). We select, as an approximation, a function which possesses those elementary properties which we know must be present in (2.1.15) and which is consistent with the data. Many demand studies have found that constant-elasticity demand schedules do not conflict with the data.[9]

When the possibility of making capital gains from inventory speculation is properly taken into account, however, the generalization to conditions of imperfect competition has a more essential influence on the equations of the system. If the relation between anticipated prices and output is given by the formula in (2.1.15), the logical interpretation of the formula for an $\dot{p}$ (associated with capital gains) would be

$$\text{an } \dot{p} = \frac{\partial p}{\partial x}\dot{x} \tag{2.1.18}$$

From the production function (2.1.1), $\dot{x}$ is given by

$$\dot{x} = \frac{\partial f}{\partial n}\dot{n} + \frac{\partial f}{\partial d}\dot{d} + \frac{\partial f}{\partial t}$$

The maximization of profits (2.1.3) subject to (2.1.1), (2.1.15), and (2.1.18) is a true problem in the calculus of variations because both derivatives and levels of variables with respect to which profits are to be maximized enter the profit integral. The only derivative under the integral sign for perfect competition is an $\dot{p}$, which is taken as given in the maximization process for the single firm.

The well-known Euler equations for the maximum of an integral in the calculus of variations take the form

$$\text{an } p\frac{\partial x}{\partial n}\left(1 - \frac{1}{\eta}\right) - \text{an } w - \frac{\partial p}{\partial x}\frac{\partial x}{\partial n}(\dot{h} - \rho h) = 0 \tag{2.1.19}$$

$$\text{an } p\frac{\partial x}{\partial d}\left(1 - \frac{1}{\eta}\right) - \text{an } q - \frac{\partial p}{\partial x}\frac{\partial x}{\partial d}(\dot{h} - \rho h) = 0 \tag{2.1.20}$$

$$\frac{\partial[\delta(h, u_3)]}{\partial h} = \text{an } \dot{p} \tag{2.1.6}$$

[9] A typical example is found in a paper by J. R. N. Stone, "The Analysis of Market Demand," *Journal of the Royal Statistical Society*, Vol. CVIII, parts III–IV, 1945, pp. 286–391.

for the problem of maximizing (2.1.3) subject to (2.1.1), (2.1.15) and (2.1.18). The first-order condition with respect to inventory variation remains the same, but an essentially new term appears in the first-order conditions associated with labor and capital variation. The size of the new term depends on the divergence between the change in inventories and a percentage of the stock of inventories, the percentage being the discount rate. The more specific form of the production function (2.1.1a) and the elasticity transformation enable us to write (2.1.19) and (2.1.20) in the forms

$$\text{(2.1.19a)} \qquad \text{an}\; wn = \alpha_1 \left(1 - \frac{1}{\eta}\right) \text{an}\; px + \frac{\alpha_1}{\eta}\, \text{an}\; p\, (\dot{h} - \rho h)$$

$$\text{(2.1.20a)} \qquad d = \alpha_2 \left(1 - \frac{1}{\eta}\right) \text{an}\, \frac{px}{q} + \frac{\alpha_2}{\eta}\, \text{an}\, \frac{p}{q}\, (\dot{h} - \rho h)$$

These equations can be written in terms of observed aggregates [10] by an application of preceding methods. In the empirical work, equations of the form (2.1.19a) and (2.1.20a) with the extra terms on the right-hand side have not yet been used. The main obstacle for empirical work here is to find satisfactory data on the size of the discount rate ρ.

While equation (2.1.6) remains in the same form as in the perfectly competitive theory, the form to be used in statistical calculation is changed because the function used to convert an $\dot{p}$ into observed quantities is different. It is given by (2.1.18) instead of (2.1.10). In aggregative discrete variables a linear approximation will now be

$$\text{(2.1.6c*)} \quad H = \eta_0{}^* + \eta_1{}^* p_{-1} + \eta_2{}^* p_{-2} + \eta_3{}^* X + \eta_4{}^* X_{-1} + u_9{}^*$$

which can be considered as an alternative to (2.1.6c).

There is yet another variable which was assumed not to vary in the maximization process, namely ρ, the discount rate. Kalecki [11] has developed a very interesting theory concerning the relation between ρ and the decision to invest in capital goods. He has assumed that risk grows as the size of the investment grows, or perhaps as the size of the investment in relation to the size of the firm grows, or some other relation which shows, in general, the extent to which the entrepreneur risks his assets against the repayment of borrowed funds. The risk of losing all his assets will grow as the amount of money borrowed for capital expansion grows, relative to the size of the firm. We may have an equation

[10] Observed values of ρ are not easy to determine, however.

[11] "The Principle of Increasing Risk," *Economica*, New Series, Vol. IV, November, 1937, pp. 440–447.

$$\rho = \rho(qd, k_{-1}, u_{11}) \tag{2.1.21}$$

The variable d shows how much capital will be used up and therefore indicates how much borrowing will have to be undertaken for future operations. The variable k_{-1} shows the size of the firm. The appropriate variable may be d/k_{-1} with $\partial\rho/[\partial(d/k_{-1})] > 0$. Instead of k_{-1}, we may have as a variable in (2.1.21) the liquidity position of the firm (current assets minus current liabilities) to show how much of the capital needed can be covered by internal funds. Or, alternatively, ρ may depend solely on the stock of net liquid assets existing at the time of planning. Since the existing assets are predetermined from past history, the discount rate would not enter as a variable in the maximization process for this case. We are concerned only with the case like that envisaged in (2.1.21) above. Otherwise, there is no change in the model as developed thus far.

The case in which the risk factor depends upon liquidity is interesting because the relation between liquidity and investment is being stressed so much today. Let us write

$$\rho = \phi \frac{qd}{\mathsf{l}_{-1}} + u_{12} \qquad \phi > 0 \tag{2.1.21a}$$

where l_{-1} = value of current assets minus current liabilities existing at the beginning of the planning period.

The profit-maximizing equations, (2.1.4) and (2.1.6), are not affected by (2.1.21a), because this last relationship does not depend upon n or h, but there is a significant alteration in (2.1.5). Upon differentiating (2.1.3) with respect to d subject to (2.1.1a), (2.1.15), and (2.1.21a), we obtain

$$\begin{aligned} &\text{an}\, p\left(1 - \frac{1}{\eta}\right)\alpha_2 \frac{x}{d} - \text{an}\, q \\ &\quad + [\text{an}\, px - \text{an}\, wn - \text{an}\, qd + \text{an}\, \dot{p}h - \delta(h, u_3)] \cdot \left(-\theta\phi \frac{\text{an}\, q}{\mathsf{l}_{-1}}\right) = 0 \end{aligned} \tag{2.1.5a**}$$

This equation must hold for all θ in the interval $[0, T]$. Choose some θ_0 for which (2.1.5a**) is true. Algebraic manipulation leads immediately to

$$d = \frac{\alpha_2\left(1 - \dfrac{1}{\eta}\right)\text{an}\,\dfrac{px}{q}}{1 - \left[\text{an}\,\dfrac{px}{q} - \text{an}\,\dfrac{wn}{q} - d + \text{an}\,\dfrac{\dot{p}h - \delta}{q}\right]\left(\dfrac{\theta_0\phi}{\dfrac{\mathsf{l}_{-1}}{\text{an}\,\dfrac{}{q}}}\right)}$$

We can expand, linearly, the right-hand fraction to get

$$(2.1.5\text{b}^{**}) \quad d = \beta_0{}^{**} + \beta_1{}^{**} \,\text{an}\, \frac{px}{q} + \beta_2{}^{**} \left[\text{an}\, \frac{px}{q} - \text{an}\, \frac{wn}{q} - d + \text{an}\, \frac{\dot{p}h - \delta}{q} \right] + \beta_3{}^{**} \,\text{an}\, \frac{\mathfrak{l}_{-1}}{q} + u_{13}$$

The remaining steps to convert anticipated values into lagged values, microvalues into macrovalues, d into v or i, etc., are straightforward and can be carried out as before.

The equations derived from the principle of increasing risk bring no new variables except $\mathfrak{l}_{-1}$, into the system, but they change the form of the structural equations very much. We shall present in the next chapter some statistical tests of equations like (2.1.5b**), although our final complete models do not yet contain equations based on the principle of increasing risk. This theory is presented here mainly for the sake of completeness.

Utility Maximization

Thus far we have assumed that the traditional theory of profit maximization is correct, but now we progress to more general principles. Economists have recently been claiming that business firms are as much concerned about the structure of their assets as about the size of their profits. In the words of Marschak, this means that firms behave so as to have the best possible profit and loss statement and the best possible balance sheet. The nature of the firm's decision may be such that the process of obtaining an optimum profit and loss statement is independent of the process of obtaining an optimum balance sheet, although there is no fundamental principle which enables us to assume this independence. Thus far, our theory of profit maximization has been essentially based on such independence, but now we alter our assumptions.

We shall construct a model on the assumption that the firm has a preference scale according to which it chooses among different types of assets and profits

$$(2.2.1) \qquad \text{an}\, u = u(\text{an}\, \underset{0}{\overset{T}{\pi}}, \underset{0}{\overset{T}{k}}, \underset{0}{\overset{T}{h}}, \underset{0}{\overset{T}{\mathfrak{m}}}, \underset{0}{\overset{T}{\mathfrak{s}}})$$

The notation remains the same as before, but we add the variables u = utility,[12] $\mathfrak{m}$ = cash balances, $\mathfrak{s}$ = net value of securities (assets

[12] In this and later discussions u = utility is not to be confused with u = random disturbance. The latter variables will always carry a subscript.

minus liabilities in the form of securities). The variable an u is a functional, calculated over the entire horizon $(0,T)$.

For corporate enterprise we could refine (2.2.1) further by splitting profits into two categories, corporate savings and dividends. However, most of the final structural equations that we use for statistical purposes will be unaffected by such a decomposition; therefore we retain total profits as the relevant variable, although we recognize that, as a further approximation, it will be necessary to take into account the division between corporate savings and dividends.

Since (2.2.1) is a functional, the maximization equations will be very complex; in fact, too complex for our statistical theories. The statistical theory is much simpler for discrete random processes than for the continuous random processes that would be derived from the functionals. Hence we simplify the problem by approximating the variables in (2.2.1) with discrete variables. The modified form is

$$\text{(2.2.1*)} \quad \text{an}\, u = u^*(\text{an}\, \pi_1, \cdots, \text{an}\, \pi_T, k_1, \cdots, k_T, h_1, \cdots, h_T, \mathsf{m}_1, \cdots, \mathsf{m}_T, \mathsf{s}_1, \cdots, \mathsf{s}_T)$$

The variable an π is a flow and is measured during a period of time, but the other independent variables of (2.2.1*) are stocks and are measured at a point of time. Our convention is that k_θ refers to the stock of capital at the end of the period θ. The problem now is to maximize an u subject to various restraints.[13]

The restraints are

$$\text{(2.2.2)} \qquad \mathsf{m}_{\theta-1} + (1 + \text{an}\, r_\theta)\mathsf{s}_{\theta-1} + \text{an}\, \pi_\theta = \mathsf{m}_\theta + \mathsf{s}_\theta \qquad \theta = 1, 2, \cdots, T$$

$$\text{(2.2.3)} \quad \text{an}\, \pi_\theta = \text{an}\, p_\theta x_\theta - \text{an}\, w_\theta n_\theta - \text{an}\, q_\theta d_\theta + \text{an}\, (\Delta p_\theta) h_\theta - \delta(h_\theta, u_3) \qquad \theta = 1, 2, \cdots, T$$

$$\text{(2.2.4)} \qquad x_\theta = x(n_\theta, d_\theta, t, u_1) \qquad \theta = 1, 2, \cdots, T$$

$$\text{(2.2.5)} \qquad d_\theta = d(i_\theta, k_{\theta-1}, u_2) \qquad \theta = 1, 2, \cdots, T$$

$$\text{(2.2.6)} \qquad \Delta k_\theta = i_\theta \qquad \theta = 1, 2, \cdots, T$$

[13] The new theory can be instructively compared with the previous theory of profit maximization. Formerly we chose among the profits of different periods over the future horizon according to the discount factor $e^{-\rho\theta}$. By maximizing an $\pi = \int_0^T \pi(\theta)e^{-\rho\theta}\, d\theta$, we assumed that the entrepreneur exchanged the profit of one period with that of another period at the rate ρ. Now we choose among the profits or assets of different periods according to a general preference scale, u. An integral discounted at the rate, ρ, could be considered as a special case of u.

The first constraint (2.2.2) states that the liquid assets in the firm's possession at the beginning of any period plus the interest earned on securities during the period plus the earnings on operations during the period equal the liquid assets in the firm's possession at the end of the period.[14] Of course, if dividends are paid out, they must be subtracted from an π_θ in (2.2.2), but we are ignoring this complication in the present model. In (2.2.2) r = interest rate. The next constraint is simply a definition of anticipated profits for the period θ. It corresponds with the definition of profits in the old theory. Equations (2.2.2) and (2.2.3) actually hold only if we ignore the profits or losses due to net capital gains on securities held. Let us assume that there are no capital gains, except on inventories. The restraints (2.2.4) and (2.2.5) are merely representations of (2.1.1) and (2.1.2) which appeared in the previous theory of profit maximization. Finally, (2.2.6) defines net investment in terms of changes in the stock of fixed capital.[15]

In order to maximize an u subject to (2.2.2)–(2.2.6) we form the function

$$\begin{aligned}\Phi = u^* &+ \sum_{\theta=1}^{T} \lambda_{1\theta}[\mathsf{m}_{\theta-1} + (1 + \text{an}\, r_\theta)\mathsf{s}_{\theta-1} + \text{an}\, \pi_\theta - \mathsf{m}_\theta - \mathsf{s}_\theta] \\ &+ \sum_{\theta=1}^{T} \lambda_{2\theta}[\text{an}\, \pi_\theta - \text{an}\, p_\theta x_\theta(n_\theta, d_\theta, t, u_1) - \text{an}\, w_\theta n_\theta - \text{an}\, q_\theta d_\theta \\ &\qquad\qquad + \text{an}\, (\Delta p_\theta) h_\theta - \delta(h_\theta, u_3)] \\ &+ \sum_{\theta=1}^{T} \lambda_{3\theta}[d_\theta - d(i_\theta, k_{\theta-1}, u_2)] + \sum_{\theta=1}^{T} \lambda_{4\theta}(\Delta k_\theta - i_\theta)\end{aligned}$$

We have substituted (2.2.4) everywhere for x_θ and thus eliminated this variable directly. The function Φ consists of the function to be maximized and all the constraints appropriately combined with the Lagrange multipliers, $\lambda_{i\theta}$.

The necessary conditions for maximization can easily be written as

$$\text{(2.2.7)} \qquad \frac{\partial \Phi}{\partial\, \text{an}\, \pi_\theta} = \frac{\partial u^*}{\partial\, \text{an}\, \pi_\theta} + \lambda_{1\theta} + \lambda_{2\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$

[14] Instead of writing one constraint for each future time period, we could discount each constraining equation to the time of planning and add the discounted (present) values of each constraint. This would give one discounted constraining equation for the whole planning period. The final statistical equations are, however, not affected by this modification.

[15] We are assuming again that the decisions of the individual firm do not influence the market variables such as prices, wages, interest rate. If we assume imperfect competition, more constraints must be added in the form of demand equations for products and supply equations for factors of production. However, this refinement has no essential influence on the results.

(2.2.8)
$$\frac{\partial \Phi}{\partial n_\theta} = \lambda_{2\theta}\left(-\text{an } p_\theta \frac{\partial x}{\partial n_\theta} + \text{an } w_\theta\right) = 0 \qquad \theta = 1, 2, \cdots, T$$

(2.2.9)
$$\frac{\partial \Phi}{\partial d_\theta} = \lambda_{2\theta}\left(-\text{an } p_\theta \frac{\partial x}{\partial d_\theta} + \text{an } q_\theta\right) + \lambda_{3\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$

(2.2.10)
$$\frac{\partial \Phi}{\partial i_\theta} = -\lambda_{3\theta} \frac{\partial d}{\partial i_\theta} - \lambda_{4\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$

(2.2.11)
$$\frac{\partial \Phi}{\partial k_\theta} = \frac{\partial u^*}{\partial k_\theta} - \lambda_{3,\theta+1} \frac{\partial d}{\partial k_\theta} + \lambda_{4\theta} - \lambda_{4,\theta+1} = 0$$
$$\lambda_{3,T+1} = 0 \qquad \lambda_{4,T+1} = 0 \qquad \theta = 1, 2, \cdots, T$$

(2.2.12)
$$\frac{\partial \Phi}{\partial \mathsf{m}_\theta} = \frac{\partial u^*}{\partial \mathsf{m}_\theta} + \lambda_{1,\theta+1} - \lambda_{1\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$
$$\lambda_{1,T+1} = 0$$

(2.2.13)
$$\frac{\partial \Phi}{\partial \mathsf{s}_\theta} = \frac{\partial u^*}{\partial \mathsf{s}_\theta} + \lambda_{1,\theta+1}(1 + \text{an } r_{\theta+1}) - \lambda_{1\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$
$$\lambda_{1,T+1} = 0$$

(2.2.14)
$$\frac{\partial \Phi}{\partial h_\theta} = \frac{\partial u^*}{\partial h_\theta} + \lambda_{2\theta}\left(-\text{an } \Delta p_\theta + \frac{\partial \delta}{\partial h_\theta}\right) = 0 \qquad \theta = 1, 2, \cdots, T$$

Equations (2.2.7)–(2.2.14) along with (2.2.2)–(2.2.6) enable us to solve for all the relevant variables and Lagrange multipliers in terms of the initial conditions and the market variables which are taken as given by the individual firm. The variables which are influenced by the decision of the firm are an π_θ, n_θ, d_θ, i_θ, k_θ, m_θ, s_θ, h_θ. The relevant initial conditions are k_0, m_0, s_0, the assets in existence at the beginning of the planning period, and the market variables are an p_θ, an w_θ, an q_θ, an r_θ. Since the set (2.2.7)–(2.2.14) involves non-measurable magnitudes such as the utility function, we must solve the system for measurable variables in terms of other measurable variables or of non-measurable variables for which measurable variables may be substituted. This eliminates the utility function from the system of statistical equations. Our results are

$$\text{(2.2.15)} \quad \text{an}\ \pi_\theta = f_1(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.16)} \quad n_\theta = f_2(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.17)} \quad d_\theta = f_3(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.18)} \quad i_\theta = f_4(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.19)} \quad k_\theta = f_5(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.20)} \quad \mathsf{m}_\theta = f_6(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.21)} \quad \mathsf{s}_\theta = f_7(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

$$\text{(2.2.22)} \quad h_\theta = f_8(\text{an}\,\tilde{p}_\theta,\ \text{an}\,\tilde{w}_\theta,\ \text{an}\,\tilde{q}_\theta,\ \text{an}\,\tilde{r}_\theta,\ k_0,\ \mathsf{m}_0,\ \mathsf{s}_0,\ u_1,\ u_2,\ u_3)$$

The $\sim$ sign represents a vector with components $1, 2, \cdots, T$. These are the structural demand equations for the individual firm. The initial conditions and market variables are such that $\mathsf{m}_0 + (1 + \text{an}\ r_1)\mathsf{s}_0$ will always enter as a single variable in equations (2.2.15)–(2.2.22). Again, we may replace the anticipated values by the lagged and current values; we may take linear approximations to the f_i-functions; we may aggregate the linear functions over the entire economy. The result is

$$\text{(2.2.15a)} \quad \text{an}\ \Pi = \alpha_{10} + \alpha_{11}p + \alpha_{12}p_{-1} + \alpha_{13}w + \alpha_{14}q + \alpha_{15}r + \alpha_{16}K_0 + \alpha_{17}\mathsf{S}_0 + \alpha_{18}\mathsf{M}_0 + u_4$$

$$\text{(2.2.16a)} \quad N = \alpha_{20} + \alpha_{21}p + \alpha_{22}p_{-1} + \alpha_{23}w + \alpha_{24}q + \alpha_{25}r + \alpha_{26}K_0 + \alpha_{27}\mathsf{S}_0 + \alpha_{28}\mathsf{M}_0 + u_5$$

$$\text{(2.2.17a)} \quad D = \alpha_{30} + \alpha_{31}p + \alpha_{32}p_{-1} + \alpha_{33}w + \alpha_{34}q + \alpha_{35}r + \alpha_{36}K_0 + \alpha_{37}\mathsf{S}_0 + \alpha_{38}\mathsf{M}_0 + u_6$$

$$\text{(2.2.18a)} \quad I = \alpha_{40} + \alpha_{41}p + \alpha_{42}p_{-1} + \alpha_{43}w + \alpha_{44}q + \alpha_{45}r + \alpha_{46}K_0 + \alpha_{47}\mathsf{S}_0 + \alpha_{48}\mathsf{M}_0 + u_7$$

$$\text{(2.2.19a)} \quad K = \alpha_{50} + \alpha_{51}p + \alpha_{52}p_{-1} + \alpha_{53}w + \alpha_{54}q + \alpha_{55}r + \alpha_{56}K_0 + \alpha_{57}\mathsf{S}_0 + \alpha_{58}\mathsf{M}_0 + u_8$$

$$\text{(2.2.20a)} \quad \mathsf{M} = \alpha_{60} + \alpha_{61}p + \alpha_{62}p_{-1} + \alpha_{63}w + \alpha_{64}q + \alpha_{65}r + \alpha_{66}K_0 + \alpha_{67}\mathsf{S}_0 + \alpha_{68}\mathsf{M}_0 + u_9$$

$$\text{(2.2.21a)} \quad \mathsf{S} = \alpha_{70} + \alpha_{71}p + \alpha_{72}p_{-1} + \alpha_{73}w + \alpha_{74}q + \alpha_{75}r + \alpha_{76}K_0 + \alpha_{77}\mathsf{S}_0 + \alpha_{78}\mathsf{M}_0 + u_{10}$$

$$\text{(2.2.22a)} \quad H = \alpha_{80} + \alpha_{81}p + \alpha_{82}p_{-1} + \alpha_{83}w + \alpha_{84}q + \alpha_{85}r + \alpha_{86}K_0 + \alpha_{87}\mathsf{S}_0 + \alpha_{88}\mathsf{M}_0 + u_{11}$$

The final set of equations for this model is written in a very general form. All the market variables need not appear in each equation. Some equations may be more sensitive than others to changes in wages or prices. In one equation lags may be of higher order than in another; even higher-order lags than we have written may appear for some of the variables in (2.2.15a)–(2.2.22a). We can tell the quantitative importance of the several variables in particular equations only after we have examined the relevant data. Only a few of the variables will appear in any single equation. Our theory merely indicates the possibilities of certain variables appearing in a specific equation. It does not tell us the quantitative importance of the variables.

There are some close relationships between (2.2.7)–(2.2.14) and (2.1.4)–(2.1.6). In the first place, (2.1.4) and (2.2.8) are identical. As far as variation of labor-input is concerned, profit maximization and utility maximization lead to the same demand equations. This fact is a result of the particular structure of the utility function. The u^*-function depends upon n only through its dependence upon an π. If u^* depended upon n explicitly as well as implicitly (i.e., via an π), the two theories would no longer be the same. However, there is no reason to assume that u^* depends upon n directly as well as through an π. It also follows that (2.1.5) and (2.2.9) differ only because u^* depends explicitly upon k. We should find that (2.1.5) and (2.2.9) would be identical if $\lambda_{3\theta} \equiv 0$ for all $\theta = 1, 2, \cdots, T$. From (2.2.10) this would imply $\lambda_{4\theta} \equiv 0$ for all $\theta = 1, 2, \cdots, T$. These two equations with (2.2.11) would, in turn, lead to $\partial u^*/\partial k_\theta \equiv 0$ for all $\theta = 1, 2, \cdots, T$, which is equivalent to stating that u^* does not depend upon k. If we take into account the structure of assets (i.e., the balance sheet items k, m, s, h), there is a significant difference between the theory of profit maximization and that of utility maximization.

According to the theory of utility maximization, the entrepreneur periodically revises his plans. The periodic revisions come long before the end of any horizon is reached; hence, the demand equations hold as *observable* behavior patterns only for θ small, say six months or one year. For any statistical calculations, we use the values of six months or one year ago for the initial conditions and the current value for the variables without lag.

We could modify the theory of utility maximization by assuming imperfect competition and increasing risk; however, the general form of the equations would not be seriously affected if the additional restrictions were the same as those introduced into the theory of profit maximization.

Taxation

We have thus far set up a model of entrepreneurial behavior without introducing the concept of business taxes. Is our theory unrealistic if it does not take taxes into account? Most economists would answer, yes; in fact, in many private discussions with the author, economists have nearly always raised the question of taxes, especially in relation to investment. A great amount of space in the current literature is devoted to an analysis of the effect of taxation on investment and other business planning. Nearly always the conclusion is reached that lower taxes stimulate investment.

First let us adopt the model of profit maximization and add the proposition that business firms pay income or profit taxes. If the income tax is a flat rate, say k per cent, we should find immediately that the structure of the profit-maximizing equations is unaffected by the size of $k(0 \leq k < 100$ per cent). What difference does it make whether firms maximize profits or $(100 - k)$ per cent of profits? The maximizing equations remain invariant. A more general proposition is: *If the marginal tax rate is always less than unity, the profit-maximizing equations are independent of the level of profit taxes.* Formally we have

$$\text{an}\ \pi - \psi(\text{an}\ \pi) = \text{an}\ \pi^* \tag{2.3.1}$$

where $\psi(\text{an}\ \pi)$ = the tax schedule showing how taxes vary with business income, and an π^* = the anticipated profits *after* taxes. For a maximum an π^* we have

$$d(\text{an}\ \pi^*) = d(\text{an}\ \pi) - \frac{d\psi}{d(\text{an}\ \pi)}\, d(\text{an}\ \pi) = 0 \tag{2.3.2}$$

or

$$\left(1 - \frac{d\psi}{d(\text{an}\ \pi)}\right) d(\text{an}\ \pi) = 0$$

If

$$\frac{d\psi}{d(\text{an}\ \pi)} < 1$$

then

$$d(\text{an}\ \pi) = 0$$

Furthermore, as long as $d\psi/[d(\text{an}\ \pi)] < 1$, the second-order conditions for a maximum of an π^* imply second-order conditions for a maximum of an π. The inequality states precisely that the marginal tax rate is less than unity; i.e., out of an extra dollar of profits, less than an extra dollar goes to the public treasury in the form of taxes. Our tax system has this property, of course.

If we develop the demand for labor, the demand for capital goods, the demand for inventories, etc., from the principle of profit maximization, these demand equations do not depend explicitly upon profit taxes. This does not mean that profit taxes have no influence on investment or on the system, but it does mean that a tax variable does not appear in the behavior equation relating to investment. Taxes may (they even do) enter into other equations of the system, such as the demand for consumer goods. Consequently, a change in the tax rates will have some influence on the system and will affect investment as well as other variables. This means that a reduced form, as distinguished from a structural behavior equation, may show a relationship between investment and taxes, whereas the profit-maximizing equations of the system do not show any such relationship.

What is true of profit taxes is not true of excise taxes. The latter taxes amount, in effect, to an increase in prices which the entrepreneur must transfer to the government. Consequently the variable px in the demand equations for labor or capital should be replaced by $px - \mathbf{e}$, where $\mathbf{e}$ = excise taxes. Otherwise the analysis is unchanged.

Taxes have a different influence on the system if we adopt the principle of utility maximization instead of the principle of profit maximization. We have shown above that, except for special cases, the theory of utility maximization does not imply profit maximization. Particularly, the demand equation for new fixed capital is changed in the theory of utility maximization, and this change makes it necessary to introduce profit taxes explicitly into the behavior equation expressing the demand for new fixed capital. In (2.2.1*) and (2.2.2) we must write $\text{an}\ \pi_\theta - \psi(\text{an}\ \pi_\theta)$ in place of $\text{an}\ \pi_\theta$. The parameters of ψ are the autonomously set tax rates determined by legislative action. We shall classify these parameters as exogenous. We previously solved the set of maximizing equations (2.2.7)–(2.2.14) and the constraints (2.2.2)–(2.2.6) for all the endogenous variables in terms of predetermined variables (initial conditions) and market variables (given to the individual firm). Now we have one new set of exogenous elements in the system, the parameters of $\psi(\text{an}\ \pi_\theta)$, which will also appear on the right-hand side of (2.2.7a)–(2.2.14a) and (2.2.7*b*)–(2.2.14b). Tax rates will enter the behavior equations as exogenous variables. Otherwise the analysis is unchanged.

Alternative Theories

The above frameworks of profit maximization and utility maximization may seem formidable as compared with actual human behavior patterns. It may be more satisfactory to ascribe much simpler rules of behavior to entrepreneurs. We may be inspired by some intuitive ideas which lead to simple models that can be tested against the factual data. For example, the capitalist system that we are trying to describe in the statistical equations is often called the profit system. Capitalism and private profits are considered synonymous. Profits are considered the prime mover in the system; hence, when expectations of profits are high, we should expect entrepreneurs to expand, and, when expectations of profits are low, we should expect entrepreneurs to contract. This can be formulated mathematically by writing

$$i = f(\text{an } \pi) \tag{2.4.1}$$

where f is a non-decreasing function of an π.[16] Let us, as before, relate an π to the past history of π and disturbances, such as in

$$\text{an } \pi = g(\pi, \pi_{-1}, \pi_{-2}, \cdots, u_1) \tag{2.4.2}$$

If we substitute (2.4.2) into (2.4.1), we get

$$i = f(\pi, \pi_{-1}, \pi_{-2}, \cdots, u_1) \tag{2.4.1a}$$

This theory is simple to test.

Another version of this theory has often appeared in the form

$$i = h\left[\frac{\pi}{k}, \left(\frac{\pi}{k}\right)_{-1}, \left(\frac{\pi}{k}\right)_{-2}, \cdots, u_2\right] \tag{2.4.3}$$

where the rate of profit, π/k, rather than the total profit, π, is the mainspring of activity in the system.

Marschak has suggested a very reasonable rationalization for these intuitive models. Instead of assuming that the single entrepreneur takes market prices, wages, and other costs as given (unaffected by his own action) and then adjusts the factors of production so as to acquire the largest possible profit, let us assume that he takes the market rate of profit as given and then adjusts the size of his firm so as to acquire the largest possible profit consistent with the prevailing market rate. Suppose that the market is such that one can expect to earn an ρ per cent on capital; i.e., if the size of the fixed capital is k, the total anticipated profit is given by

$$\text{an } \pi = \text{an } \rho k \tag{2.4.4}$$

We shall, however, accomplish nothing with this model if we assume that the entrepreneur now proceeds to maximize an π with an ρ given by the state of the market and uninfluenced by the decisions of the

[16] In this section π will be regarded as "real" profits.

single firm. We find that the bigger the firm's capital (k = capital), the bigger the anticipated profits. No firm of finite size in this model will yield maximum profits. Some limitational factors must be introduced. Why does a single firm stop investing, at a finite limit, usually very small? The obvious answer is the existence of risk. An entrepreneur cannot expect to obtain the rate an ρ indefinitely and over an arbitrarily wide market. The larger the investment outlay at the profit rate, an ρ, the more risky is the venture and the less chance the entrepreneur has of realizing his expectations. We write this symbolically as

$$\text{risk} = f(i) \qquad f'(i) > 0 \tag{2.4.5}$$

Let the risk be measured in the same units as profits, i.e., as deductions from profits. Hence (2.4.4) must be altered by making allowance for (2.4.5) to get

$$\text{an } \pi = \text{an } \rho(k) - f(i) \qquad \Delta k = i \tag{2.4.6}$$

or

$$\text{an } \pi = \text{an } \rho(k_{-1} + i) - f(i)$$

Let us now maximize an π with respect to i. Our result is

$$\frac{\partial \text{ an } \pi}{\partial i} = \text{an } \rho - f'(i) = 0 \tag{2.4.7}$$

$$\frac{\partial^2 \text{ an } \pi}{\partial i^2} = -f''(i) < 0 \tag{2.4.8}$$

The second condition is a special case of the principle of increasing risk. A simple form of (2.4.5) is

$$f(i) = \alpha_0 + \alpha_1 i + \alpha_2 i^2 \qquad \alpha_2 > 0 \tag{2.4.5a}$$

Upon substituting (2.4.5a) into (2.4.7) we get

$$\text{an } \rho - \alpha_1 - 2\alpha_2 i = 0 \tag{2.4.7a}$$

or

$$i = \alpha_1{}' + \alpha_2{}' \text{ an } \rho$$

The anticipated market return on capital will now be written as a linear combination of current and past return:

$$\text{an } \rho = \beta_0 + \beta_1\rho + \beta_2\rho_{-1} + u_3 \tag{2.4.9}$$

We may combine (2.4.7a) and (2.4.9) to get

$$i = \beta_0{}' + \beta_1{}'\rho + \beta_2{}'\rho_{-1} + u_4 \tag{2.4.7b}$$

This is our investment schedule for the single firm. If the observed market rates of profit are the same for all firms in all industries, we can sum both sides of (2.4.7b) to get the equation for the economy as a whole:

$$\Sigma i = I = \Sigma\beta_0' + \rho\Sigma\beta_1' + \rho_{-1}\Sigma\beta_2' + \Sigma u_4 \tag{2.4.7c}$$

$$I = \bar{\beta}_0 + \bar{\beta}_1\rho + \bar{\beta}_2\rho_{-1} + u_5$$

In this case, the market rate of profit, ρ, can be written as the aggregate (or average) profit divided by the aggregate (or average) capital. Thus (2.4.7c) will be of the form

$$I = \bar{\beta}_0 + \bar{\beta}_1 \frac{\Pi}{K} + \bar{\beta}_2 \left(\frac{\Pi}{K}\right)_{-1} + u_5 \tag{2.4.7c}$$

There are good data on all the variables of (2.4.7c), and tests can easily be made of the validity of this model.

If the observed market rates of profit vary for each industry $1, 2, \cdots, n$, the aggregation of (2.4.7c) will appear as

$$I = \sum_{jl} (\beta_0')_{jl} + \sum_{j=1}^{n} \rho_j \sum_{l=1}^{m_j} (\beta_1')_{jl} + \sum_{j=1}^{n} (\rho_{-1})_j \sum_{l=1}^{m_j} (\beta_2')_{jl} + \sum_{jl} (u_4)_{jl} \tag{2.4.7d}$$

where j denotes the industry and l denotes the firm. There are m_j firms in the jth industry. The term $\sum_{j=1}^{n} \rho_j \sum_{l=1}^{m_j} (\beta_1')_{jl}$ is a linear combination of all the different profit rates of each industry. It is thus proportional to a weighted average of all the individual rates. The term Π/K, where Π = aggregate profits and K = aggregate capital, is a weighted average of the individual profit rates.[17] We expect that the relation

$$\sum_{j=1}^{n} \rho_j \sum_{l=1}^{m_j} (\beta_1')_{jl} = \lambda \frac{\Pi}{K} \tag{2.4.10}$$

will hold, at least approximately. If (2.4.10) does hold, then (2.4.7c) and (2.4.7d) are equivalent.

[17] $\dfrac{\sum \left(\dfrac{\pi_i}{k_i}\right) k_i}{\Sigma k_i} = \dfrac{\Sigma \pi_i}{\Sigma k_i} = \dfrac{\Pi}{K}.$

Various modifications of this model can still be made. We may change the risk function to

$$\text{risk} = f^*\left[i, \left(\frac{\mathsf{l}}{q}\right)_{-1}\right] \tag{2.4.5*}$$

where l_{-1} is the value of the stock of liquid assets on hand at the time of planning. l_{-1} is a predetermined variable which shows the extent to which the investment outlay, i, can be covered by internal funds. A high liquidity position has the effect of lessening the risk factor. If we were to introduce (2.4.5*) with the convenient linearizations, we should get, in place of (2.4.7c),

$$I = \bar{\beta}_0^* + \bar{\beta}_1^* \frac{\Pi}{K} + \bar{\beta}_2^* \left(\frac{\Pi}{K}\right)_{-1} + \bar{\beta}_3^* \left(\frac{\mathsf{L}}{q}\right)_{-1} + u_5^* \tag{2.4.7c*}$$

$$\bar{\beta}_3^* > 0$$

Professor Schumpeter has remarked that tests of (2.4.7c) or (2.4.7c*) should not use data on aggregate Π where negative profits are balanced against positive profits. He suggests calculation of the profit rate by use of the profits and capital of only those firms making positive profits. He believes that entrepreneurs, in deciding upon an expansion of their plant (or on a contraction), look only at the bright side of market possibilities. This is again a new model.

The simplest theory of investment is that it is largely an exogenous variable. This theory is based on the business-cycle writings of Tugan-Baranovski, Spiethoff, Schumpeter, Keynes, and Hansen. These authors have taken investment to be the truly dynamic element in the economic process, which induces fluctuations in the other variables of the system. The fluctuations in investment are related to highly volatile, capricious, unpredictable exogenous forces such as population movements, wars, inventions, and psychological judgments. These are not variables which are related in any simple way to wages, prices, profits, or other endogenous economic magnitudes. This theory is tempting; it leads to very simple models; but it does not add much to our knowledge of the working of the system. If we do not explain the fluctuations of investment, we do not explain the behavior of national income or employment. If the economist leaves the explanation of investment to other social scientists, he is not assuming his responsibilities. In this chapter we have attempted to develop theories of investment which explain as much as is considered possible in terms of market variables, but we always leave a random disturbance in our equations. This random disturbance which we cannot explain represents our judgment of the part of investment that is exogenous. It should be pointed out that government investment, the foreign balance, and investment of non-profit institutions are classified as exogenous variables. Thus a great many non-consumer expenditures are classified in our models as exogenous, but we do not go so far as some economists in classifying all investment as exogenous.

Rules of Thumb and Rational Behavior

We have, in this chapter, set forth specific patterns of behavior for entrepreneurs, such as profit maximization or utility maximization. Many economists will claim that our versions of these theories are too complex, that businessmen merely follow certain customary rules of thumb. For example, it is often claimed that the individual firm sets its prices as average variable cost plus a percentage markup (dictated by the market) and sells all that it can at that price. Why does the entrepreneur follow this rule? Evidently because it works. He does not think that some other mechanism would be more profitable. The mere fact that this rule is mechanical does not mean that it does not maximize profits. If marginal cost for the firm is constant in the range of operations available to the firm, the method of percentage markup may be equivalent to profit maximization because profit maximization leads to the result that price equals marginal cost (= average variable cost) divided by 1 minus the reciprocal of demand elasticity. The rule of thumb and rational behavior have the same general form. This fact explains why a particular rule of thumb is adopted in preference to some other. Most cost studies have found the total cost curve to be linear and the marginal cost curve to be a constant in the range of observed operations.

Another example is the demand for labor. A simple rule to be followed is: the entrepreneur sets aside a certain fraction of the value of total output as labor's share, the size of this fraction being dictated by the market. An entrepreneur finds that this rule works, in the sense that it brings his operations to a profitable level, a more profitable level than he can attain by giving any other reward to his labor in view of the existing institutional framework. It is also well known [18] that, if the production function is of the Cobb-Douglas type, the optimum wage bill is a constant fraction of the value of total output. The Cobb-Douglas production function has been found to hold approximately in the relevant range of observed output and input; hence the rule of thumb coincides with the rational behavior pattern of profit maximization. For these reasons we should not be misled by those economists who

[18] For example, see equation (2.1.4a) above.

insist that entrepreneurs do not know the meaning of partial derivatives and hence do not behave so as to maximize profits or psychic income of some type.

THE THEORY OF THE HOUSEHOLD

The theory of consumer behavior has been developed at length by many writers—Pareto, Slutsky, Hicks, Allen, Samuelson, Wilson, Mosak, and others. The modern theorists usually proceed as follows: They assume that the individual consumer possesses a utility function, $u = u(x_1, \cdots, x_n)$, which depends upon the consumption of the n commodities $x_1, \cdots, x_n$. The individual behaves so as to maximize this function subject to the constraint that total expenditures equal income, $\sum_{i=1}^{n} p_i x_i = y$. They then show that all the important results of this maximization procedure are invariant under a monotonic transformation of u; i.e., they will get the same results if they maximize $F(u)$ instead of u, where $F'(u) > 0$. This amendment changes the theory from one of cardinal utility to one of ordinal utility, and it is claimed that the assumptions of ordinal utility are weaker than those of cardinal utility.

We shall first begin with this simple theory, except for one modification, and then proceed to more complex theories which combine the principle of consumer demand for commodities and the liquidity preferences of individuals.

In the theory of the firm, profit or utility maximization is introduced in order to form a theoretical framework from which to develop the demand for factors of production and the holding of various types of assets. Similarly, in the theory of household behavior, utility maximization is introduced in order to provide a foundation for the development of the laws of demand for consumer goods, some of the most important components of national income. As in the theory of the firm, we start with the individual unit and then aggregate over the entire system in order to achieve a model of macroeconomics.

The most unsatisfactory thing about the usual forms of the theory of consumer behavior is their implicit assumption that the total income is spent on present consumer goods. However, to get away from this implicit assumption is more a matter of reinterpretation of known results than of the derivation of new results. Let us classify goods into two categories: (1) goods consumed today $(x_1, x_2, \cdots, x_m)$; (2) goods to be consumed in the future $(x_{m+1}, x_{m+2}, \cdots, x_n)$. The total amount of income spent on goods consumed today will be called consumption expenditures, and the total amount of income set aside for future spending will be called savings (income not spent now). In this model, the household will behave so as to maximize

$$u = u(x_1, \cdots, x_m, x_{m+1}, \cdots, x_n) \tag{2.5.1}$$

subject to

$$\sum_{i=1}^{m} \text{an } p_i x_i + \sum_{j=m+1}^{n} \text{an } p_j x_j = \text{an } \mathsf{y} \tag{2.5.2}$$

where u = utility, p_i = price of the ith commodity, y = income. The prices of the future goods are the same as the prevailing prices of the corresponding goods today, discounted at the going interest rate from the present time to the time at which it is expected that the goods will be purchased. The maximization will be carried out by forming the function

$$\Phi = u - \lambda\left(\sum_{i=1}^{m} \text{an } p_i x_i + \sum_{j=m+1}^{n} \text{an } p_j x_j - \text{an } \mathsf{y}\right)$$

and setting the derivatives with respect to the x_i $(i = 1, 2, \cdots, n)$ equal to zero.

$$\frac{\partial \Phi}{\partial x_i} = \frac{\partial u}{\partial x_i} - \lambda \text{ an } p_i = 0 \qquad i = 1, 2, \cdots, n \tag{2.5.3}$$

The market prices and income are taken as given by the household; hence these variables are treated as constants in the differentiation. Equations (2.5.3) are the individual demand equations for the goods $x_1, \cdots, x_n$. Following the arguments of Professor Wilson,[19] we may write

$$\lambda = \frac{\sum_{i=1}^{n} \frac{\partial u}{\partial x_i} x_i}{\text{an } \mathsf{y}} \tag{2.5.4}$$

by combining (2.5.2) and (2.5.3). We may then eliminate λ between (2.5.3) and (2.5.4) to get

$$\frac{\partial u}{\partial x_i} = \frac{\text{an } p_i}{\text{an } \mathsf{y}} \sum_{i=1}^{n} \frac{\partial u}{\partial x_i} x_i \qquad i = 1, 2, \cdots, n \tag{2.5.5}$$

The system (2.5.5) enables us to solve for each of the x_i $(i = 1, 2, \cdots, n)$ in terms of the prices and income, but the solution is of a particular type. We notice that equiproportional changes in the prices and income leave the solution $x_1, \cdots, x_n$ in (2.5.5) unchanged. Thus

[19] See E. B. Wilson, "Notes on Utility Theory and Demand Equations," *Quarterly Journal of Economics*, Vol. LX, May, 1946, pp. 453–460.

the model is homogeneous of degree zero in prices and incomes. The solution for each of the x_i takes the form

$$x_i = x_i\left(\text{an}\frac{p_1}{p_i}, \cdots, \text{an}\frac{p_n}{p_i}, \text{an}\frac{Y}{p_i}\right) \quad i = 1, 2, \cdots, n \tag{2.5.6}$$

Equations (2.5.6) are the demand equations of microeconomics. They express the demand for each consumer good as a function of all prices and income. We desire, however, to obtain the demand for total consumption of $x_1, \cdots, x_n$ for our equations of macroeconomics. Suppose, for simplicity, that all the functions on the right-hand side of (2.5.6) are linear or that they can be approximated by linear functions (at least in the neighborhood of the equilibrium solution). Furthermore, let us consider only the demand for present consumer goods $x_1, \cdots, x_m$. We have

$$x_i = \sum_{j=1}^{n} \alpha_{ij}\,\text{an}\frac{p_j}{p_i} + \beta_i\,\text{an}\frac{Y}{p_i} \qquad i = 1, 2, \cdots, m \tag{2.5.6a}$$

The price of any good in the category $x_{m+1}, \cdots, x_m$ can be written as

$$p_j = \frac{p_k}{(1+i)^t} \qquad \begin{array}{l} j = m+1, \cdots, n \\ k \neq m+1, \cdots, n \end{array} \tag{2.5.7}$$

where i = interest rate; consequently (2.5.6a) can be rewritten as

$$x_i = \sum_{j=1}^{m} \alpha_{ij}'\,\text{an}\frac{p_j}{p_i} + \beta_i\,\text{an}\frac{Y}{p_i} \qquad i = 1, 2, \cdots, m \tag{2.5.6b}$$

where some of the α_{ij}' may not be constants but will depend upon i (= interest rate) and t.

We must now transform from anticipated prices to observable prices. In the preceding section this transformation was carried out for business firms by relating anticipated prices to past and current prices. Businessmen, however, must plan much more carefully than the typical household. Few households maintain the elaborate records that are customary in business planning. Households react to market variables almost immediately because they do not have to plan operations in detail over a long horizon. Consequently we say that anticipated prices for the household are the same as present prices plus a random error.[20] The preceding equation thus becomes

[20] In our example we shall use $\text{an}\frac{p_j}{p_i} = \frac{p_j}{p_i} + v_j$ and $\text{an}\frac{Y}{p_i} = \frac{Y}{p_i} + v_i$ as the transformation, where all the v's are independent.

$$x_i = \sum_{j=1}^{m} \alpha_{ij}' \frac{p_j}{p_i} + \beta_i \frac{\mathsf{Y}}{p_i} + v_i' \qquad i = 1, 2, \cdots, m \tag{2.5.6c}$$

Next we multiply both sides of (2.5.6c) by p_i and sum to get

$$\sum_{i=1}^{m} p_i x_i = \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_{ij}' p_j + \mathsf{Y} \sum_{i=1}^{m} \beta_i + \sum_{i=1}^{m} v_i' p_i \tag{2.5.6d}$$

The first term on the right-hand side is a linear combination of prices, and, as before, we assume that the following equation holds:

$$\sum_{i=1}^{m} \sum_{j=1}^{m} (\alpha_{ij}' + v_j')p_j = \alpha_0 p + v' p \tag{2.5.8}$$

where p is a general price index. Wilks' theorem (*loc. cit.*) does not apply separately to the expression $\sum_{j=1}^{m} v_j' p_j$ in (2.5.8) because the weights have zero means. Equation (2.5.8) will not hold exactly if we require that α_0 be a constant. But to assume that this equation holds approximately does not seem to be a large error, since both $\sum_{j=1}^{m} \sum_{i=1}^{m} \alpha_{ij}' p_j$ and p are linear combinations. The coefficients α_{ij}' depend upon i and t, as pointed out above. This fact introduces a further aberration from our assumption about the constancy of α_0.

We define $c = \sum_{i=1}^{m} p_i x_i / p$ as aggregate consumer expenditures for the individual, calculated in constant prices (real terms). We can now combine (2.5.6d) and (2.5.8) to obtain

$$c = \alpha_0 + \alpha_1 \frac{\mathsf{Y}}{p} + v' \tag{2.5.9}$$

This is a possible version of an individual's aggregate consumption function.

If we define aggregate savings as $s = \sum_{i=m+1}^{n} p_i x_i / p$, then (2.5.2) and (2.5.9) lead to

$$s = -\alpha_0 + (1 - \alpha_1) \frac{\mathsf{Y}}{p} - v' \tag{2.5.10}$$

The personal consumption function and the budget equation define the savings function. In our system we may retain consumption and income, or savings and income, as endogenous variables; but we cannot regard all three as endogenous variables.

We recall now that (2.5.9) refers only to the individual household. If there are N households in the economy, the community consumption function will be

$$\sum_{i=1}^{N} c^{(i)} = \sum_{i=1}^{N} \alpha_0{}^{(i)} + \sum_{i=1}^{N} \alpha_1{}^{(i)} \left(\frac{\mathsf{Y}}{p}\right)^{(i)} + \sum_{i=1}^{N} (v')^{(i)} \tag{2.5.9a}$$

$$C = \bar{\alpha}_0 + \bar{\alpha}_1 \frac{\mathsf{Y}}{p} + v$$

where

$$C = \sum_{i=1}^{N} c^{(i)} \qquad \alpha_0 = \sum_{i=1}^{N} \alpha_0{}^{(i)}$$

$$\alpha_1 = \frac{\sum_{i=1}^{N} \alpha_1{}^{(i)} \left(\frac{\mathsf{Y}}{p}\right)^{(i)}}{\sum_{i=1}^{N} \left(\frac{\mathsf{Y}}{p}\right)^{(i)}} \qquad \frac{\mathsf{Y}}{p} = \sum_{i=1}^{N} \left(\frac{\mathsf{Y}}{p}\right)^{(i)}$$

We shall identify C as total community consumption in constant prices and Y/p as total community income in constant prices. For the later statistical work, the community to be studied will be the nation as a whole. If the distribution of real income changes drastically, $\bar{\alpha}_1$ will not be a constant even if the $\alpha_1{}^{(i)}$ are. But, for moderate changes in the income distribution, we shall treat $\bar{\alpha}_1$ as constant. Also $\bar{\alpha}_0$ will vary as N, population, varies unless there are compensating fluctuations in the individual $\alpha_0{}^{(i)}$. Barring the latter possibility, we should have

$$\bar{\alpha}_0 = \bar{\alpha}_0(N) \tag{2.5.11}$$

There are various ways of dealing with $\bar{\alpha}_0(N)$. This parameter shows the level to which consumption would fall if income were zero; we might call it the minimum subsistence level of consumption. Naturally the community's subsistence level will be larger when the population is larger. We could write

$$\bar{\alpha}_0(N) = \beta_0 + \beta_1 N \tag{2.5.11a}$$

as the simplest representation of the hypothesis. This would mean that N would appear as an additional variable in the community consumption function (2.5.9a). If $\beta_0 = 0$, we have the choice of introducing N as an additional variable in (2.5.9a) and omitting the constant term in the equation, or of working entirely with per capita data for the consumption function as follows:

$$C = \bar{\alpha}_0(N) + \bar{\alpha}_1 \frac{Y}{p} + v = \beta_1 N + \bar{\alpha}_1 \frac{Y}{p} + v \tag{2.5.9a}$$

$$\frac{C}{N} = \beta_1 + \bar{\alpha}_1 \frac{\frac{Y}{p}}{N} + \frac{v}{N}$$

This form of the consumption function relates per capita consumption to per capita income. It may be considered the average of a distribution of consumption-income patterns that is taken to represent any particular element of the distribution. Our basic equation referring to the individual consumption-income pattern is given by (2.5.9). The per capita form of (2.5.9a) is the average corresponding to the individual element (2.5.9). This result has been obtained because the individual consumption function was assumed to be linear. Had it been assumed to be non-linear, other characteristics of the distribution pattern (e.g., the second and higher moments, if the assumed individual consumption function is taken to be a polynomial) would have to be used in addition to the average.

There is yet another way of handling the population variable in the consumption function. Population has grown very gradually and smoothly for the past few decades in this country. The growth may be highly non-linear over a long period of time, but any two or three decades may be reasonably well approximated by a linear trend. Hence, we may substitute a linear trend term for (2.5.11) and introduce time as a new variable in (2.5.9a). The reason for using the proxy variable, time, is that the variation of $\bar{\alpha}_0$ may be a result of other influences besides population growth. The subsistence level of income may gradually increase as population shifts from farm to city, as the advertising industry grows and cultivates new tastes for consumers, etc. These other variables affecting $\bar{\alpha}_0$ cannot always be measured objectively, as can population; hence we lump them all together and call the result a general trend in the function. Many economists object to trends as variables because they are too general and not sufficiently specific. However, if it is believed that a large number of gradually changing, non-measurable variables are influencing the structure of the system, we can do little but introduce time trend as an explicit proxy variable. In any case, the results differ little between systems in terms of per capita variables and systems in terms of aggregates and trends.

Thus far our theory of household behavior is practically the same as the standard theory found in the literature on consumer behavior. However, there are various plausible alternatives that must be considered. In the theory of the firm, the customary assumption of profit maximization can be generalized to a theory of utility maximization. In the present theory of the household we can also make some generalizations that are similar in nature. The utility function need not depend solely upon variables that are flows per unit of time, such as the consumption of present and future commodities. Utility may depend on the holding of assets also. The household may be interested in maintaining the best possible balance sheet or structure of assets. In the Keynesian system, the behavior of households is actually generalized in this way. The essence of the Keynesian theory is that individuals make two types of decisions: (1) They decide, on the basis of their income, either to spend or to save today. (2) They decide, on the basis of the level of the market interest rate, to hold their accumulated savings in the form of either money or securities. This theory could be developed by first maximizing utility, as a function of commodities consumed, subject to the budget constraint, and then maximizing a different utility, as a function of holdings of securities and money, subject to the constraint that total asset holdings equal accumulated savings. In this section we have already carried out the first type of maximization because it represents the theory of consumer behavior as it is usually presented. However, we shall now present a theory which combines the two separate decisions of the Keynesian system into a single decision. The case of two separate sets of decisions becomes a special case of the more general theory. Let us write

$$(2.5.12)\quad u = u(x_{11}, \cdots, x_{n1}, \cdots, x_{1T}, \cdots, x_{nT}, \mathsf{b}_{11}, \cdots, \mathsf{b}_{r1}, \cdots, \mathsf{b}_{1T}, \cdots, \mathsf{b}_{rT}, \mathsf{m}_1, \cdots, \mathsf{m}_T)$$

(2.5.13)

$$\mathsf{m}_{t-1} + \sum_{j=1}^{r} (1 + \text{an}\, i_{j,t})\mathsf{b}_{j,t-1} + \text{an}\, \mathsf{y}_t - \sum_{j=1}^{n} \text{an}\, p_{jt}x_{jt} = \mathsf{m}_t + \sum_{j=1}^{r} \mathsf{b}_{jt}$$

$$t = 1, 2, \cdots, T$$

where x_{jt} = consumption of the jth commodity in the tth period of the future, b_{jt} = the holding of jth security at the end of the tth period, m_t = the holding of cash balances at the end of the tth period of the future, i_{jt} = the interest rate corresponding to the jth security at the end of the $(t-1)$th period of the future, p_{jt} = the price of the jth commodity in the tth period of the future.

As in the theory of the firm, the utility function depends upon the asset structure as well as upon the current flows of commodities. In order to simplify the analysis this function is constructed with the independent variables as discrete functions of time. The next set of equations (2.5.13) is very similar to the set of balance equations introduced in the theory of the firm. An equation of this set states merely that the holdings of cash and securities at the beginning of any period, plus the interest earned on securities held, plus the other income *not* spent on commodities in the period (savings), equal cash and securities held at the end of the period. In other words, it states that the rate of change of wealth equals savings, or that total wealth equals historically accumulated savings.

Some assumptions which are buried in equations (2.5.13) must be pointed out to the reader. In the first place, income must be defined to include capital gains, if (2.5.13) is to be true. In these equations the value of wealth at the beginning of the period is on the left-hand side, and the value of wealth at the end of the period is on the right-hand side. The change in the value of wealth could be a result of price increases as well as of physical quantity increases. The price increases appear in capital gains. The equation will not hold true unless y is defined to include capital gains. Similarly, durable consumer goods are treated differently from durable producer goods. The balance equation in the theory of the firm assigns to π, profits, the role that $\mathsf{y} - \sum_{j=1}^{n} p_j x_j$, savings, plays here. Profits are net of depreciation charges on durable producer goods, but personal savings are not computed according to depreciation accounting for consumer durable goods.

The problem now is to maximize u subject to (2.5.13). First we form the function

$$\Psi = u + \sum_{t=1}^{T} \lambda_t \left[\mathsf{m}_{t-1} + \sum_{j=1}^{r} (1 + \text{an } i_{j,t})\mathsf{b}_{j,t-1} + \text{an } \mathsf{y}_t - \sum_{j=1}^{n} \text{an } p_{jt}x_{jt} - \mathsf{m}_t - \sum_{j=1}^{r} \mathsf{b}_{jt} \right]$$

Next we derive necessary conditions for a maximum:

$$\frac{\partial \Psi}{\partial x_{jt}} = \frac{\partial u}{\partial x_{jt}} - \lambda_t \text{ an } p_{jt} = 0 \tag{2.5.14}$$

$$j = 1, 2, \cdots, n; t = 1, 2, \cdots, T$$

$$\frac{\partial \Psi}{\partial \mathsf{b}_{jt}} = \frac{\partial u}{\partial \mathsf{b}_{jt}} + \lambda_{t+1}(1 + \text{an } i_{j,t+1}) - \lambda_t = 0 \tag{2.5.15}$$

$$j = 1, 2, \cdots, r; t = 1, 2, \cdots, T$$

$$\frac{\partial \Psi}{\partial \mathsf{m}_t} = \frac{\partial u}{\partial \mathsf{m}_t} + \lambda_{t+1} - \lambda_t = 0 \tag{2.5.16}$$

$$t = 1, 2, \cdots, T$$

$$\lambda_{T+1} = 0 \text{ by definition}$$

The set of equations (2.5.14)–(2.5.16), along with the constraints in (2.5.13), enables us to solve for all the x's, b's, and m's in terms of the prices, interest rates, income, and the initial conditions. These solutions are the demand equations for the flows of commodities and the stocks of liquid assets.

The system that we have now derived has homogeneity properties different from those of the previous simplified model. Previously we have found that, if all prices and incomes are changed in the same proportion, there is no effect on the quantities of goods demanded. However, if we merely change all prices and incomes by a fixed proportion in the current model, we shall find that the demand has changed. The homogeneity principle will now have to be reformulated. If we change all prices, income, the values of securities and money by a constant proportion, if $\partial u/\partial x_{jt}$ are homogeneous of degree one in the m's and b's, and if $\partial u/\partial \mathsf{b}_{jt}$ and $\partial u/\partial \mathsf{m}_t$ are homogeneous of degree zero in the m's and b's, there will be no influence on the demand for commodities or liquid assets.

We may indicate the solution of the system of maximizing equations and restraints of the general model (2.5.13)–(2.5.16) as

$$(2.5.17)\quad x_{jt} = x_{jt}(\text{an } p_{11}, \cdots, \text{an } p_{nT}, \text{an } i_{11}, \cdots, \text{an } i_{rT}, \text{an } \mathsf{y}_1, \cdots, \text{an } \mathsf{y}_T, \mathsf{m}_0, \mathsf{b}_{10}, \cdots, \mathsf{b}_{r0})$$

$$(2.5.18)\quad \mathsf{b}_{jt} = \mathsf{b}_{jt}(\text{an } p_{11}, \cdots, \text{an } p_{nT}, \text{an } i_{11}, \cdots, \text{an } i_{rT}, \text{an } \mathsf{y}_1, \cdots, \text{an } \mathsf{y}_T, \mathsf{m}_0, \mathsf{b}_{10}, \cdots, \mathsf{b}_{r0})$$

$$(2.5.19)\quad \mathsf{m}_t = \mathsf{m}_t(\text{an } p_{11}, \cdots, \text{an } p_{nT}, \text{an } i_{11}, \cdots, \text{an } i_{rT}, \text{an } \mathsf{y}_1, \cdots, \text{an } \mathsf{y}_T, \mathsf{m}_0, \mathsf{b}_{10}, \cdots, \mathsf{b}_{r0})$$

The form of the constraints are such that the separate variables y_1, $\mathsf{m}_0, \mathsf{b}_{10}, \cdots, \mathsf{b}_{r0}$, an $i_{11}, \cdots$, an i_{r1} always occur in the particular form

$$(2.5.20)\qquad \mathsf{m}_0 + \sum_{j=1}^{r} (1 + \text{an } i_{j1})\mathsf{b}_{j0} + \text{an } \mathsf{y}_1$$

If we assume, as before, that all anticipated values can be expressed as a linear combination of the observed values of the variables in question, we may rewrite the demand equations (2.5.17)–(2.5.19) in terms of the observable variables. In this case the initial conditions will not be as in (2.5.20) because of the appearance of anticipated values there. Suppose that (2.5.17)–(2.5.19) are linear, and also suppose that we have aggregated over the entire community as has been done repeatedly in this study. We should have

$$\text{(2.5.17a)} \qquad \mathsf{C} = \alpha_{10} + \alpha_{11}p + \alpha_{12}i + \alpha_{13}\mathsf{Y} + \alpha_{14}\mathsf{L}_{-1} + u_1$$

$$\text{(2.5.18a)} \qquad \mathsf{B} = \alpha_{20} + \alpha_{21}p + \alpha_{22}i + \alpha_{23}\mathsf{Y} + \alpha_{24}\mathsf{L}_{-1} + u_2$$

$$\text{(2.5.19a)} \qquad \mathsf{M} = \alpha_{30} + \alpha_{31}p + \alpha_{32}i + \alpha_{33}\mathsf{Y} + \alpha_{34}\mathsf{L}_{-1} + u_3$$

where C = total consumer expenditures, B = total value of security holdings at the end of the period, M = total value of the stock of cash balances at the end of the period, p = price of consumer goods, i = interest rate, Y = total community income, L_{-1} = total value of liquid assets at the beginning of the period $(\mathsf{B}_{-1} + \mathsf{M}_{-1})$.

If all the equations are homogeneous in p's, y's, m's, b's, we will have the alternative formulation

$$\text{(2.5.17b)} \qquad \frac{\mathsf{C}}{p} = \alpha_{11}' + \alpha_{12}'i + \alpha_{13}'\frac{\mathsf{Y}}{p} + a_{14}'\frac{\mathsf{L}_{-1}}{p} + u_1'$$

$$\text{(2.5.18b)} \qquad \frac{\mathsf{B}}{p} = \alpha_{21}' + \alpha_{22}'i + \alpha_{23}'\frac{\mathsf{Y}}{p} + a_{24}'\frac{\mathsf{L}_{-1}}{p} + u_2'$$

$$\text{(2.5.19b)} \qquad \frac{\mathsf{M}}{p} = \alpha_{31}' + \alpha_{32}'i + \alpha_{33}'\frac{\mathsf{Y}}{p} + a_{34}'\frac{\mathsf{L}_{-1}}{p} + u_3'$$

The two sets of macroequations are merely examples of plausible results that may follow from our model. There may be a number of lags in these equations for p, i, or Y. Furthermore, all the variables will not have the same importance in each equation. The existence of lags and the relative importance of the different variables will be discussed more fully in the sections dealing with empirical results.

In the constraining equations of (2.5.13), the income from interest on securities was separated from y. However, in solving the maximizing equations and (2.5.13) for the set (2.5.17)–(2.5.19), we could have solved for the x's, b's, m's in terms of an $\mathsf{y}_t + \sum_{j=1}^{r}$ an $i_{j,t}\mathsf{b}_{j,t-1}$ instead of in terms of an y_t. Hence we may consider the variable, Y, of our macrosystem to be total income, inclusive of interest income.

Also, it should be pointed out that the available income data will not include capital gains, although the theoretical income variable of our system does include capital gains.

We remarked above that the consumption function and the savings function need not both be introduced into the model since the savings function is defined as income minus the consumption function. Similarly, if the demand for money is one of the equations of the system (liquidity preference), the demand for securities may be omitted. It is not necessary to introduce both (2.5.18a) and (2.5.19a) [21] as equations of the model; either one is sufficient. The reason why it is possible to dispense with one of the demand equations for assets is that the constraints (2.5.13) define holdings of securities and cash in terms of known variables that are either predetermined or given from other equations of the system. The initial stocks of securities and cash held at the beginning of the planning period are predetermined and hence known. Income not currently spent (savings) is also known from other parts of the system, because consumption and income enter as endogenous variables. To follow customary procedure, we should use the consumption function, (2.5.17a) or (2.5.17b), and the liquidity preference function, (2.5.19a) or (2.5.19b), but not the savings function or the demand function for securities. This choice is arbitrary and convenient and has no essential influence upon any results that may be derived. It is a property of the equation-systems methods of statistical estimation that identities and definitions may be eliminated from the system in any mathematically legitimate way without having any influence on the estimation of the parameters.

This concludes the discussion of the theory of household behavior. It is not implied that we have exhausted all plausible variants of this theory. We could still modify the utility function in a number of ways and get different results. Or we could even introduce some alternative theory to utility maximization. However, we have presented a plausible theory which has large-scale support among economists, and it will be worth while to test this theory against the data.

THE MARKET EQUATION

It is not sufficient to present a theory of behavior of business firms and a theory of behavior of households. Each unit arranges its behavior pattern in terms of the market variables such as prices, wages, interest rates. It is necessary to present, in addition, a theory of the determination of these market variables to formulate a complete theory. There are, on the one hand, optimal behavior patterns of business firms and, on the other hand, optimal behavior patterns of the households; we must now study the interaction of these various units in the market

[21] Or (2.5.18b) and (2.5.19b).

place. This interaction serves to determine the observed levels of prices, wages, and interest rates.

The simplest theory of price determination follows from the well-known "law of supply and demand." A version of this law follows. From the theory of consumer behavior there is a market demand schedule for any product as a function of the price of that product.[22] From the theory of the firm there is a market supply schedule for any product as a function of the price of that product.[23] Whenever market supply exceeds market demand, price falls, and, whenever market demand exceeds market supply, price rises. A mathematical model [24] is

$$q^D = q^D(p) + u_1 \tag{2.6.1}$$

$$q^S = q^S(p) + u_2 \tag{2.6.2}$$

$$\dot{p} = f(q^S - q^D) + u_3 \tag{2.6.3}$$

The usual formulation of the law of supply and demand also adds that the price movements always tend to restore equilibrium; i.e., that a price drop tends to wipe out excess supply and bring about an equilibrium of market forces. In the same way, it is assumed that a price rise tends to wipe out excess demand and restore equilibrium. This additional assumption implies that there is stability in the system, that price movements as a result of excess demand or excess supply will not degenerate into hyper-inflation or hyper-deflation, respectively. This assumption also imposes certain restrictions on the form of the functions q^D, q^S, and f.

The system (2.6.1)-(2.6.3) is a first-order differential equation in p as a function of time. This can easily be seen by substituting from (2.6.1) and (2.6.2) into (2.6.3) to get

$$\begin{aligned} \dot{p} &= f[q^S(p) - q^D(p) + u_2 - u_1] + u_3 \\ \dot{p} &= f^*(p, u_2 - u_1) + u_3 \end{aligned} \tag{2.6.3*}$$

If all the functions are linear or if we take linear expansions of the function in the neighborhood of the equilibrium solution, it is easy to solve

[22] A more general case is that in which demand is a function of all prices in the system, rather than merely the price of the product in question.

[23] More generally, supply will be a function of several prices, wages, interest rates, etc.

[24] Several models of this type of behavior are analyzed by Paul A. Samuelson in "The Stability of Equilibrium: Comparative Statics and Dynamics," *Econometrica*, Vol. 9, April, 1941, pp. 97-120.

the linear differential equation of the first order and specify the equilibrium conditions from the requirement

$$\lim_{t \to \infty} p(t) = \text{constant} \tag{2.6.4}$$

where $p(t)$ is the solution of the differential equation.

This theory can be tested because the variables q^D and q^S are sometimes measurable separately. The quantity $q^S - q^D$ is defined as the rate of change of inventories and is available for many markets and for the economy as a whole. We might think that the total stock of inventories is more relevant for the theory of price determination than just the rate of change of inventories. This modification can be taken care of by writing

$$\dot{p} = g\left[\int_{-\infty}^{t} (q^S - q^D)d\theta\right] + u_4 \tag{2.6.5}$$

in place of (2.6.3). Oscillatory solutions of such systems have been studied by Dresch and Samuelson.

Hicks, Samuelson, Lange, and Metzler have studied the problem of the relationship between excess demand or supply and price fluctuations when there is a complete system of equations and when every demand or supply function depends upon all the prices in the system. However, the basic equation, which states that the rate of change of the price of each good is a function of excess demand or excess supply of that good, still holds. The only modification in this theory is that equation (2.6.3*) would have to be written

$$\dot{p}_i = f^{**}(p_1, \cdots, p_n, u_{2i} - u_{1i}) + u_{3i} \tag{2.6.3**}$$

if there are n prices in the system. The first theory falls under the heading of partial equilibrium analysis, and the second theory falls under the heading of general equilibrium analysis. Since our objective is a complete system, we must operate with the theory of general equilibrium.

Even though we adopt the general equilibrium approach, in which all prices enter into the demand and supply functions, we still retain this one basic principle, namely, that the rate of change of price of any commodity is a function of the excess demand or supply of that commodity. Let us write, therefore, in discrete form

$$\Delta p_i = f_i(h_i) + u_{5i} \qquad \frac{df_i}{dh_i} < 0 \qquad i = 1, 2, \cdots, n \tag{2.6.6}$$

where h_i = inventory stock of the ith commodity. In our above notation we would have

$$q_i^S - q_i^D = \Delta h_i \qquad i = 1, 2, \cdots, n$$

provided q_i^S and q_i^D are measured as flows of commodities during a period of time.

If in the theory of consumer behavior we define the price aggregate as $\sum_{i=1}^{n} \alpha_i p_i$, and if f_i are linear, we can write

$$\text{(2.6.6a)} \qquad \sum_{i=1}^{n} \alpha_i[p_i - (p_i)_{-1}] = \sum_{i=1}^{n} \alpha_i\beta_{0i} + \sum_{i=1}^{n} \alpha_i\beta_{1i}h_i + u_{5i}'$$

or

$$\Delta p = \beta_0 + \beta_1 H + u_5'$$

where $\sum_{i=1}^{n} \alpha_i p_i = p$, $\sum_{i=1}^{n} \alpha_i\beta_{0i} = \beta_0$, and $\sum_{i=1}^{n} \alpha_i\beta_{1i}h_i = \beta_1 H$. We have again applied the customary principles of aggregation in linear systems and have again relied on the principle that variations in the system of weights have little influence on the magnitude of the average or aggregate. The latter assumption is necessary because H may appear in other equations of the system where parameters relating to the firm rather than to the household are used as weights in aggregating the components h_i.

Equation (2.6.6a) is our principal theory of price determination in competitive markets. It is the statement of the "law of supply and demand" in aggregative form.

A fuller explanation needs to be given concerning the measurement of the variable H, used in these price-formation equations. The variable that we desire to have in this equation as the moving force behind price fluctuations is excess supply or excess demand. All inventories, however, are not excess supply. Some inventories are held by entrepreneurs so that they have enough goods on hand to cover their current sales; other inventories are purposely held by entrepreneurs for price speculation. Those inventories which entrepreneurs desire to hold for rational reasons are not excess supply; they are merely factors of production; i.e., working capital demanded by the firm for the smooth operation of business. From the theory of the firm, we have developed demand equations for each of the factors of production including labor, fixed capital, and inventories. But entrepreneurs may hold more or less inventories than are covered by their demand for inventories, which is based on the theory of profit maximization. If they misjudge the market, they may hold too much or too little in the form of inventories, compared with that amount of inventories that they would demand according to their profit-maximizing equations. Inventories should thus be split into two components: (1) desired inventories held for purely rational motives; (2) undesired inventories held because the market demand is different from that which was anticipated.

The categories of desired and undesired inventory holdings are not objectively measurable classifications unless we make some stronger assumptions. We shall adopt the following procedure: Assume that the market is always cleared except for random disturbances in our inventory demand equation. This random disturbance will be called undesired inventories. If the random error is negative there is excess demand in the market, and if the error is positive there is excess supply. In the first case, inventories are less than desired, presumably because purchasers are overly willing to take goods from the market and deplete supplies below the desired level. In the second case, inventories pile up in the entrepreneur's hands, presumably because purchasers are not willing to buy readily on the market under existing conditions.

A simple mathematical representation of this model is

$$H = f_1(p, p_{-1}, X - \Delta H, H_{-1}) + u_6 \tag{2.6.7}$$

$$\Delta p = f_2[u_6, (u_6)_{-1}] + u_7 \qquad \frac{\partial \Delta p}{\partial u_6} < 0 \tag{2.6.8}$$

where H = aggregate stock of inventories, p = price level, X = aggregate production, $X - \Delta H$ = aggregate production minus the change in the aggregate stock of inventories = sales, H_{-1} = initial conditions, u_6 = random disturbance, u_7 = random disturbance. Equation (2.6.7) states that the demand for inventories is a function of the price level and its rate of change, the volume of sales, and the initial conditions plus a random disturbance. This is one version of the theory presented in the section of this chapter dealing with the theory of the firm. The f_1-function is the demand for inventories based on the rational, profit-maximizing motives of speculation and the convenience of carrying on transactions. This function represents desired inventories. The disturbance u_6 represents the amount of inventories held above or below the desired amount, f_1. This variable represents undesired inventories; it will be identified with excess supply when it is positive and excess demand when it is negative. We have assumed that supply and demand are always in balance except for random perturbations. Equation (2.6.8) states that price falls when u_6 is positive (excess supply) and rises when u_6 is negative (excess demand). This equation is also subject to random error; i.e., the adjustment process does not operate exactly; hence we add the disturbance u_7.

The theory so far has been based on the assumption of a competitive market where individual firms and households adjust their behavior patterns to market prices and then interact in the market to determine prices in accordance with supply and demand. In this theory prices are always the equilibrating mechanism, for they always adjust to maintain an equilibrium between demand and supply. But in an imperfectly competitive market, business firms do not always adjust their behavior patterns according to market prices as parameters of action. Monopolistic firms have a direct influence on the market prices and do not accept them as given by forces outside the firm. Furthermore, it has been noted throughout the literature that monopolistic or oligopolistic firms do not like price fluctuations. They like to maintain stable prices and allow output to carry the brunt of adjustment. Witness the phenomenon of rigid steel prices during the depression and extreme fluctuations in output. Instead of relating price fluctuations to excess supply and demand, we may relate output fluctuations to excess supply and demand, especially when there is a lack of competition in the system.

The theory of output fluctuations to wipe out excess supply or demand derives from the controversies which arose over the equality of savings and investment. Suppose that savings are greater than investment. This means that all income which is not consumed is not being offset by investment; hence all the production which gave rise to the realized income is not being demanded in the form of consumption or investment. This means that undesired inventories will pile up in the warehouses of entrepreneurs by just that amount by which savings exceed (are not offset by) investment. If we define undesired inventory accumulation forced on entrepreneurs by excess savings as investment (forced investment in inventories), the savings-investment equality is not broken. However, if we include in investment only the desired inventory accumulation and exclude undesired inventory accumulation, we find a discrepancy between savings and investment equal to undesired inventory accumulation and equal, in turn, to excess supply. In a dynamical process, we would continue this example by pointing out that entrepreneurs would find themselves with excess stocks of goods (undesired inventories) and would be able to fulfill some of the future demand by selling from stocks on hand; hence they would be led to curtail production. Here we have an output adjustment to wipe out excess supply. By a similar method of reasoning we obtain the result that an increase in output responds to excess demand or to investment greater than savings.

A mathematical model will clarify these ideas and make them more precise. We shall define income as the sum of the output of consumer goods and producer goods. Let us write

$$C + I + \Delta H_1 + \Delta u_6 = Y \tag{2.6.9}$$

where C = consumption, I = investment in plant and equipment, ΔH_1 = the change in desired inventories, Δu_6 = the change in undesired inventories, Y = total income. If H = total inventories, we have the relation $H_1 + u_6 = H$. By convention we shall call $I + \Delta H_1$ total investment. This is the amount spent by business firms on the plant and equipment or on desired inventories. We can rewrite (2.6.9) as

$$I + \Delta H_1 + \Delta u_6 = Y - C = S \tag{2.6.9}$$

where S = savings is defined as income not spent on consumer goods. We now see that the difference between investment and savings is Δu_6, or the change in undesired inventories. If entrepreneurs adjust output according to excess supply or demand we will have, in addition, the equation

$$\Delta Y = f(\Delta u_6) + u_8 \tag{2.6.10}$$

This is the output adjustment equation. It could also be written as

(2.6.10a) [25]
$$\Delta Y = f(S - I - \Delta H_1) + u_8$$

Since we have defined $I + \Delta H_1$ as investment, this theory states that the difference between savings and investment causes fluctuations in the level of output, or that output is adjusted to equilibrate savings and investment. In the theory of partial equilibrium, supply and demand are functions of price, and price is the adjustment variable which maintains the equilibrium between supply and demand. In the theory of employment, savings and investment are functions of income, and income is the adjustment variable which maintains the equilibrium between savings and investment. The first exposition of this theory in unambiguous mathematical models was given by Samuelson.[26]

[25] Alternative formulations would be

$$\Delta Y = g[(u_6)_{-1}] + u_9 \qquad \Delta Y = g\left[\sum_{-\infty}^{t-1} (S - I - \Delta H_1)\right] + u_9$$

[26] See Paul A. Samuelson, *loc. cit.*

In the statistical models to be presented later, we shall introduce the market adjustment equations of both types, the price adjustment type and the output adjustment type. In the housing market where there is a large number of competing landlords, we shall have recourse to the price adjustment equations relating the rate of change of the rent level to the vacancies of housing space (undesired inventories). For the economy as a whole, exclusive of housing, we shall introduce the theory of output adjustment as a function of excess supply or demand, defined as undesired inventories.

CHAPTER III

STATISTICAL MODELS

Chapters I and II provide a theoretical framework for the empirical work that we actually want to carry out. We now have, before us, the principles for constructing economic models and for estimating the relevant parameters of these models by correct statistical methods. In this chapter we shall present various models and the statistical tests applied to each model.

MODEL I, A SIMPLE THREE-EQUATION SYSTEM

Naturally, we should commence to study the least complicated type of models in the first applications. A completely determined system containing three statistical equations (i.e., three equations involving random terms and unknown parameters) plus some definitions or identities can easily be derived from our preceding theories.

The macroeconomic system will be considered to be made up of consumer goods and two factors of production: (1) producer goods, and (2) human labor power. Our three equations will thus be the three demand schedules for the three goods in the system. The fact that our three-equation system is composed entirely of three demand schedules does not mean that the supply side is neglected. The equations of demand for factors of production determine the supply of commodities. The supply schedules are merely the production function at its equilibrium value of profit maximization. But the equations of profit maximization are the demand schedules for factors of production; hence we have not neglected the supply side if we have included the demand for labor and capital. The arguments will become clear in the following example. The notation is preserved from Chapter II.[1]

(3.1.1) $x = f(n, d)$ production function

(3.1.2) $d = d(i, k_{-1})$ equation of the use of capital

(3.1.3) $\dfrac{\partial f}{\partial n} = \dfrac{w}{p}$ marginal productivity—demand for labor

(3.1.4) $\dfrac{\partial f}{\partial d} = \dfrac{q}{p}$ marginal productivity—demand for capital

[1] The variable i = investment in this discussion.

If we substitute (3.1.3) and (3.1.4) into (3.1.1), we obtain

$$x = f^*\left(\frac{w}{p}, \frac{q}{p}\right) \tag{3.1.1*}$$

This reduced equation states that output is a function of the real wage rate and the real price of capital services. Equation (3.1.1*) is the supply schedule of output, x, in the sense in which we usually employ the term. Thus a combination of (3.1.1)–(3.1.4) is equivalent to the supply equation, and any model that takes account of the decisions involved in (3.1.1)–(3.1.4) has not neglected the supply side of the market.

The first equation of our model is the consumption function, which we write as

$$C = \alpha_0 + \alpha_1 W + \alpha_2 \Pi + u_1 \tag{3.1.5}$$

where C = consumption in constant dollars, W = wage bill in constant dollars, and Π = non-wage income (profits) in constant dollars. This equation is derived from exactly the same principles from which we derived the consumption demand equations of Chapter II. However, there we did not distinguish between two types of incomes, wages and profits. Instead, we lumped all types of income into a single aggregate. If we believe, though, that there is a significant difference between α_1 and α_2, we gain more information by specifying a consumption function like (3.1.5).

It is very simple to obtain (3.1.5) by the same methods used for the consumption function, which depends on aggregate income alone. Suppose that we have for the jth individual

$$c_j = \alpha_{0j} + \alpha_{1j} y_j + u_j \tag{3.1.6}$$

Suppose, further, that income receivers get income from only one of two sources, wages or profits. Denote wage earners by 1, 2, $\cdots$, N_1 and profit recipients by $N_1 + 1$, $\cdots$, N. Aggregate (3.1.6) in two stages:

$$\sum_{j=1}^{N_1} c_j = \sum_{j=1}^{N_1} \alpha_{0j} + \sum_{j=1}^{N_1} \alpha_{1j} y_j + \sum_{j=1}^{N_1} u_j \tag{3.1.7}$$

or

$$C_W = \alpha_0^{(1)} + \alpha_1 W + u_1^{(1)}$$

$$\sum_{j=N_1+1}^{N} c_j = \sum_{j=N_1+1}^{N} \alpha_{0j} + \sum_{j=N_1+1}^{N} \alpha_{1j} y_j + \sum_{j=N_1+1}^{N} u_j \tag{3.1.8}$$

or

$$C_\Pi = \alpha_0^{(2)} + \alpha_2 \Pi + u_1^{(2)}$$

We may also add (3.1.7) and (3.1.8) to get

$$C_W + C_\Pi = C = \alpha_0 + \alpha_1 W + \alpha_2 \Pi + u_1 \tag{3.1.9}$$

which is the same as (3.1.5).[2]

It is not strictly correct to assume that income receivers get income from only one of the two categories, wages or profits, but it is a very good approximation. It may be better to split income into categories according to size, such as incomes less than $5000 per year and incomes of $5000 per year or larger. However, most wage earners are in the low-income classes, and most profit recipients are in the high-income classes. It happens that there are good data on wages and profits (national income by distributive shares) and fragmentary data on aggregate income by size classes.

The second equation of the model is

$$I = \beta_0 + \beta_1 \Pi + \beta_2 \Pi_{-1} + \beta_3 K_{-1} + u_2 \tag{3.1.10}$$

where I = net investment in constant dollars, Π = profits in constant dollars, K_{-1} = stock of capital at the beginning of the year. This equation expresses the heuristic principle, mentioned in Chapter II, that profits are the mainspring of economic action in a capitalist society. Entrepreneurs expand when profits are anticipated to be high and contract when profits are anticipated to be low. However, not only the absolute size of profits but also their relation to the existing stock of capital is important; hence the variable K_{-1} is introduced. We could also write (3.1.10) with Π/K and $(\Pi/K)_{-1}$, the rate of profit, as independent variables. This modification would produce a nonlinear system, and, for a simple beginning model, it is desirable to retain the linearized form of (3.1.10).[3]

Heuristic principles are not the sole methods of deriving equations like (3.1.10). We could carry further the division of the economy into two groups, workers and capitalists or wage earners and profit recipients. Let us assume that the workers attempt to maximize their satisfactions, which depend upon the current and future consumption of household goods and services. This maximization is carried out subject to the customary budget constraint, and it leads to equations like (3.1.7). Capitalists also try to maximize satisfactions, but their satisfactions

[2] The income variables of (3.1.9) may be lagged since anticipated profits and wages are likely to have some influence on consumer expenditures. It seems particularly appropriate to introduce lagged profits because profit recipients must often base their behavior on expectations of future market conditions.

[3] For an argument in support of the use of linear systems see J. Tinbergen, *op. cit.*, Vol. II, pp. 11–12.

depend upon the use of current and future consumer goods and services and also upon the consumption of producer goods in their possession. In a capitalist society, the most important single feature is the fact that workers own no producer goods, and capitalists (and the government) alone own these goods. The income of capitalists can be used for two purposes, to spend on consumer goods and to spend on producer goods. They derive "pleasure" from both types of spending. The maximization of utility by capitalists will lead to equations like (3.1.8) for the demand for consumer goods, but it will also lead to equations like

$$d_j = \beta_{0j} + \beta_{1j}\pi_j + u_j \tag{3.1.11}$$

for the demand for producer goods. If we substitute a relation like (3.1.2) into (3.1.11), the result is

$$i_j = \beta_{0j}^* + \beta_{1j}^*\pi_j + \beta_{2j}^*(k_j)_{-1} + u_j \tag{3.1.11*}$$

The variable π_j may be considered as representative of anticipated profits, in which case there may be lagged as well as current profits in (3.1.11*). If there is a lag, and if (3.1.11*) is aggregated over all firms, the result is an equation of the same form as (3.1.10). We should expect a lag of *some* order in equations dealing with business planning, like (3.1.10), and should not expect to omit Π_{-1} from this equation, although we cannot say, in advance, whether the time lag should be greater or less than one time unit.

Finally, we have the equation expressing the demand for labor

$$W = \gamma_0 + \gamma_1 Y + \gamma_2 Y_{-1} + \gamma_3 t + u_3 \tag{3.1.12}$$

where W = wage bill in constant dollars, Y = output in constant dollars, t = time. Equations of this type have already been developed in Chapter II by setting the derivative of profits with respect to labor input equal to zero. This particular form of the marginal-productivity equation follows if the production function is of the constant-elasticity type, a form of the production function which is known to fit the data well in many cases. This convenient assumption is probably not far from the truth.

Equation (3.1.12) will also hold true if the capitalists, as above, attempt to make satisfactions as large as possible subject to a budget constraint. The constraint equates capitalist spending plus saving with profits. The latter term is equal to the value of output (which is itself a function of input) minus the wage bill. It can be shown that maximization of the utility function subject to this constraint implies also the maximization of capitalist income (profit) with respect to labor. The implied maximization leads directly to equations of the form of

(3.1.12) if the production function is linear in logarithms. The trend variable in equation (3.1.12) is there to reflect an institutional phenomenon, namely, the growing bargaining strength of labor. Not only will labor's income fluctuate with the fluctuations of output as determined by strict profit maximization, but it will also increase gradually as the strength of the organized labor movement grows. Increasing unionization means a persistent, gradual shift of the demand function, and this shift is reflected in the variable, t.

There are also three definitional equations that are not subject to random error and that have no unknown parameters. These equations complete our system.

$$C = \alpha_0 + \alpha_1 W + \alpha_2 \Pi + u_1 \tag{3.1.5}$$

$$I = \beta_0 + \beta_1 \Pi + \beta_2 \Pi_{-1} + \beta_3 K_{-1} + u_2 \tag{3.1.10}$$

$$W = \gamma_0 + \gamma_1 Y + \gamma_2 Y_{-1} + \gamma_3 t + u_3 \tag{3.1.12}$$

$$C + I + G = Y \tag{3.1.13}$$

$$\Pi + W = Y \tag{3.1.14}$$

$$\Delta K = I \tag{3.1.15}$$

The new equation (3.1.13) states that the total output is the sum demanded by consumers (C) plus goods demanded by business firms (I) plus goods demanded by the government and foreigners[4] (G). The net change in inventories (whether desired or not) is included in I; hence demand and output cannot differ. The left and right sides of (3.1.13) must balance. The next equation (3.1.14) states that total output (income) is the sum of profits and wages. The last equation (3.1.15) defines investment as the rate of change of the capital stock. There are now six equations in the endogenous variables C, I, W, Π, K, Y and the exogenous variables t and G. We call this system a three-equation model because we can eliminate any three endogenous variables by substitution from the exact relations (3.1.13), (3.1.14), (3.1.15). The estimates of the unknown parameters by the equation-systems methods are unaffected by this elimination.

The basic flaw of such a small system is that it is too aggregative. In order to keep the number of variables small, we must perform many undesirable aggregations. The variable I includes several categories of investment that could well be separated. More information would be gained from such separations, but at the same time new computational problems would be created. We might split investment, for

[4] The foreign demand is net of imports.

example, into (1) expenditures on business plant and equipment, (2) expenditures on net additions to inventories, (3) expenditures on residential construction. Introducing the first category may make it necessary to introduce the price of output and the price of capital goods as separate variables; hence new equations would be required to explain price determination. The second category may also entail prices and the third category, rents and construction costs, as new variables for which additional equations would be needed. Similarly, the expenditures on consumer goods could be segregated into such classes as services, non-durables, and durables, but the demand for each class of consumer goods would probably depend upon the price ratios between the several classes. Again we would need more equations to determine these prices. Thus in order to keep this model as simple as possible, we have given inadequate treatment to the various categories of producer and consumer goods. Later we shall discuss more complex models in which some of these subdivisions are actually introduced.

It will also be noted that our simple model is purely a "real" system; the quantity of money plays no role in equations (3.1.5)–(3.1.15). There are various ways of incorporating the quantity of money in the system. Demand for investment goods may very well depend upon the interest rate, which in turn is related to the demand and supply of money; or consumers may be directly affected by the quantity of liquid assets which they possess in making their decisions whether to spend or save. During the course of this book, some of the hypotheses concerning the money market will be tested, and at the present time we shall proceed with the simplifying assumption that the quantity of money is not important in its influence on the system. We shall show that our three-equation system, which is independent of the quantity of money, is consistent with the observed data.

Many economists will recognize the resemblance between the three-equation model, Kalecki's models of the business cycle, and some of the doctrines of Marxist economics. This model could actually be called a Marxian theory of effective demand. It is possible to develop this model, as we have shown, from the un-Marxian principles of utility and profit maximization, but it is also possible to develop this model from purely Marxian principles.[5] The same model can be consistent with a multiplicity of hypotheses. The problem of developing models from Marxian principles is of great interest from the point of view of the history of economic thought, but is not an essential problem of this book, which is concerned mainly with quantifying a true description

[5] See Lawrence R. Klein, "Theories of Effective Demand and Employment," *Journal of Political Economy*, Vol. LV, April, 1947, pp. 108–131.

of the structure of United States economy. We mention this relationship only in passing, as a point of general interest.

Of what use is the statistical treatment of this simple model? It is agreed that this model is very simplified, very aggregative, but, at the same time, it is more than a mere demonstration of various statistical methodologies. The calculations made for this model serve as a test of certain economic hypotheses. If the data were to refute this model, we should have grounds for questioning the validity of the Marxian theory of effective demand. Since the data do not refute the model, we cannot conclude that this theory stands as proved, but we can have more faith in it, or in any other theory which would produce this model, than would be the case if we made no tests at all. In the same way, we can test the hypothesis that investment depends upon profits as opposed to other theories of the investment schedule. We shall find that there are several theories of the investment schedule that are not refuted by the data, and this is information worth having. Other cases will be demonstrated, however, in which certain popular hypotheses can be rejected in the sense that they are not consistent with the data.

Before we come to the statistical analysis of this simple model, we should make one modification. We have assumed thus far that goods are demanded by households, business firms, and the government, and that goods are produced only by business firms operating for profit. We must allow, in addition, for the productive efforts of government. The government contribution to total output is not based on principles of profit or utility maximization; it is considered, instead, to be exogenous. We shall measure government activity by the payments that it makes to factors of production. Customarily these payments of the government are listed in the national income statistics as wages and interest on the public debt. It is certainly true that government wage payments represent a payment for productive effort and should be considered as part of the government's contribution to national income. The role of interest payments is, on the other hand, questionable. If the public debt were used to finance the construction of income-producing government capital, interest payments would also represent a contribution to national income on the part of the government. But much of the public debt on which interest payments are made is built up as a result of war expenditure. It is very debatable whether the interest payments on the war debt are a contribution to income by the government. We shall adopt the convention that government wage payments alone represent income originating in the government; hence the government contribution to production will be represented by W_2, the government wage bill measured in constant dollars. The private wage bill will be represented [6] by W_1.

On the side of production, we shall measure national output at market prices; i.e., what is sometimes called net national product. On the side of income, we shall measure national income at factor costs; i.e., what is sometimes called net national income. The difference between the two concepts is roughly business taxes, represented by the symbol T. Business taxes are also treated as an exogenous variable.

In the consumption function, it would be desirable to measure wages and profits as wages after personal taxes and profits after personal taxes. This would place these variables in the category of disposable income instead of total income paid to productive factors. However, available data are not adequate to measure disposable wage income and disposable profit income. Disposable income is really the appropriate variable for the consumption function. When we deal later with consumption functions that do not distinguish between wages and profits, we shall use disposable income rather than factor payments as the appropriate variable.

It should also be pointed out that our profit variable includes corporate savings. Disposable income is practically always calculated net of corporate savings because it is reasoned that consumers do not have these savings in their hands, available for expenditures on consumption goods. We are following Kalecki's suggestions, and leaving corporate savings in the profit variable of both the consumption function and the investment function. There is no question that corporate savings might influence the demand by corporations for producer goods, but there is some question whether or not these retained earnings influence the household consumption of capitalists. Capitalists will be assumed here to consider income as their own whether it is in their personal possession or in the possession of their corporations. The corporation is viewed as a particular instrumentality of accumulation created by the capitalists.

After these modifications our model will be:

$$C = \alpha_0 + \alpha_1(W_1 + W_2) + \alpha_2\Pi + u_1 \tag{3.1.16}$$

$$I = \beta_0 + \beta_1\Pi + \beta_2\Pi_{-1} + \beta_3K_{-1} + u_2 \tag{3.1.17}$$

$$W_1 = \gamma_0 + \gamma_1(Y + T - W_2) + \gamma_2(Y + T - W_2)_{-1} + \gamma_3t + u_3 \tag{3.1.18}$$

[6] It is obvious that disposable income in the hands of consumers is independent of the treatment of government interest payments as a transfer or as a payment to a factor of production. The definition of total output is, however, affected.

$$(3.1.19) \qquad Y + T = C + I + G$$

$$(3.1.20) \qquad Y = W_1 + W_2 + \Pi$$

$$(3.1.21) \qquad \Delta K = I$$

This is the final form of our simple model.[7] We shall consider, in the course of our statistical investigations, the possibilities of adding or dropping lags and trends in the first three equations, but the basic structure will remain the same.

Identification

The first step in any attempt to undertake structural estimation in an economic system is to examine the conditions for identifiability. Can we make statistical estimates of the α's, β's, and γ's individually, or can we estimate only certain functions of the parameters? The system (3.1.16)–(3.1.21) is an identified system, but there exist special cases of this model in which it is not possible to identify all the parameters. The exogenous variables of government activity and the lags and trends actually insure the identification here. Suppose, for example, that all government variables are identically zero, and that there are no lags or trends in the system; then we should have for our system

$$(3.1.16^*) \qquad C = \alpha_0 + \alpha_1 W + \alpha_2 \Pi + u_1$$

$$(3.1.17^*) \qquad I = \beta_0 + \beta_1 \Pi + \beta_3 K_{-1} + u_2$$

$$(3.1.18^*) \qquad W = \gamma_0 + \gamma_1 Y + u_3$$

$$(3.1.19^*) \qquad Y = C + I$$

$$(3.1.20^*) \qquad Y = \Pi + W$$

$$(3.1.21^*) \qquad \Delta K = I$$

If we eliminate C, W, Y from (3.1.16*), (3.1.18*), (3.1.19*) and (3.1.20*), we get

$$(3.1.22^*) \qquad \alpha_0(\gamma_1 - 1) + \gamma_0(1 - \alpha_1) + (1 - \alpha_2\gamma_1 - \alpha_1\gamma_1 - \alpha_2)\Pi + (\gamma_1 - 1)I = (1 - \gamma_1)u_1 + (\alpha_1 - 1)u_3$$

A linear combination of (3.1.17*) and (3.1.22*) will produce a new linear relation of the form

$$(3.1.23^*) \qquad \delta_0 + \delta_1 \Pi + \delta_2 I + \delta_3 K_{-1} = \delta_4 u_1 + \delta_5 u_2 + \delta_6 u_3$$

[7] Alternatively, (3.1.16) could be written

$$C = \alpha_0 + \alpha_1(W_1 + W_2) + \alpha_2 \Pi + \alpha_3 \Pi_{-1} + u_1$$

We shall then have no way of distinguishing between (3.1.23*) and (3.1.17*). We shall not know whether statistical estimates of parameters in a linear relation connecting I, Π, K_{-1} are estimates of the structural parameters of (3.1.17*) or estimates of the parameters of a linear combination of several equations of the model, say (3.1.23*). But let us replace (3.1.19*) by (3.1.19). Then G will enter linearly into (3.1.22*). In this event we shall be able to distinguish between (3.1.17*) and a linear combination of (3.1.17*) with (3.1.22*) because there will be a unique and different predetermined or exogenous variable in each of the equations (3.1.17*) and (3.1.22*), K_{-1} in the former and G in the latter. Any linear combination of the two equations will contain two predetermined variables, K_{-1} and G; hence we can distinguish this linear combination from (3.1.17*) which has K_{-1} as its only predetermined variable. By the same methods it can be shown that the presence of government spending, government output, taxes, lags, and trends in the equations makes it impossible to derive any equation of the model from linear combinations of other equations; hence we have identifiability in (3.1.16)–(3.1.21) and can proceed with statistical estimation.

Statistical Estimates

From the data of the interwar period we can measure all the variables of the three-equation model. The values of all variables for each year 1921–1941 and the methods of construction of each such time series are given in the Appendix.

The method of maximum likelihood (using all available a priori information) has been applied to (3.1.16)–(3.1.21) for the special case in which it is assumed that

$$\rho_{u_1u_2} = \frac{Eu_1u_2}{\sqrt{Eu_1^2 \cdot Eu_2^2}} = 0$$

$$\rho_{u_1u_3} = \frac{Eu_1u_3}{\sqrt{Eu_1^2 \cdot Eu_3^2}} = 0$$

$$\rho_{u_2u_3} = \frac{Eu_2u_3}{\sqrt{Eu_2^2 \cdot Eu_3^2}} = 0$$

The conditions state that there is no correlation between the disturbances in any of the structural equations. This assumption greatly simplifies the computational work and is convenient to make, provided it is justified. We shall present below estimates of the ρ's made from more general models in which they are not assumed to be zero.

The steps involved in the computations are explained in detail in other publications of the Cowles Commission.[8] The estimates of the structural parameters are

$$C = 16.78 + 0.02\Pi + 0.23\Pi_{-1} + 0.80(W_1 + W_2) + u_1' \tag{3.1.23}$$

$$I = 17.79 + 0.23\Pi + 0.55\Pi_{-1} - 0.15K_{-1} + u_2' \tag{3.1.24}$$

$$W_1 = 1.60 + 0.42(Y + T - W_2) + 0.16(Y + T - W_2)_{-1} + 0.13(t - 1931) + u_3' \tag{3.1.25}$$

$$Y + T = C + I + G \tag{3.1.19}$$

$$Y = \Pi + W_1 + W_2 \tag{3.1.20}$$

$$\Delta K = I \tag{3.1.21}$$

All variables except t are measured in billions of 1934 dollars. The time variable is measured by the years; i.e., 1921, 1922, 1923, etc.; u_i' represents the estimate of each year's value of the random disturbance.

All the point estimates of the parameters given above look very reasonable; i.e., they seem not to contradict the general evidence that would be based on experience other than that contained in the data used, and all have the signs that we should expect on the basis of economic theory. No measures of reliability are given here. Yet, the point estimates, alone, are not sufficient. It is much more satisfactory to give an entire confidence interval which, with a very high probability, would include the true value of the parameter. At a later stage, it is hoped that measures of reliability in the form of confidence intervals will be available. We can, however, give, as a statistical test on the specified form of the equations, estimates of each year's values of the disturbances u_1, u_2, u_3. The variation of the disturbances also shows how much of the behavior of the system can be estimated from structural relationships and how much must be attributed to chance. The smaller the variation of the disturbances, the more closely we can estimate the endogenous variables on the basis of known values of the exogenous or predetermined variables. The estimates of each year's

[8] The reader can obtain an account of available information on methods of computation of maximum likelihood estimates in systems of simultaneous equations from Cowles Commission Monograph No. 10 entitled *Statistical Inference in Dynamic Economic Models*, John Wiley & Sons, Inc., 1950, edited by T. C. Koopmans. See especially T. C. Koopmans, H. Rubin, and R. B. Leipnik, "Measuring the Equation Systems of Dynamic Economics," Section 4. The estimation of the parameters in (3.1.23)–(3.1.25) will be explained, in detail, in another publication.

values of the disturbances are obtained by calculating the difference between the values of C, I, W_1, as obtained from (3.1.23), (3.1.24), (3.1.25) and the observations on these variables. Table I gives the

TABLE I

ESTIMATES OF ANNUAL VALUES OF DISTURBANCES BY METHOD OF MAXIMUM LIKELIHOOD: MODEL I (ASSUMING DIAGONAL COVARIANCE MATRIX)

(Billions of 1934 dollars)

	u_1'	u_2'	u_3'		u_1'	u_2'	u_3'
1921	−0.55	−1.01	−1.29	1931	−1.07	−0.60	0.41
1922	−0.68	0.18	0.37	1932	−1.33	−0.59	−0.12
1923	−1.36	0.95	1.31	1933	0.63	1.04	0.40
1924	−0.32	−1.53	−0.15	1934	−0.13	−0.16	0.30
1925	0.16	0.30	−0.40	1935	0.03	0.11	0.04
1926	0.84	1.28	−0.44	1936	2.05	1.56	−0.76
1927	1.41	0.92	−0.74	1937	−0.55	−0.13	0.98
1928	1.31	−0.07	0.32	1938	−0.21	−3.11	−0.61
1929	−0.41	1.62	1.22	1939	1.47	0.05	−0.29
1930	−0.66	−0.65	−0.29	1940	1.11	−0.27	−1.04
				1941	−1.75	0.11	0.77

results of this computation. The root-mean-squares of the estimates given in the table, adjusted for degrees of freedom, are

$$\bar{S}_1 = 1.14 \qquad \bar{S}_2 = 1.19 \qquad \bar{S}_3 = 0.78 \tag{3.1.26}$$

It is obvious that these estimates of the disturbances are very small compared with the order of magnitude of the important terms in the respective equations. Thus our structural relations account for the major part of the variations in the system.

Two basic assumptions are made concerning the behavior of the variables u_1, u_2, u_3. We assume (1) that each of the three variables is non-autocorrelated; i.e., there is no correlation between its successive values, and (2) that $\rho_{u_1u_2} = 0$, $\rho_{u_1u_3} = 0$, $\rho_{u_2u_3} = 0$. The latter assumption introduces particular computational simplifications in the method of maximum likelihood (without being essential to it), but the former assumption is made for all the commonly known methods of estimation. We shall present below another method of estimation which makes the first assumption without making the second. Both assumptions should be tested to see how closely they hold for the estimated values of u_1, u_2, u_3.[9]

We can test the first assumption by calculating the ratio of the mean square successive differences to the variances of the estimates. The formula for this ratio is

$$\left(\frac{\delta^2}{S^2}\right)_i = \frac{\sum_{t=1922}^{1941} (u_{it}' - u_{i,t-1}')^2}{\sum_{t=1921}^{1941} (u_{it}')^2} \frac{21}{20} \qquad i = 1, 2, 3 \tag{3.1.27}$$

If in this quantity the estimated values of the disturbances u_{it}' are replaced by the disturbances u_{it}, the distribution of the ratio so obtained is known and has been tabulated for samples of various sizes.[10] If the latter ratio is greater than 1.25 and less than 3, for samples of size 20, the probability is 5 per cent or greater that the sample estimate could have come from a population with no autocorrelation. Assuming that we can approximately disregard the effect of using u_{it}' instead of u_{it}, we get from our data

$$\left(\frac{\delta^2}{S^2}\right)_1 = 1.54 \qquad \left(\frac{\delta^2}{S^2}\right)_2 = 2.18 \qquad \left(\frac{\delta^2}{S^2}\right)_3 = 2.20 \tag{3.1.28}$$

Apparently, the data do not contradict the hypothesis that each of the u's is non-autocorrelated.

The second assumption should be tested from a model in which no assumption is made about the correlations among the several disturbances. However, it is interesting to examine the estimated values of

[9] The matter is, in reality, more complicated than has been mentioned in the text. Not only is it assumed that each of the disturbances is non-autocorrelated, but it is also assumed that there is no lag correlation between any two different disturbances. Rather than test each estimated disturbance series separately for autocorrelation, it is better to test all series simultaneously. This can be done by constructing a matrix of all lagged correlations in the estimated disturbance series, where the highest-order lag depends upon the order of lags in the system. The elements of this matrix include both autocorrelations and lagged correlations between different series. Since the value of the determinant of this matrix follows a known distribution pattern, the probability of a determinant of any given size can be calculated under the null hypothesis that there is no lagged correlation in the population. A simultaneous test of this type will be given below for the method of reduced forms, but the computations are lengthy and have not been carried out for every method.

[10] See J. von Neumann, "Distribution of the Ratio of the Mean Square Successive Difference to the Variance," *Annals of Mathematical Statistics*, Vol. 12, December, 1941, pp. 367–395; B. I. Hart and J. von Neumann, "Tabulation of the Probabilities for the Ratio of the Mean Square Successive Difference to the Variance," *Annals of Mathematical Statistics*, Vol. 13, June, 1942, pp. 207–214.

these correlations even in the present model. The formula for the estimate is

$$r_{u_iu_j} = \frac{\sum_{t=1921}^{1941} u_{it}'u_{jt}'}{\sqrt{\sum_{t=1921}^{1941} (u_{it}')^2 \sum_{t=1921}^{1941} (u_{jt}')^2}} \tag{3.1.29}$$

The numerical results are

$$r_{u_1'u_2'} = 0.32 \qquad r_{u_1'u_3'} = -0.51 \qquad r_{u_2'u_3'} = 0.30 \tag{3.1.30}$$

The method of reduced forms (limited information maximum likelihood method) has also been applied to this three-equation system. The theory and techniques underlying this method can be found in papers by T. W. Anderson, Jr., and H. Rubin.[11] In applying this method no assumption was made about the covariance matrix of the disturbances u_i, but it was assumed that there are no lagged correlations among these disturbances. The estimates of the parameters by the second method leads to the system [12]

$$C = 17.71 + \underset{(0.07)}{0.02}\Pi + \underset{(0.04)}{0.87}(W_1 + W_2) + u_1'' \tag{3.1.31}$$

$$I = 22.59 + \underset{(0.23)}{0.08}\Pi + \underset{(0.21)}{0.68}\Pi_{-1} - \underset{(0.04)}{0.17}K_{-1} + u_2'' \tag{3.1.32}$$

$$W_1 = 1.53 + \underset{(0.03)}{0.43}(Y + T - W_2) + \underset{(0.03)}{0.15}(Y + T - W_2)_{-1} + \underset{(0.02)}{0.13}(t - 1931) + u_3'' \tag{3.1.33}$$

$$Y + T = C + I + G \tag{3.1.19}$$

$$Y = \Pi + W_1 + W_2 \tag{3.1.20}$$

$$\Delta K = I \tag{3.1.21}$$

The numbers in parentheses below the coefficients are asymptotic standard errors. There is no change of sign in any of the parameters as compared with the maximum likelihood estimates, but there are some significant differences in magnitude. It should be pointed out

[11] The method has been presented in a paper given by Anderson at the August, 1946, meetings of the Institute of Mathematical Statistics, Ithaca, New York. There is a chapter on the method in Cowles Commission Monograph No. 10 (John Wiley & Sons, Inc., 1950, T. C. Koopmans, editor) and an article, "Estimation of Parameters of a Single Equation in a Complete System of Stochastic Equations," by Anderson and Rubin, *Annals of Mathematical Statistics*, Vol. 20, March, 1949, pp. 46–63.

[12] Two variants, with and without lagged profits in the consumption equation, are presented for the reduced-form estimates.

that, whereas the standard error of the coefficient of Π in (3.1.32) is relatively large, the standard error of the sum of the coefficients of Π and Π_{-1} in this equation is much smaller. Taking account of the covariance as well as the variance, we obtain 0.09 for the standard error of the sum. In many economic problems, such as the investigation of stability conditions and the properties of static systems, the sum of the coefficients is more important than the separate components.

TABLE II

ESTIMATES OF ANNUAL VALUES OF DISTURBANCES BY METHOD OF REDUCED FORM (LIMITED-INFORMATION METHOD): MODEL I

(Billions of 1934 dollars)

	u_1''	u_2''	u_3''		u_1''	u_2''	u_3''
1921	−0.50	−1.61	−1.29	1931	−1.08	−1.00	0.54
1922	−0.96	0.33	0.32	1932	−1.96	−1.18	0.04
1923	−0.95	0.77	1.22	1933	−0.98	1.55	0.44
1924	0.43	−1.65	−0.14	1934	−0.97	−0.15	0.29
1925	1.03	0.22	−0.45	1935	−0.74	0.17	0.02
1926	1.72	1.14	−0.47	1936	1.34	1.93	−0.82
1927	2.12	1.01	−0.73	1937	−0.69	−0.25	0.99
1928	1.98	0.28	0.33	1938	−0.28	−3.46	−0.51
1929	0.39	1.96	1.20	1939	0.70	0.51	−0.35
1930	0.50	−1.23	−0.19	1940	0.94	0.05	−1.08
				1941	−2.03	0.60	0.64

The most unreliable estimate in the computed system is that of capitalists' marginal propensity to consume. An alternative consumption function will be presented in order to attempt to improve the estimate of this parameter. This is an important parameter of the system, and as reliable an estimate as possible should be obtained.

The estimates of the disturbances are given in Table II. The root-mean-squares of these estimates, adjusted for degrees of freedom, are

$$\bar{S}_1 = 1.30 \qquad \bar{S}_2 = 1.43 \qquad \bar{S}_3 = 0.77 \tag{3.1.34}$$

The values for the test of autocorrelation of the time series of the estimated disturbances are

$$\left(\frac{\delta^2}{S^2}\right)_1 = 0.98 \qquad \left(\frac{\delta^2}{S^2}\right)_2 = 2.18 \qquad \left(\frac{\delta^2}{S^2}\right)_3 = 2.10 \tag{3.1.35}$$

The time series u_1'' are seen to be definitely autocorrelated.

No assumption was made about the correlations between the disturbances in the populations, and the estimates of these correlations show them to be

$$(3.1.36)\qquad r_{u_1''u_2''} = 0.27 \qquad r_{u_1''u_3''} = -0.47 \qquad r_{u_2''u_3''} = 0.26$$

None of the correlations is large compared to correlations often obtained from economic data with about twenty annual observations, but to regard -0.47 as insignificantly different from zero does not seem reasonable.

In order to attempt to get a better estimate of the marginal propensity to consume out of capitalist income, we introduce Π_{-1} as a variable in the consumption equation. Since Π_{-1} is a predetermined variable already in the system, it can be introduced in the consumption equation without affecting any of the estimates in the other equations. This is a property of the method of reduced forms. The result of the computations is

$$(3.1.31^*)\qquad C = 17.15 - \underset{(0.19)}{0.22}\Pi + \underset{(0.17)}{0.40}\Pi_{-1} + \underset{(0.05)}{0.82}(W_1 + W_2) + u_1^*$$

$$(\bar{S})_1{}^* = 1.55 \qquad \left(\frac{\delta^2}{S^2}\right)_{1^*} = 1.56$$

The residual variation is decidedly less autocorrelated than u_1'', but the capitalists' marginal propensity to consume is not reliably estimated. The coefficient of Π_{-1} has a relatively smaller standard error than does the coefficient of Π, but the standard error of the sum of the two coefficients is 0.11.

Many of the statistical properties of (3.1.31*) are superior to those of (3.1.31). If we calculate the determinant of the matrix of all possible lagged correlations in the estimated disturbances with equations (3.1.31) and (3.1.31*) alternatively in the system, we find that, when the population correlations are zero, the probability of getting a determinant as large as or larger than that which results from our sample is approximately 0.123 in the case of equation (3.1.31) and 0.861 in the case of equation (3.1.31*). Thus (3.1.31*) is for practical purposes consistent with the assumption that the u's are random with respect to time, whereas (3.1.31) is not. Another test of assumptions that can be applied is the test of independence between exogenous variables and disturbances. It is implied in the hypotheses adopted that there is no correlation between the exogenous variables and the disturbances. We compute the determinant of the matrix of all correlations between exogenous variables and the estimated disturbances. The result is that the probability is approximately 0.075 of getting a determinant as large as or larger than that observed, with equation (3.1.31) in the system, when the hypothesis is one of no correlation. The corresponding probability for the system with (3.1.31*) is 0.997.[13]

There seems to be little doubt, on statistical grounds, that the system with (3.1.31*) is preferable to the system with (3.1.31). On economic grounds, some of the implications of (3.1.31*) may be questioned. This equation states that there is negative correlation between capitalists' current income and their consumption. We may also interpret this equation as claiming that the level of profits has a positive influence on consumption, whereas the rate of change has a negative influence. This can be seen by rewriting (3.1.31*) as

$$(3.1.31^*) \quad C = 17.15 + 0.18\Pi - 0.40\Delta\Pi + 0.82(W_1 + W_2) + u_1{}^*$$

In the static case, in which profits do not change, the relation between consumption and profits is, of course, positive. Another way of stating the relationship expressed by (3.1.31*) is to say that the average consuming lag for capitalists is greater than one year. If the coefficients of Π and Π_{-1} were both positive, the average lag would be less than one year.

The fact that the rate of change of profits has a negative influence on consumption does not exclude (3.1.31*), for increasing profits may lead to such large expenditures on fixed capital that consumption receives an inverse effect. The superior statistical properties of (3.1.31*) are certainly strong points in its favor. The negative effect of the rate of change of income is found in other studies. Estimates of the influence of the rate of change of total income on total consumption show the same result in model II below. The results of models I and II are consistent if one takes the distribution between wages and other income into account. Family-budget data show the same phenomenon.[14]

Finally, we have calculated the parameters of our system by the classical *single-equation method of least squares*. In this approach, we have treated each equation independently of all others in the system. We have also had to select a dependent variable in each of the three equations along whose axis the sum of squares of residuals is to be minimized. We are able to present a unique set of estimates for this method only because of an *arbitrary* choice of dependent variables for each of the three equations. Any other choice, of which there are

[13] H. Rubin is responsible for the suggestion of the test criteria and the carrying out of the appropriate computations given in the text paragraph.

[14] See Ruth P. Mack, "The Direction of Change and the Consumption Function," *The Review of Economic Statistics*, Vol. XXX, November, 1948, pp. 239–258.

several possibilities, would lead to different estimates. The least-squares results are

$$C = 16.43 + \underset{(0.07)}{0.25}\Pi + \underset{(0.04)}{0.80}(W_1 + W_2) + u_1''' \tag{3.1.37}$$

$$I = 10.13 + \underset{(0.10)}{0.48}\Pi + \underset{(0.10)}{0.33}\Pi_{-1} - \underset{(0.03)}{0.11}K_{-1} + u_2''' \tag{3.1.38}$$

$$W_1 = 1.50 + \underset{(0.03)}{0.44}(Y + T - W_2) + \underset{(0.04)}{0.15}(Y + T - W_2)_{-1} + \underset{(0.03)}{0.13}(t - 1931) + u_3''' \tag{3.1.39}$$

$$Y + T = C + I + G \tag{3.1.19}$$

$$Y = \Pi + W_1 + W_2 \tag{3.1.20}$$

$$\Delta K = I \tag{3.1.21}$$

Compared with the estimates obtained by the more appropriate statistical techniques, the single-equation estimates of the parameters show the same signs, but the orders of magnitude are different. By the equation-system methods of estimation, a higher weight is given to Π_{-1} than to Π in the investment equation, whereas the single-equation method attributes a greater weight to Π than to Π_{-1}. But it should be pointed out that, in all three cases, the sum of the coefficients of Π and Π_{-1} is about the same. The most important discrepancy occurs in the estimates of the capitalists' marginal propensity to consume. It is also interesting to consider the steady-state marginal propensity to spend out of capitalist income. This marginal propensity to spend is the sum of the coefficients of Π and Π_{-1} in (3.1.37) and (3.1.38). The method of reduced forms leads to estimates less than unity; the method of maximum likelihood (using all available information), to estimates slightly greater than unity. The least-squares estimate, on the other hand, is farther above unity. The least-squares estimate seems, on common-sense grounds, to be farther from the truth than do the estimates obtained by the other methods.

We shall complete the statistical description with the presentation of the estimates of the values of the disturbances (Table III), their root-mean-square, and the test for autocorrelation.

$$\bar{S}_1 = 1.05 \qquad \bar{S}_2 = 1.01 \qquad \bar{S}_3 = 0.77 \tag{3.1.40}$$

$$\left(\frac{\delta^2}{S^2}\right)_1 = 1.34 \qquad \left(\frac{\delta^2}{S^2}\right)_2 = 1.92 \qquad \left(\frac{\delta^2}{S^2}\right)_3 = 1.96 \tag{3.1.41}$$

TABLE III

ESTIMATES OF ANNUAL VALUES OF THE DISTURBANCES BY METHOD OF LEAST SQUARES: MODEL I

(Billions of 1934 dollars)

	u_1'''	u_2'''	u_3'''		u_1'''	u_2'''	u_3'''
1921	−0.30	−0.07	−1.30	1931	0.03	0.04	0.58
1922	−1.54	−0.05	0.29	1932	−0.15	0.37	0.10
1923	−1.57	1.25	1.18	1933	−0.14	0.22	0.45
1924	−0.42	−1.35	−0.14	1934	−0.22	−0.17	0.28
1925	0.12	0.42	−0.47	1935	−0.22	0.01	0.01
1926	1.05	1.40	−0.48	1936	1.34	0.97	−0.85
1927	1.46	0.79	−0.73	1937	−0.40	0.05	1.00
1928	1.11	−0.63	0.34	1938	0.35	−2.56	−0.47
1929	−0.47	1.08	1.20	1939	0.71	−0.67	−0.38
1930	0.83	0.28	−0.15	1940	0.69	−0.76	−1.09
				1941	−2.28	−0.64	0.59

Stability Conditions and Multipliers

The solution for Y of the dynamical system (3.1.31)–(3.1.33) and (3.1.19)–(3.1.21) leads to the difference equation

$$
\begin{aligned}
(3.1.42)\quad 0.5689Y = {} & 0.9675Y_{-1} - 0.5834Y_{-2} + 0.1020Y_{-3} + 3.2247 \\
& + 0.0188(t - 1931) + 0.4389W_2 \\
& - 0.1796(W_2)_{-1} - 0.0966(W_2)_{-2} \\
& - 0.1020(W_2)_{-3} - 0.6689T + 0.3841T_{-1} \\
& + 0.0966T_{-2} + 0.1020T_{-3} + G - 0.83G_{-1} \\
& + u_1'' - 0.83(u_1'')_{-1} + u_2'' - (u_2'')_{-1} \\
& + 0.77u_3'' - 1.3055(u_3'')_{-1} + 0.68(u_3'')_{-2}
\end{aligned}
$$

The general solution of this difference equation expresses Y as a function of time. It is equal to the general solution of the homogeneous equation (i.e., the equation obtained by suppressing the terms that do not contain Y) plus a particular solution of the complete equation. It is not, in general, possible to say in advance whether or not the general solution is damped, but we can examine the stability properties of the solution of the homogeneous equation. The general solution of the homogeneous equation is damped if all the roots of the following equation are less than unity in absolute value:

$$0.5689\lambda^3 - 0.9675\lambda^2 + 0.5834\lambda - 0.1020 = 0 \tag{3.1.43}$$

The solution yields two conjugate complex roots and one real root. The largest root in absolute value is 0.92, which gives a damped solution for the homogeneous equation. The complex roots of (3.1.43) introduce periodicity into the model and thus give a representation of the business cycle. The damping is not great, and the fluctuations of the exogenous variables superimposed upon the solution of the homogeneous equation could lead to an explosion of the system.

The root of (3.1.43) is a function of the estimates of the parameters of the system and is therefore subject to sampling fluctuations. It is not known whether a confidence interval, for a reasonable confidence coefficient, would cover values of $|\lambda| > 1$.

Similar calculations can be carried out with the other estimates of Model I. Equations (3.1.42) and (3.1.43) are given merely as examples of methods of analyzing the properties of the models.

We are also able to calculate various "multipliers" from the models. Let us take for an example the government spending multiplier when taxes and government wages remain unchanged. The multipliers obtained from (3.1.42) differ significantly from the traditional concept of multipliers in that the exogenous variable which is subjected to autonomous impulses, G in this case, occurs with a lag as well as with its current value. In an equation of the form

$$Y = \sum_{i=1}^{p} \mu_i Y_{-i} + \nu G \tag{3.1.44}$$

the multiplier is

$$\frac{d\overline{Y}}{dG} = \frac{\nu}{1 - \sum_{i=1}^{p} \mu_i} \tag{3.1.45}$$

where $\overline{Y}$ represents the stationary solution of (3.1.44). We find this multiplier by substituting $Y_{-i} = \overline{Y}$; $i = 0, 1, 2, \cdots, p$, solving for $\overline{Y}$ in terms of G, and differentiating $\overline{Y}$ with respect to G. The multiplier thus shows the change in the stationary solution (equilibrium level) for Y as a result of a change in G.

Instead of (3.1.44) we may have an equation of the form

$$Y = \sum_{i=1}^{p} \mu_i Y_{-i} + \nu_1 G + \nu_2 G_{-1} \tag{3.1.46}$$

In this case, it can be shown that the multiplier is

$$\frac{d\overline{Y}}{d\overline{G}} = \frac{\nu_1 + \nu_2}{1 - \sum_{i=1}^{p} \mu_i} \tag{3.1.47}$$

This multiplier is identical with that obtained by inserting $Y_{-i} = \overline{Y}$, $i = 0, 1, 2, \cdots, p$, and $G = G_{-1} = \overline{G}$ into (3.1.46), solving for $\overline{Y}$ in terms of $\overline{G}$, and differentiating $\overline{Y}$ with respect to $\overline{G}$.

Equation (3.1.42) yields a multiplier of the form (3.1.47) with $p = 3$. Furthermore, it can be seen from (3.1.42) that $\nu_1 > 0$ and $\nu_2 < 0$; consequently the lag in government spending occurring in this model has a depressing effect on the multiplier. The structural characteristic of the three-equation model which gives rise to the term G_{-1} in equation (3.1.42) is the presence of the term $\beta_3 K_{-1} (\beta_3 < 0)$ in the investment equation. If there were no effect of capital accumulation on investment ($\beta_3 = 0$), there would be no lag in G in (3.1.42), and the multiplier would be greater than it would be otherwise. We can easily compute the multiplier from our system in terms of the original structural parameters in (3.1.16)–(3.1.21) (the α's, β's, γ's) under two hypotheses: (1) $\beta_3 < 0$, (2) $\beta_3 = 0$. Under the first hypothesis, we get

$$\frac{d\overline{Y}}{d\overline{G}} = \frac{1}{1 - \alpha_2(1 - \gamma_1 - \gamma_2) - \alpha_1(\gamma_1 + \gamma_2)} \tag{3.1.48}$$

Under the second hypothesis, we get

$$\frac{d\overline{Y}}{d\overline{G}} = \frac{1}{\left\{\begin{matrix} 1 - \alpha_2(1 - \gamma_1 - \gamma_2) - \alpha_1(\gamma_1 + \gamma_2) \\ - \beta_1(1 - \gamma_1 - \gamma_2) - \beta_2(1 - \gamma_1 - \gamma_2) \end{matrix}\right\}} \tag{3.1.49}$$

For the usual range of variation of the parameters $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2$, the second multiplier will be greater than the first. The introduction of the variable K_{-1} in the investment equation has a depressing influence on the multiplier effects of government spending. If capital accumulation is a part of the model, the multiplier effect of government spending is given by (3.1.48) and is equal to the reciprocal of one minus the *marginal propensity to consume*. If capital accumulation is not a part of the model, the multiplier effect of government spending is given by (3.1.49) and is equal to the reciprocal of one minus the *marginal propensity to spend*. The marginal propensity to spend covers both consumer and producer goods and is always greater than the marginal propensity to consume, assuming reasonable ranges of variation for

α_1, α_2, β_1, β_2, γ_1, γ_2, of course. There is a very simple explanation for these differences in the multiplier. In any stable dynamical system the multiplier is equivalent to the derivative of the stationary level of income with respect to the change in an exogenous variable or parameter. But the stationary solution to a system in which K occurs as a variable implies $K = K_{-1}$, which, in turn, implies $I = I_{-1} = 0$. If investment is zero in the stationary solution, the marginal propensity to invest will not affect the multiplier which shows the change in the stationary solution with respect to a change in government spending, for example.

One of the major defects of some previous econometric models with endogenous investment has been the fact that they impose very large multipliers on the system.[15] If investment is not considered entirely exogenous, the marginal propensity to spend is often greater than 0.9 as calculated from least-squares regressions of C on Y and of I on Y. This has always bothered economists who have thought intuitively that the multiplier should be in the range 2.0 to 3.0, implying a marginal propensity to spend of about 0.5 to 0.7. However, the previous econometric investigations have omitted the variable K_{-1} from the investment equation and have generated forecast equations like (3.1.44) rather than like (3.1.46).

The multiplier in our three-equation system can easily be calculated from the estimates of our parameters in (3.1.31)–(3.1.33). The value is

$$\frac{dY}{dG} = 1.9 \tag{3.1.50}$$

This multiplier appears to be reasonable.

A multiplier can be viewed as the added income caused by a unit impulse of government spending, where the added income is cumulated over all future time to $t = \infty$. A more satisfactory concept may be that of the truncated multiplier in which the added income is summed over a *finite* number of future years. It is not elucidating to write the general formula for the truncated multiplier of n years because it is so complex algebraically. However, we can easily calculate this multiplier for two or three years. For two years (in the model with $\gamma_2 = 0$) it is

$$\frac{\left\{\begin{matrix} 1 - \alpha_2(1-\gamma_1) - \alpha_1\gamma_1 - \beta_1(1-\gamma_1) \\ + \beta_2(1-\gamma_1) + \beta_3\beta_1(1-\gamma_1) \end{matrix}\right\}}{1 - \alpha_2(1-\gamma_1) - \alpha_1\gamma_1 - \beta_1(1-\gamma_1)} \tag{3.1.51}$$

[15] See, e.g., Mordecai Ezekiel, "Statistical Investigations of Saving, Consumption and Investment," *The American Economic Review*, Vol. XXXII, March, 1942, pp. 22–49, and June, 1942, pp. 272–307.

It is evident from (3.1.51) that the truncated multiplier of two years varies inversely with the absolute value of β_3 provided $\beta_3 < 0$, $\beta_1 > 0$, $(1 - \gamma_1) > 0$, and the denominator is positive, as is reasonable. The greater becomes the effect of capital accumulation in the investment function, the smaller becomes the truncated multiplier, other parameters remaining constant.

MODEL II, A REDUCED FORM [16]

We shall now turn to a different type of model, the validity of which will serve as a test of other hypotheses; this model will demonstrate another statistical method. We shall now be interested primarily in testing whether the stock of cash balances or lagged income or both have a significant influence on consumption. We shall also compare the multiplier effect of increasing the amount of money with the multiplier effect of increasing the government deficit.

There are several alternative economic structures in which the amount of money can enter as a variable. Let us examine first the following simple model:

$$\text{(3.2.1)} \qquad \frac{\mathsf{C}}{pN} = \alpha_0 + \alpha_1 \frac{\mathsf{Y}}{pN} + \alpha_2 \left(\frac{\mathsf{Y}}{pN}\right)_{-1} + \alpha_3 \left(\frac{\mathsf{M}}{pN}\right)_{-1} + u$$

$$\text{(3.2.2)} \qquad \mathsf{GNP} = \mathsf{C} + \mathsf{I}' + \mathsf{G}$$

$$\text{(3.2.3)} \qquad \mathsf{GNP} = \mathsf{Y} + \mathsf{T}$$

where C = consumption in current dollars; Y = disposable income in current dollars; M = money supply in current dollars; I' = gross investment in current dollars; G = government expenditure plus foreign balance in current dollars; GNP = gross national product in current dollars; p = cost-of-living index; N = population of continental United States; T = government receipts plus corporate savings plus business reserves minus transfer payments minus inventory profits, all measured in current dollars; u = random disturbance.

The endogenous variables are C, Y, GNP, and the exogenous variables are I'/pN, G/pN, T/pN. M is an exogenous variable, and $(\mathsf{M}/pN)_{-1}$ is a predetermined variable.

[16] A discussion of the use of this model and the model of the next section in connection with the formation of economic policies is contained in the author's paper "The Use of Econometric Models as a Guide to Economic Policy," *Econometrica*, Vol. 15, April, 1947, pp. 111–151. For a complete discussion of methods of estimation of the parameters in models similar to those of this section, see T. Haavelmo, "Methods of Measuring the Marginal Propensity to Consume," *Journal of the American Statistical Association*, Vol. 42, March, 1947, pp. 105–122.

It is important to note that the variables of (3.2.1) are all per capita. If it is assumed that the important trends in the system are identified with population growth, it becomes unnecessary to introduce specific trend variables in this system.

If we combine (3.2.2) and (3.2.3) we get

$$\text{(3.2.4)} \qquad \mathsf{Y} = \mathsf{C} + (\mathsf{I}' + \mathsf{G} - \mathsf{T})$$

in which the variable $\mathsf{I}' + \mathsf{G} - \mathsf{T}$ shows the net injection of purchasing power into the economy by business firms, government, and foreigners. It is the sum of spending by business firms for capital goods (including inventories and residences) plus spending by government for goods and services plus spending by foreigners for domestic goods and services less the retention of profits and depreciation reserves by business firms less the collection of taxes and other receipts by government less purchases of goods and services from abroad plus transfer payments by the government. We have so far regarded G/pN and T/pN as exogenous variables, but in this particular model we are also considering I'/pN an exogenous variable. There are theories of economic fluctuations to support the treatment of private investment as exogenous, although the general approach elsewhere in this volume is to classify it as endogenous.

By combining (3.2.1)–(3.2.4), we easily obtain

$$\text{(3.2.5)} \qquad \frac{\mathsf{Y}}{pN} = \frac{\alpha_0}{1-\alpha_1} + \frac{\alpha_2}{1-\alpha_1}\left(\frac{\mathsf{Y}}{pN}\right)_{-1} + \frac{\alpha_3}{1-\alpha_1}\left(\frac{\mathsf{M}}{pN}\right)_{-1} + \frac{1}{1-\alpha_1}\left(\frac{\mathsf{I}' + \mathsf{G} - \mathsf{T}}{pN}\right) + \frac{1}{1-\alpha_1}u$$

This is a reduced-form equation for which the method of least squares is unbiased in sufficiently large samples because all the variables on the right-hand side are either predetermined or exogenous. The model is such that, if we can estimate the parameters of the reduced form, we can very easily transform back to the parameters of (3.2.1).

The statistical results for the years 1922–1941 are

$$\text{(3.2.6)} \qquad \frac{\mathsf{Y}}{pN} = 186.53 + \underset{(0.13)}{0.30}\left(\frac{\mathsf{Y}}{pN}\right)_{-1} + \underset{(0.10)}{0.13}\left(\frac{\mathsf{M}}{pN}\right)_{-1} + \underset{(0.35)}{2.36}\left(\frac{\mathsf{I}' + \mathsf{G} - \mathsf{T}}{pN}\right) + u'$$

$$\bar{S} = 20.69 \qquad \frac{\delta^2}{S^2} = 1.28$$

All the variables are measured in terms of 1935–1939 dollars per person. The general cost-of-living index was used as a deflator for all price changes. The numbers below the regression coefficients in parentheses

TABLE IV

ESTIMATES OF ANNUAL VALUES OF DISTURBANCES OF (3.2.5)

(1935–1939 dollars per person)

	u'		u'		u'		u'
1922	4.5	1926	8.1	1931	−34.0	1936	25.3
1923	−12.8	1927	−0.7	1932	−48.9	1937	4.3
1924	8.4	1928	23.5	1933	−5.2	1938	1.1
1925	−7.8	1929	5.2	1934	8.3	1939	23.4
		1930	−8.1	1935	10.0	1940	18.7
						1941	−22.8

are standard errors (valid in large samples), and $\bar{S}$ is the standard error of estimate adjusted for degrees of freedom.

The estimates of the parameters of (3.2.1) are given by

(3.2.7) $$a_1 = 0.58 \qquad a_2 = 0.13 \qquad a_3 = 0.06$$

We are interested in testing whether or not α_3 is significantly different from zero. The Keynesian theory of employment has often been based implicitly on the assumption that α_3 is near zero. If the Keynesian assumption is correct, policies of wage cuts or increases in the money supply will not be suitable for increasing the level of employment. It is important to know in times of depression whether to follow policies of deficit spending or policies of easy money. We could make analogous remarks for a study of policies to combat inflation.

We can form a confidence interval for the coefficient of $(M/pN)_{-1}$ in (3.2.6) of the form $0.13 \pm (0.10)t_\alpha$, where t_α is taken from the t-distribution at the α per cent significance level. For any significance level of 5 per cent or less, we find that the confidence interval covers the value zero; thus we cannot reject the hypothesis that the true value of the parameter is zero. If the coefficient $\alpha_3/(1 - \alpha_1)$ is not significantly different from zero, it follows that α_3 is not significantly different from zero because $1/(1 - \alpha_1)$ is significantly different from zero. The coefficient of $[(I' + G - T)/pN]$ in (3.2.5) is $1/(1 - \alpha_1)$, and the very small standard error of the estimate of this coefficient implies that it is significantly non-zero. We conclude that the data do not refute the theory that consumer demand is independent of the stock of cash balances.

If we assume $\alpha_3 = 0$ in advance of statistical estimation, our reduced form calculation will be

$$(3.2.8)\quad \frac{\mathsf{Y}}{pN} = 202.54 + \underset{(0.12)}{0.37}\left(\frac{\mathsf{Y}}{pN}\right)_{-1} + \underset{(0.28)}{2.39}\left(\frac{\mathsf{I}' + \mathsf{G} - \mathsf{T}}{pN}\right) + u''$$

$$\bar{S} = 21.21 \qquad \frac{\delta^2}{\bar{S}^2} = 1.14$$

Equation (3.2.8) is, from a statistical point of view, as good for the prediction of Y on the basis of predetermined or exogenous variables as is (3.2.6), although the estimated disturbances in the second formulation have more serial correlation. These examples show clearly that lagged income is a significant variable since its coefficient has a relatively low standard error in both (3.2.6) and (3.2.8). In neither case can we accept at the 5 per cent level of significance the hypothesis that $\alpha_2/(1 - \alpha_1)$ is zero.

We can calculate multipliers in the system corresponding to the point estimates, but we must remember that the point estimates for α_3 are covered in a very big confidence interval. The multipliers for the equilibrium solution of the system in (3.2.6) are

$$(3.2.9)\qquad \frac{d\left(\dfrac{\bar{\mathsf{Y}}}{pN}\right)}{d\left(\dfrac{\mathsf{I}' + \mathsf{G} - \mathsf{T}}{pN}\right)} = 3.37$$

$$(3.2.10)\qquad \frac{d\left(\dfrac{\bar{\mathsf{Y}}}{pN}\right)}{d\left(\dfrac{\bar{\mathsf{M}}}{pN}\right)} = 0.19$$

An extra dollar of deficit spending (money supply constant) creates \$3.37 additional disposable income, and an extra dollar of money supplied (deficit constant) creates \$0.19 additional disposable income.

It should be obvious to the reader that (3.2.1)–(3.2.3) is by no means the only set of structural equations that will lead to a reduced form like (3.2.5), i.e., to a linear form in the same variables. For example, the interest rate may be introduced as an endogenous variable in (3.2.1) and also in a new equation showing the demand for money. This type of system would lead to a reduced form in the same variables as those of (3.2.5). The reader is free to experiment with other systems which lead to the same reduced form and therefrom test new hypotheses or make new tests of the specific hypotheses considered above.

MODEL III, A LARGE STRUCTURAL MODEL

The two statistical examples of model building discussed thus far in this chapter were used to illustrate techniques and methodology; consequently some sacrifices were made in each model. In Model I, the three-equation system, we sacrificed details of economic behavior patterns in order to illustrate different methods of structural estimation in dynamical economic systems. Since the computations by the newer methods of statistical inference are complicated and present a number of technical problems, it has been, therefore, very instructive to experiment, in a preliminary way, with a small system where the number of parameters is more manageable. In Model II we estimated those parameters that were necessary for purposes of certain types of forecasting; in addition, it was not difficult, in that case, to estimate the structural parameters by simple algebraic transformations of the least-squares estimates of the parameters of the reduced forms. This was possible because Model II possesses just as many [but not more] identifying conditions as [than] are necessary to make the structural parameters identifiable. However, if the system had possessed more identifying conditions, i.e., if it had been overidentified, the transformation from the estimates of the reduced-form parameters to estimates of the structural parameters would have been more complicated. The simplifications of the system led to relatively easy methods of statistical estimation and showed to what extent, and in what way, we can estimate economic behavior patterns by the method of least squares on a single-equation basis.

On the basis of the theoretical developments in Chapter II, we can formulate a more complete aggregative model of the economic system. It does not follow that we shall be able to find enough data to test such a model; however, we shall first proceed to write down a particular model without considering the available data, and then point out where the data are lacking and proceed to make compromises.

From the theory of the firm [17] we can develop the demand for producer goods, labor, and inventories. The demand for labor is of the form

$$\text{(3.3.1)} \qquad \mathsf{W} = \alpha_0 + \alpha_1 pX + \alpha_2 (pX)_{-1} + u_1$$

[17] See Chapter II, pp. 14–40.

where W = aggregate wage bill, and pX = aggregate value of privately produced output.

In the profit-maximization theory we have a demand equation for producer goods (net of depreciation):

$$I = \beta_0 + \beta_1 \frac{pX}{q} + \beta_2 \left(\frac{pX}{q}\right)_{-1} + \beta_3 K_{-1} + u_2 \tag{3.3.2}$$

where I = net investment (in constant prices), q = price of producer goods, K = end-of-period stock of capital (in constant prices).

The demand for inventories, as we developed this theory in Chapter II, actually dealt only with speculative stocks, but there are certain "pipeline" or "transactions" holdings of goods which are purely a result of the institutional setting. Stores must have a certain amount of stocks on hand to show customers and out of which to make sales; these stocks are called transactions stocks and should be added to speculative stocks in order to get total stocks. The transactions inventories are purely a function of sales. We may also divide our total inventory stocks according to another classification, namely, consumer goods and producer goods. Denoting real stocks of consumer goods held at the end of the period by H_1 and real stocks of producer goods held at the end of the period by H_2, we obtain the two inventory demand equations:

$$H_1 = \gamma_0 + \gamma_1 C + \gamma_2 p_c + \gamma_3 (p_c)_{-1} + u_3 \tag{3.3.3}$$

$$H_2 = \delta_0 + \delta_1 I + \delta_2 q + \delta_3 q_{-1} + u_4 \tag{3.3.4}$$

In equations (3.3.3) and (3.3.4), C = sales of consumer goods (in constant prices), I = sales of producer goods (in constant prices); p_c = price of consumer goods, q = price of producer goods.

On the consumers' side of the market, we have the demand equation for consumer goods as developed from the theory of household behavior. This equation is

$$C = \epsilon_0 + \epsilon_1 \frac{pX}{p_c} + u_5 \tag{3.3.5}$$

The consumers and producers interact in the market to determine prices from the market adjustment equations:

$$\Delta p_c = \zeta_0 + \zeta_1 (u_3)_{-1} + u_6 \tag{3.3.6}$$

$$\Delta q = \eta_0 + \eta_1 (u_4)_{-1} + u_7 \tag{3.3.7}$$

In (3.3.6) or (3.3.7) or both we can use the rate of change of output instead of the rate of change of price, depending upon the characteristics of the market being considered.

In addition to the seven behavior equations (3.3.1)–(3.3.7) there are three definitional equations which close the system:

$$Iq + Cp_c + p_c \,\Delta H_1 + q\,\Delta H_2 + p_c G = pX \tag{3.3.8}$$
[18]

$$I + C + \Delta H_1 + \Delta H_2 + G = X \tag{3.3.9}$$

$$\Delta K = I \tag{3.3.10}$$

The variable G in (3.3.9) represents government expenditures for goods and services plus the net foreign balance.

The above system contains ten equations; ten endogenous variables: I, q, C, p_c, H_1, H_2, p, X, K, W; and one exogenous variable, G. Data are available for the measurement of all the variables of this system except two, H_1 and H_2. There are numerous data on aggregate inventories and on inventories in many separate industries such as manufacturing, mining, agriculture, but there are no satisfactory data on inventories classified according to producer-goods industries and consumer-goods industries. If we want to include also the production functions for the two categories of industries, one for the producer-goods industries and one for the consumer-goods industries, we shall neet with the same difficulties in attempting to divide the input of labor and capital according to these two industrial groupings.

We shall be forced to compromise at this stage and develop systems that do not have enough equations to determine the separate prices of consumer and producer goods. Instead, we shall use an inventory demand equation and a market adjustment equation that refer to both industries. We shall have enough equations to determine a combination of the separate price levels but not enough to determine each price level individually.

At the same time that we are reducing the number of equations as a result of the lack of data, we shall also add some new equations which explain certain markets in greater detail. In particular, we shall introduce money and housing as separate markets. This will give us more information on important sectors of the economy. Furthermore, there are certain refinements which must be introduced into the various equations to take account of excise taxes, government production, etc. We shall thus reconsider all the equations of the system in order to obtain the final form for purposes of statistical estimation.

[18] All purchases included in G have arbitrarily been evaluated at the market prices of consumer goods. It would, of course, be possible to divide G between consumer and producer goods.

Taxes and the Theory of the Firm

We have shown, in Chapter II, that the theory of profit maximization for the individual firm implies that the behavior equations representing the demand for factors of production are independent of the particular income tax schedule in operation. Such independence does not hold when considerations broader than profit maximization (e.g., utility maximization) are taken into account. For the model of this section, we shall assume only profit maximization. At the end of this chapter, though, we shall test some of the hypotheses which follow from the more general theories.

Excise taxes must be treated differently from income taxes. In the theory of profit maximization, the particular excise tax schedule does have some influence on the behavior equations. Excise taxes can be considered a deduction from sales value which the entrepreneur must turn over to the government. In place of the aggregate value of privately produced output as a variable in the investment and wage equations, we shall use the aggregate value of privately produced output less excise taxes. We shall also allow for trends in the system and introduce t as an explicit variable representing time. Our equations now take the form

$$(3.3.1a) \quad \mathsf{W} = \alpha_0 + \alpha_1(pX - \mathsf{E}) + \alpha_2(pX - \mathsf{E})_{-1} + \alpha_3 t + u_1$$

$$(3.3.2a) \quad I = \beta_0 + \beta_1 \frac{pX - \mathsf{E}}{q} + \beta_2 \left(\frac{pX - \mathsf{E}}{q}\right)_{-1} + \beta_3 K_{-1} + \beta_4 t + u_2$$

where p = price level of output as a whole, X = aggregate privately produced output (in constant prices), E = excise tax payments.

Inventories

The inventory demand equation is not influenced by the introduction of excise taxes into the model because the sales variable in that equation is introduced principally to show the demand for transactions holding; for this demand we need a measure of the real value of sales including excise taxes, since the price deflators used to get "real" sales from the value of sales is a measure of the prices paid by the final user of the goods. Such prices include excise taxes. But in addition to our transactions and speculative demand for inventories, there is one other type of demand which should be superimposed upon the system. There is an "inertia" factor in the holding of various types of assets. It is not easy to switch immediately from holding a large amount of assets to holding a small amount (or vice versa). People tend to hold assets in any particular period in some relation to the assets of the past period. There is some persistency in the structure of the system merely because ownership cannot change without delay and friction. We shall thus introduce an inventory equation made up of three demand variables: (1) transactions, (2) speculation, (3) inertia. The equation is

$$(3.3.3a)\quad H = \gamma_0 + \gamma_1(X - \Delta H) + \gamma_2 p + \gamma_3 p_{-1} + \gamma_4 H_{-1} + \gamma_5 t + u_3$$

The variable $X - \Delta H$ equals production less the net change in stocks, which is the same thing as sales.

Consumption

The consumption function is derived from the theory of household behavior. In most versions of that theory income enters the consumption function via the budget constraint, subject to which utility is maximized. If the government takes away income from households in the form of personal taxes or if business firms retain some profits in the form of corporate savings, these funds are not available for spending or saving, and they are not included in the budget constraint. Similarly, if the government makes transfer payments (relief, pensions, etc.) out of its revenues, disposable income is added back into the consumers' budget constraint. The appropriate variable for the consumption function is thus *disposable* income which is defined as net national income less personal taxes less social security payments less corporate savings plus transfers. It is also equal to net national product less business taxes less personal taxes less social security payments less corporate savings plus transfers.

If a consumer owns assets which change in value, the capital gain or loss should have some influence on consumption patterns, especially if the capital gain or loss is realized by selling the asset at its market price. In the most general theory of consumer behavior of Chapter II, the form of the budget constraint makes it necessary to include capital gains and losses in disposable income. Only the lack of good data on capital gains leads us to exclude them from the consumption function, but it is an interesting hypothesis that should be tested.

In (3.3.5), pX denotes the income variable of the consumption function. In our revised and expanded version of the model we shall let Y denote disposable income in *constant* prices ("real" income). Our consumption function will now appear as

$$(3.3.5a)\quad C = \delta_0 + \delta_1 Y + \delta_2 t + u_4$$

The reader should notice the difference between this consumption function and that of the above section on the simple three-equation system. In that system we distinguish between wage income and other (profit) incomes, but here we lump all income into the variable Y. Also, corporate savings are treated differently in the two models.

The trend term in (3.3.5a) accounts for shifts in the consumption function by changing the constant term. For each period of time we get a new relation between C and Y which is everywhere δ_2 units above or below the consumption function of the previous period. But the shifts in the consumption function need not be parallel shifts. The slope as well as the constant term may change gradually with time. An alternative form of the consumption function which allows for shifts in the marginal propensity to consume as well as in the constant term is

$$C = \delta_0' + \delta_1'Y + \delta_2'Yt + \delta_3't + u_4' \tag{3.3.5b}$$

In (3.3.5b) the marginal propensity to consume changes according to a linear trend

$$\frac{\partial C}{\partial Y} = \delta_1' + \delta_2't$$

Consumption functions of type (3.3.5a) and of type (3.3.5b) have been tested in the model to be presented below in full.

Housing

There is one good in the system which is sometimes called a consumer good and sometimes called a producer good. This good, which is a borderline case, is the new construction of residential housing. There are many particular factors which distinguish housing from other parts of the economy: extreme durability, a close relationship to population variables, long cycles, and others. Rather than lump housing entirely with investment [19] or entirely with consumption, our practice has been to separate the housing market as a subset of the equations of our system. Since residential construction is a major component of private capital formation, it is useful to separate it out because we then get more specific information about the components of consumption and investment. In other words, if the relationship between private producers' expenditure on plant and equipment and income is not obscured by the relationship between housing and income, we may be able to get a better estimate of the marginal propensity to spend, especially since we shall find that housing is related to a different form of the income

[19] In the simple, three-equation model, housing is included in investment.

variable[20] than are private producers' expenditures on plant and equipment.

Our first step in the analysis of the housing market is to divide the demand for residential construction into two categories: (1) the new housing demanded by owner-occupiers and (2) the new housing demanded by landlords. The importance of making separate classifications for the two types of owners should be obvious. People who purchase a house to live in behave more like households demanding any other durable consumer good, and people who purchase a house to rent to others behave more like business firms producing services. We shall rely very much on the theory of household behavior to explain the demand on the part of owner-occupiers and on the theory of the firm to explain the demand on the part of landlords.

Let us consider, first, the demand for new residential construction for direct occupancy by the owner. The theory of consumer behavior of the preceding chapter states that the prospective buyer of a new house will take into account all the relative prices of goods in his budget, his income, and possibly his liquid wealth. All these variables need not influence his demand, although they are all potentially important. As a first approximation for our model we shall select those relative prices which should be most important in making a decision whether or not to purchase a new house for occupancy. The price of new houses and the price of housing services which would have to be paid if the purchaser were to rent from a landlord are certainly two of the most important prices to be considered. We shall represent the price of new houses by an index of construction costs, and the price of housing services by an index of rents. Other prices which may be relevant are the prices of old houses, the mortgage interest rates, and an index of the prices of other items in the consumer's budget. The prices of old houses and the mortgage interest rate are two important time series which have never been prepared and which leave a serious gap in our business annals. It may be considered that the price of old houses is closely related to rent, so that our rent variable already accounts for the variations in the former variable. In order to avoid expanding the size of our system beyond manageable proportions, we have not included the prices of other consumer goods. The introduction of a new endogenous variable always calls for the introduction of a new equation to maintain the completeness of the system.

The income variable of the housing demand equation (owner-

[20] Disposable income is the correct variable for the demand equation of owner-occupied housing, but national product is the correct variable for the demand equation of plant and equipment.

occupiers) will be the same as the income variable of the consumption function, namely, disposable income. However, since the down payment on a new house represents a relatively large consumer outlay, the income of one year may not be an adequate variable for the demand equation. Instead the disposable income of the past three to five years may be a more appropriate variable. By introducing the cumulated income of the past three years we have taken account of the effect of both income and liquid wealth, because wealth is accumulated from past income.

We have discussed the economic demand factors that follow from the theory of household behavior, but there is at least one variable that is peculiar to the housing market which must receive explicit attention in addition to that given the economic demand factors. The additional variable is population growth. There is an institutional relation dictated by our social structure which must hold approximately:

Every family must have a separate dwelling unit.

The economic factors of rent, construction cost, and income will, of course, show why families double up in depression or own both town and country dwellings in prosperity, but over long periods of time there will tend to be just as many dwelling units as families. The number of new families created per period shows the *need* for housing and will be introduced as a variable into our demand equation, which will now take the form

$$(3.3.11)\quad D_1 = \epsilon_0 + \epsilon_1\left(\frac{r}{q_1}\right) + \epsilon_2(Y + Y_{-1} + Y_{-2}) + \epsilon_3\,\Delta F + u_5$$

where D_1 = gross expenditures on new residential construction for owner-occupancy measured in constant prices, r = index of rents, q_1 = index of construction costs, Y = disposable income measured in constant prices, ΔF = number of new families. It should be stressed that ΔF is not simply another way of writing D_1. In the first place, the two variables are measured in different units. This is as it should be, for new families can demand housing units of various qualities and sizes. In the second place, D_1 refers to only a part of the housing market, whereas ΔF refers to the entire market. New families may either buy or rent a dwelling, and only by introducing ΔF as a separate variable can we measure the extent to which they buy or rent.

Two other qualifications must be made with respect to (3.3.11). First, the variable D_1 represents gross rather than net construction. We have been able to calculate total housing depreciation but not that on rented residences and owner-occupied residences respectively. We have been able to use total housing depreciation in order to get net total residential construction and net national product, but we have not been able to obtain estimates of *net* expenditures on the construction of owner-occupied residences. Second, equation (3.3.11) refers only to the non-farm sector of the population. Most of the variables refer to the non-farm sector: r = rents on non-farm dwellings, q_1 = construction costs in urban areas, ΔF = number of new non-farm families, D_1 = expenditures on non-farm residential construction of owner-occupied residences, but Y = total disposable income. There is a slight inconsistency in the case of Y, but all other variables refer specifically to the non-farm sector.

We turn now to the development of the demand equation for new rented residences. In order to develop this equation, we shall rely on the theory of the firm as presented in Chapter II. In the housing market, houses for rent play the role of fixed capital goods which produce housing services. The analogue of the ordinary investment equation in the housing market would be something like

$$(3.3.12) \quad D_2 = \zeta_0' + \zeta_1' \frac{rS_H}{q_1} + \zeta_2' \left(\frac{rS_H}{q_1}\right)_{-1} + \zeta_3'(K_H)_{-1} + u_6'$$

where D_2 = gross expenditures on new residential construction for rental purposes measured in constant prices, S_H = real output of rental housing services, K_H = stock of capital in rental housing. The form of equation (3.3.12) is not convenient for our model. Instead of introducing S_H, the output of rental housing services, as a new variable, we can rewrite (3.3.12) in a way which corresponds more closely to (3.3.11). We can write

$$(3.3.13) \qquad D_2 = D_2[r, r_{-1}, q_1, (q_1)_{-1}, w_1, (w_1)_{-1}, (K_H)_{-1}]$$

where w_1 = wage rate paid to labor producing housing services (janitors' wages, for example). Equation (3.3.13) is obtained by solving the production function and all the profit-maximizing equations for output and the factor inputs in terms of market variables alone such as r, q_1, w_1, and predetermined variables like $(K_H)_{-1}$. But even equation (3.3.13) will have to be modified. From a priori considerations it seems to be a useful simplification to omit the variable w_1 and assume that it has a negligible influence in this equation. The main factor of production in the provision of housing services is the dwelling, and wages paid to attendants who take care of the dwelling should be of second-order importance. This is certainly true of (unattended) one- and two-family rental units. Another alteration in (3.3.12) is forced upon us through lack of data. Since we are not able to split residential dwelling depreciation into depreciation on owner-occupied units and on rental units, we cannot calculate K_H. We can calculate the total stock of capital in housing covering both types of units, but we cannot calculate the separate components of the stock of capital. Consequently, we must omit the variable K_H from (3.3.13).

The demand for new dwelling units for owner occupancy and for rental purposes does not exhaust the total demand for new dwelling units. There is yet another category of demand which is called speculative demand. Speculative building occurs especially in boom periods. Speculative building will depend largely on movements in the price of houses. If the price of new houses is expected to rise, speculators will demand new houses to sell on the rising market and thus make capital gains. We shall represent anticipated movements in the price of new houses by lags in q_1, which is the variable representing the cost of new houses. The speculator will not only consider price movements in the new housing market; he will also consider the alternative advantage of investing his funds in the security markets. The rate of interest on corporate securities, as well as movements in q_1, will be taken as an indicator of the speculative demand for new houses. The need for houses as shown by the formation of new families will also be a factor in the market for new rental houses. Landlords will demand new units in order to supply those new families who are not owner-occupiers.

A linear approximation to the demand equation for new non-owner-occupied units is

$$(3.3.14)\quad D_2 = \zeta_0 + \zeta_1 r_{-1} + \zeta_2(q_1)_{-1} + \zeta_3(q_1)_{-2} + \zeta_4 i + \zeta_5(\Delta F)_{-1} + u_6$$

where i = corporate bond yield. The lags introduced in (3.3.14) are purely the result of empirical study. We do not know in advance the length of lags in the system, and we accept those that are found in the data. Often it has been found in preliminary and provisional investigations that the equations describing household behavior do not have lags, but that those describing business firm behavior often have significant lags. Except for the lagged income terms in (3.3.11), the housing market adheres to this observed pattern. Like (3.3.11), equation (3.3.14) refers to non-farm housing only.

Equations (3.3.11) and (3.3.14) are to be introduced into our model, but they carry along with them new variables. The model of the housing market is not complete unless more equations are introduced to explain the level of rents, which is one of the new endogenous variables introduced in (3.3.11) and (3.3.14).

To explain the level of rents it is necessary to consider a market broader than the market for new dwelling units. Rents are paid on both new and existing units; hence the determination of rent must be based on an analysis of supply and demand for all dwelling space. We shall denote the demand for dwelling space by N^D and the supply by N^S. The demand for housing services, like the demand for any commodity in the ordinary consumer's budget, will be a function of income and prices. We write our equation as

$$N^D = \eta_0' + \eta_1' r + \eta_2' Y + \eta_3' t + u_7' \tag{3.3.15}$$

It is certainly conceivable that the prices of other goods in the budget could enter as variables in (3.3.15), but these prices have never been found to be statistically significant. The reader should understand clearly the difference between (3.3.15) and (3.3.11). Equation (3.3.11) is the demand equation for the ownership of a new house, whereas equation (3.3.15) is the demand equation for shelter.

The supply of dwelling units, new and old, is not explained in a separate equation of our system. This supply is the result of 50 to 100 years of past history of the system and will be regarded as a predetermined variable represented by the number of existing housing units.

There are data on the percentage of dwelling units occupied and the percentage of those vacant. We shall represent the percentage occupied by v. On dividing (3.3.15) by N^S, we get

$$\frac{N^D}{N^S} = v = \frac{\eta_0' + \eta_1' r + \eta_2' Y + \eta_3' t + u_7'}{N^S} \tag{3.3.16}$$

Let us denote the mean of v by $\bar{v}$ and the mean of N^S by $\bar{N}^S$ and write $v = V + \bar{v}$, $n^S = N^S - \bar{N}^S$. Now (3.3.16) becomes

$$(\bar{v} + V)(n^S + \bar{N}^S) = \eta_0' + \eta_1' r + \eta_2' Y + \eta_3' t + u_7'$$

or

$$v = \bar{v} + \frac{1}{\bar{N}^S}(\eta_0' + \eta_1' r + \eta_2' Y + \eta_3' t - \bar{v}N^S - Vn^S + u_7')$$

If $Vn^S = 0$, we can rewrite (3.3.16) in the form

$$v = \eta_0 + \eta_1 r + \eta_2 Y + \eta_3 t + \eta_4 N^S + u_7 \tag{3.3.17}$$

N^S does not have great variation; hence it is a good approximation to assume $Vn^S = 0$.

The final equation in the housing market will be designed to show what causes fluctuations in the level of rents. For this equation we rely on the theory of market behavior, the fluctuations of prices in response to supply and demand. The hypothesis is made that vacant houses represent excess supply of housing space and cause rents to fall. When vacancies fall below some minimum frictional level, rents begin to rise. The variable v in (3.3.17) shows the ratio of demand to supply and will be taken as the measure of excess demand or supply in the housing market.

We do not know from theoretical consideration alone whether the absolute or percentage rate of change of rents is related to v. It seems plausible to assume that it is easier for rents to fall than to rise when they are high and that it is easier for rents to rise than to fall when they are low. This suggests that the absolute change in the level of rents should be a function of v and of the initial position of the rent levels. Furthermore, landlords will find it easier to raise rents as a consequence of excess demand if incomes are high than if incomes are low. Perhaps the disposable income should also be a variable in the rent adjustment equation. These considerations lead to the hypothesis

$$\Delta r = \theta_0 + \theta_1 v_{-1} + \theta_2 Y + \theta_3 \frac{1}{r_{-1}} + u_8 \tag{3.3.18}$$

which is a formulation of the rent adjustment equation.

Money

We have discussed the demand for goods thus far in our more comprehensive system. But households and business firms may also hold money or securities. We shall not need to develop the demand for both money and securities since, if one is given, the other will follow as a residual in a complete system. If we know the total demand for goods and for money, we can always obtain the demand for securities from a definitional equation relating initial asset holdings, demand for goods, income, and final asset holdings. We shall follow the procedure of Keynes and deal with the money market rather than the securities market, although this decision has no essential influence on our results.

Money may be considered a medium of exchange and a store of value. Money in its function as a medium of exchange is purely an instrument of convenience which enables our economy to run more smoothly than a barter economy. Households and firms hold supplies of money in order to carry out the various exchanges that are apt to occur in everyday life; these supplies are called transactions balances. They are the same as pipeline stocks, and they are the analogue of the pipeline inventories of goods discussed above. It is usually assumed that people hold a stable fraction of their income in the form of transactions balances, the exact values of the fraction depending upon the frequency of wage-salary disbursements and certain habits. In addition to holding balances necessary for everyday transactions, people hold balances for contingencies that may occur in the future. Households put away cash for the rainy day when sickness or accident may come, when unemployment may be their lot, etc. In the same way, business firms hold balances for contingencies that may arise in the world of affairs. The precautionary or contingency balances may be considered a lump sum below which people cannot allow their cash holdings to go, or they may be considered to depend upon income. In any case there will be a demand equation for active cash balances (transactions plus precautionary) of the form

$$M_1^D = \iota_0' + \iota_1' p(Y + T) + \iota_2' t + u_9' \tag{3.3.19}$$

where M_1^D = circulating currency plus demand deposits, $p(Y + T)$ = net national product measured in billions of current dollars.

Equation (3.3.19) has not been developed as the result of maximizing profit or utility; it is the result of the particular structure of our society.

The parameters of (3.3.19) depend upon such social characteristics as the frequency of payments to workers and upon consumers' tastes. There is no reason to suppose that these tastes and habits change only linearly over a period of time. If we interpret ι_0' as the precautionary balances and ι_1' as the reciprocal of the average number of times a dollar turns over in expenditure on new goods and services, we see that each parameter depends upon different types of habits. The linear trend implies that the only habits subject to change are those connected with ι_0', but we know that ι_1' may also change. We can introduce a trend term showing the interaction between $p(Y + T)$ and t, as in the consumption function, to allow for structural changes over time in ι_1'. Our demand equation will then take the form

$$M_1^D = \iota_0 + \iota_1 p(Y + T) + \iota_2 t + \iota_3 p(Y + T)t + u_9 \tag{3.3.19a}$$

with

$$\frac{\partial M_1^D}{\partial p(Y + T)} = \iota_1 + \iota_3 t$$

In equation (3.3.19) the variable M_1^D is measured as circulating currency plus demand deposits. Chapter IV, on the adequacy of data for the entire model, will consider in some detail whether or not M_1^D is a good measure of active cash balances. It is possible that M_1^D will include some idle as well as active balances, in which case it will be appropriate to introduce additional variables into (3.3.19) to account for those factors which influence idle balances. In the statistical estimation of the parameters of the system, we shall attempt to test the hypothesis that other variables can enter into (3.3.19).

We come now to the demand for idle balances. In the generalized theory of the household, it is found that the demand for cash, like the demand for goods, depends upon the several prices, interest rates, incomes, and initial conditions facing the household. All these variables are eligible for consideration in the demand equation for idle balances. The prices and interest rates show the relative advantages of holding goods or securities as opposed to money. The original formulation of the liquidity-preference theory by Keynes considers only the alternatives of holding money or securities and omits the possibility of holding accumulated savings in the form of goods. Accordingly the interest rate, but not the price level, has been introduced as a variable in the liquidity preference function. A. J. Brown [21] has found the price level to be a statistically significant variable in the liquidity preference equation in England, although no such significant correlations have yet been found for the United States.

If we adopt the utility maximization principle for business firms, we find that the demand for idle cash (and other assets) is a function of prices received for output, prices paid for capital goods, wage rates, interest rates, and initial conditions. The available data compel us to lump the demand for idle cash by firms and by households in one equation. This procedure modifies the demand equation of households by the addition of factor prices and new initial conditions. All or some of these new variables may be statistically significant.

We have adopted the theory of profit maximization rather than utility maximization in order to derive the investment, labor demand, and inventory demand equations for business firms. If money and securities are introduced into the profit function through terms showing the interest income on securities and an appropriate risk factor for liquidity, we shall have a different profit function from the one used in Chapter II.

The profit function will include that income and outlay due to productive operations plus a term showing the net return on the holding of securities minus a term corresponding to the risk of becoming illiquid (holding goods or securities instead of money).[22] In discrete form the profit function will be

[21] See A. J. Brown, "Interest, Prices, and the Demand Schedule for Idle Money," *Oxford Economic Papers*, No. 2, May, 1939, pp. 46–69.

[22] This approach is similar to that introduced by L. Hurwicz, "Theory of the Firm and of Investment," *Econometrica*, Vol. 14, April, 1946, pp. 109–136.

$$\text{(3.3.20)} \quad \text{an}\,\pi = \sum_{\theta=1}^{T} [\text{an}\,p_\theta x_\theta - \text{an}\,w_\theta n_\theta - \text{an}\,q_\theta d_\theta + \text{an}\,\Delta p_\theta h_\theta + \text{an}\,i_\theta \mathsf{b}_{\theta-1} - \epsilon(h_\theta, \mathsf{m}_{\theta-1}, \mathsf{b}_{\theta-1}, u_\theta)] \frac{1}{(1+\rho)^\theta}$$

There will also be a constraint (in addition to the production function) of the form

$$\text{(3.3.21)} \qquad \mathsf{m}_{\theta-1} + \mathsf{b}_{\theta-1} + \text{an}\,\mathsf{s}_\theta = \mathsf{m}_\theta + \mathsf{b}_\theta \qquad \theta = 1, 2, \cdots, T$$

where b = net holding of securities (end of period market value), i = yield on securities, ϵ = storage cost and risk of becoming illiquid, u = disturbance, s = business savings (profits less dividends). All other variables are as defined in Chapter II in the theory of the firm.

The maximization conditions are now

$$\text{(3.3.22)} \qquad \frac{\partial\,\text{an}\,\pi}{\partial n_\theta} + \lambda_\theta \frac{\partial\,\text{an}\,\mathsf{s}_\theta}{\partial n_\theta} = \frac{\partial\,\text{an}\,\pi}{\partial n_\theta} + \lambda_\theta \frac{\partial\,\text{an}\,\pi}{\partial n_\theta} = 0$$

or

$$\frac{\partial\,\text{an}\,\pi}{\partial n_\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$

$$\text{(3.3.23)} \qquad \frac{\partial\,\text{an}\,\pi}{\partial d_\theta} + \lambda_\theta \frac{\partial\,\text{an}\,\mathsf{s}_\theta}{\partial d_\theta} = \frac{\partial\,\text{an}\,\pi}{\partial d_\theta} + \lambda_\theta \frac{\partial\,\text{an}\,\pi}{\partial d_\theta} = 0$$

or

$$\frac{\partial\,\text{an}\,\pi}{\partial d_\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$

$$\text{(3.3.24)} \qquad \frac{\partial\,\text{an}\,\pi}{\partial h_\theta} + \lambda_\theta \frac{\partial\,\text{an}\,\mathsf{s}_\theta}{\partial h_\theta} = \frac{\partial\,\text{an}\,\pi}{\partial h_\theta} + \lambda_\theta \frac{\partial\,\text{an}\,\pi}{\partial h_\theta} = 0$$

or

$$\frac{\partial\,\text{an}\,\pi}{\partial h_\theta} = 0 \qquad \theta = 1, 2, \cdots, T$$

$$\text{(3.3.25)} \quad \frac{\partial\,\text{an}\,\pi}{\partial \mathsf{m}_\theta} + \lambda_{\theta+1} \frac{\partial\,\text{an}\,\mathsf{s}_{\theta+1}}{\partial \mathsf{m}_\theta} - \lambda_\theta + \lambda_{\theta+1} = \frac{\partial\,\text{an}\,\pi}{\partial \mathsf{m}_\theta} + \lambda_{\theta+1} \frac{\partial\,\text{an}\,\pi}{\partial \mathsf{m}_\theta} - \lambda_\theta + \lambda_{\theta+1} = 0$$

$$\lambda_{T+1} = 0 \qquad \theta = 1, 2, \cdots, T$$

$$\text{(3.3.26)} \quad \frac{\partial\,\text{an}\,\pi}{\partial \mathsf{b}_\theta} + \lambda_{\theta+1} \frac{\partial\,\text{an}\,\mathsf{s}_{\theta+1}}{\partial \mathsf{b}_\theta} - \lambda_\theta + \lambda_{\theta+1} = \frac{\partial\,\text{an}\,\pi}{\partial \mathsf{b}_\theta} + \lambda_{\theta+1} \frac{\partial\,\text{an}\,\pi}{\partial \mathsf{b}_\theta} - \lambda_\theta + \lambda_{\theta+1} = 0$$

$$\lambda_{T+1} = 0 \qquad \theta = 1, 2, \cdots, T$$

Since s_θ = profits of the θth period less dividends of the θth period and since dividends are considered to be exogenous, the derivative of s with respect to any variable is the same as the derivative of an π with respect to any variable.

From (3.3.22) and (3.3.23), we see that the profit-maximizing equations for labor and capital are unchanged from the previous formulation. Equations (3.3.25) and (3.3.26) can be written as

$$\frac{\dfrac{\partial \text{ an } \pi}{\partial \mathsf{m}_\theta}}{\dfrac{\partial \text{ an } \pi}{\partial \mathsf{b}_\theta}} = 1 \qquad \theta = 1, 2, \cdots, T \tag{3.3.27}$$

Choose $\theta = 1$. Then (3.3.21), (3.3.24), (3.3.27) are three equations in

$$\text{an } i_1, \text{ an } \Delta p_1, \mathsf{m}_0, \mathsf{b}_0, \mathsf{s}_1, u_1, \mathsf{b}_1, \mathsf{m}_1, h_1$$

Solve for m_1, b_1, h_1 in terms of an i_1, an Δp_1, m_0, b_0, s_1, u_1 to get

$$\mathsf{m}_1 = \mathsf{m}_1(\text{an } i_1, \text{ an } \Delta p_1, \mathsf{s}_1, \mathsf{m}_0, \mathsf{b}_0, u_1) \tag{3.3.28}$$

$$\mathsf{b}_1 = \mathsf{b}_1(\text{an } i_1, \text{ an } \Delta p_1, \mathsf{s}_1, \mathsf{m}_0, \mathsf{b}_0, u_1) \tag{3.3.29}$$

$$h_1 = h_1(\text{an } i_1, \text{ an } \Delta p_1, \mathsf{s}_1, \mathsf{m}_0, \mathsf{b}_0, u_1) \tag{3.3.30}$$

Equations (3.3.28)–(3.3.30) are the complete analogues of the liquidity preference equations for the household. The variable s, corporate savings, is to the firm what disposable income is to the household; similarly we have the initial conditions (holding of assets), interest rates, and rate of change of prices.[23]

Combining the liquidity preference equation for the firm and the household, we should find that the demand for idle balances is a function of the interest rate, prices, income (disposable income plus corporate savings), net holdings of cash and securities of the previous period. The net holdings of securities should add up to zero in a closed system (except for government securities) because everybody's asset is someone else's liability. This is certainly true of bonds, although it is questionable for stocks. We shall represent the total initial assets by time deposits plus the public debt. This leaves an error to the amount of the net foreign debt and *possibly* the value of stocks (as distinct from bonds).

Our demand for idle balances will be written as

$$\mathsf{M}_2{}^D = \kappa_0' + \kappa' i + \kappa_2' i_{-1} + \kappa_3' \mathsf{L}_{-1} + \kappa_4' t + u_{10}' \tag{3.3.31}$$

[23] The inventory demand equation need not be changed since we can always appeal to the possibility that several of the variables in (3.3.30) may not be statistically significant, as is the case.

where M_2^D = time deposits, i = bond yield, L = time deposits plus public debt. There is no theoretical reason why prices should not be included as a variable in (3.3.31), but they have not been found to be statistically significant. The same is true of income.

We have also modified (3.3.31) to the form

$$(3.3.31a) \quad M_2^D = \kappa_0 + \kappa_1 i + \kappa_2 i_{-1} + \kappa_3 (M_2^D)_{-1} + \kappa_4 t + u_{10}$$

This is not the same equation as that which follows from our theory because L_{-1} has been modified to be only $(M_2^D)_{-1}$. This form, however, has been found to fit the interwar data very closely.

The supply of money is often treated as an exogenous variable in the macroeconomic models of business-cycle theories. It is assumed that the central banking system can make the supply of money any amount it pleases within limits that exceed the observed range, and that there is no theory of behavior to explain the action of the authorities of the central banking system. It is further assumed that the supply of money which the central banking authorities decide to create is just equal to the amount of money that people decide to hold. In other words, the equilibrium condition of supply equals demand is assumed. This is one possible solution to the problem of supply and demand in the money market; however, a close analysis of the powers of the central banking authorities suggests another solution.

The Federal Reserve Board of Governors has available three methods for control of the supply of money. The Board can change the rediscount rate; it can alter the required reserve ratios within limits set by Congress; or it can engage in open-market operations. Neglecting the first method as relatively unimportant, we can say that the supply of money is defined as the amount that people hold in the form of currency and deposits plus a term under the control of the Federal Reserve Board. The term under control is a multiple of excess reserves. The multiplier is the inverse of the reserve ratio set by the Board. The Board can obviously set the reserve ratio at any desired level (within limits set by Congress) and can force excess reserves to practically any desired level through open-market operations and manipulation of the reserve ratios. Excess reserves are defined as total reserves minus legally required reserves. Open-market operations influence the total reserves, lowering them when the Federal Reserve Banks sell securities on the open market and raising them when the Federal Reserve Banks buy securities on the open market. Manipulation of the reserve ratios causes direct changes in the legally required reserves and thus influences excess reserves.

The supply of money regarded as the sum of currency, deposits, and a multiple of excess reserves is the *potential* supply of money; it shows how much money could be created under the conditions of the existing laws and reserves. We shall regard the excess reserves as inventories of money available to the borrowing public at the going interest rate. When the banks accumulate large excess reserves they find that there is a glut of the money market and that the price of money, interest rates, must fall. Similarly, when excess reserves fall to very low levels, money becomes scarce, and its price must rise. We should find an adjustment equation in the money market of the form

$$\Delta i = \lambda_0 + \lambda_1 \lambda \mathsf{E}_R + \lambda_2 i_{-1} + \lambda_3 t + u_{11} \qquad (3.3.32)$$

where i = interest rate, λ = reciprocal of the reserve ratio, E_R = excess reserves.

The laws concerning reserve ratios are very complex; they vary for different types of banks and different types of deposits. There has not been calculated a single average which would be representative of all the reserve ratios; hence we have omitted λ from the statistical measurements of (3.3.32) and have used simply the term $\lambda_1 \mathsf{E}_R$ instead. E_R is a good indicator of the presence or absence of inventories of money, although it, alone, does not show the full extent to which the money supply can be expanded at any point of time.

The alternative theory that the supply of money is an exogenous variable and equal to demand leads to the substitution of

$$\mathsf{M}^S = \mathsf{M}_1{}^D + \mathsf{M}_2{}^D \qquad (3.3.32a)$$

for equation (3.3.32).

Total Market Adjustment

We are lacking one final equation in our system. In the housing and money markets which were introduced separately into the model, we constructed adjustment equations to show the process by which the market is cleared of a glut. Equations (3.3.18) and (3.3.32) are the adjustment equations for these markets. In the housing market rent was taken as the variable which fluctuates in order to bring supply and demand into balance, and in the money market the interest rate was taken as the variable which fluctuates in order to bring supply and demand into balance. Since in both these markets, especially in housing, there is much competition, the classical law of supply and demand operates in its traditional form. But it would be incorrect to assume competitive behavior for the rest of the economy viewed as one single market. Instead of taking price as the adjustment variable here, we take output. Our fundamental hypothesis will be the following: Whenever excess inventories accumulate, entrepreneurs decrease production and sell from stock. Whenever excess inventories are depleted, entrepreneurs increase production in order to fill sales.

What are excess inventories? The transaction inventories which entrepreneurs demand as pipeline stocks are certainly not excess inventories. Similarly, inventories that are demanded for any other purposes, such as speculation or "inertia," cannot be excess inventories. According to equation (3.3.3a), all inventories held are demanded as a result of a definite behavior pattern except for the amount u_3, which is the random disturbance. We call u_3 undesired or excess inventories which the entrepreneur holds because he misjudged the market. We shall assume, for the economy as a whole, that supply and demand balance except for a random disturbance. Immediately entrepreneurs see *undesired* inventories accumulating, they decrease production, and immediately they see *desired* inventories depleted, they step up production. Such behavior is not possible in the housing and money markets. In these markets excess inventories cannot be considered a random variable. Since houses are durable goods that last many decades and take some time to build, the supply (amount standing) cannot be immediately altered to take care of excess demand or supply. The excess supply in the money market is an exogenous variable under the control of the central bank authorities, who do not behave in a random pattern.

Our adjustment equation is

$$\Delta X = \mu_0 + \mu_1 (u_3)_{-1} + \mu_2 \, \Delta p + u_{12} \tag{3.3.33}$$

We have added the term Δp to show that price fluctuations as well as excess inventories may have a bearing on variations in output.

The Complete System

We have the following variables:

I = net investment in private producers' plant and equipment, measured in billions of constant dollars
q = price index of capital goods
p = price index of output as a whole
X = output of the private sector of the economy (excluding housing services), measured in billions of constant dollars
E = excise taxes, measured in billions of current dollars

K = stock of business fixed capital, measured at the end of the year in billions of constant dollars

H = stock of inventories, measured at the end of the year in billions of constant dollars

W_1 = private wage-salary bill, measured in billions of current dollars

Y = disposable income, measured in billions of constant dollars

C = consumer expenditures, measured in billions of constant dollars

D_1 = gross construction expenditures on owner-occupied, single-family, non-farm residences, measured in billions of constant dollars

r = index of rent

q_1 = index of construction costs

ΔF = thousands of new non-farm families

D_2 = gross construction expenditures on rented, non-farm residences, measured in billions of constant dollars

i = average corporate bond yield

v = percentage of non-farm housing units occupied at the end of the year

N^S = millions of available non-farm housing units at the end of the year

M_1^D = demand deposits + circulating currency, averaged during the year, measured in billions of current dollars

M_2^D = time deposits, averaged during the year, measured in billions of current dollars

L = time deposits, averaged during the year, measured in billions of current dollars, + public debt, at the end of the year, measured in billions of current dollars

E_R = excess reserves, averaged during the year, measured in millions of current dollars

T = government revenues + corporate savings − transfer payments − government interest payments, all measured in billions of constant dollars

G = government expenditures on goods and services + net exports + net investment of non-profit institutions, all measured in billions of constant dollars

D_3 = gross construction expenditures on farm residences, measured in billions of constant dollars

D'' = depreciation on all residences (farm and non-farm), measured in billions of constant dollars

W_2 = government wage-salary bill, measured in billions of current dollars

R_1 = non-farm rentals, paid and imputed, measured in billions of current dollars
R_2 = farm rentals, paid and imputed, measured in billions of current dollars
ρ_0 = base year rent level, measured in thousands of dollars per annum
u_i = random disturbance

The equations are

(3.3.1a) $$\mathsf{W}_1 = \alpha_0 + \alpha_1(pX - \mathsf{E}) + \alpha_2(pX - \mathsf{E})_{-1} + \alpha_3 t + u_1$$
(demand for labor)

(3.3.2a) $$I = \beta_0 + \beta_1\left(\frac{pX - \mathsf{E}}{q}\right) + \beta_2\left(\frac{pX - \mathsf{E}}{q}\right)_{-1} + \beta_3 K_{-1} + \beta_4 t + u_2$$
(demand for private producers' plant and equipment)

(3.3.3a) $$H = \gamma_0 + \gamma_1(X - \Delta H) + \gamma_2 p + \gamma_3 p_{-1} + \gamma_4 H_{-1} + \gamma_5 t + u_3$$
(demand for inventories)

(3.3.5a) $$C = \delta_0 + \delta_1 Y + \delta_2 t + u_4$$
(demand for consumer goods)

(3.3.11) $$D_1 = \epsilon_0 + \epsilon_1\left(\frac{r}{q_1}\right) + \epsilon_2(Y + Y_{-1} + Y_{-2}) + \epsilon_3\,\Delta F + u_5$$
(demand for owner-occupied housing)

(3.3.14) $$D_2 = \zeta_0 + \zeta_1 r_{-1} + \zeta_2(q_1)_{-1} + \zeta_3(q_1)_{-2} + \zeta_4 i + \zeta_5\,\Delta F_{-1} + u_6$$
(demand for rental housing)

(3.3.17) $$v = \eta_0 + \eta_1 r + \eta_2 Y + \eta_3 t + \eta_4 N^S + u_7$$
(demand-supply for dwelling space)

(3.3.18) $$\Delta r = \theta_0 + \theta_1 v_{-1} + \theta_2 Y + \theta_3 \frac{1}{r_{-1}} + u_8$$
(rent adjustment equation)

(3.3.19a) $$\mathsf{M}_1^D = \iota_0 + \iota_1 p(Y + T) + \iota_2 t + \iota_3 p(Y + T)t + u_9$$
(demand for active balances)

$$\mathsf{M}_2{}^D = \kappa_0 + \kappa_1 i + \kappa_2 i_{-1} + \kappa_3(\mathsf{M}_2{}^D)_{-1} + \kappa_4 t + u_{10} \tag{3.3.31a}$$

(demand for idle cash balances)

$$\Delta i = \lambda_0 + \lambda_1 \mathsf{E}_R + \lambda_2 i_{-1} + \lambda_3 t + u_{11} \tag{3.3.32}$$

(interest rate adjustment equation)

$$\Delta X = \mu_0 + \mu_1(u_3)_{-1} + \mu_2 \,\Delta p + u_{12} \tag{3.3.33}$$

(output adjustment equation)

$$Y + T = I + \Delta H + C + D_1 + D_2 + D_3 - D'' + G \tag{3.3.34}$$

(definition of net national product)

$$X = \frac{p(Y + T) - \mathsf{W}_2 - \mathsf{R}_1 - \mathsf{R}_2}{p} \tag{3.3.35}$$

(definition of private output exclusive of housing)

$$\Delta K = I \tag{3.3.36}$$

(definition of stock of capital)

$$\mathsf{R}_1 = \rho_0 r \left(\frac{vN^S}{100} + \frac{v_{-1}N_{-1}{}^S}{100}\right)\frac{1}{2} \tag{3.3.37}$$

(definition of rent payments)

The last five equations are definitions that hold without disturbance. The endogenous variables are W_1, p, X, I, K, H, C, Y, D_1, r, D_2, i, v, $\mathsf{M}_1{}^D$, $\mathsf{M}_2{}^D$, R_1. The exogenous variables are E, t, q, q_1, ΔF, N^S, T, E_R, D_3, D'', G, W_2, R_2. The variables q and q_1 certainly do not belong in the exogenous category, but the compromises forced on the econometrician through lack of data lead to such classification. It seems reasonable to suppose that the resulting error is not large.

Identification

Before presenting the results of statistical calculations for this system, we should investigate the problem of identification to see whether or not it is possible to estimate those parameters that we wish to estimate.

Necessary and sufficient conditions for identifiability in linear systems can be formulated in terms of the rank of matrices of the structural parameters.[24] Since we do not know enough of the properties of the

[24] All the theorems on identification used here can be found in the work of T. Koopmans, H. Rubin, and R. B. Leipnik. See Cowles Commission Monograph No. 10, *Statistical Inference in Dynamic Economic Models,* T. C. Koopmans, editor, John Wiley & Sons, 1950.

structural parameters, a priori, it will not be fruitful to apply these necessary and sufficient conditions. A weaker condition, which is only necessary, can, however, be easily applied here. Restrictions on the form of the system which lead to identification may be of various types. Two important types are the specification as to which variables enter into which equation and the specification of certain known functional relationships among the different parameters of the system. Let us consider first the specifications as to which variables enter into which equation. If the only restrictions imposed upon the system are of this type, a necessary condition for the identification of the jth equation in a linear system is that the number of variables in the entire complete system minus the number of variables appearing in the jth equation be at least as great as the number of equations in the system less one. The total number of variables in the system is equal to the number of endogenous variables plus the number of lagged endogenous variables plus the number of exogenous variables. By a complete system we mean, of course, a system in which the number of endogenous variables is equal to the number of equations.

We shall substitute the last four equations of our system, the definitions, into the behavior equations. We then will have a system of twelve equations in the twelve endogenous variables: W_1, p, K, H, C, D_1, D_2, r, i, v, $\mathsf{M}_1{}^D$, $\mathsf{M}_2{}^D$; and in the twenty-eight exogenous or lagged endogenous variables: E, p_{-1}, X_{-1}, E_{-1}, t, q, q_{-1}, K_{-1}, H_{-1}, q_1, Y_{-1}, Y_{-2}, ΔF, r_{-1}, $(q_1)_{-1}$, $(q_1)_{-2}$, $(\Delta F)_{-1}$, N^S, v_{-1}, T, i_{-1}, $(\mathsf{M}_2{}^D)_{-1}$, E_R, H_{-2}, p_{-2}, $(D_3 - D'' + G)$, $(\mathsf{W}_2 - \mathsf{R}_2)$, $N_{-1}{}^S$.

We shall proceed as though we had a linear system of twelve equations; then in order to fulfill the necessary condition for identification for all equations we must have at most twenty-nine variables in each equation. In the equation of demand for private producers' plant and equipment there are, after substitution of the definitions, more variables present than in any other equation. There are twenty-two variables in this equation after substitution of the definitions of ΔK, X, Y, R_1. The twenty-two variables are K, p, K_{-1}; H, H_{-1}, C, D_1, D_2, r, v, N^S, v_{-1}, $N_{-1}{}^S$, $(\mathsf{W}_2 - \mathsf{R}_2)$, $(D_3 - D'' + G)$, E, q, p_{-1}, X_{-1}, E_{-1}, q_{-1}, t. Since there are not more than twenty-two in any single equation, the necessary condition is fulfilled. In addition, there are two other types of restrictions present in the system which have not been mentioned yet. These additional restrictions make identifiability even more likely. The endogenous variables which enter into the first of the definitional equations, that for $Y + T$, appear as a sum, yet they were enumerated as separate variables among the twenty-two after this definition was substituted into the investment function. The fact that the variables enter the investment equation in a known combination is another type of restriction. This type of restriction prescribes ratios between the parameters corresponding to different variables in the same equation. The other type of restriction that we must point out is that of non-linearities of known functional forms in the system. For example, our theory of economic behavior leads us to write the variables p, X, q in the form pX/q in the investment equation. This non-linearity, when known in advance, is another restriction which gives additional information about the structure of the system and thus aids identification. In sum, it seems safe to conjecture that each equation of our model is identified when we take all types of restrictions into account.

There are some variants of the model which we shall also consider, and these too satisfy the necessary conditions for identifiability. One variant would place the sum $\mathsf{M}_1^D + \mathsf{M}_2^D$ in the exogenous category $(= \mathsf{M}^S)$ and omit the interest rate adjustment equation (3.3.32) along with the variable E_R = excess reserves. These changes do not violate any of the necessary conditions for identification. We may also find that some of the variables in our equations, such as p_{-1} in the inventory demand equation, are not statistically significant; however, such minor changes do not have any important influence on the properties of identification, since we have so many extra restrictions on our system.

The Method of Reduced Form (Limited-Information Method)

A modified version of the reduced-form method of estimation has been applied to the large system of Model III. The first step in the application of the reduced-form method is to solve for each of the endogenous variables in the system in terms of the exogenous and lagged endogenous variables. This brings the system into its reduced form. The method of least squares is then used to estimate the parameters of the reduced form. But there are so many variables in Model III that there will be few or no degrees of freedom for the estimation of the parameters of the reduced form since there are only twenty or twenty-one observations. Fortunately, it can be shown [25] that the reduced-form method can be applied to incomplete systems, i.e., systems in which all the equations are not used. In an incomplete system only a part of the exogenous and lagged endogenous variables need to be used; hence there will be more degrees of freedom in the estimation of the parameters of the reduced forms. In a much larger sample, a loss of information is likely to be involved in this approach, but the reverse is true for a sample of the size we are dealing with.

[25] See T. W. Anderson, Jr., and H. Rubin, *op. cit.*

The grouping in incomplete systems was done as follows: (1) all the equations of the firm, (3.3.1a), (3.3.2a), (3.3.3a), (3.3.33), were lumped together. (2) All the equations of household demand, (3.3.5a), (3.3.11), (3.3.17), (3.3.18), were lumped together.[26] (3) The equations of the money market and the demand for rental housing, (3.3.19a), (3.3.31a), (3.3.14), were lumped together. (4) The interest-rate adjustment equation was estimated separately because it is already in reduced form. In groups (1) and (2) twenty annual observations were used, and in groups (3) and (4) twenty-one annual observations were used. It was found that the reduced-form estimates of the parameters of equation (3.3.14) are very sensitive to the omission or presence of a single year (1921), but this is not true for the other equations in these groups. The estimated parameters of (3.3.14) have signs different from a priori expectations if the years 1922–1941 are used, and they must therefore be rejected. The results look much more reasonable if the period 1921–1941 is used instead. The equations of groups (1) and (2) were estimated from the observations for 1922–1941, and all others from the observations for 1921–1941. All constant-dollar variables are expressed in terms of 1934 dollars, and all indexes are on a 1934 base. The observed time series used are tabulated in the Appendix, where the sources for these data are also given. Each variable on the right-hand side of the particular equation has been multiplied by its coefficient, and the resulting time series plotted in the charts of the Appendix as solid-line curves. The sum of these time series equals the "computed value" of the variable occurring on the left-hand side of the equation and is indicated by the superscript c. It is plotted as a solid-line curve. The observed values of the variable occurring on the left-hand side of the equation have also been plotted, in dotted lines. The differences between the observed and the computed values are the estimates of residuals and are plotted at the bottom of each chart.

The estimated equations[27] are

$$(3.3.38)\quad \mathsf{W}_1 = 5.04 + 0.41\,(pX - \mathsf{E}) + 0.17\,(pX - \mathsf{E})_{-1} + 0.17\,(t - 1931) + u_1'$$

$$\frac{\delta^2}{S^2} = 1.89 \qquad \bar{S} = 1.00$$

[26] Three exogenous variables, $(q_1)_{-1}$, $(q_1)_{-2}$, and $(\Delta F)_{-1}$, were used in the reduced forms of group (2).

[27] In the author's article "The Use of Econometric Models as a Guide to Economic Policy," *Econometrica*, Vol. 15, April, 1947, pp. 111–151, there was a slight computing error which was pointed out by a reader. Corrected estimates of the parameters affected by this error are given here

(3.3.39) $$I = 2.59 + 0.12\left(\frac{pX - \mathsf{E}}{q}\right) + 0.04\left(\frac{pX - \mathsf{E}}{q}\right)_{-1} - 0.10K_{-1} + u_2'$$

$$\frac{\delta^2}{S^2} = 1.59 \qquad \bar{S} = 0.47$$

(3.3.40) $$H = 1.17 + 4.60p + 0.12\,(X - \Delta H) + 0.50H_{-1} + u_3'$$

$$\frac{\delta^2}{S^2} = 2.26 \qquad \bar{S} = 0.55$$

(3.3.41) $$C = 11.87 + 0.73Y + 0.04(t - 1931) + u_4'$$

$$\frac{\delta^2}{S^2} = 1.20 \qquad \bar{S} = 1.36$$

(3.3.42) $$D_1 = -9.03 + 3.74\left(\frac{r}{q_1}\right) + 0.02(Y + Y_{-1} + Y_{-2}) + 0.0043\,\Delta F + u_5'$$

$$\frac{\delta^2}{S^2} = 2.26 \qquad \bar{S} = 0.21$$

(3.3.43) $$D_2 = -2.14 + 2.81r_{-1} + 0.02(q_1)_{-1} - 0.44(q_1)_{-2} + 0.0016(\Delta F)_{-1} - 0.18i + u_6'$$

$$\frac{\delta^2}{S^2} = 2.07 \qquad \bar{S} = 0.26$$

(3.3.44) $$v = 178.01 + 0.29Y - 2.62r + 1.42(t - 1931) - 3.76N^S + u_7'$$

$$\frac{\delta^2}{S^2} = 1.52 \qquad \bar{S} = 0.79$$

(3.3.45) $$\Delta r = -2.15 + 0.02v_{-1} + 0.00071Y + 0.17\frac{1}{r_{-1}} + u_8'$$

$$\frac{\delta^2}{S^2} = 1.04 \qquad \bar{S} = 0.03$$

(3.3.46) $$\mathsf{M}_1^D = 8.45 + 0.24p(Y + T) + 0.03p(Y + T)(t - 1931) - 1.43(t - 1931) + u_9'$$

$$\frac{\delta^2}{S^2} = 1.33 \qquad \bar{S} = 1.26$$

$$(3.3.47)\qquad \mathsf{M}^D{}_2 = 15.37 + 0.28i - 1.90i_{-1} + 0.74(\mathsf{M}_2{}^D)_{-1} - 0.18(t - 1931) + u_{10}'$$

$$\frac{\delta^2}{S^2} = 1.49 \qquad \bar{S} = 0.67$$

$$(3.3.48)\qquad \Delta i = 2.00 - 0.17\mathsf{E}_R - 0.37i_{-1} - 0.0052(t - 1931) + u_{11}'$$

$$\frac{\delta^2}{S^2} = 1.77 \qquad \bar{S} = 0.47$$

$$(3.3.49)\qquad \Delta X = 2.55 - 4.46(u_3)_{-1} + 82.76\,\Delta p + u_{12}'$$

$$\frac{\delta^2}{S^2} = 1.83 \qquad \bar{S} = 2.61$$

$$(3.3.34)\qquad Y + T = I + \Delta H + C + D_1 + D_2 + D_3 - D'' + G$$

$$(3.3.35)\qquad X = \frac{p(Y + T) - \mathsf{W}_2 - \mathsf{R}_1 - \mathsf{R}_2}{p}$$

$$(3.3.36)\qquad \Delta K = I$$

$$(3.3.37)\qquad \mathsf{R}_1 = 0.278r\left(\frac{vN^S}{100} + \frac{v_{-1}N_{-1}{}^S}{100}\right)\frac{1}{2}$$

In terms of the time series actually used, equation (3.3.37) is not really exact, but it is a very close approximation.

The estimated disturbances are given in Table V.

Some variations of the model can be obtained. Additional variables, already in the system, are introduced into equations in which they have not previously appeared. The reduced-form estimates of the parameters in equations that are not being changed are not affected by this change.

The variable $\mathsf{M}_1{}^D$ is supposed to represent active balances, or, in other words, balances held for transactions and precautionary reasons. It has been claimed that many business firms hold idle balances in the form of demand deposits, which are a major component of $\mathsf{M}_1{}^D$; hence it is also claimed that the variables affecting the holding of idle balances, in the theoretical sense, should be introduced into equation (3.3.19). If the corporate bond yield, current and lagged, is introduced into this equation, the statistical result is

$$(3.3.46\text{a})\qquad \mathsf{M}_1{}^D = -23.40 + 0.44p(Y + T) - 0.20i + 3.45i_{-1} + 1.07t + u_9''$$

TABLE V

ESTIMATES OF ANNUAL VALUES OF DISTURBANCES BY METHOD OF REDUCED FORM (LIMITED-INFORMATION METHOD): MODEL III

	u_1'	u_2'	u_3'	u_4'	u_5'	u_6'	u_7'	u_8'	u_9'	u_{10}'	u_{11}'	u_{12}'
1921			−0.74			−0.07			−0.62	−0.75	0.55	
1922	−0.12	−0.37	−0.31	−0.64	0.13	−0.06	0.80	−0.009	−0.27	−0.13	−0.50	3.60
1923	1.80	0.36	0.76	−1.34	0.03	0.02	−0.04	−0.021	0.39	−0.34	0.27	0.10
1924	0.17	0.07	−0.61	0.02	0.34	0.04	0.40	0.010	0.65	−0.02	−0.02	1.20
1925	−0.86	0.25	−0.30	−0.12	0.11	0.18	−0.04	−0.011	1.21	0.41	−0.20	−3.51
1926	−1.05	−0.13	0.24	0.39	−0.21	0.19	−0.47	−0.010	0.54	0.30	−0.25	−0.72
1927	−1.20	−0.17	−0.16	0.87	−0.09	0.05	−0.38	0.004	−0.02	0.57	−0.32	1.45
1928	0.22	−0.23	−0.67	2.37	0.15	0.04	0.01	0.015	−0.45	0.85	−0.19	−3.62
1929	1.36	0.63	0.76	1.09	−0.44	0.37	−0.47	0.030	−1.74	−0.14	0.11	−2.78
1930	−0.06	0.44	0.52	−0.44	−0.05	−0.74	−0.30	0.042	−1.56	0.47	−0.17	−1.00
1931	0.19	−0.10	−0.10	−2.03	−0.17	−0.23	−0.12	0.019	−0.37	−0.56	0.63	1.75
1932	−0.19	0.21	−0.19	−2.00	0.20	−0.15	0.06	−0.033	1.01	−1.60	1.28	−1.91
1933	0.16	−0.07	0.57	1.04	−0.15	0.11	1.20	−0.056	0.39	0.44	−0.32	1.11
1934	0.76	0.14	−0.02	0.15	−0.12	0.04	0.37	−0.006	−0.52	0.86	−0.45	−0.23
1935	−0.67	0.06	−0.49	0.11	0.12	−0.16	0.14	0.013	0.67	0.61	−0.21	−1.03
1936	−1.37	−0.03	0.61	−0.42	−0.11	−0.02	−1.95	0.009	0.46	0.12	−0.47	2.87
1937	0.65	0.62	0.00	1.19	−0.18	−0.01	−1.00	0.028	−0.69	−0.58	−0.25	−1.67
1938	−1.06	−0.16	−1.16	1.70	0.25	−0.10	0.61	0.021	1.56	−0.58	0.18	−3.39
1939	−0.07	−0.13	−0.35	0.39	0.08	0.31	−0.30	−0.011	0.83	0.22	−0.07	1.69
1940	−0.36	0.04	0.23	0.35	0.17	0.16	0.82	−0.013	0.67	0.07	0.31	0.65
1941	1.70	−1.42	0.67	−2.68	0.04	0.04	0.67	−0.019	−3.12	−0.23	0.07	5.42

Since these estimates do not agree with the economic theory that the demand for cash varies inversely with the interest rate, we are led to reject the hypothesis that interest rates belong in the equation of the demand for M_1^D.

The theory underlying the development of the demand for idle balances shows that income may also be a pertinent variable in equation (3.3.31a). Reduced-form estimates have been obtained for a modification of this equation to

$$\text{(3.3.47a)} \quad M_2^D = \text{const.} + 0.02i - 1.96i_{-1} + 0.76(M_2^D)_{-1} - 0.22(t - 1931) - 0.01p(Y + T) + u_{10}''$$

Again the sign of the coefficient of $p(Y + T)$, besides being relatively insignificant, is not in agreement with theoretical considerations, and we are led to reject the hypothesis that income is a variable in the equation of demand for M_2^D.

Other variants of the model have also been obtained by other methods of statistical estimation. However, this concludes the discussion of the estimates made by the method of reduced forms. No calculations of the reliability of the estimated parameters have been made for this model.

The *single-equation, least-squares* estimates for Model III are given below. The left-hand variables of each equation have been arbitrarily selected as the "dependent variables."

$$\mathsf{W}_1 = 4.70 + \underset{(0.03)}{0.47}(pX - \mathsf{E}) + \underset{(0.03)}{0.12}(pX - \mathsf{E})_{-1} + \underset{(0.04)}{0.19}(t - 1931) + u_1'' \tag{3.3.50}$$

$$\frac{\delta^2}{S^2} = 1.27 \qquad \bar{S} = 0.96$$

$$I = 5.24 + \underset{(0.02)}{0.08}\left(\frac{pX - \mathsf{E}}{q}\right) + \underset{(0.02)}{0.07}\left(\frac{pX - \mathsf{E}}{q}\right)_{-1} - \underset{(0.02)}{0.12}K_{-1} + u_2'' \tag{3.3.51}$$

$$\frac{\delta^2}{S^2} = 1.81 \qquad \bar{S} = 0.42$$

$$H = 1.06 + \underset{(1.15)}{4.66}p + \underset{(0.02)}{0.13}(X - \Delta H) + \underset{(0.08)}{0.48}H_{-1} + u_3'' \tag{3.3.52}$$

$$\frac{\delta^2}{S^2} = 2.17 \qquad \bar{S} = 0.57$$

$$C = 9.70 + \underset{(0.04)}{0.77}Y - \underset{(0.004)}{0.01}Y(t - 1931) + \underset{(0.26)}{0.76}(t - 1931) + u_4'' \tag{3.3.53}$$

$$\frac{\delta^2}{S^2} = 1.46 \qquad \bar{S} = 1.17$$

$$D_1 = -7.49 + \underset{(0.52)}{3.14}\left(\frac{r}{q_1}\right) + \underset{(0.003)}{0.02}(Y + Y_{-1} + Y_{-2}) + \underset{(0.0005)}{0.0039}\,\Delta F + u_5'' \tag{3.3.54}$$

$$\frac{\delta^2}{S^2} = 1.80 \qquad \bar{S} = 0.21$$

$$D_2 = -1.99 + \underset{(0.39)}{2.93}r_{-1} + \underset{(0.53)}{0.16}(q_1)_{-1} - \underset{(0.52)}{0.44}(q_1)_{-2} + \underset{(0.0006)}{0.0013}(\Delta F)_{-1} - \underset{(0.07)}{0.25}i + u_6'' \tag{3.3.55}$$

$$\frac{\delta^2}{S^2} = 2.16 \qquad \bar{S} = 0.25$$

$$(3.3.56)\quad v = 181.62 + \underset{(0.05)}{0.24}Y - \underset{(3.32)}{1.05}r + \underset{(0.33)}{1.57}(t - 1931) - \underset{(0.42)}{3.88}N^S + u_7''$$

$$\frac{\delta^2}{S^2} = 1.22 \qquad \bar{S} = 0.76$$

$$(3.3.57)\quad \Delta r = -2.13 + \underset{(0.002)}{0.02}v_{-1} + \underset{(0.0007)}{0.0013}Y + \underset{(0.04)}{0.16}\frac{1}{r_{-1}} + u_8''$$

$$\frac{\delta^2}{S^2} = 1.09 \qquad \bar{S} = 0.03$$

$$(3.3.58)\quad \mathsf{M}_1^D = 9.55 + \underset{(0.02)}{0.23}p(Y + T) + \underset{(0.003)}{0.02}p(Y + T)(t - 1931) - \underset{(0.24)}{0.95}(t - 1931) + u_9''$$

$$\frac{\delta^2}{S^2} = 1.37 \qquad \bar{S} = 1.05$$

$$(3.3.59)\quad \mathsf{M}_2^D = 14.25 - \underset{(0.18)}{1.00}i - \underset{(0.18)}{0.92}i_{-1} + \underset{(0.03)}{0.84}(\mathsf{M}_2^D)_{-1} - \underset{(0.02)}{0.26}(t - 1931) + u_{10}''$$

$$\frac{\delta^2}{S^2} = 2.13 \qquad \bar{S} = 0.33$$

$$(3.3.60)\quad \Delta i = 2.00 - \underset{(0.10)}{0.17}\mathsf{E}_R - (1 - \underset{(0.17)}{0.63})i_{-1} - \underset{(0.0353)}{0.0052}(t - 1931) + u_{11}''$$

$$\frac{\delta^2}{S^2} = 1.77 \qquad \bar{S} = 0.47$$

$$(3.3.61)\quad \Delta X = 2.64 - \underset{(1.17)}{4.41}(u_3'')_{-1} + \underset{(11.36)}{80.64}\,\Delta p + u_{12}''$$

$$\frac{\delta^2}{S^2} = 1.78 \qquad \bar{S} = 2.68$$

$$(3.3.34)\quad Y + T = I + \Delta H + C + D_1 + D_2 + D_3 - D'' + G$$

$$(3.3.35)\quad X = \frac{p(Y + T) - \mathsf{W}_2 - \mathsf{R}_1 - \mathsf{R}_2}{p}$$

$$(3.3.36)\quad \Delta K = I$$

$$(3.3.37)\quad \mathsf{R}_1 = 0.278r\left(\frac{vN^S}{100} + \frac{v_{-1}N_{-1}{}^S}{100}\right)\frac{1}{2}$$

Data for the years 1921–1941 were used to obtain the least-squares estimates for all equations in this group, except (3.3.51), (3.3.52), (3.3.57), and (3.3.61). Data for 1922–1941 were used for these four equations.

There are few general comments that can be made upon a comparison between the two different methods of estimating the parameters of Model III. It is left to the reader to form his own judgments on the comparison.

The estimated values of the disturbances, for the method of least squares, are given in Table VI.

TABLE VI

ESTIMATES OF ANNUAL VALUES OF DISTURBANCES BY METHOD OF LEAST SQUARES: MODEL III

	u_1''	u_2''	u_3''	u_4''	u_5''	u_6''	u_7''	u_8''	u_9''	u_{10}''	u_{11}''	u_{12}''
1921	−2.27		−0.43	0.38	0.20	0.04	0.59		0.13	0.01	0.55	
1922	0.15	−0.45	−0.26	0.81	0.09	−0.01	0.62	−0.002	0.28	−0.73	−0.50	4.58
1923	1.66	0.23	0.76	−0.98	0.00	0.09	−0.03	−0.018	0.50	0.24	0.27	0.30
1924	0.60	−0.16	−0.58	0.31	0.33	0.08	0.26	0.013	0.76	0.01	−0.02	1.05
1925	−0.88	0.32	−0.31	−0.23	0.16	0.16	−0.10	−0.010	1.15	0.17	−0.20	−3.36
1926	−0.97	−0.08	0.22	−0.04	−0.13	0.16	−0.44	−0.012	0.44	−0.05	−0.25	−0.83
1927	−0.84	−0.22	−0.16	0.38	−0.03	0.01	−0.32	0.001	−0.07	0.08	−0.32	1.19
1928	0.44	−0.15	−0.68	1.94	0.19	−0.02	0.03	0.012	−0.46	0.46	−0.19	−3.71
1929	1.46	0.74	0.74	0.51	−0.38	0.32	−0.34	0.026	−1.71	−0.26	0.11	−2.89
1930	0.70	0.36	0.58	−0.88	−0.12	−0.78	−0.28	0.039	−1.56	−0.02	−0.17	−1.32
1931	1.24	−0.36	0.03	−2.31	−0.14	−0.23	−0.23	0.018	−0.60	0.06	0.63	1.65
1932	0.82	−0.15	0.00	−2.07	0.16	−0.06	−0.28	−0.029	0.38	−0.17	1.28	−1.66
1933	0.40	−0.10	0.70	0.89	−0.24	0.13	0.87	−0.050	−0.45	−0.05	−0.32	1.80
1934	0.56	0.22	0.04	−0.17	−0.25	0.08	0.20	0.000	−1.28	0.48	−0.45	0.40
1935	−0.85	0.17	−0.49	−0.30	0.00	−0.11	0.02	0.018	0.03	0.58	−0.21	−0.78
1936	−1.70	0.15	0.55	−0.71	−0.16	−0.02	−1.78	0.009	0.17	−0.22	−0.47	2.82
1937	0.48	0.58	−0.06	0.94	−0.23	−0.01	−1.04	0.028	−0.76	−0.28	−0.25	−1.99
1938	−0.54	−0.43	−1.16	1.39	0.19	−0.11	0.34	0.023	1.11	−0.04	0.18	−3.86
1939	−0.42	0.07	−0.45	0.51	0.06	0.24	−0.31	−0.015	0.84	0.01	−0.07	1.60
1940	−0.70	0.19	0.09	0.90	0.18	0.09	0.98	−0.019	2.19	0.02	0.31	0.15
1941	0.68	−0.93	0.43	−1.11	0.12	−0.06	1.24	−0.032	−1.03	−0.32	0.07	4.85

SOME TESTS OF ALTERNATIVE HYPOTHESES

Models I, II, and III of this chapter are all based on different hypotheses. In Model I the major hypotheses are that the distribution of income between wages and profits has some influence on consumption and that the volume of investment depends upon the level and rate of change of profits. The observed data are not inconsistent with these hypotheses. In Model II the major hypotheses are that investment is exogenous and that the stock of cash balances is a factor which affects consumers' behavior patterns. There is no contradiction between the data and the former hypothesis, but it is doubtful whether or not the latter hypothesis can be maintained in the light of the data from the interwar period. In Model III there is a variety of hypotheses, which are too numerous to list here again and which have all been subjected to test. It has been possible to change the variables in some of the equations from time to time in order to test various alternative theories. In spite of all the tests that have been carried out in the work on these three models, there remain some important hypotheses which require further examination.

Just as it has become a current fashion to argue that liquidity of individuals should be an important variable in the consumption function, so has it been argued that liquidity of business firms should be an important factor in the determination of investment decisions. This hypothesis can best be tested for particular industries in which there are good data on total assets and liabilities. Since the Interstate Commerce Commission publishes balance sheet data for railroads in a convenient form, it was decided to make some calculations for this industry. Thus we have selected a demand equation for plant and equipment of railroads just like the corresponding demand equation of Model III, except for the fact that a liquidity variable is added. Define

I_R = net investment of railroads, measured in billions of 1934 dollars

K_R = end-of-year stock of plant and equipment in railroads, measured in billions of 1934 dollars

L_R = end-of-year current assets − current liabilities of railroads, measured in billions of current dollars

$p_R X_R$ = index of income originating in railroads, 1934: 1.00

q_R = index of prices paid for plant and equipment by railroads, 1934: 1.00

Since we have not constructed an entire model of the railroad industry, we did not use equation systems methods [28] of statistical estimation. Instead, we computed the parameters by the method of least squares. Although the results thus computed are known to be biased, they will, at least, show whether or not the several variables are correlated, which is the goal of our present inquiry. The least-squares results for 1920–1941 are

$$(3.4.1)\quad I_R = \text{const.} + \underset{(0.21)}{0.20}\left(\frac{p_R X_R}{q_R}\right) + \underset{(0.21)}{0.56}\left(\frac{p_R X_R}{q_R}\right)_{-1} - \underset{(0.05)}{0.11}(K_R)_{-1}$$
$$- \underset{(0.0729)}{0.0004}\left(\frac{\mathsf{L}_R}{q_R}\right)_{-1} + \underset{(0.125)}{0.029}(t - 1931) + u' \qquad \bar{R} = 0.90 \qquad \bar{S} = 0.10$$

[28] For example, maximum likelihood using full or limited information.

$\bar{R}$ = the multiple correlation coefficient adjusted for degrees of freedom. If the variable representing liquidity is entirely omitted from the calculations, the other parameters remain unchanged. This means that the results are insensitive to the presence or absence of the liquidity variable; hence we conclude that the data do not contain evidence that liquidity influences investment in the railroad industry.[29]

F. Lutz has published [30] data on liquid assets held by a sample of manufacturing corporations for the period 1914–1943. If it is assumed that the asset holdings of this sample can be taken as representative of the asset holdings of all manufacturing enterprises, it is possible to make a test of the influence of asset holdings on investment similar to that carried out for railroads. It must be remarked, however, that in the previous case liquidity is represented by the difference between current assets and current liabilities, whereas in the present case liquidity will be represented by asset items alone. Define

L_M = end-of-year cash balances and marketable securities of a sample of large manufacturing corporations, measured in billions of current dollars

$\mathsf{L}_M{}^*$ = end-of-year "free" liquid funds [31] of a sample of large manufacturing corporations, measured in billions of current dollars

The subscript M will denote variables pertaining to the manufacturing industry. In this example, investment, capital, and output are measured in millions of dollars.

The least-squares regression equations (1921–1940) are

$$\begin{aligned}(3.4.2)\qquad I_M = \text{const.} &+ \underset{(0.02)}{0.09}\left(\frac{p_M X_M}{q_M}\right) + \underset{(0.03)}{0.04}\left(\frac{p_M X_M}{q_M}\right)_{-1} \\ &- \underset{(0.10)}{0.16}(K_M)_{-1} + \underset{(0.32)}{0.18}\left(\frac{\mathsf{L}_M}{q_M}\right)_{-1} \\ &- \underset{(0.020)}{0.006}(t - 1931) + u' \qquad \bar{R} = 0.90 \qquad \bar{S} = 0.20\end{aligned}$$

[29] At present the author is engaged in a more detailed study of investment behavior and other aspects of an econometric model for this industry. In this more recent study it is found that better data can be obtained and that the investment function can be improved in various ways, but the broad conclusions of this section of the present volume still hold.

[30] F. Lutz, *Corporate Cash Balances, 1914–43, Manufacturing and Trade,* National Bureau of Economic Research, New York, 1945.

[31] "Free" liquid funds are equal to total liquid funds minus a fraction of total payments. The particular value of the fraction is estimated from the ratio of total liquid funds to total payments in a past period. For the statement of the method used to obtain this series see F. Lutz, *op. cit.*, pp. 40–49.

$$(3.4.3) \qquad I_M = \text{const.} + \underset{(0.02)}{0.09}\left(\frac{p_M X_M}{q_M}\right) + \underset{(0.03)}{0.04}\left(\frac{p_M X_M}{q_M}\right)_{-1}$$

$$- \underset{(0.09)}{0.14}(K_M)_{-1} + \underset{(0.35)}{0.17}\left(\frac{\mathsf{L}_M{}^*}{q_M}\right)_{-1}$$

$$+ \underset{(0.0133)}{0.0005}(t - 1931) + u' \qquad \bar{R} = 0.90 \qquad \bar{S} = 0.20$$

The coefficients of the liquidity variables seem plausible, but their corresponding standard errors are large; consequently the estimates are not very reliable. The correlation is little affected by the omission or inclusion of these variables. In fact the standard error of the coefficient of K_M is reduced if the liquidity variables are omitted. It seems that the admission of these variables makes the estimate of the relationship less reliable.

The data used in (3.4.2) and (3.4.3) require a few words of explanation. All variables other than L (= liquidity) cover both manufacturing and mining, but manufacturing dominates mining in magnitude. The classification of the data into mining and manufacturing was made for other purposes, and the data were used for this particular investigation. Adequate data have not been found for the construction of a price index of capital goods in manufacturing and mining; hence the variable q_M is taken as the average price paid by all types of business firms for capital goods. This was the best approximation that we could find. Also, the year 1941 was excluded from these calculations. Manufacturing investment was actually large in that year to match the large volume of business activity, but a very large portion of the manufacturing investment was undertaken by the government as part of the defense program. Private capital formation is small for 1941; it would have been larger if it had not been included in exogenous government investment. The same phenomenon did not occur in other industries.

The best that we can say at this stage is that the influence of liquidity on investment remains doubtful.

The estimations in this volume have been criticized on the grounds that they use annual data. Monthly or quarterly observations have been suggested as more desirable. Many problems are involved in obtaining models from monthly or quarterly data: (1) the data are not all available by months or quarters; (2) more work is needed on methods of dealing with seasonal variation; (3) the assumption of non-autocorrelated successive disturbances becomes less permissible as the observation interval becomes smaller. We have not entirely neglected the criticism of the use of annual data, however. The equation of demand for inventories has been estimated (least-squares) from quarterly data for comparison with the results obtained from annual data. This particular equation has been selected because it contains terms to represent demand for purposes of price speculation. Speculation in inventories is a short-run phenomenon, and if annual data lead generally to incorrect results they should show up in the inventory equations.

Let

H = end-of-quarter inventories measured in billions of fourth-quarter-1934 dollars

X = total quarterly outlay (Barger)[32] measured in billions of fourth-quarter-1934 dollars

p = quarterly wholesale price index, fourth-quarter-1934: 1.00

The least-squares regression equation for the period 1921–1938 is

$$(3.4.4)\quad H = -0.71 + \underset{(0.04)}{0.16}(X - \Delta H) + \underset{(0.32)}{1.10}p + \underset{(0.04)}{0.88}H_{-1} + u'$$

$$\bar{R} = 0.98 \qquad \bar{S} = 0.41 \qquad \frac{\delta^2}{S^2} = 1.68$$

Barger's data have been adjusted for seasonal variation so that no seasonal variables are used in (3.4.4).

By successive substitution for different time periods in (3.4.4) we can derive

$$\begin{aligned}(3.4.5)\quad H = -0.71[1 + 0.88 + (0.88)^2 + (0.88)^3] &+ 0.16[X - \Delta H \\ &+ 0.88(X - \Delta H)_{-1} + (0.88)^2(X - \Delta H)_{-2} \\ &+ (0.88)^3(X - \Delta H)_{-3}] + 1.10[p + 0.88p_{-1} \\ &+ (0.88)^2p_{-2} + (0.88)^3p_{-3}] + (0.88)^4H_{-4} + v'\end{aligned}$$

The relation between the annual and quarterly variables is

End-of-year inventories = end-of-fourth-quarter inventories

$$\text{Annual sales} = X - \Delta H + (X - \Delta H)_{-1} + (X - \Delta H)_{-2} + (X - \Delta H)_{-3}$$

where $X - \Delta H$ = fourth-quarter sales

$$\text{Annual price index} = \tfrac{1}{4}(p + p_{-1} + p_{-2} + p_{-3})$$

[32] The quarterly data on inventories and outlay come from H. Barger, *Outlay and Income in the United States, 1921–1938*, National Bureau of Economic Research, 1942.

where p = fourth-quarter price index. In terms of annual variables, the inventory demand equation is

$$H = a_0 + a_1(X - \Delta H) + a_2 p + a_3 H_{-1} + u \qquad (3.4.6)$$

If we make the simplifying (though not justified) assumption that the quarterly observations have the following properties: $X - \Delta H = (X - \Delta H)_{-1} = (X - \Delta H)_{-2} = (X - \Delta H)_{-3}$, $p = p_{-1} = p_{-2} = p_{-3}$, where $X - \Delta H$ = fourth-quarter sales and p = fourth-quarter price index, (3.4.5) should imply the following estimates of a_0, a_1, a_2, a_3:

$$a_0 = -0.71[1 + 0.88 + (0.88)^2 + (0.88)^3] = -2.3$$

$$4a_1 = 0.16[1 + 0.88 + (0.88)^2 + (0.88)^3] = 0.53$$

$$a_2 = 1.10[1 + 0.88 + (0.88)^2 + (0.88)^3] = 3.66$$

$$a_3 = (0.88)^4 = 0.60$$

The least-squares estimates obtained directly from annual data, 1921–1941, lead to

$$\begin{aligned} a_0 &= 1.06 \\ 4a_1 &= 0.52 \\ a_2 &= 4.66 \\ a_3 &= 0.48 \end{aligned}$$

There is fair agreement for all terms except a_0, which is not so important as some of the other parameters. There should not be perfect agreement, though, because the assumption that the data remain the same for the four quarters of each year is not correct, and the time series used for the quarterly estimates are based on different sources and include different years from those used for the annual estimates. However, the same variables turn out to be statistically significant for the quarterly as well as for the annual estimates, and the latter are probably not so bad as some statisticians believe.

The equation of Model III which is the most questionable, from the point of view of theoretical basis, is the output adjustment equation, (3.3.33). This equation closes the system for the determination of the general price level. If this equation is excluded there are not enough equations to determine all the endogenous variables. One more variable must then be classified as exogenous, the most logical choice, in this case, being the general price level. There is still another alternative. We can drop (3.3.33) from the system and, at the same time, add some new equations which might be regarded as determining the price level.

Two equations used in the theoretical derivation of the system that do not appear explicitly in Model III are (2.1.1) and (2.1.2), which refer to the production process. Equation (2.1.1) is a technological input-output relationship, and (2.1.2) is an equation showing how the services of fixed capital will be used from fixed capital of different age groups. The introduction of macrorepresentations of these two equations would introduce two new variables, the input of labor and the input of capital; therefore no direct progress has been made on the problem of getting a more completely determined system. It should be added, however, that valuable information is provided in these equations, especially for the study of technological change. Although their introduction may not solve the problem of completing the system it may improve the model. In equation (3.3.1a), the variable W_1 (= wage bill) appears. The alteration of the model to include equations like (2.1.1) and (2.1.2) would introduce total employment as an endogenous variable. If the wage bill and employment are endogenous, the wage rate (= ratio of the wage bill to employment) is determined as an endogenous variable in the system.

The wage rate, like other prices in the system, should also appear in a wage adjustment equation. The adjustment equation for the labor market should make the rate of change of the wage rate a function of excess supply (= unemployment). Excess supply in the labor market is the difference between the labor force and employment. One new variable appears in this new equation, namely, the labor force. Classically the labor force has been regarded as an endogenous variable, but it is a reasonable alternative to assume that the labor force is determined by exogenous and especially by demographic factors. If the labor force is exogenous, the wage-adjustment equation adds one new equation and no new endogenous variables; hence we have enough equations again to classify the general price level as an endogenous variable, and the output-adjustment equation has been dropped from the system.

The new variables are

D'' = depreciation of private producers' plant and equipment, measured in billions of 1934 dollars

N_E = employment, measured in millions of persons

N_L = labor force, measured in millions of persons

w = annual average wage rate, measured in current dollars per employee

The least-squares estimates of the new equations are

$$(3.4.7)\quad \log X = -1.95 + \underset{(0.13)}{2.32} \log N_E - \underset{(0.12)}{0.04} \log D'' + \underset{(0.0005)}{0.0035}(t - 1931) + \underset{(0.00011)}{0.00052}(t - 1931)^2 + u_1'$$

$$\bar{R} = 0.99 \qquad \bar{S} = .012 \qquad \frac{\delta^2}{S^2} = 1.96$$

$$(3.4.8)\quad D'' = -3.36 + \underset{(0.03)}{0.21}I + \underset{(0.01)}{0.08}K_{-1} + u_2'$$

$$\bar{R} = 0.92 \qquad \bar{S} = 0.19 \qquad \frac{\delta^2}{S^2} = 2.36$$

$$(3.4.9)\quad \Delta w = 1056.20 - \underset{(3.20)}{25.08}(N_L - N_E) - \underset{(5.76)}{7.90}(N_L - N_E)_{-1} - \underset{(0.09)}{0.64}w_{-1} + \underset{(1.88)}{9.76}(t - 1931) + u_3'$$

$$\bar{R} = 0.93 \qquad \bar{S} = 32.26 \qquad \frac{\delta^2}{S^2} = 1.79$$

The first two equations are calculated from the period 1921–1941, and the last from the period 1920–1941. There is a slight discrepancy between the definitions of N_E in (3.4.7) and (3.4.9); however, the difference is not important for this problem. In the Appendix, the reader can find the exact definitions of all variables in these equations. The reasons why the discrepancy exists is that the data were used at different stages of this study for different purposes, and it does not seem worth while to recompute the equations with the slightly changed data. The coefficient of capital input is not reliable in the estimated production function, and its point value has a sign which is contrary to a priori expectations. We do not have a satisfactory measure of capital input, and, as explained in the following paragraph, there is a strong possibility of bias in the least-squares results. In other production studies it has been found that better estimates of *all* parameters can be obtained by changing from employment data on a person basis to data on a man-hour basis.

From the theoretical discussion in Chapter II it can be seen that there are definite relationships between the parameters of (3.4.7), (3.4.8), the demand-for-labor equation, and the investment equation. If (3.4.7) and (3.4.8) are to be made part of the model and if the parameters of these two equations are to be correctly estimated by the equation-systems methods of estimation, the relationships among the parameters of the several equations will have to be taken into account. This will complicate the estimation procedure and make the computations even more difficult than they already are.

In this section we have presented merely a few of the numerous plausible alternatives. Various tests have been applied to such alternatives as the effect of interest rates on investment, the effect of relative prices on consumption,[33] the effect of capital gains on consumption. In some cases the available data were not adequate, and in others the result of the test was negative in the sense that it did not call for revision of the models. It is desired to impress upon the reader that the models of this volume are put forth in full knowledge of the existence of the problem of multiplicity of hypotheses.

[33] For a discussion of the effect of relative prices on consumption and some statistical estimates of the effect, see L. R. Klein, "A Post-Mortem on Transition Predictions of National Product," *Journal of Political Economy*, Vol. LIV, August, 1946, pp. 299–302.

Chapter IV

ADEQUACY OF THE AVAILABLE DATA

The intensive use of econometric methods of economic research is much newer than the construction of national income statistics, index numbers, and other types of economic time series. Economic statistics have been prepared on the basis of intuitive concepts without regard to specific models of the system from which the data are derived; consequently, there is serious lack of coordination between the econometrician and the national income statistician. The readily available economic time series are almost never in a form suitable for immediate use in econometric studies. Until the time series shall have been prepared in a form designed to be used in the analysis of specific models there must necessarily be a large degree of processing of the data preparatory to its use by the econometrician. This extensive processing involves many assumptions and questionable steps; hence we must embark upon a qualitative description of the data to show the extent to which we have measured what we intended to measure. The quantitative details of the construction of the time series are given in the Appendix.

It will be recalled that one of the main features of the new methods used in econometrics is the estimation of parameters from systems of equations that have as many relationships as endogenous variables. It is very easy and straightforward to construct simple models in which we have the appropriate number of equations and variables, but it is difficult to find data for the estimation of the relevant parameters. The data are usually prepared in such a way that they are not suitable for measuring some of the relationships of the model, and the econometrician will eventually have to make a compromise by working with a partially complete system. For example, we have seen in the previous chapter that in order to estimate the parameters of a complete system we have to have the following information: (1) labor and capital used in the producer-goods industry, (2) labor and capital used in the consumer-goods industry, (3) inventories in the producer- and consumer-goods industry. We can get estimates of the labor, capital, and inventories used in industry as a whole, but we have not been able to construct estimates of the use of these factors in the consumer- and producer-goods industries separately. Estimates of labor, capital assets, depreciation, etc., can be found in detail for farming, manufacturing, construction, finance, and other sectors of the economy; however, the construction of mathematical models shows that the particular industrial divisions customarily used are not those which are especially significant in the econometric models. It is more important, for purposes of statistical inference, to have the data broken down by consumer- and producer-goods industries than by the industries found in national income statistics. If we could classify all data by the consumer- and producer-goods industries, it would be much easier to construct complete econometric models. Instead we are forced to compromise by calling certain variables exogenous which are really endogenous. We have not been able to give a complete picture of the supply relationships in the two industries separately (consumer and producer goods); hence we have not been able to get enough equations to determine the separate price levels in each industry. We have had to be content with the determination of the price level of output as a whole rather than its separate components.

Some information is useful in building economic hypotheses which is not operational; i.e., not capable of being tested in the world in which we live. The principal example of such information which is not operational is utility. It is not the fault of the national income statistician that there are no time series of utility. But the variables that are required for our models are certainly measurable. There is nothing subjective about the number of man-hours used in the consumer-goods industry. These data are capable of being measured, but it so happens that nobody has ever measured them.

We were not successful in preparing measurements of some of the variables that enter into a complete system, but, for those cases where we were successful, some words of qualification are necessary.

Stock of Capital, Depreciation and Investment

One of the most important variables in our models has been the stock of capital. This variable is important in order to estimate the influence of capital accumulation on the level of economic activity. In order to measure this variable we need to know the stock of capital at some fixed period of time, the gross investment, and depreciation. We can then use the formulas

$$I = I' - D'' \tag{4.1.1}$$

$$K_t = K_0 + \sum_{i=1}^{t} I_i \qquad t > 0 \tag{4.1.2}$$

$$K_t = K_0 - \sum_{i=t+1}^{0} I_i \qquad t < 0$$

where I = net real investment, I' = gross real investment, D'' = real depreciation, K = end-of-period stock of real capital.

The data on each of these variables are not all of the same quality. The best data are those pertaining to gross investment. From business records it is possible to compute periodic (usually annual) estimates of expenditures on durable producers' goods, or fixed capital assets. In order to get these expenditures in real terms, we can deflate these expenditures by index numbers of prices of capital goods which are available in great detail. The next step in the preparation of our series is the construction of data on depreciation. The difficulties inherent in this step are well known [1] to economists. We prepared our own estimates of depreciation of residential dwellings, but for business capital we relied, as is the custom, on the accounting estimates of business firms. This is admittedly unsatisfactory because business records are not intended to show what the econometrician desires to measure. Depreciation policies of business firms may be influenced by tax considerations and other matters which have nothing to do with the consumption of fixed capital. It is generally felt that accounting records on depreciation are far from perfect but that they do give some indication of the amount of capital used up in the productive process. Subject to numerous qualifications, business records of depreciation charges are accepted as the most reliable measure of time series on depreciation. It must be pointed out to the reader, though, that our estimates of net investment are no better than the business records on depreciation.

Net investment, itself, is used as a separate variable in the models, but it is also used in the measurement of capital stock according to formula (4.1.2). After we calculate net investment, subject to the qualifications on depreciation charges, our next step in the estimation of the stock of capital is the construction of an initial stock of capital at some arbitrary time period. There is a great abundance of information on the flow of income but very little data on the stock of wealth. The former censuses of wealth (U.S.A.) have been discontinued. The most satisfactory source for our estimate of an initial stock of wealth has again been the business records. Except for agriculture, where there has been a special wealth census, we have relied mainly on the aggregate balance sheets of United States corporations to obtain an initial figure for total capital assets, exclusive of land.

For many parts of our theory an error in the measurement of K_0, the initial stock of capital, is not very serious. Consider, for example, the investment function where the stock of capital enters as an important variable. In the notation of Chapter III, we have

[1] An exhaustive study of this problem is that of S. Fabricant, *Capital Consumption and Adjustment*, National Bureau of Economic Research, N. Y., 1938.

$$(4.1.3)\quad I = \beta_0 + \beta_1 \frac{pX - E}{q} + \beta_2 \left(\frac{pX - E}{q}\right)_{-1} + \beta_3 K_{-1} + \beta_4 t + u$$

We can rewrite K_{-1} as

$$(4.1.4)^{2}\quad K_{-1} = K_0 + \sum_{i=1}^{t-1} I_t$$

Substitution of (4.1.4) in (4.1.3) leads to

$$(4.1.5)\quad I = (\beta_0 + \beta_3 K_0) + \beta_1 \frac{pX - E}{q} + \beta_2 \left(\frac{pX - E}{q}\right)_{-1} + \beta_3 \left(\sum_{i=1}^{t-1} I_t\right) + \beta_4 t + u$$

We can now find an estimate of the parameter β_3 independent of the size of K_0. If we make an error in the measurement of K_0, the constant term may be estimated incorrectly, but the coefficient of the stock of capital may not. In the measurement of multipliers and other problems the important parameter is β_3 and not the constant term. For these measurements, we need not be concerned with the accuracy of the initial stock of wealth, but only with the accuracy of the measurement of net investment, particularly the measurement of depreciation, which is subtracted from gross investment to get net investment.

INVENTORIES

Another important component of national wealth, in addition to fixed capital, is working capital, or stocks of goods that have not yet reached their final destination, where they will be consumed. In inventories, we include only those stocks of goods that are held by business firms; hence accounting records are the main source of information. In these records, we find the end-of-year valuation of the stocks of goods on hand, but such data alone are not suitable for the statistics of national income or econometric models. The net change in the year-end stocks of goods represents a net addition to the current flow of output and should consequently be a component of national income. However, the annual changes in the year-end valuation of stocks of goods found in accounting records are made up of two types of change, the change due to variations in the quantity of goods and the change

[2] For simplicity, let us assume that the 0-point is the origin of our time series measurements so that we are always dealing with the case $t > 0$.

due to variations in the price of goods. For small changes we have

$$\frac{d(pX)}{dt} = p\frac{dX}{dt} + X\frac{dp}{dt} \tag{4.2.1}$$

where p = price and X = quantity. The first term of the right-hand side of (4.2.1) measures the value of the change of X (inventories) as a result of changes in the quantity of inventories. This term will be called the *value of change*. The second term represents the capital gains which result from the price changes. The sum of both terms will be called the *change in value*. Since the national income concept excludes all income represented by capital gains, the appropriate inventory component of national income is the value of the change rather than the change in value. The first differences in the year-end valuations of commodity stocks, as found in accounting records, give the change in value, whereas it is desired to obtain the value of the change. It is necessary to subtract from the change in value of stocks the net inventory profit of all business firms. This inventory profit is estimated from data on changes in price indexes during the period under consideration. Compared with other economic time series, the series on net inventory profits (capital gains) are unreliable. Accounting practices in the valuation of year-end commodity stocks vary from firm to firm; this fact makes it very difficult to obtain a measure of capital gains appropriate to the aggregate of firms.

For the purposes of our models, we need not only the net change in inventories (a flow), but also the stock of inventories in a constant price unit. We apply the same methods used for the stock of fixed capital. We deflate the net change in inventories measured in current prices by a current price index to get the net change in constant prices. Knowing an initial value, measured in the same prices as the deflated net change in inventories, we can obtain an estimate of year-end stocks in real terms by

$$H_t = H_0 + \sum_{i=1}^{t} \Delta H_i \qquad t > 0 \tag{4.2.2}$$

$$H_t = H_0 - \sum_{i=t+1}^{0} \Delta H_i \qquad t < 0$$

where H = inventories.

It is obvious from this discussion that the major stumbling block in the measurement of the flow or stock of inventories is the measurement of capital gains. Our data on inventories are only as reliable as are the estimates by the national income statisticians of the net inventory profits.

THE GOVERNMENT SECTOR OF THE ECONOMY

In models which attempt to describe economic behavior patterns, it is very desirable to separate the government from the private units (households and firms). This separation is desirable because it is believed that private units are motivated by factors different from those which motivate governments. For example, it does not seem possible to ascribe any maximizing behavior to governments in their demands for goods or factors of production. The government variables are usually placed in the exogenous category.

Government purchases of goods and services from private business firms are obtained from published budgets. On the demand side of the market, the data pertaining to government present no problem. But the supply side presents real difficulties.[3] How are we to measure the value of services (and goods) provided by the government? We know the payments of the government to labor (wages and salaries) and to rentiers (interest). We regard the government as a non-profit institution; hence the supply of services by government, represented by total factor payments, is the sum of wages, salaries, and interest payments. It seems legitimate to measure the services of government employees by wages and salaries; however, interest payments do not seem to be a good measure of the services of government capital. If all government bonds were issued in order to acquire funds for the purpose of constructing government buildings, for example, the interest payments on these bonds could be taken as a measure of the services provided by these buildings. But the fact remains that a very large portion of the government interest disbursements is paid on bonds issued to finance wars or other projects not represented by tangible fixed capital. Do the interest payments for postwar years on bonds issued to finance a war provide a measure of the value of government services during the postwar years? It can be said, in the event of victory, that the money borrowed to finance the war made it possible to carry on economic activity after the war by giving the winning country freedom from supervision by an enemy country. However, such arguments appear to stretch matters too far.

The policy adopted for the models of this volume has excluded all government interest payments from national production.[4] The value of services supplied by government has been taken to be represented by the remuneration of labor alone. This policy is not entirely satisfac-

[3] This point has been previously mentioned in Chapter III, p. 64.

[4] Government interest payments are included in the disposable income of households, however.

tory, because it excludes the supply of some services which should be included in national production; but it was adopted as the lesser evil since, in the future, government interest payments on war bonds will be a very large proportion of total interest payments by government.

Government receipts, as well as payments for goods, services, and factors of production, enter in an important way into the model. The appropriate income variable for households which shows how much money is available in the form of income for spending on consumer goods is: total national product, less payments by the private units to the government, less corporate savings, plus transfer payments by the government; in other words, disposable income.

There are some difficult problems in measuring the payments of individuals to governments in order to get disposable income. It is evident that all taxes paid to the government are not available for spending purposes by households; thus all tax payments should be deducted from national product. But there are non-tax payments by individuals to governments, and it must be decided whether or not to deduct these payments also. If individuals purchase fishing licenses from a government, it seems reasonable to regard these as consumption expenditures which should be dealt with like other consumer expenditure. Unfortunately, all non-tax payments are not of this one type, which is a payment for consumer goods. Fines for traffic violations are an example. We have treated all non-tax payments, like tax payments, as a deduction from household incomes in order to get disposable income, although the above examples show that some errors are committed in such a procedure.

It is also a problem to choose between tax payments actually sent by individuals or firms to governments, and tax liabilities that have accrued. If accrued taxes are set aside and not used for spending between the time they are incurred and the time they are paid, an accrual basis is to be preferred to a payment basis. However, if total liquid resources on hand influence spending plans, actual payments to governments are to be preferred. We have used taxes paid as the appropriate measure in our models although the usual estimates of disposable income deduct taxes accrued. For most of the interwar period, taxes paid and taxes accrued differ by very little; but in the last years before the war, during the defense program, accrued taxes differed significantly from paid taxes.[5]

[5] Perhaps a superior compromise would be to deduct taxes paid from household income and taxes accrued from business income. The difficulty in making this compromise, however, is that tax statements for unincorporated businesses are a part of personal income tax returns. The Department of Commerce now treats corporations on an accrual basis and individuals on a payment basis in the national accounts.

Housing

For the data on the housing market, one of the most important problems is to split new construction expenditures into two categories, expenditures by owner-occupants and expenditures by landlords. Good data are available for total expenditures; so the problem is one of finding the expenditures by either of the two above categories and computing the other type of expenditure as a residual.

Unfortunately, annual housing statistics are not compiled according to tenure. Since different behavior patterns are adopted by owner-occupiers and by landlords, it is important to attempt to process the data to bring out the most important characteristics of this market.

Available housing statistics are classified by type of structure, however, and this classification may enable us to get an approximation of our desired result. Newly constructed units can be classified as single-family units or multiple-family units. Owner-occupiers have both types of units, but owner-occupiers of multiple-family units rent at least part of their dwelling space to tenants; therefore they are classified as landlords. Our problem is to determine expenditures by owner-occupiers on non-farm, single-family dwellings. We know the total number of single-family units constructed annually, and we have to determine the proportions in which this total is divided between owner-occupiers and landlords. From the 1940 census, it is possible to obtain the percentage of single-family, non-farm units built in each of the periods 1920–1925, 1925–1930, 1930–1935, 1935–1940 that were owner-occupied in 1940. The percentage figure for 1935–1940 was 63 per cent; this is taken as our basic proportion because there was little change in tenure, it seems plausible to assume, in the short period elapsing between the average date of construction (1935–1940) and 1940. We do know that there was a change in tenure of great proportions during the depression years; hence the percentages of single-family dwellings, built in the earlier five-year intervals, which were owner-occupied in 1940, should be much smaller, as the calculated percentages show. The percentage figures calculated for 1920–1925, 1925–1930, 1930–1935 should thus be corrected upward to show the percentage taken by owner-occupiers at the time of construction. We do not know the precise upward correction to apply in order to account for the fact that many owners became tenants during the 'thirties, but we assume that the figure of 63 per cent, applicable to the new construction of 1935–1940, is also applicable to the earlier periods. This correction is, in any case, in the right direction. Thus it must be pointed out that our major assumption here is that a constant percentage (63 per cent) of the single-family units were purchased by owner-occupiers in the entire interwar period.

The estimation of the total expenditures by owner-occupiers in single-family, non-farm dwelling units is carried out in two stages. The first stage, the estimation of the physical volume, has been described in the preceding paragraph. The second stage is the estimation of the appropriate price series. The available data on building permit valuations provide the best indicator of unit value, but not without improvement by correction of the figures for certain omissions. Total value is understated in the permit valuation because many building costs are not included. The undervaluation is commonly reckoned at between 15 and 20 per cent for the *interwar period.* Following the procedure of Wickens,[6] we have used an upward correction of 18 per cent. This figure is undoubtedly not constant over the business cycle, but, since we do not know the annual variations, we have used a constant.

We must make another correction to account for the fact that owner-occupied, single-family dwellings are on the average more expensive than other single-family dwellings. This correction factor is taken as the ratio of the average rental value of owner-occupied, single-family, non-farm dwellings to the average rental value of all single-family, non-farm dwellings constructed 1935–1940. There should perhaps be another correction to account for the time lag between the filing of building permits and the actual construction expenditures. However, surveys[7] show that in 1931 the average time that elapsed between the issuance of a permit and the start of work was about 6 days, and that the average time taken to complete buildings in 1931 was about 3½ months. These time lags are much larger during the postwar construction bottleneck, but all the data in this book refer to the interwar period. It is believed that the correction for time lag is minor.

Finally, we must make some additional assumptions in order to derive a percentage figure which will show how repair and maintenance expenditures should be divided according to our tenure classifications. We know from Department of Commerce data how repairs are divided between owner-occupied non-farm units and rented non-farm units in 1930 and 1940, but we need, for our series, the division of repairs by tenure between owner-occupied, single-family, non-farm units and all other non-farm units. We assume that the fraction of repairs going to all owner-occupied non-farm units is in the same ratio to the fraction

[6] D. Wickens, *Residential Real Estate,* National Bureau of Economic Research, New York, 1941.

[7] "Elapsed Time and Cost in Residential Construction," *Monthly Labor Review,* Vol. 63, September, 1946, pp. 346–354.

of repairs going to owner-occupied, single-family, non-farm units as the fraction of all non-farm units represented by owner-occupied non-farm units is to the fraction of all non-farm units represented by owner-occupied, single-family, non-farm units. The formula for this complicated expression is given in the Appendix.

A further complication of the data on repairs is that only major repairs are included in the time series. Only those repairs for which it is necessary to take out a building permit appear in the data. Small repairs are not included here, but they do contribute to national income and will be implicit elsewhere in the system, namely, in the sections explaining consumer expenditures. In our model small repairs are treated like ordinary consumption expenditures and are a part of the consumption function instead of the demand equation for housing.

Another problem in connection with housing data is the estimation of the depreciation of residences. In the case of producers' plant and equipment, we had extensive accounting records of business firms, which serve as a guide to the estimation of depreciation charges, but in the case of housing we do not have the advantage of accounting records. We must prepare our own estimates of depreciation in the housing market. At the suggestion of C. Rapkin, formerly of the National Housing Agency, we assumed that dwellings depreciate at the rate of 3 per cent compounded annually. This figure was arrived at by Rapkin by calculating the time rate at which houses move from a higher rental class to a lower class. If we know the value of the stock of housing for a particular period and gross residential construction in the prices of the same period we can calculate the annual depreciation by the well-known formulas of compound interest. The resulting figures for depreciation will be in the same constant prices as gross real residential construction and the stock of housing. The formulas for this calculation are given in the Appendix.

CASH BALANCES

The theories of liquidity preference, from which the demand equations for money are derived, are based on certain classifications of the stock of cash balances which are not given precisely in the existing data. According to the liquidity-preference theory there are basically two types of cash balances: (1) those held for speculative purposes, (2) those held for transactions and precautionary purposes. Speculative balances are sometimes called *idle* balances, and transaction-precautionary balances are sometimes called *active* balances. The data on cash balances are classified as circulating currency, checking deposits, and savings deposits. Our assumptions have been that savings deposits are idle balances and that checking deposits plus circulating currency are active balances. It is probably correct to assume that savings deposits are idle balances, i.e., not intended to be used on short notice, but it is questionable to assume that checking accounts are active balances exclusively. There may be some checking accounts that are held as idle balances, particularly the checking accounts of business firms. The checking accounts that are held as precautionary balances appear idle in the sense that they remain unused for a long period of time, but they are not held for speculative reasons and will be used without delay in case of an emergency. It will be recalled that the demand equation for active balances contains a constant term as well as a term depending upon income. This constant term includes the precautionary balances.

On the basis of the available data our classification seems to be the best to adopt.

WAGES AND PROFITS

In the example of the simple, three-equation system, we found it very important to distinguish between profits and wages. Profits are supposed to represent the incomes accruing to those individuals who own the factors of production, and wages are supposed to represent the incomes of those individuals who work for the profit receivers. There are some individual entrepreneurs who do not hire labor and do all their own work. Is the reward they pay to themselves wages or profit? The best solution would be to classify all entrepreneurial incomes below a certain amount as wage incomes and all entrepreneurial incomes above this amount as profit income, but there is no way of dividing the entrepreneurial income between wage income and profit income, except for agriculture. D. Gale Johnson of the University of Chicago has permitted the author to use his unpublished study [8] in which he allocates farm income between wages and profits. Thus for farmers we have been able to separate entrepreneurial income in our data; but all other entrepreneurial income is included in profit income, although many entrepreneurs behave more like wage earners than profit recipients in their consumption habits.

Not only should we transfer certain entrepreneurial income to wage income, but also we should transfer certain salary income to profit income. Statistics are not adequate to segregate high salaries of business executives from the wage-salary time series. We have included all salaries in the wage category; thus our time series on wages includes income that should be treated like profit income.

[8] This study has since been published as "Allocation of Agricultural Income," *Journal of Farm Economics*, Vol. XXX, November, 1948, pp. 724–749.

The net result is that we make compensating errors, but we do not know the extent of compensation. In each category we include income that belongs in the other category, but we can say no more.

In this chapter we have illustrated some of the difficulties that beset the econometrician in carrying out statistical measurements for the real world. Many assumptions have to be made in order to get numerical answers, and the reader must judge for himself whether or not we have actually measured what we set out to measure. It hardly seems possible, though, that our data can be too much in error. Time and again we find that the data we do have follow systematic patterns that are developed in advance on purely theoretical grounds. If the data were very poor, we should undoubtedly have found many more inconsistencies than we have already uncovered.

FUTURE COLLECTION OF DATA

The discussion in this chapter shows the need for improving the statistical data, especially in adapting it for use in econometric models. Under ideal circumstances, a coordinating statistical body should decide in advance what kind of models would be most needed and then encourage the collection of data which would be necessary to estimate the parameters of this model.

Undoubtedly we shall need to have all output and input series classified by consumer- and producer-goods industries. These industries can be divided further according to the traditional industrial classifications such as manufacturing, construction, transportation. All assets must also be similarly classified according to the type of industry which holds them. Housing data must be segregated according to tenure status of the occupier; vacant units should be classified on an annual basis as vacant for sale or vacant for rent. Time series on the mortgage interest rate and price level of old houses should be developed. The consumption function could be greatly improved if we had annual data on income by size. These are only some of the most obvious examples of ways in which the data should be reclassified or extended. But it is certain that, if we are to get better statistical models, we must get better data.

APPENDIX

MODEL I: TIME SERIES

	C	I	G	$Y+T$	K_{-1}	Π	W_1	W_2	Y
1920	39.8	2.7	4.6	47.1	180.1	12.7	28.8	2.2	43.7
1921	41.9	−0.2	6.6	48.3	182.8	12.4	25.5	2.7	40.6
1922	45.0	1.9	6.1	53.0	182.6	16.9	29.3	2.9	49.1
1923	49.2	5.2	5.7	60.1	184.5	18.4	34.1	2.9	55.4
1924	50.6	3.0	6.6	60.2	189.7	19.4	33.9	3.1	56.4
1925	52.6	5.1	6.5	64.2	192.7	20.1	35.4	3.2	58.7
1926	55.1	5.6	6.6	67.3	197.8	19.6	37.4	3.3	60.3
1927	56.2	4.2	7.6	68.0	203.4	19.8	37.9	3.6	61.3
1928	57.3	3.0	7.9	68.2	207.6	21.1	39.2	3.7	64.0
1929	57.8	5.1	8.1	71.0	210.6	21.7	41.3	4.0	67.0
1930	55.0	1.0	9.4	65.4	215.7	15.6	37.9	4.2	57.7
1931	50.9	−3.4	10.7	58.2	216.7	11.4	34.5	4.8	50.7
1932	45.6	−6.2	10.2	49.6	213.3	7.0	29.0	5.3	41.3
1933	46.5	−5.1	9.3	50.7	207.1	11.2	28.5	5.6	45.3
1934	48.7	−3.0	10.0	55.7	202.0	12.3	30.6	6.0	48.9
1935	51.3	−1.3	10.5	60.5	199.0	14.0	33.2	6.1	53.3
1936	57.7	2.1	10.3	70.1	197.7	17.6	36.8	7.4	61.8
1937	58.7	2.0	11.0	71.7	199.8	17.3	41.0	6.7	65.0
1938	57.5	−1.9	13.0	68.6	201.8	15.3	38.2	7.7	61.2
1939	61.6	1.3	14.4	77.3	199.9	19.0	41.6	7.8	68.4
1940	65.0	3.3	15.4	83.7	201.2	21.1	45.0	8.0	74.1
1941	69.7	4.9	22.3	96.9	204.5	23.5	53.3	8.5	85.3

C: Consumption, measured in billions of 1934 dollars.

$$C = \frac{(1) + (2)}{(3)}$$

(*1*) = Consumer expenditures, Mary S. Painter, "Estimates of Gross National Product, 1919–1928," *Federal Reserve Bulletin*, Vol. 31, September, 1945, pp. 872–873. Official estimates of the U. S. Department of Commerce are given in this article for the period 1929–1941.

(2) = Imputed net rents on owner-occupied residences, S. Kuznets, *National Income and Its Composition*, Vol. II, p. 735. These data are given for 1919–1938. For 1939, 1940, 1941 the estimates of gross imputed rents have been extended by Kuznets' method (estimated number of owner-occupied units multiplied by average monthly rental; converted to net rents by use of free-hand regression between net and gross imputed rents, 1933–1938):

1939: \$1.5 billion
1940: \$1.5 billion
1941: \$1.6 billion

(3) = Price index implicit in the adjustment of consumers' outlay, 1934: 1.00, S. Kuznets, *op. cit.*, Vol. I, p. 145, col. 3, converted to 1934 base year. These data are given for 1919–1938. For 1939, 1940, 1941 the estimates have been extended by Kuznets' method (weighted average of Bureau of Labor Statistics cost-of-living index and U. S. Department of Agriculture index of prices paid by farmers for subsistence, weights being proportionate to ratio of urban and of rural populations respectively):

1939: 1.026
1940: 1.034
1941: 1.093

I: Net investment, measured in billions of 1934 dollars.

$$I = \frac{(4)}{(5)} + \frac{(6)}{(7)} - \frac{(8)}{(9)} - \frac{(10)}{(7)} + \frac{(11)}{(12)} + \frac{(13)}{(14)} + \frac{(15)}{(16)} - D'' + \frac{(17) - 0.1}{(14)}$$

$$D'' = \left[67.6 - \sum_{i=t+1}^{1933} \left(\frac{(13)}{(14)} + \frac{(15)}{(16)}\right)_i + \sum_{i=t+1}^{1933} D_i''\right] \frac{0.03}{0.97} - \left(\frac{(13)}{(14)} + \frac{(15)}{(16)}\right)_t \frac{0.015}{0.97} \qquad t < 1933$$

$$D'' = 67.6\,\frac{0.03}{0.97} - \left(\frac{(13)}{(14)} + \frac{(15)}{(16)}\right)_t \frac{0.015}{0.97} \qquad t = 1933$$

$$D'' = 67.6(0.03) + \left(\frac{(13)}{(14)} + \frac{(15)}{(16)}\right)_t 0.015 \qquad t = 1934$$

$$D'' = 67.6(0.97)^{t-1934}\cdot 0.03 + \sum_{i=1934}^{t-1} \left(\frac{(13)}{(14)} + \frac{(15)}{(16)}\right)_i 0.985(0.97)^{t-1-1934}\cdot 0.03 + \left(\frac{(13)}{(14)} + \frac{(15)}{(16)}\right)_t 0.015 \qquad t > 1934$$

(4) = Gross expenditures on private producers' non-agricultural plant and equipment, G. Terborgh, *Federal Reserve Bulletin*, Vol. 25, September, 1939, p. 732, for 1919–1936, Vol. 26, February, 1940, p. 116, for 1937, 1938; F. Dirks, *Federal Reserve Bulletin*, Vol. 28, April, 1942, p. 318, for 1939–1941.

(*5*) = Price index of business capital goods, 1934: 1.00, S. Fabricant, *Capital Consumption and Adjustment*, pp. 178–179; converted to 1934 base year. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence.

(*6*) = Expenditures on farm service buildings (excluding dwellings) + expenditures on farm machinery + expenditures on farm trucks + 40 per cent of expenditures on farm automobiles other than trucks, *Income Parity for Agriculture*, Part II, Section 5, p. 25, col. 1, and *Net Farm Income and Parity Report*, U. S. Department of Agriculture, 1943, p. 27. (The series on expenditures on farm service buildings has been adjusted from 1933–1941.)

(*7*) = Price index of farm capital goods, 1934: 1.00. Weighted average of prices paid by farmers for building materials (*Agricultural Statistics*, 1943, p. 394), for farm machinery (*ibid.*), and for motor vehicles (*Income Parity for Agriculture*, U. S. Department of Agriculture, Part III, section 4, p. 11, for 1919–1938 and extended for 1939–1941 by adopting the same percentage changes as those of the wholesale price index of motor vehicles of the U. S. Department of Commerce). The weights are, respectively, current-dollar expenditures for farm service buildings (excluding dwellings), for farm machinery, and for farm trucks plus 40 per cent of other farm motor vehicles. The sources are the same as those for item (*6*). The weighted average was then converted to a 1934 base.

(*8*) = Depreciation charges on private producers' non-agricultural plant and equipment. The basic data for 1919–1928 come from S. Fabricant, *op. cit.*, pp. 260–261, and for 1929–1941 from J. Mosak, "Forecasting Postwar Demand: III," *Econometrica*, Vol. 13, 1945, pp. 45–53. In splicing these two series, industry by industry, certain adjustments had to be made to get comparable data for the residual group which excluded agriculture, public utilities, transportation, mining, and manufacturing. Depreciation on rented residences, in particular, was subtracted from Mosak's data; then the figures for 1919–1928 were estimated by constructing a series having the same year-to-year percentage changes as Fabricant's series but having Mosak's adjusted value for 1929. Mosak's adjusted figures were used for 1929–1941. Also depreciation on transportation in common carrier buses, motor-trucking, taxis, air and harbor craft was shifted from the transportation category to the miscellaneous category. This change introduced an error of rounding in adding up the components of total depreciation.

(*9*) = Price index underlying business depreciation charges, 1934 base, S. Fabricant, *op. cit.*, p. 183, converted to 1934 base year by use of Fabricant's general index of the prices of business capital goods, pp. 178–179, Table 32. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence. The ratio (*8*)/(*9*) gives a somewhat oversimplified formula for the calculation of depreciation on private producers' non-agricultural plant and equipment in 1934 dollars because the series actually used is the sum of constant(1934)-dollar series for each of three industry groups, which is not exactly the same thing as total depreciation divided by an aggregate price index. However, the two different methods of calculation lead to results that differ only slightly.

(*10*) = Depreciation charges on farm plant and equipment (exclusive of dwellings), calculated at replacement costs. *Net Farm Income and Parity Report*, U. S. Department of Agriculture, 1943, p. 23, and *Income Parity for Agriculture*, U. S. Department of Agriculture, Part II, Section 5, p. 25.

(*11*) = Net inventory change, Mary S. Painter, *loc. cit.*

(*12*) = Wholesale price index 1934: 1.00, Supplement to the *Survey of Current Business*, 1942, p. 18, converted to 1934 base year.

(*13*) = Gross expenditures on construction of private non-farm residences, *The Construction Industry in the United States*, U. S. Bureau of Labor Statistics, Bulletin 786, Table 1, p. 4.

(*14*) = American Appraisal Co. index of construction costs (national average), 1934: 1.00, Supplement to the *Survey of Current Business*, 1942, p. 25, converted to 1934 base year.

(*15*) = Gross expenditures on construction of farm residences, *The Construction Industry in the United States*, U. S. Bureau of Labor Statistics, Bulletin 786, Table 1, p. 4.

(*16*) = Index of farm construction costs, 1934: 1.00, *Income Parity for Agriculture*, U. S. Department of Agriculture, Part II, Section 5, p. 27. These data are for 1919–1940. The 1941 figure was taken from a free-hand regression between (*16*) and the index of the cost of building materials taken from *Agricultural Prices*, February 29, 1944, col. 7, p. 25. All indexes are converted to 1934 base year.

D'' = Depreciation of residential dwellings, measured in billions of 1934 dollars.

67.6 = The estimated value, January 1, 1934, of the stock of residential dwellings in the United States. It is 80 per cent of Wickens' (*Residential Real Estate*, National Bureau of Economic Research, Table A8, p. 93) figure for the value of non-farm land and dwellings plus 14.3 per cent of the value of total farm real estate, *Net Farm Income and Parity Report*, U. S. Department of Agriculture, 1943, Table 18, p. 29. The figures of 80 per cent and 14.3 per cent are taken as the proportions of non-farm residential and farm real estate, respectively, that are represented by dwellings alone.

(*17*) = Gross expenditures on construction by non-profit institutions, G. Terborgh, *loc. cit.*, and F. Dirks, *loc. cit.*

0.1 = Annual depreciation charges in billions of current dollars attributed to the plant of non-profit institutions. This figure is about 3 per cent of the 1934, end-of-year stock of capital owned by these institutions.

G: Exogenous investment, measured in billions of 1934 dollars.

$$G = \frac{(18)}{(12)} + \frac{(19)}{(14)} + \frac{(20) - (21) - (19)}{(22)}$$

(*18*) = Net exports and monetary use of gold and silver, Mary S. Painter, *loc. cit.*

(*19*) = Public construction expenditures, including work-relief construction, *The Construction Industry in the United States*, U. S. Bureau of Labor Statistics, Bulletin 786, Table 1, p. 4, and *Statistical Abstract of the United States*, 1942, p. 992.

(*20*) = Government expenditures for goods and services, Mary S. Painter, *loc. cit.*

(*21*) = Government interest payments, S. Kuznets, *op. cit.*, Vol. II, p. 813. These data are for 1919–1928. J. Mosak, *op. cit.*, p. 51. These data are for 1929–1941.

(*22*) = Index of wholesale prices of non-farm products, 1934: 1.00. *Agricultural Statistics*, 1942, p. 649, converted to 1934 base year.

$Y + T$: Net national product, measured in billions of 1934 dollars.

$$Y + T = C + I + G$$

K: End-of-year stock of capital, measured in billions of 1934 dollars.

$$K = 199.0 - \sum_{i=t+1}^{1934} I_i \qquad t < 1934$$

$$K = 199.0 \qquad t = 1934$$

$$K = 199.0 + \sum_{i=1935}^{t} I_i \qquad t > 1934$$

199.0 = Value of private producers' plant and equipment + non-profit institutions' plant and equipment + residential dwellings + in-inventories, all measured at the end of the year 1934. The stock of capital, excluding land, of all private corporations for end-of-year 1934 are given in S. Fabricant, *op. cit.*, p. 271. These data are converted to estimates for both corporate and non-corporate enterprises by dividing the figure for each industry by the ratio of corporate output to total output. These ratios are

Mining and Quarrying	0.96
Manufacturing	0.92
Construction	0.60
Transportation and Public Utilities	1.00
Trade	0.63
Service	0.30
Finance and Real Estate	0.84

These ratios are found in *The Structure of the American Economy*, National Resources Planning Board, pp. 375–376. The capital for the service industries includes the capital for non-profit institutions. Agricultural capital exclusive of livestock and dwellings is taken from *The Structure of the American Economy*, p. 377.

The total value of dwellings at the end of 1934 is obtained as given above under the explanation of the variable I. It is the value for January 1, 1934 (\$67.6 billion), plus the net investment in housing during 1934 (−\$1.3 billion).

Business inventories and inventories of agricultural corporations come from *The Survey of Current Business*, Vol. 22, September, 1942, p. 18. Total agricultural inventories come from *Agritural Statistics*, 1942 (value of crops for sale + value of livestock).

Business inventories end of 1934	$17,913 million
Less: Inventories of agricultural corporations	196
	17,717
Plus: Total agricultural inventories	4,087
	21,804
Private plant and equipment	$110.9 billion
Housing	66.3
Inventories	21.8
Stock of capital, end-of-year 1934	199.0

Π: Profits, measured in billions of 1934 dollars.

$$\Pi = \frac{(23) + (24) + (25) + (2) + (26) - (21) + (27) - (28) + (29)}{p}$$

$$p = \frac{\left\{\begin{matrix}(1) + (2) + (4) + (6) - \dfrac{(8)}{(9)}(5) - (10) + (11) + (13) - (14)D'' \\ + (15) + (17) - 0.1 + (18) + (20) - (21)\end{matrix}\right\}}{Y + T}$$

(23) = Dividends, M. Hoffenberg, "Estimates of National Output, Distributed Income, Consumer Spending, Saving, and Capital Formation," *Review of Economic Statistics*, Vol. 25, May, 1943, p. 156 (Department of Commerce estimates).

(24) = Corporate net savings, M. Hoffenberg, *loc. cit.* (Department of Commerce estimates).

(25) = Rent, M. Hoffenberg, *loc. cit.* (Department of Commerce estimates).

(26) = Interest, M. Hoffenberg, *loc. cit.* (Department of Commerce estimates).

(27) = Entrepreneurial net income, M. Hoffenberg, *loc. cit.* (Department of Commerce estimates).

(28) = Return to labor in agriculture, D. Gale Johnson, *Allocation of Agricultural Income*, unpublished manuscript, University of Chicago.

(29) = Total cash wages and value of board, lodging, and other perquisites paid to hired labor in agriculture, *Farm Wage Rates, Farm Employment and Related Data*, U. S. Department of Agriculture, January, 1943, p. 177.

W_1: Labor income originating in private employment, measured in billions of 1934 dollars.

$$W_1 = \frac{(30) - (31) + (28) - (29)}{p}$$

(30) = Total compensation of employees, M. Hoffenberg, *loc. cit.* (Department of Commerce estimates).

(31) = Government wages and salaries including work relief wages, *Survey of Current Business*, Vol. 23, March, 1943, p. 22. These data are for 1929–1941. The figures for 1919–1928 are estimated from a free-hand regression between the U. S. Department of Commerce estimates (1929–1938) and Kuznets' estimates (1929–1938). See S. Kuznets, *op. cit.*, Vol. II, p. 811.

Kuznets' figures are known for 1919–1928; hence the corresponding Department of Commerce figures can be estimated from the regression.

W_2: Labor income originating in government, measured in billions of 1934 dollars.

$$W_2 = \frac{(31)}{p}$$

Y: Net national income, measured in billions of 1934 dollars.

$$Y = \Pi + W_1 + W_2$$

MODEL II: TIME SERIES

	C	I′	G	Y	GNP	T	p	N	M
1921	50.5	9.2	8.5	52.8	68.2	15.4	1.277	0.1085	37.8
1922	52.0	10.0	7.9	57.0	69.9	12.9	1.197	0.1101	39.0
1923	57.8	15.3	8.5	66.1	81.6	15.5	1.219	0.1120	42.7
1924	60.9	12.3	8.8	66.7	82.0	15.3	1.222	0.1141	44.5
1925	63.0	15.7	9.3	70.6	88.0	17.4	1.254	0.1158	48.3
1926	66.3	16.4	9.5	73.2	92.3	19.1	1.264	0.1174	50.6
1927	66.0	14.9	10.0	73.5	90.9	17.4	1.240	0.1190	52.2
1928	68.8	14.3	10.6	75.4	93.7	18.3	1.226	0.1205	54.7
1929	70.8	17.6	11.0	79.6	99.4	19.8	1.225	0.1218	55.2
1930	64.9	12.1	11.2	70.7	88.2	17.5	1.194	0.1238	54.4
1931	54.2	6.4	11.5	59.6	72.1	12.5	1.087	0.1248	52.9
1932	43.0	2.2	10.2	45.6	55.4	9.8	0.976	0.1256	45.4
1933	42.4	3.3	9.1	44.5	54.8	10.3	0.924	0.1263	41.7
1934	47.7	5.3	10.8	51.0	63.8	12.8	0.957	0.1271	46.0
1935	52.2	6.7	11.9	56.3	70.8	14.5	0.981	0.1280	49.9
1936	59.1	10.0	12.6	65.2	81.7	16.5	0.991	0.1289	55.1
1937	62.5	11.6	13.6	69.2	87.7	18.5	1.027	0.1296	57.3
1938	58.5	7.7	14.4	62.9	80.6	17.7	1.008	0.1307	56.6
1939	61.7	10.9	16.0	67.7	88.6	20.9	0.994	0.1317	60.9
1940	65.7	14.8	16.7	72.9	97.1	24.2	1.002	0.1328	67.0
1941	74.6	19.1	26.5	88.7	120.2	31.5	1.052	0.1340	74.2

C: Consumer expenditures on goods and services, measured in billions of current dollars, U. S. Department of Commerce.

I′: Private gross capital formation, measured in billions of current dollars, U. S. Department of Commerce.

G: Government expenditures for goods and services, measured in billions of current dollars, U. S. Department of Commerce.

Y: Disposable income, measured in billions of current dollars, U. S. Department of Commerce.

GNP: Gross national product, measured in billions of current dollars, U. S. Department of Commerce.

T: GNP − Y.

(The six above series do not take account of the revisions published in July, 1947.)

p: Cost-of-living index, 1935–1939: 1.00, U. S. Bureau of Labor Statistics.

N: Population of the continental United States, measured in billions of persons, U. S. Bureau of the Census.

M: Total deposits adjusted and currency outside banks, measured in billions of current dollars on June 30 of each year, Board of Governors of the Federal Reserve System.

MODEL III: TIME SERIES

	C	I	q	ΔH	D_1	q_1	D_2	D_3	D''	G
1920	39.8	1.1	1.459	2.4	0.4	1.758	0.5	0.2	2.0	4.7
1921	41.9	0.1	1.169	0.1	0.8	1.342	0.5	0.1	1.9	6.7
1922	45.0	0.6	1.067	0.6	1.3	1.242	1.0	0.2	1.9	6.3
1923	49.2	2.0	1.173	2.1	1.3	1.391	1.4	0.2	2.0	5.9
1924	50.6	1.6	1.154	−0.2	1.5	1.379	1.7	0.2	2.0	6.9
1925	52.6	2.4	1.133	0.7	1.7	1.348	1.8	0.2	2.0	6.9
1926	55.1	2.5	1.132	1.2	1.5	1.348	1.9	0.2	2.1	7.0
1927	56.2	2.2	1.130	0.2	1.4	1.348	1.8	0.2	2.1	8.1
1928	57.3	2.0	1.118	−0.3	1.4	1.348	1.6	0.2	2.2	8.3
1929	57.8	2.8	1.139	1.3	1.0	1.348	1.7	0.2	2.2	8.4
1930	55.0	1.6	1.092	−0.3	0.8	1.242	0.6	0.2	2.2	9.7
1931	50.9	−0.7	1.016	−2.1	0.7	1.106	0.6	0.1	2.2	10.9
1932	45.6	−2.2	0.944	−2.7	0.4	0.963	0.2	0.1	2.1	10.3
1933	46.5	−2.7	0.926	−0.8	0.3	0.932	0.1	0.1	2.1	9.3
1934	48.7	−1.6	1.000	−0.1	0.4	1.000	0.2	0.1	2.0	10.0
1935	51.3	−0.6	1.013	0.2	0.7	1.006	0.2	0.2	2.0	10.5
1936	57.7	0.5	1.034	2.0	0.9	1.056	0.4	0.2	2.0	10.4
1937	58.7	1.4	1.129	1.0	0.8	1.230	0.5	0.2	2.0	11.1
1938	57.5	−0.4	1.133	−1.2	0.9	1.236	0.5	0.2	2.0	13.1
1939	61.6	0.2	1.131	0.9	1.1	1.242	0.8	0.2	2.0	14.5
1940	65.0	1.2	1.159	1.7	1.3	1.267	0.8	0.2	2.0	15.5
1941	69.7	1.3	1.222	3.0	1.6	1.354	0.7	0.3	2.0	22.4

C: Consumption, measured in billions of 1934 dollars.

$$C = \frac{(1) + (2)}{(3)}$$

(*1*) = Consumer expenditures, Mary S. Painter, "Estimates of Gross National Product, 1919–1928," *Federal Reserve Bulletin*, Vol. 31, September, 1945, pp. 872–873. Official data of the U. S. Department of Commerce are given in this article for the period 1929–1941.

(*2*) = Imputed net rents on owner-occupied residences, S. Kuznets, *National Income and Its Composition*, Vol. II, p. 735. These data are given for 1919–1938. For 1939, 1940, 1941 the estimates for gross imputed rent have been extended by Kuznets' method (estimated number of owner-occupied units multiplied by average monthly rental; converted to net

MODEL III: TIME SERIES (*Continued*)

	$Y+T$	Y	p	W_2	W_1	R_1	R_2	r	ΔF	v *
1920	47.1	42.3	1.611	3.5	41.4	6.5	0.9	1.279	612	99.1
1921	48.3	44.2	1.317	3.6	32.0	8.1	0.8	1.468	561	100.4
1922	53.0	47.0	1.251	3.6	34.1	8.4	0.7	1.512	582	100.7
1923	60.1	53.7	1.281	3.7	40.6	9.3	0.8	1.551	630	100.1
1924	60.2	53.7	1.267	3.9	40.1	10.1	0.8	1.606	515	98.8
1925	64.2	56.6	1.294	4.1	41.6	10.2	0.8	1.612	529	97.6
1926	67.3	59.3	1.293	4.3	44.3	10.0	0.8	1.596	542	96.4
1927	68.0	60.1	1.260	4.5	44.3	10.1	0.8	1.571	465	95.2
1928	68.2	59.5	1.267	4.7	45.8	10.0	0.8	1.534	419	94.3
1929	71.0	61.9	1.259	5.0	48.1	10.4	0.8	1.498	474	93.4
1930	65.4	60.1	1.211	5.1	43.1	10.0	0.8	1.457	306	92.7
1931	58.2	56.6	1.088	5.2	35.4	9.3	0.8	1.380	234	92.0
1932	49.6	49.2	0.966	5.1	26.6	8.2	0.7	1.238	112	91.1
1933	50.7	46.2	0.931	5.2	24.6	7.2	0.6	1.067	380	92.5
1934	55.7	50.4	1.000	6.0	28.5	6.9	0.6	1.000	553	94.1
1935	60.5	54.0	1.040	6.3	31.2	7.0	0.6	0.998	551	95.2
1936	70.1	63.5	1.047	7.7	35.3	7.2	0.6	1.021	572	95.7
1937	71.7	62.6	1.096	7.3	41.0	7.7	0.6	1.069	585	95.8
1938	68.6	60.2	1.047	8.1	37.0	8.0	0.6	1.103	456	95.8
1939	77.3	67.6	1.035	8.1	40.0	8.2	0.6	1.105	520	95.8
1940	83.7	72.3	1.049	8.4	44.0	8.5	0.7	1.108	508	96.3
1941	96.9	82.9	1.121	9.5	55.1	9.0	0.7	1.122	518	97.2

* A figure greater than 100 indicates that demand exceeds physical supply.

rents by use of free-hand regression between net and gross imputed rents, 1933–1938):

1939: \$1.5 billion
1940: \$1.5 billion
1941: \$1.6 billion

$(\mathcal{P})$ = Price index implicit in the adjustment of consumers' outlay, 1934: 1.00, S. Kuznets, *op. cit.*, Vol. I, p. 145, col. 3, converted to 1934 base year. These data are given for 1919–1938. For 1939, 1940, 1941 the estimates have been extended by Kuznets' method (weighted average of Bureau of Labor Statistics cost-of-living index and U. S. Department of Agriculture index of prices paid by farmers for subsistence, weights being proportionate to ratio of urban and of rural populations respectively):

MODEL III: TIME SERIES (*Continued*)

	N^a	i	M_1^D	M_2^D	E_R	K	H	E	u_3	X
1920	19.0	7.08	23.7	15.8	28	97.2	22.1	0.6		40.3
1921	19.4	7.04	20.8	16.6	29	97.3	22.2	0.6	−0.74	38.8
1922	19.9	5.95	21.4	17.4	64	97.9	22.8	0.6	−0.31	42.8
1923	20.7	6.04	22.7	19.7	35	99.9	24.9	0.6	0.76	49.3
1924	21.5	5.80	23.5	21.3	16	101.5	24.7	0.6	−0.61	48.5
1925	22.3	5.47	25.2	23.1	11	103.9	25.4	0.7	−0.30	52.5
1926	23.1	5.21	25.7	24.7	−35	106.4	26.6	0.7	0.24	55.6
1927	23.9	4.97	25.8	26.4	5	108.6	26.8	0.8	−0.16	55.8
1928	24.6	4.94	26.3	28.2	13	110.6	26.5	0.9	−0.67	56.0
1929	25.3	5.21	26.4	28.5	43	113.4	27.8	1.0	0.76	58.1
1930	25.8	5.09	25.3	28.6	55	115.0	27.5	1.1	0.52	52.3
1931	26.2	5.81	23.3	27.9	89	114.3	25.4	1.2	−0.10	44.1
1932	26.4	6.87	20.8	25.1	256	112.1	22.7	1.4	−0.19	35.1
1933	26.6	5.89	19.8	22.6	528	109.4	21.9	1.8	0.57	36.7
1934	26.7	4.96	21.4	22.6	1564	107.8	21.8	2.2	−0.02	42.2
1935	27.0	4.46	25.1	23.8	2469	107.2	22.0	2.6	−0.49	47.1
1936	27.4	3.87	29.0	24.8	2512	107.7	24.0	3.0	0.61	55.3
1937	27.9	3.94	30.4	25.8	1220	109.1	25.0	3.2	0.00	57.5
1938	28.5	4.19	30.4	26.3	2522	108.7	23.8	3.3	−1.16	52.6
1939	29.2	3.77	33.8	26.7	4392	108.9	24.7	3.5	−0.35	61.0
1940	30.1	3.55	39.0	27.4	6326	110.1	26.4	3.9	0.23	66.9
1941	31.0	3.34	45.5	27.8	5324	111.4	29.4	4.5	0.67	79.8

1939: 1.026
1940: 1.034
1941: 1.093

I: Net investment in private producers' plant and equipment, measured in billions of 1934 dollars.

$$I = \frac{(4)}{(5)} + \frac{(6)}{(7)} - \frac{(8)}{(9)} - \frac{(10)}{(7)}$$

(*4*) = Gross expenditures on private producers' non-agricultural plant and equipment, G. Terborgh, *Federal Reserve Bulletin*, Vol. 25, September, 1939, p. 732, for 1919–1936, Vol. 26, February, 1940, p. 116, for 1937–1938; F. Dirks, *Federal Reserve Bulletin*, Vol. 28, April, 1942, p. 318, for 1939–1941.

(*5*) = Price index of business capital goods, 1934: 1.00, S. Fabricant, *Capital Consumption and Adjustment*, pp. 178–179, converted to 1934 base year.

These data are for 1919–1935. The figures for 1936–1941 have been supplied by Fabricant in private correspondence.

(6) = Expenditures on farm service buildings (excluding dwellings) + expenditures on farm machinery + expenditures on farm trucks + 40 per cent of expenditures on farm automobiles other than trucks, *Income Parity for Agriculture,* Part II, Section 5, p. 25, col. 1, and *Net Farm Income and Parity Report,* U. S. Department of Agriculture, 1943, p. 27. The series on expenditures on farm service buildings has been adjusted from 1933–1941.

(7) = Price index of farm capital goods, 1934: 1.00. Weighted average of prices paid by farmers for building materials (*Agricultural Statistics,* 1943, p. 394), for farm machinery (*ibid.*), and for motor vehicles (*Income Parity for Agriculture,* U. S. Department of Agriculture, Part III, Section 4, p. 11, for 1919–1938 and extended for 1939–1941 by adopting the same percentage changes as those of the wholesale price index of motor vehicles of the U. S. Department of Commerce). The weights are, respectively, current-dollar expenditures for farm service buildings (excluding dwellings), for farm machinery, and for farm trucks plus 40 per cent of other farm motor vehicles. The sources are the same as those for item *(6)*. The weighted average was then converted to a 1934 base.

(8) = Depreciation charges on private producers' non-agricultural plant and equipment. The basic data for 1919–1928 come from S. Fabricant, *op. cit.,* pp. 260–261, and for 1929–1941 from J. Mosak, "Forecasting Postwar Demand: III," *Econometrica,* Vol. 13, 1945, pp. 45–53. In splicing these two series, industry by industry, certain adjustments had to be made to get comparable data for the residual group which excluded agriculture, public utilities, transportation, mining, and manufacturing. Depreciation on rented residences, in particular, was subtracted from Mosak's data; then the figures for 1919–1928 were estimated by constructing a series having the same year-to-year percentage changes as Fabricant's series but having Mosak's adjusted value for 1929. Mosak's adjusted figures were used for 1929–1941. Also depreciation on transportation in common carrier buses, motor-trucking, taxis, air and harbor craft was shifted from the transportation category to the miscellaneous category. This change introduced an error of rounding in adding up the components of total depreciation.

(9) = Price index underlying business depreciation charges, 1934 base, S. Fabricant, *op. cit.,* p. 183, converted to 1934 base year by use of Fabricant's general index of the prices of business capital goods, pp. 178–179, Table 32. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence. The ratio *(8)*/*(9)* gives a somewhat oversimplified formula for the calculation of depreciation on private producers' non-agricultural plant and equipment in 1934 dollars because the series actually used is the sum of constant(1934)-dollar series for each of three industry groups, which is not exactly the same thing as total depreciation divided by an aggregate price index. However, the two different methods of calculation lead to results that differ only slightly.

(10) = Depreciation charges on farm plant and equipment (exclusive of dwellings), calculated at replacement costs. *Net Farm Income and Parity Re-*

port, U. S. Department of Agriculture, 1943, p. 23, and *Income Parity for Agriculture*, U. S. Department of Agriculture, Part II, Section 5, p. 25.

q: Price index of private producers' plant and equipment, 1934: 1.00.

$$q = (5)\frac{(4)}{(4) + (6)} + (7)\frac{(6)}{(4) + (6)}$$

ΔH: Net change in inventories, measured in billions of 1934 dollars.

$$\Delta H = \frac{(11)}{(12)}$$

(11) = Net inventory change, Mary S. Painter, *loc. cit.*
(12) = Wholesale price index, 1934: 1.00, Supplement to the *Survey of Current Business*, 1942, p. 18, converted to 1934 base year.

D_1: Gross expenditures on construction of owner-occupied, single-family, non-farm residences, measured in billions of 1934 dollars.

$$D_1 = \frac{(13)(1.18)(1.126)(14)(0.63) + (15)(0.23)}{(16)}$$

(13) = Average permit valuation of single-family, non-farm residences, *The Construction Industry in the United States*, Bureau of Labor Statistics, Bulletin 786, Table 11, p. 21.
1.18 = Correction for undervaluation of building permits, taken from D. Wickens, *Residential Real Estate.*
1.126 = Ratio of the average rental value of owner-occupied, single-family, non-farm residences (constructed 1935–1940) to the average rental value of all single-family, non-farm residences (constructed 1935–1940), *Census of Housing*, 1940, Vol. III, Part I, Table A4.
(14) = Number of single-family, non-farm dwelling units constructed annually, Bureau of Labor Statistics, *op. cit.*, Table 18, p. 35.
0.63 = Fraction of single-family, non-farm dwelling units constructed 1935–1940, that were owner-occupied in 1940, *Census of Housing, loc. cit.*
(15) = Expenditures for alterations, additions, and repairs, Twentieth Century Fund, *American Housing*, Table 12, p. 367. These data are for 1919–1940. For 1941 it was assumed that maintenance was the same percentage of total housing expenditures as in 1940.
$0.23 = 0.32\,\dfrac{0.357}{0.524}$, where 0.32 = the ratio of owner-occupied single-family, non-farm dwelling units to total non-farm dwelling units in 1940.
0.357 = Fraction of repairs estimated by the U. S. Department of Commerce to be attributable to owner-occupied, non-farm residences (average of 1930 and 1940 figures).
0.524 = Ratio of owner-occupied non-farm units to total non-farm units, 1930.
(16) = American Appraisal Co. index of construction costs (national average), 1934: 1.00, Supplement to the *Survey of Current Business*, 1942, p. 25, converted to 1934 base year.

q_1: Index of construction costs, 1934: 1.00.

$$q_1 = (16)$$

D_2: Gross expenditures on construction of rented, non-farm residences, measured in billions of 1934 dollars.

$$D_2 = \frac{(17)}{(16)} - D_1$$

(17) = Gross expenditures on construction of private non-farm residences, Bureau of Labor Statistics, *op. cit.*, Table 1, p. 4.

D_3: Gross expenditures on construction of farm residences, measured in billions of 1934 dollars.

$$D_3 = \frac{(18)}{(19)}$$

(18) = Gross expenditures on construction of farm residences, Bureau of Labor Statistics, *op. cit.*, Table 1, p. 4.

(19) = Index of farm construction costs, 1934: 1.00, *Income Parity for Agriculture*, U. S. Department of Agriculture, Part II, Section 5, p. 27. These data are for 1919–1940. The 1941 figure was taken from a free-hand regression between (19) and the index of the cost of building materials taken from *Agricultural Prices*, February 29, 1944, col. 7, p. 25. All indexes are converted to 1934 base year.

D'': Depreciation of all residences (farm and non-farm), measured in billions of 1934 dollars.

$$D'' = \left[67.6 - \sum_{i=t+1}^{1933} (D_1 + D_2 + D_3)_i + \sum_{i=t+1}^{1933} D_i''\right]\frac{0.03}{0.97} - (D_1 + D_2 + D_3)_t\frac{0.015}{0.97} \qquad t < 1933$$

$$D'' = 67.6\frac{0.03}{0.97} - (D_1 + D_2 + D_3)_t\frac{0.015}{0.97} \qquad t = 1933$$

$$D'' = 67.6(0.03) + (D_1 + D_2 + D_3)_t(0.015) \qquad t = 1934$$

$$D'' = (67.6)(0.97)^{t-1934}(0.03) + \sum_{i=1934}^{t-1} (D_1 + D_2 + D_3)_i(0.985)(0.97)^{i-1-1934}(0.03) + (D_1 + D_2 + D_3)_t(0.015) \qquad t > 1934$$

67.6 = Estimated value, January 1, 1934, of the stock of residential dwellings in the United States. It is 80 per cent of Wickens' (*Residential Real Estate*, National Bureau of Economic Research, Table A8, p. 93) figure for the value of non-farm land and dwellings + 14.3 per cent of the value of total farm real estate, *Net Farm Income and Parity Report*, U. S. Department of Agriculture, 1943, Table 18, p. 29. The figures of 80 per cent and 14.3 per cent are taken as the proportions of non-farm residential and farm real estate, respectively, that are represented by dwellings alone.

G: Government expenditures on goods and services (exclusive of government interest payments) + net exports + net investment of non-profit institutions, all measured in billions of 1934 dollars.

$$G = \frac{(20) - (21) - (22)}{(23)} + \frac{(22)}{(16)} + \frac{(24)}{(12)} + \frac{(25) - 0.1}{(16)}$$

(20) = Government expenditures for goods and services, Mary S. Painter, *loc. cit.*

(21) = Government interest payments, S. Kuznets, *op. cit.*, Vol. II, p. 813. These data are for 1919–1928. J. Mosak, *op. cit.*, p. 51. These data are for 1929–1941.

(22) = Public construction expenditures, including work-relief construction, Bureau of Labor Statistics, *op. cit.*, Table 1, p. 4, and *Statistical Abstract of the United States*, 1942, p. 992.

(23) = Index of wholesale prices of non-farm products, 1934: 1.00, *Agricultural Statistics*, 1942, p. 649, converted to 1934 base year.

(24) = Net exports and monetary use of gold and silver, Mary S. Painter, *loc. cit.*

(25) = Gross construction expenditures by non-profit institutions, G. Terborgh, *loc. cit.*, and F. Dirks, *loc. cit.*

0.1 = Annual depreciation charges attributed to the plant of non-profit institutions. This figure is about 3 per cent of the 1934, end-of-year stock of capital owned by these institutions.

$Y + T$: Net national product, measured in billions of 1934 dollars.

$$Y + T = C + I + \Delta H + D_1 + D_2 + D_3 - D'' + G$$

Y: Disposable income, measured in billions of 1934 dollars.

$$Y = \frac{\left\{\begin{array}{l}(1) + (2) + (4) + (6) - \dfrac{(8)}{(9)}(5) - (10) + (11) + (17) + (18) - (16)D'' \\ + (20) + (24) + (25) - 0.1 - (26) - (27) - (28) + (29)\end{array}\right\}}{(3)}$$

(26) = Federal government receipts, *Annual Report of the Secretary of the Treasury.*

(27) = State and local government receipts, National Industrial Conference Board, *Economic Almanac*, 1944–1945, p. 102. These data are given for fiscal years and are converted to a calendar-year basis by taking a two-year moving average of a three-year moving average.

(28) = Net corporate savings, M. Hoffenberg, "Estimates of National Output, Distributed Income, Consumer Spending, Saving and Capital Formation," *Review of Economic Statistics*, Vol. 25, May, 1943, p. 156 (Department of Commerce estimates).

(29) = Transfer payments, *Survey of Current Business*, Vol. 22, May, 1942, p. 12, and Vol. 23, August, 1943, p. 13. These data are for 1929–1941. The data for 1919–1928 are set at a constant rate of \$0.7 billion per annum. This constant figure is justified on the grounds that veterans' payments by the federal government were stable at \$0.4 billion for the entire period of the 1920's and that state relief expenditures were also stable

at \$0.1 billion for the same period. The remaining transfer payments, those of local governments, were assumed to be stable also. Hence the figure \$0.7 billion, observed for 1929 and 1930, was extended to the previous years.

p: Price index of output as a whole, 1934: 1.00.

$$p = \frac{\left\{\begin{array}{l}(1) + (2) + (4) + (6) - \frac{(8)}{(9)}(5) - (10) + (11) + (17) + (18) \\ - (16)D'' + (20) - (21) + (24) + (25) - 0.1\end{array}\right\}}{Y + T}$$

W_1: Private wage-salary bill, measured in billions of current dollars.

$$\mathrm{W}_1 = (30) - (31)$$

(30) = Total employee compensation, M. Hoffenberg, *loc. cit.* (Department of Commerce estimates).

(31) = Government wages and salaries, including work relief wages, *Survey of Current Business*, Vol. 23, March, 1943, p. 22. These data are for 1929–1941. The figures for 1919–1928 are estimated from a free-hand regression between the U. S. Department of Commerce estimates (1929–1938) and Kuznets' estimates (1929–1938). See S. Kuznets, *op. cit.*, Vol. II, p. 811. Kuznets' figures are known for 1919–1928; hence the corresponding Department of Commerce figures can be estimated from the regression.

W_2: Government wage-salary bill, measured in billions of current dollars

$$\mathrm{W}_2 = (31)$$

R_1: Non-farm rentals, paid and imputed, measured in billions of current dollars.

$$\mathrm{R}_1 = (32)$$

(32) = Gross non-farm rentals, paid and imputed, for residential dwellings, special release, U. S. Department of Commerce, 1929–1941. Estimated 1921–1928 from free-hand regression between Department of Commerce gross rentals (1929–1938) and Barger's data on gross rentals (1929–1938). See H. Barger, *Outlay and Income in the United States, 1921–1938*, pp. 226–227. Barger's figures are known for 1921–1928; hence the corresponding Department of Commerce figures can be estimated from the regression. Estimated 1919–1920 from $\mathrm{R}_1 = 0.278r\left(\frac{vN^s}{100} + \frac{v_{-1}N_{-1}^s}{100}\right)\frac{1}{2}$.

See below for definititons of v and N^s.

R_2: Farm rentals, paid and imputed, measured in billions of current dollars.

$$\mathrm{R}_2 = (33)$$

(33) = Gross rentals, paid and imputed, for farm dwellings, *Agricultural Statistics*, 1942, p. 660.

r: Index of rents, 1934: 1.00.

$$r = (34)$$

(*34*) = Rent component of Bureau of Labor Statistics cost-of-living index, 1934: 1.00, converted to 1934 base year, *Statistical Abstract of the United States*, 1943, pp. 404–405.

ΔF: Thousands of new non-farm families.

$$\Delta F = (35)$$

(*35*) = Annual number of families added in non-farm areas, Twentieth Century Fund, *op. cit.*, Table 40, p. 418. These data are for 1919–1939. For 1940 and 1941, the data are the first differences in the beginning-of-the-year number of families given in the *Statistical Abstract of the United States*, 1943, Table 40, p. 46, multiplied by 0.795, the ratio of non-farm families to all families in 1940.

v: Percentage of non-farm housing units occupied at the end of the year.

$$v = (36) \qquad 1919\text{–}1927$$

$$v = (37) \qquad 1928\text{–}1941$$

(*36*) = Ratio of non-farm families to available non-farm dwelling units, L. Chawner, *Residential Building*, Housing Monograph Series, No. 1, Table VI, p. 16.

(*37*) = Percentage of urban dwelling units occupied, K. Johnson, "Residential Vacancies in Wartime United States," *Survey of Current Business*, Vol. 22, December, 1942, Table 1, p. 19.

N^s: Millions of available non-farm dwelling units at the end of the year.

$$N^s = 24.6 - \sum_{i=t+1}^{1928} \Delta(38)_i \qquad t = 1919\text{–}1927$$

$$N^s = (39) \qquad 1928\text{–}1941$$

(*38*) = Available non-farm dwelling units, at the end of the year, L. Chawner, *op. cit.*, Table VI, p. 16.

24.6 = Available non-farm dwelling units on January 1, 1929, "Dwelling Units in the United States, 1929–1942," special release of the U. S. Department of Commerce.

(*39*) = Available non-farm dwelling units, at the end of the year, "Dwelling Units in the United States, 1929–1942."

i: Average corporate bond yield.

$$i = (40)$$

(*40*) = Average corporate bond yield, Board of Governors of the Federal Reserve System, *Banking and Monetary Statistics*, p. 468.

M_1^D: Demand deposits adjusted + currency outside banks, averaged during the year, measured in billions of current dollars.

$$M_1^D = (41)$$

(*41*) = Demand deposits adjusted + currency outside banks, average of end of previous year, middle of current year, and end of current year figures are used, Board of Governors of the Federal Reserve System, *op. cit.*, p. 34.

M_2^D: Time deposits, averaged during the year, measured in billions of current dollars.

$$M_2^D = (42)$$

(*42*) = Time deposits, average of end of previous year, middle of current year, and end of current year figures (except 1919–1923, for which middle of current year figures are used), Board of Governors of the Federal Reserve System, *op. cit.*, p. 34.

E_R: Excess reserves, averaged during the year, measured in millions of current dollars.

$$E_R = (43) \qquad 1919\text{–}1928$$

$$E_R = (44) \qquad 1929\text{–}1941$$

(*43*) = Member bank, excess reserve balances, average of last call date of previous year and call dates of current year, Board of Governors of the Federal Reserve System, *op. cit.*, p. 395.

(*44*) = Member bank, excess reserve balances, annual averages of daily figures, Board of Governors of the Federal Reserve System, *op. cit.*, p. 368.

K: End-of-year stock of private producers' plant and equipment, measured in billions of 1934 dollars.

$$K = 107.8 - \sum_{i=t+1}^{1934} I_i \qquad t < 1934$$

$$K = 107.8 \qquad 1934$$

$$K = 107.8 + \sum_{i=1935}^{t} I_i \qquad t > 1934$$

107.8 = End-of-1934 stock of private producers' plant and equipment. Net capital assets, excluding land, for corporations, end-of-1934, are taken from S. Fabricant, *op. cit.*, p. 271. These data are converted to estimates for both corporate and non-corporate enterprises by dividing the figure for each industry by the ratio of corporate output to total output. These ratios are

Mining and Quarrying	0.96
Manufacturing	0.92
Construction	0.60
Transportation and Public Utilities	1.00
Trade	0.63
Service	0.30
Finance and Real Estate	0.84

These ratios are found in *The Structure of the American Economy*, National Resources Planning Board, pp. 375–376. The capital for the service industries is corrected by subtracting the capital of non-profit institutions. Agricultural capital, exclusive of livestock, land, and dwellings, is taken from *The Structure of the American Economy*, p. 377.

H: End-of-year stock of inventories, measured in billions of 1934 dollars.

$$H = 21.8 - \sum_{i=t+1}^{1934} (\Delta H)_i \qquad t < 1934$$

$$H = 21.8 \qquad 1934$$

$$H = 21.8 + \sum_{i=1935}^{t} (\Delta H)_i \qquad t > 1934$$

21.8 = End-of-1934 stock of inventories.

Business inventories, end of 1934	$17,913 million
Less: Inventories of agricultural corporations	196
	17,717
Plus: Total agricultural inventories	4,087
	$21,804 million

Business inventories and inventories of agricultural corporations come from *Survey of Current Business*, Vol. 22, September, 1942, p. 18. Agricultural inventories come from *Agricultural Statistics*, 1942 (value of crops for sale, plus value of livestock).

E: Excise taxes, measured in billions of current dollars.

$$E = (45)$$

(45) = State sales taxes + state insurance taxes + state public utility taxes + federal taxes on alcoholic beverages, tobacco, and motor fuel + other manufacturers' excises. *Federal, State, and Local Government Fiscal Relations*, Senate Document No. 69, 78th Congress, first session, 1943, pp. 320, 337, 340–343. Fiscal-year data are converted to calendar-year data by taking a two-year moving average of a three-year moving average.

MISCELLANEOUS TIME SERIES

	I_R	$p_R X_R$	q_R	K_R	L_R	I_M	$p_M X_M$	q_M	K_M
1920	0.202	2.173	1.445	19.927	−0.199	0.160	21.691	1.493	26.638
1921	0.145	1.711	1.148	20.072	−0.132	−0.257	12.196	1.174	26.381
1922	0.186	1.714	1.022	20.258	0.027	−0.247	16.000	1.067	26.134
1923	0.681	1.990	1.150	20.939	−0.133	−0.202	20.237	1.187	25.932
1924	0.536	1.913	1.114	21.475	0.086	−0.278	18.407	1.163	25.654
1925	0.381	2.012	1.086	21.856	0.071	0.119	20.076	1.141	25.773
1926	0.515	2.093	1.089	22.371	−0.037	0.350	20.870	1.140	26.123
1927	0.360	2.002	1.114	22.731	−0.117	0.098	19.974	1.138	26.221
1928	0.310	2.010	1.086	23.041	0.167	0.245	21.014	1.124	26.466
1929	0.439	2.086	1.119	23.480	0.047	0.695	22.816	1.147	27.161
1930	0.484	1.736	1.073	23.964	−0.082	−0.104	16.832	1.097	27.057
1931	0.051	1.334	1.006	24.015	−0.308	−0.719	11.355	1.017	26.338
1932	−0.025	0.935	0.928	23.990	−0.384	−0.953	6.756	0.942	25.385
1933	−0.157	0.951	0.921	23.833	−0.518	−0.875	9.015	0.923	24.510
1934	−0.040	1.000	1.000	23.793	−0.711	−0.410	11.793	1.000	24.100
1935	−0.105	1.080	1.030	23.688	−1.060	−0.088	13.815	1.014	24.012
1936	0.022	1.271	1.041	23.710	−1.007	0.349	17.009	1.036	24.361
1937	0.187	1.339	1.126	23.897	−1.294	0.591	19.745	1.139	24.952
1938	−0.054	1.102	1.120	23.843	−1.573	−0.268	14.788	1.139	24.684
1939	−0.017	1.269	1.133	23.826	−1.747	0.016	18.313	1.140	24.700
1940	0.111	1.388	1.172	23.937	0.173	0.604	21.795	1.170	25.304
1941	0.148	1.722	1.277	24.085	0.335	−0.058	30.443	1.234	25.246

I_R: Net investment in plant and equipment of railroads, measured in billions of 1934 dollars.

q_R: Index of prices paid for plant and equipment by railroads, 1934: 1.00.

$$I_R = \frac{(1)}{q_R} - \frac{(2)}{(3)}\frac{(4)}{(5)}$$

$$q_R = \tfrac{1}{11}[(6) + (7) + 2(8) + 5(9) + 2(10)] \qquad 1920\text{–}1940$$

$$q_R = \tfrac{1}{9}[(6) + (7) + 2(8) + 5(9)] \qquad 1941$$

(1) = Expenditures for plant and equipment by railroads, G. Terborgh, *Federal Reserve Bulletin*, Vol. 25, September, 1939, p. 732, for 1919–1936, Vol. 26, February, 1940, p. 116, for 1937–1938; F. Dirks, *Federal Reserve Bulletin*, Vol. 28, April, 1942, p. 318, for 1939–1941.

MISCELLANEOUS TIME SERIES (*Continued*)

	i_M	i_M^*	H	$X - \Delta H$	p	D''	N_E	$N_L - N_E$	w
1920	0.859	−0.145				4.8	37.4	0.558	1468.38
1921	0.812	0.111				4.4	33.7	4.754	1335.03
(1)			20.817	10.780	1.398				
(2)			21.376	11.415	1.255				
(3)			21.658	11.550	1.219				
(4)			20.737	11.110	1.223				
1922	0.839	0.093				5.0	35.4	2.917	1341.06
(1)			21.501	11.561	1.206				
(2)			21.241	11.928	1.243				
(3)			20.975	11.772	1.294				
(4)			21.275	12.384	1.309				
1923	0.877	−0.086				4.9	38.1	0.749	1431.20
(1)			21.690	12.030	1.349				
(2)			22.432	12.779	1.332				
(3)			23.297	13.176	1.287				
(4)			23.329	13.754	1.287				
1924	0.936	0.041				5.1	37.8	2.034	1438.52
(1)			23.329	13.994	1.296				
(2)			23.945	13.914	1.253				
(3)			23.896	13.870	1.261				
(4)			23.361	14.088	1.300				
1925	1.009	−0.100				5.0	38.5	0.817	1465.03
(1)			23.634	13.848	1.354				
(2)			23.671	13.980	1.334				
(3)			23.721	13.790	1.356				
(4)			23.921	14.554	1.355				
1926	1.227	−0.109				5.8	39.8	0.464	1492.08
(1)			24.493	14.677	1.330				
(2)			24.573	15.007	1.311				
(3)			24.548	15.088	1.298				
(4)			24.816	15.403	1.287				
1927	1.267	0.005				5.6	39.9	1.620	1497.30
(1)			24.613	15.856	1.249				
(2)			24.562	15.941	1.228				
(3)			25.125	15.258	1.244				
(4)			24.951	15.400	1.258				
1928	1.531	0.160				6.0	40.1	1.857	1540.90
(1)			25.330	16.046	1.252				
(2)			24.684	15.785	1.265				
(3)			24.640	15.877	1.278				
(4)			25.052	16.532	1.255				

MISCELLANEOUS TIME SERIES (*Continued*)

	ι_M	ι_M^*	H	$X - \Delta H$	p	D''	N_E	$N_L - N_E$	w
1929	1.288	−0.054				6.2	41.8	0.429	1569.47
(1)			25.066	16.685	1.251				
(2)			25.262	16.641	1.241				
(3)			25.708	16.670	1.257				
(4)			26.264	16.274	1.227				
1930	1.352	0.103				6.2	39.6	2.896	1518.97
(1)			26.893	16.264	1.193				
(2)			27.596	16.294	1.155				
(3)			28.245	16.124	1.102				
(4)			27.910	15.458	1.061				
1931	1.371	0.423				5.9	36.1	7.037	1419.38
(1)			27.406	16.320	1.005				
(2)			27.404	16.257	0.958				
(3)			27.487	15.674	0.937				
(4)			26.917	15.314	0.910				
1932	1.280	0.619				5.3	32.8	11.385	1255.25
(1)			26.227	15.074	0.868				
(2)			25.940	14.574	0.843				
(3)			24.534	13.570	0.849				
(4)			23.775	12.912	0.830				
1933	1.254	0.592				5.2	32.8	11.842	1184.42
(1)			23.532	13.356	0.787				
(2)			23.301	13.943	0.818				
(3)			23.188	13.298	0.910				
(4)			22.445	12.586	0.927				
1934	1.214	0.389				5.1	35.2	9.761	1142.88
(1)			22.316	14.306	0.956				
(2)			22.616	13.736	0.965				
(3)			22.196	14.212	0.996				
(4)			21.804	13.982	1.000				
1935	1.256	0.260				5.0	36.4	9.092	1183.53
(1)			21.806	13.188	1.034				
(2)			21.671	13.087	1.044				
(3)			21.685	14.064	1.047				
(4)			22.011	14.403	1.053				
1936	1.182	0.034				5.1	38.2	7.386	1225.77
(1)			22.144	14.469	1.048				
(2)			22.524	17.137	1.034				
(3)			22.932	15.643	1.060				
(4)			23.800	16.326	1.080				

MISCELLANEOUS TIME SERIES (*Continued*)

	L_M	L_M^*	H	$X - \Delta H$	p	D''	N_E	$N_L - N_E$	w
1937	1.022	−0.326				5.4	40.0	6.403	1331.68
(1)			24.070	14.939	1.132				
(2)			24.759	15.589	1.142				
(3)			26.059	15.578	1.144				
(4)			25.898	15.908	1.090				
1938	1.335	0.105				5.4	37.5	9.796	1292.93
(1)			25.554	14.557	1.046				
(2)			25.033	14.335	1.023				
(3)			24.491	15.550	1.023				
(4)			24.315	17.204	1.011				
1939	1.534	0.289				5.4	39.0	8.786	1338.42
1940	1.851	0.414				5.5	40.0	7.607	1398.38
1941	2.308	0.303				5.9	42.7	3.276	1541.07

(*2*) = Depreciation charges of class I railways, *Statistics of Railways of the United States*, Interstate Commerce Commission, 1920–1941.

(*3*) = Price index underlying business depreciation charges, 1934 base, S. Fabricant, *Capital Consumption and Adjustment*, p. 183, converted to 1934 base year. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence.

(*4*) = Mileage owned by all classes of railways, *Statistics of Railways of the United States*, Interstate Commerce Commission, 1920–1941.

(*5*) = Mileage owned by class I railways, *Statistics of Railways of the United States*, Interstate Commerce Commission, 1920–1941.

(*6*) = Index of railroad construction costs, roadway machines, S. Fabricant, *op. cit.*, pp. 178–179. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence.

(*7*) = Index of railroad construction costs, roadway small tools, S. Fabricant, *op. cit.*, pp. 178–179. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence.

(*8*) = Index of railroad construction costs, shop machinery, road, S. Fabricant, *op. cit.*, pp. 178–179. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence.

(*9*) = Index of railroad construction costs, equipment, S. Fabricant, *op. cit.*, pp. 178–179. These data are for 1919–1935. The figures for 1936–1941 have been supplied by S. Fabricant in private correspondence.

(*10*) = Index of railroad construction costs, road, S. Fabricant, *op. cit.*, pp. 178–179. These data are for 1919–1935. The figures for 1936–1940 have been supplied by S. Fabricant in private correspondence.

The weights used in the construction of q_R from component indexes are taken from S. Fabricant, *op. cit.*, p. 179.

$p_R X_R$: Index of income originating in railroads, 1934: 1.00.

$$p_R X_R = \frac{(11)}{2{,}230}$$

(11) = Net income originating in steam railroads, Pullman, and Railway Express, S. Kuznets, *National Income and Its Composition*, 1919–1938, Vol. II, p. 660. These data are used for 1919–1928. *Statistical Abstract of the United States*, 1944–45, p. 404. These data are used for 1929–1941.

2230 = Millions of dollars of net income originating in steam railroads, Pullman, and Railway Express, in the year 1934.

K_R: End-of-year stock of fixed capital in railroads, measured in billions of 1934 dollars.

$$K_R = 23.793 - \sum_{i=t+1}^{1934} (I_R)_i \qquad t < 1934$$

$$K_R = 23.793 \qquad t = 1934$$

$$K_R = 23.793 + \sum_{i=1935}^{t} (I_R)_i \qquad t > 1934$$

23.793 = End-of-1934 stock of fixed capital (road, equipment, miscellaneous physical property) in class I railroads, multiplied by ratio of 1934 mileage owned by all classes to 1934 mileage owned by class I railroads.

December 31, 1934 (Class I Railways)

Investment in road equipment:	
Road	13,785,831,035
Equipment	5,383,001,838
Improvement on leased railway property:	
Road	543,582,737
Equipment	12,981,382
Miscellaneous physical property:	173,365,306
Total	19,898,762,298
Accrued depreciation:	
Road	97,885,613
Equipment	2,362,003,198
Miscellaneous physical property:	7,573,099
Total	2,467,461,910
Net physical assets	17,431,300,388

Statistics of Railways in the United States, Interstate Commerce Commission, 1934, pp. S-88 and S-89. Ratio of 1934 mileage owned by all classes of railways to 1934 mileage owned by class I railways = 1.365.

17,431,300,388(1.365) = 23,793,725,030

L_R: End-of-year current assets − current liabilities of railroads, measured in billions of current dollars.

$$L_R = (12) - (13)$$

(12) = End-of-year cash + all other current assets less materials and supplies of class I railways, *Statistics of Railways in the United States*, Interstate Commerce Commission, 1920–1941.

(13) = End-of-year current liabilities of class I railways, *Statistics of Railways in the United States*, Interstate Commerce Commission, 1920–1941. (For 1935–1940, working capital of lessor roads is included.)

I_M: Net investment in plant and equipment by mining and manufacturing industries, measured in billions of 1934 dollars.

$$I_M = \frac{(14)}{(15)} - \frac{(16)}{(3)}$$

(14) = Expenditures on private producers' plant and equipment in manufacturing and mining industries, G. Terborgh, *Federal Reserve Bulletin*, Vol. 25, September, 1939, p. 732, for 1919–1936, Vol. 26, February, 1940, p. 116, for 1937–1938; F. Dirks, *Federal Reserve Bulletin*, Vol. 28, April, 1942, p. 318, for 1939–1941.

(15) = Price index of business capital goods, 1934: 1.00. S. Fabricant, *op. cit.*, pp. 178–179, converted to 1934 base year. These data are for 1919–1935. The figures for 1936–1941 have been supplied by Fabricant in private correspondence.

(16) = Depreciation charges on plant and equipment in manufacturing and mining industries, S. Fabricant, *op. cit.*, pp. 260–261. These data are used for 1919–1928. J. Mosak, "Forecasting Postwar Demand: III," *Econometrica*, Vol. 13, 1945, p. 47. These data are for 1929–1941.

$p_M X_M$: Net income originating in manufacturing and mining industries, measured in billions of current dollars.

$$p_M X_M = (17) + (18)$$

(17) = Net income originating in manufacturing, unpublished tables, U. S. Department of Commerce.

(18) = Net income originating in mining, unpublished tables, U. S. Department of Commerce.

q_M: Price index of manufacturing capital goods, 1934: 1.00.

$$q_M = (15)$$

K_M: End-of-year stock of fixed capital in manufacturing and mining industries, measured in billions of 1934 dollars.

$$K_M = 24.100 - \sum_{i=t+1}^{1934} (I_M)_i \qquad t < 1934$$

$$K_M = 24.100 \qquad t = 1934$$

$$K_M = 24.100 + \sum_{i=1935}^{t} (I_M)_i \qquad t > 1934$$

24.1 = End-of-1934 stock of fixed capital in manufacturing and mining industries. It is the value of net capital assets excluding land in corporations of each industry, divided by the ratio of corporate to total output in each industry. The data on capital assets came from S. Fabricant, *op. cit.*, p. 271, and the expansion ratios came from the *Structure of the American Economy*, National Resources Planning Board, p. 375.

$\mathfrak{l}_M$: End-of-year cash balances and marketable securities of a sample of large manufacturing corporations, measured in billions of current dollars.

$$\mathfrak{l}_M = (19)$$

(*19*) = End-of-year cash balances and marketable securities of a sample of large manufacturing corporations, F. Lutz, *Corporate Cash Balances, 1914–43*, National Bureau of Economic Research, p. 114.

$\mathfrak{l}_M{}^*$: End-of-year "free" liquid funds of a sample of large manufacturing corporations, measured in billions of current dollars.

$$\mathfrak{l}_M{}^* = (20)$$

(*20*) = End-of-year "free" liquid funds of a sample of large manufacturing corporations, F. Lutz, *op. cit.*, p. 116.

H: End-of-quarter inventories, measured in billions of fourth-quarter-1934 dollars.

$$H = 21.804 - \sum_{i=t+\frac{1}{4}}^{1934\frac{4}{4}} (\Delta H)_i \qquad t < 1934\tfrac{4}{4}$$

$$H = 21.804 \qquad t = 1934\tfrac{4}{4}$$

$$H = 21.804 + \sum_{i=1935\frac{1}{4}}^{t} (\Delta H)_i \qquad t > 1934\tfrac{4}{4}$$

21.804 = End-of-1934 inventories. For computation of this value see the appendix section for Model III.

$$\Delta H = \frac{(21)}{(22)}$$

(*21*) = Quarterly inventory change, seasonally adjusted, H. Barger, *Outlay and Income in the United States, 1921–1938*, National Bureau of Economic Research, pp. 114–119.

(*22*) = Quarterly index of wholesale prices, fourth-quarter, 1934: 1.00. Quarterly averages of monthly figures converted to fourth-quarter-1934 base, *Survey of Current Business.*

X: Total quarterly outlay, measured in billions of fourth-quarter-1934 dollars.

$$X = \frac{(23)}{(22)}$$

(*23*) = Total quarterly outlay, seasonally adjusted, H. Barger, *loc. cit.*

D'': Depreciation of private producers' plant and equipment, measured in billions of 1934 dollars.

$$D'' = \frac{(24)}{(3)} + \frac{(25)}{(26)}$$

(*24*) = Depreciation charges on private producers' non-agricultural plant and equipment. The basic data for 1919–1928 come from S. Fabricant, *op. cit.*, pp. 260–261, and for 1929–1941 from J. Mosak, "Forecasting Post-war Demand: III," *Econometrica*, Vol. 13, 1945, pp. 45–53. In splicing these two series, industry by industry, certain adjustments had to be made to get comparable data for the residual group which excluded agriculture, public utilities, transportation, mining, and manufacturing. Depreciation on rented residences, in particular, was subtracted from Mosak's data; then the figures for 1919–1928 were estimated by constructing a series having the same year-to-year percentage changes as Fabricant's series but having Mosak's adjusted value for 1929. Mosak's adjusted figures were used for 1929–1941. Also depreciation on transportation in common carrier buses, motor-trucking, taxis, air and harbor craft was shifted from the transportation category to the miscellaneous category. This change introduced an error of rounding in adding up the components of total depreciation.

(*25*) = Depreciation charges on farm plant and equipment (exclusive of dwellings) calculated at replacement costs, *Net Farm Income and Parity Report*, U. S. Department of Agriculture, 1943, p. 23, and *Income Parity for Agriculture*, U. S. Department of Agriculture, Part II, Section 5, p. 25.

(*26*) = Price index of farm capital goods, 1934: 1.00. Weighted average of prices paid by farmers for building materials (*Agricultural Statistics*, 1943, p. 394), for farm machinery (*ibid.*), and for motor vehicles (*Income Parity for Agriculture*, U. S. Department of Agriculture, Part III, section 4, p. 11, for 1919–1938 and extended for 1939–1941 by adopting the same percentage changes as those of the wholesale price index of motor vehicles of the U. S. Department of Commerce). The weights are, respectively, current-dollar expenditures for farm service buildings (excluding dwellings), for farm machinery, and for farm trucks plus 40 per cent of other farm motor vehicles. The sources are the same as those for item (*6*) of the time series for Model I. The weighted average was then converted to a 1934 base.

N_E: Employment, measured in millions of persons.

$$N_E = (27) - (28) + (29)$$

(*27*) = Total employees, S. Kuznets, *op. cit.*, Vol. I, p. 314.

(*28*) = Government employees, S. Kuznets, *op. cit.*, Vol. I, p. 314.

(*29*) = Total entrepreneurs, S. Kuznets, *op. cit.*, Vol. I, p. 316. All these series are for 1919–1938. The three observations for 1939, 1940, 1941 are obtained from a free-hand regression (1929–1938) between the series from Kuznets' data and a series represented by total agricultural employment (U. S. Department of Agriculture) + total non-agricultural employment (U. S. Department of Labor) − total government employment (U. S. Department of Labor).

$N_L - N_E$: Unemployment, measured in millions of persons.

$$N_L - N_E = (30)$$

(*30*) = Unemployment, *The Economic Almanac*, National Industrial Conference Board, 1944–45, pp. 48–49.

w: Annual average wage rate, measured in current dollars per employee.

$$w = \frac{(31)}{(32) + (33) + (34) + (35) + (36)}$$

(*31*) = Total employee compensation, M. Hoffenberg, "Estimates of National Output, Distributed Income, Consumer Spending, Saving, and Capital Formation," *Review of Economic Statistics*, Vol. XXV, 1943, p. 156 (Department of Commerce estimates).

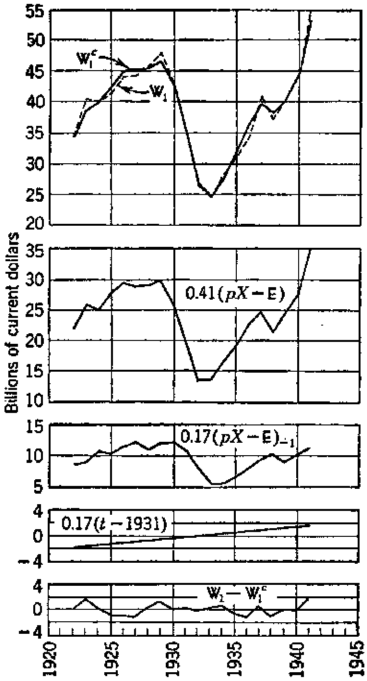
(*32*) = Ratio of total cash wage bill on farms to average annual farm wage rate (without board), *Farm Wage Rates, Farm Employment and Related Data*, U. S. Department of Agriculture, January, 1943, pp. 3 and 174.

(*33*) = Total full-time equivalent employees in manufacturing and mining, *Survey of Current Business*, June, 1945, p. 20. These data are for 1929–1941. They are extrapolated back to 1919 from a free-hand regression (1929–1938) with Kuznets' estimates of the same series. See S. Kuznets, *op. cit.*, Vol. I, pp. 314–315.

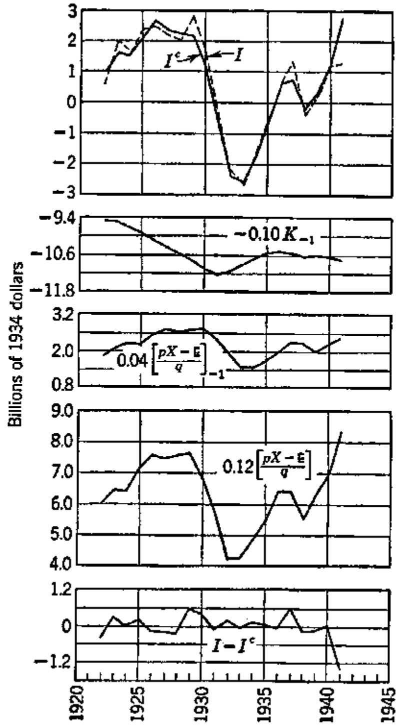
(*34*) = Total full-time equivalent employees in transportation, communication, and public utilities, *Survey of Current Business*, June, 1945, p. 20. These data are for 1929–1941. They are extrapolated back to 1919 from a free-hand regression (1929–1938) with Kuznets' estimates of the same series. See S. Kuznets, *op. cit.*, Vol. I, pp. 314–315.

(*35*) = Total full-time equivalent employees in construction, trade, finance, forestry, fisheries, insurance, real estate, service, and miscellaneous industries, *Survey of Current Business*, June, 1945, p. 20. These data are for 1929–1941. They are extrapolated back to 1919 from free-hand regression (1929–1938) with Kuznets' estimates of the same series. See S. Kuznets, *op. cit.*, Vol. I, pp. 314–315.

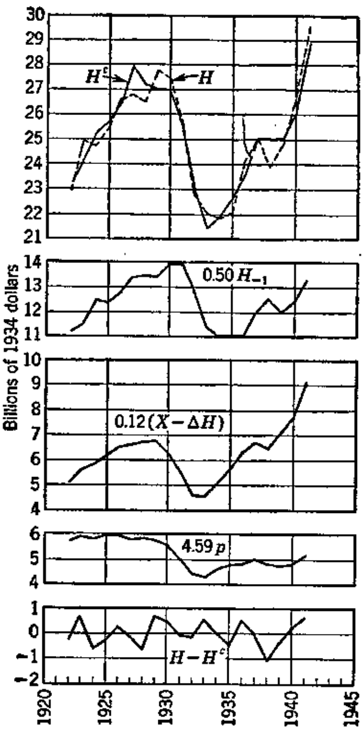
(*36*) = Total government employees, including emergency government employees. Regular government employees are taken from S. Kuznets, *op. cit.*, Vol. I, pp. 314–315. Emergency employees are taken from *Economic Almanac* (1944–1945), p. 49. Kuznets' data for the last three years, 1939, 1940, 1941, are extrapolated from a free-hand regression between his data and estimates of the same series in the *Economic Almanac* (1932–1938).



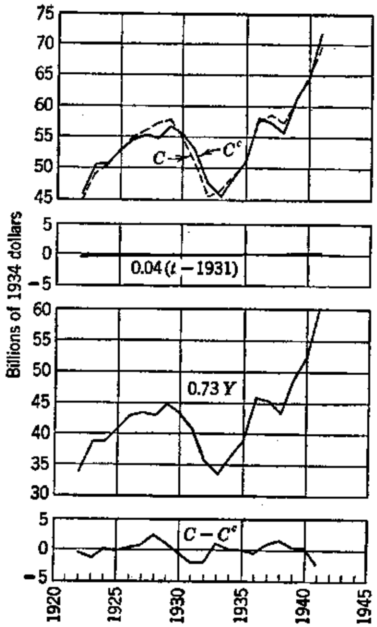
EQUATION 3.3.38



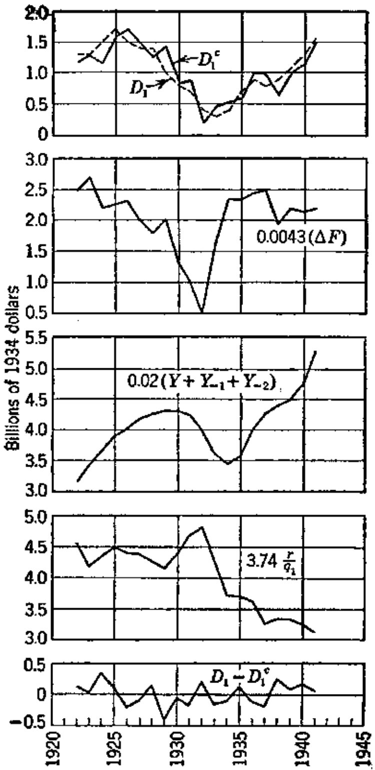
EQUATION 3.3.39



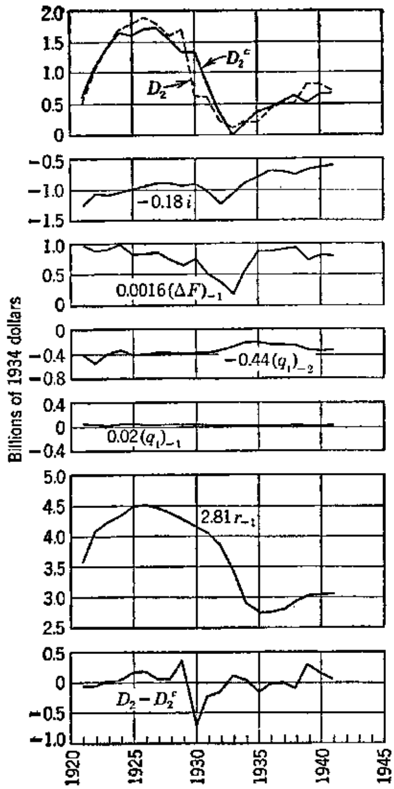
EQUATION 3.3.40



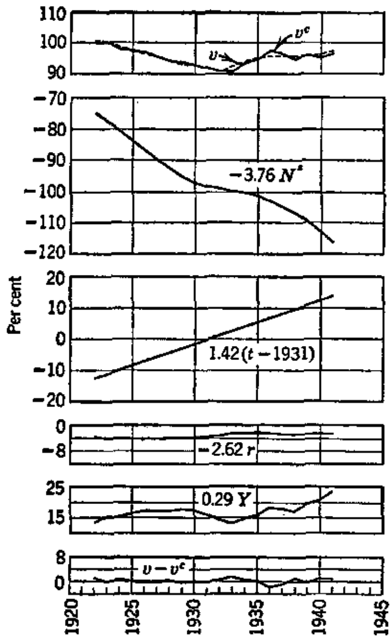
EQUATION 3.3.41



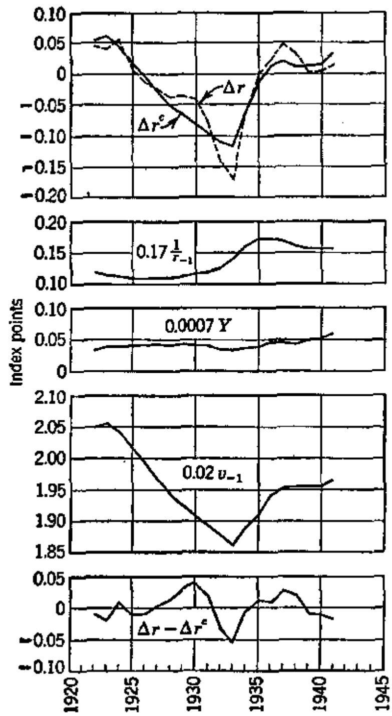
EQUATION 3.3.42



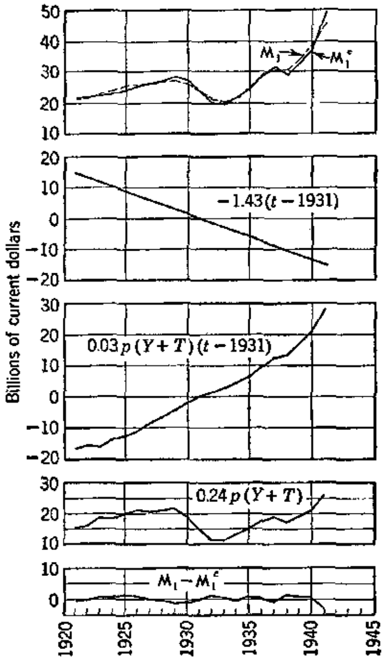
EQUATION 3.3.43



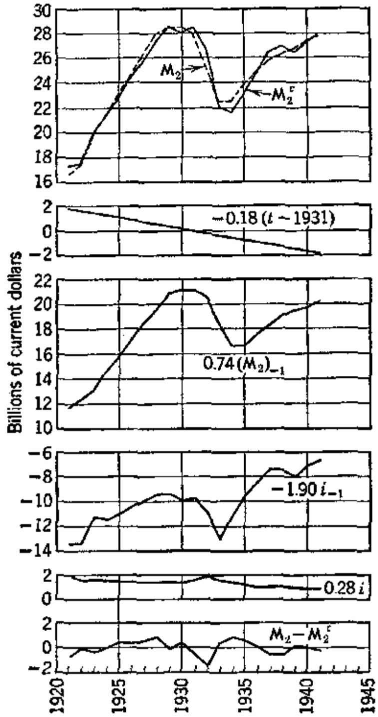
EQUATION 3.3.44



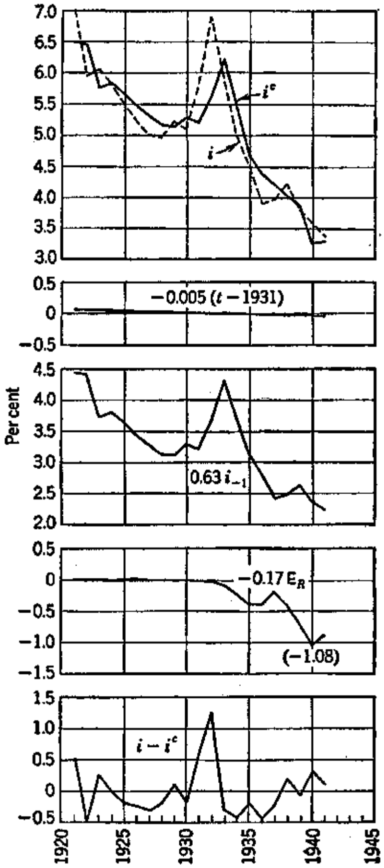
EQUATION 3.3.45



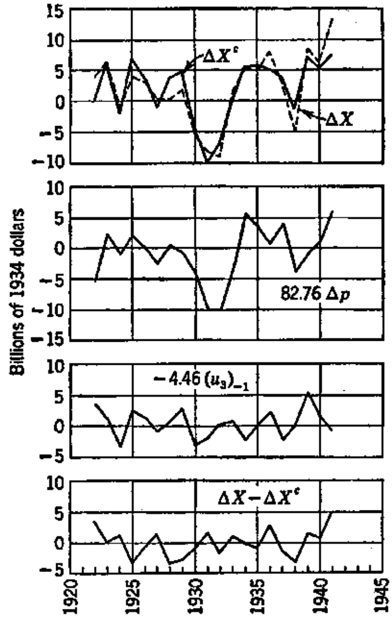
EQUATION 3.3.46



EQUATION 3.3.47



EQUATION 3.3.48



EQUATION 3.3.49

INDEX