Journal of Economic Literature
Vol. XXXII (March 1994), pp. 30–59

The Cowles Commission's Contributions to Econometrics at Chicago, 1939–1955

By CARL F. CHRIST
The Johns Hopkins University

I wish to acknowledge helpful suggestions from Stephen Blough, Arnold Zellner, and three referees. Philip Gerson assisted in translating the paper from my word processing program to the Journal's. *Remaining inadequacies are my own.*

1. *Introduction*

THE COWLES COMMISSION for Research in Economics created a revolution in econometric methods and practice during its years at the University of Chicago from 1939 to 1955. This article describes that revolution, together with some of its antecedents, and attempts to put it in perspective today, using simple examples and no matrix algebra.[1]

a. *A Brief Sketch of the Origin of the Cowles Commission*

Alfred Cowles was an investment counselor in Colorado Springs, Colorado. He told me in 1952 that after the stock market crash of 1929, he realized that he did not understand the workings of the economy, and so in 1931 he stopped publishing his market advisory letter, and began research on stock market forecasting.[2] Harold T. Davis, a mathematician who spent summers in Colorado Springs, put him in touch with Irving Fisher at Yale, president of the fledgling Econometric Society. Fisher had known Cowles' father and uncle when all three were undergraduates at Yale. Davis and Fisher proposed an economics research organization and a journal. The idea appealed to Cowles, and he agreed to provide financing. The Cowles Commission was founded in Colorado Springs in 1932, and *Econometrica* began publishing in 1933. The Commission's Articles of In-

---

[1] Two earlier retrospective works on the Cowles Commission are Christ (1952), a nonmathematical history of its first twenty years, and Clifford Hildreth (1986), a detailed and technical account of its work at the University of Chicago from 1939 to 1955. Theodore W. Anderson (1991) and James Heckman (1992) discuss Trygve Haavelmo's contributions. Christ (1985, Section VII), deals with the identification problem. Roy Epstein (1987) and Mary Morgan (1990) discuss the history of econometrics in general, including the Cowles Commission's contributions.

[2] Cowles and his associates generated random stock market advice, and random portfolio choices, and compared them with actual market newsletters and actual fire insurance company portfolios. They found that on the average the actual advice and portfolios were slightly inferior to the random ones. The best actual advice and portfolios were about as good as the best random ones, but the worst actual ones were worse than the worst random ones (Cowles 1933).

corporation give its purpose as "to advance the scientific study and development . . . of economic theory in its relation to mathematics and statistics." It moved, with Cowles himself, from Colorado Springs to Chicago in 1939.

Jacob Marschak came to Chicago as professor, and to the Cowles Commission as research director, in 1943. It was he who assembled the staff that created the econometrics revolution. He had organized an informal econometrics seminar that met in New York on weekends during 1940–1942; it included Abraham Wald, Trygve Haavelmo, Tjalling Koopmans, and others. The Cowles theoretical econometric work began in earnest with Marschak's arrival, and continued during the early years of Koopmans' research directorship, which ran from 1948 to 1954.

Seminars at Cowles were lively informal affairs. Students as well as senior members were encouraged to participate in the discussion, and did so with enthusiasm. To maintain a semblance of order, Marschak established the "Cowles Commission rule" under which discussion from the floor was prohibited until after the speaker had finished his or her prepared remarks, except for "clarifying" questions. Members soon learned to preface their interruptions by saying "I have a clarifying question," and the chair for that day had to rule as to whether the interruption was clarifying or not.

In 1955, after James Tobin of Yale had declined to move to the Cowles Commission to become director, the Commission moved to Yale instead, and Tobin became director. As Marschak put it, if Mohammed would not come to the mountain, the mountain had to go to Mohammed. The attractions included faculty appointments in Economics at Yale for Cowles research associates, and the Cowles family's willingness to provide an endowment at Yale (hence the change in name from "Commission" to "Foundation").

b. *The Main Components*

The two main components of the Cowles econometric revolution at Chicago were

- an explicit probabilistic framework, and
- the concept of a simultaneous equations model.

The Cowles program was intended to combine economic theory, statistical methods, and observed data to construct and estimate a system of simultaneous equations that could describe the workings of the economy. The aim was to learn from such a system of equations how economic policy could improve the performance of the economy.

The program required the development of theoretical methods for solving the *identification problem*, that is, for determining the conditions under which one can be sure that an equation fitted to data is actually the equation one wants, rather than a different equation, or a mixture of the desired equation with other equations. The program also required the development of *methods of estimation and hypothesis testing* suitable for identified equations in simultaneous equations models. In such models the typical equation has more than one dependent variable, so that ordinary least squares regression is biased. In these two tasks the program was enormously influential. However, the implications for applied work are still controversial, as we shall see.

The two chief originators of Cowles' theoretical econometric work were Haavelmo and Koopmans. Both have received Nobel Prizes in economics.[3]

---

[3] Other Nobel Prize recipients associated with the Cowles Commission at Chicago are Kenneth Arrow, Gerard Debreu, Lawrence Klein, Harry Markowitz, Franco Modigliani, and Herbert Simon. Only Klein was closely associated with the Cowles econometric program.

Haavelmo's pioneering contributions are contained in two fundamental papers (1943, 1944).[4] The main body of Cowles theoretical econometric results is contained in two Cowles Commission monographs, Koopmans (1950a) and William C. Hood and Koopmans (1953). Many of the papers in the former were first presented in 1945 at a conference at the University of Chicago.[5]

The rest of this paper is organized as follows. Section 2 gives an overview of the Cowles approach. Section 3 presents a simple two-equation model whose special cases will be used as vehicles to illustrate the Cowles work. Section 4 deals with the identification problem and its theoretical solution. Section 5 describes simultaneous equations estimation and testing methods and results. Section 6 deals with the Cowles assumptions. Section 7 evaluates the Cowles work in relation to recent developments. Section 8 is a brief summing up.

2. *An Overview of the Cowles Approach*

a. *The Status Quo Ante*

Already available before the Cowles Commission came to Chicago were Leon Walras' general equilibrium theory and contemporary business cycle theories, especially J. M. Keynes' *General Theory.* Simon Kuznets, Richard Stone, Raymond Goldsmith, and others were compiling data on national income and wealth, and the U.S. Department of Commerce was soon to begin annual publication of the national accounts.

Attempts to estimate simultaneous supply and demand equations are much older than the Cowles Commission. Marcel Lenoir (1913), Philip Wright (1915), Elmer J. Working (1927), and Ragnar Frisch (1933) understood that if the supply curve alone shifts, then price-quantity data trace out the demand curve, and if the demand curve alone shifts, the data trace out the supply curve. But none of these showed what to do if both curves were shifting.

The ordinary least squares method was well known to economists as a means of estimating and testing a single linear equation having an additive disturbance. Koopmans in his doctoral dissertation (1937) had generalized it by showing rigorously how to estimate a linear equation when all variables are measured with stochastic errors. He specified the joint probability distribution of the observable variables, based on assumptions about the joint distribution of the measurement errors, and obtained estimators by the

---

[4] Haavelmo used explicit stochastic simultaneous equations models and applied rigorous statistical inference theory to them. Haavelmo (1943) is a short clear demonstration, by means of simple examples, of why least squares yields biased and inconsistent estimators in simultaneous equations models, and how to get consistent estimators in special cases that we now recognize as just identified. Haavelmo (1944) is a long and rather technical paper that sets forth the approach followed by the Cowles group.

[5] Koopmans' major theoretical econometric contributions are contained in four important papers which are included in these two monographs. The first (1949) is a clear presentation of the solution of the identification problem for linear models. The second (1950b) rigorously defines exogenous and predetermined variables. The third, by Koopmans, Herman Rubin, and Roy B. Leipnik (1950a), is the Cowles Commission's theoretical econometrics magnum opus; it gives the basic Cowles theorems on identification and maximum likelihood estimation of simultaneous equations systems as a whole, in esoteric and difficult mathematical notation. The fourth, Koopmans and Hood (1953), presents some of the Cowles results, including some later ones concerning the estimation of one equation at a time, in a form that is rigorous but less technical and more readable than Koopmans, Rubin, and Leipnik (1950a).

Also important were an early paper by Henry B. Mann and Wald (1943), proving the consistency of maximum likelihood estimators for simultaneous difference equations having no exogenous variables, and three papers on the limited information maximum likelihood method for estimating a single equation in a simultaneous system, one moderately technical by Meyer A. Girshick and Haavelmo (1947), abridged in Hood and Koopmans (1953), and two quite technical by Anderson and Rubin (1949, 1950). Significant theoretical contributions were also made by Marschak, Leonid Hurwicz, and Herman Chernoff and Rubin; see the bibliography.

method of maximum likelihood. Most of the Cowles Commission's work concentrated on models whose stochastic elements were disturbances to equations, rather than errors in variables as in Koopmans' early study. Nevertheless, his explicit treatment of stochastic elements by the method of maximum likelihood was found in all the Commission's theoretical econometric work.

There had been some early successes with estimation methods for simultaneous equations systems, but they had not been generally recognized and accepted by economists. Sewall Wright, a son of Philip Wright, had devised the related method of path analysis (1925, 1934). Further, Philip Wright (1928; possibly with Sewall's help, though the book does not indicate this) and Jan Tinbergen (1929–30) gave the correct solution, apparently independently, to the problem of estimating linear supply and demand curves when both curves shift, for a special case that we now recognize as just identified, but their work was not referred to in Haavelmo's papers or in the Cowles theoretical econometric monographs.

Empirical econometric studies had become more frequent. Henry Schultz at Chicago, in his monumental demand study (1938), went far beyond his teacher Henry Moore. Wesley Mitchell's work (1913, 1927) had provided a great deal of understanding of business cycle behavior, albeit without mathematical models. The famous paper by Charles W. Cobb and Paul H. Douglas (1928) on the aggregate production function had long since appeared. The first macroeconometric models had been prepared by Koopmans' mentor for the Netherlands (Tinbergen 1937, 1959) and for the United States (1939).

Frisch in an unpublished memo (1938) had criticized Tinbergen's models for not attending to the identification problem (without using that term), and he proposed a sufficient condition for the nonidentifiability of a linear dynamic difference-equation model having no exogenous variables and no disturbances that is similar to the standard condition we use today, but he did not arrive at the definitive solution of the problem. And Milton Friedman (1940) had criticized Tinbergen's U.S. model from a different point of view, urging that it not be accepted unless and until its predictions had been tested successfully against new data.

b. *The Central Idea of the Cowles Econometric Approach, and Some Key Concepts*

The Cowles workers regarded economic behavior as the result of the simultaneous interaction of different agents, as exemplified most simply by the intersection of a supply and a demand curve. For this reason they built their econometric methods around systems of simultaneous equations.

The Cowles view was that to understand a particular aspect of economic behavior, such as the price of food, or aggregate personal consumption, one wanted a system of equations capable of describing it. These equations should contain relevant observable variables, be of known form (e.g., linear, log-linear, quadratic), and have estimatable coefficients. The Cowles program was intended to provide a method for choosing the variables relevant to a particular problem, obtaining a suitable system of equations, and estimating the values of its parameters. Little attention was given to how to choose the variables and the form of the equations; it was thought that economic theory would provide this information in each case. More about this later. The main effort was directed to estimating the equations once they had been formulated.

Experience had shown that except for definitional identities, no equation ever describes observable economic behavior exactly. This may be because of an incorrect mathematical form for the equation; errors of measurement in the data; the presence of variables that are difficult or impossible to observe, or some combination of these. Hence the equations were assumed to contain two types of variables: *systematic observable* variables that are ingredients of economic theory (prices, quantities, government expenditures, tax rates, . . .), and *nonsystematic unobservable* variables that represent random disturbances to equations.[6] The random disturbance in each equation is supposed to account for the discrepancy between the actual observed value of a variable and its theoretical value given by the systematic part of the equation. Assumptions were made about the probability distributions of the random disturbances, in order that the statistical properties of estimators of the parameters of the equations could be deduced by Neyman-Pearson methods.

The observable variables were further classified into two types: *endogenous*, whose behavior is to be explained by the system of equations, and *exogenous*, whose values are assumed to be generated outside of the system of equations and taken as given. Exogenous variables can be stochastic or nonstochastic. Examples of variables often assumed to be endogenous are quantities of inputs and outputs demanded and supplied, prices, wages, and interest rates. Examples of variables often assumed to be exogenous are weather, population, and public policy variables. The Cowles work did not have much to say about the crucial issue of how to decide whether a variable should be treated as endogenous or exogenous, although Koopmans (1950b) had presented rigorous definitions. Of course, if a policy variable is determined by some systematic process that responds to private behavior, then it must be endogenous. More about this later.

The endogenous variables at any time were further classified into two types: the *current* values at that time, and the *lagged* values. The lagged endogenous variables were said to be *predetermined* as of that time, because they are not affected by the current operation of the system. Exogenous variables are also predetermined. The current endogenous variables at any time were said to be *jointly dependent* as of that time.

The end product envisioned by the Cowles strategy for an applied economic problem was thus a system of simultaneous equations that would describe how the jointly dependent variables are determined simultaneously, given the values of the parameters, the predetermined variables, and the disturbances. Because the disturbances are unobservable and assumed to be random, they are typically replaced for forecasting purposes by the expected values of their probability distributions (usually zero).

A song, written by graduate students to the tune of "The American Patrol" march and presented in about 1949 at one of the skits that enlivened parties at the Economics Department, explained the Cowles Commission's objectives to skeptics in the Department (and there were some). Here it is:

> We must be rigorous, we must be rigorous,
> We must fulfill our role.
> If we hesitate or equivocate
> We won't achieve our goal.
> We must elaborate our systems complicate

---

[6] Two other types of variables should be considered: systematic unobservable variables representing expectations held by market participants, and random unobservable variables representing errors of measurement of systematic variables. The Cowles Commission did little work on these two, although as noted above Koopmans (1937) had dealt with the latter.

To make our models whole.
Econometrics brings about
Statistical control!
Our esoteric seminars
Bring statisticians by the score.
But try to find economists
Who don't think algebra's a chore.
Oh we must urge you most emphatically
To become inclined mathematically
So that all that we've developed
May some day be applied.

The applied econometric work of the Cowles Commission, inspired by Marschak and directed at the improvement of macroeconomic policy, had a definite Keynesian flavor. Friedman was one of the most persistent critics of the Cowles brand of econometrics, not only from the point of view of econometric methods (as exemplified by his critical review of Tinbergen's U.S. model) but also because of his skeptical view of Keynesian models and the Keynesian consumption function.[7] A couple of verses from another song, to the Gilbert and Sullivan tune of "When I Was a Lad," presented at the same party, exemplify this (remember that Friedman had been a student of Arthur F. Burns, and assigned Alfred Marshall's *Principles* to his own students):

When I was a lad I served a term
Under the tutelage of A. F. Burns.
I read my Marshall completely through
From beginning to end and backwards too.
I read my Marshall so carefully
That now I am Professor at the U. of C.
(Chorus) He read his Marshall so carefully
That now he is Professor at the U. of C.
Of Keynesians I make mince meat
Their battered arguments now line the street.
I get them in their weakest assumption:
"What do you *mean* by consumption function?"
They never gave an answer that satsifed me
So now I am Professor at the U. of C.
(Chorus) They never gave an answer that satisfied he
So now he is Professor at the U. of C.

3. *A Simple Illustrative Model and More Key Concepts*

a. *Model 1: The Generic Version of the Simple Model*

To illustrate the econometric methods devised by the Cowles Commission, the problems they solved, and some criticisms of them, we will use a simple two-equation linear model[8] that has two endogenous variables y_1 and y_2, two predetermined variables x_1 and x_2 which may be either exogenous or lagged endogenous, two random disturbances ϵ_1 and ϵ_2, and parameters denoted by β's and γ's with subscripts (the β's go with the y's, and the γ's go with the x's). The subscript t will be used to denote time. For example, if annual data for 1951 through 1990 are used, t could take values from one (for 1951) through 40 (for 1990). The two equations are

$$y_{1t} + \beta_1 y_{2t} = \gamma_1 x_{1t} + \gamma_2 x_{2t} + \epsilon_{1t} \qquad (1)$$

$$\beta_2 y_{1t} + y_{2t} = \gamma_3 x_{1t} + \epsilon_{2t}. \qquad (2)$$

This will be called *Model 1*. Though it is simple, it is flexible enough to describe a variety of situations, as we will see.

The equations of Model 1 as written do not appear to contain constant terms. We could include explicit constant terms by adding $+\gamma_4$ to (1), and $+\gamma_5$ to (2).

---

[7] Melvin Reder (1982, p. 10) refers to "a fairly intense struggle" at Chicago in the late 1940s and early 1950s involving Frank Knight and his students, including Friedman, vs. the Cowles Commission and its adherents. Friedman and Gary Becker (1957) argue that the goodness of fit of the least squares estimate of the simple Keynesian consumption function is in part a statistical illusion, because consumption is being regressed on a variable (income) of which consumption is a major part, and they propose a correct approach. Klein (1958) comments that their point deals with least squares bias, and that it had been recognized and made clear by Haavelmo (1943, 1947). Friedman and Becker (1958) reply that Klein may be right about this, but they think not. Friedman (1957) presents his own version of the consumption function.

[8] Sometimes a simple transformation will render an apparently nonlinear equation linear. For example, a Cobb-Douglas production function with a multiplicative disturbance e^u, such as $Y = AK^\alpha L^\beta e^u$, becomes linear with an additive disturbance u upon taking its logarithm.

To keep the ensuing algebra simple we won't do this. However, the equations as written can provide for constant terms, in either of two different ways. First, suppose we had included explicit constant terms as just described. We could subtract from each equation at time t its mean calculated over the sample period (for example, from $t = 1$ to $t = 40$). The constant terms would cancel out, and the equations would appear exactly like (1–2) without an explicit constant term, except that then all the variables would be defined as deviations from their sample means. Second, if one of the x's (say x_1) never varies, but is always equal to one, then the parameters attached to it become explicit constant terms.

The disturbances are assumed to come from a probability distribution, usually with zero mean. If an equation has an unknown constant term, this assumption of a zero mean disturbance imposes no restriction, because if the mean were not zero, but were, say, five, then five could be added to the constant term and subtracted from the disturbance, thus yielding an equivalent equation whose disturbance does have zero mean. But if an equation is specified to have no constant term, then the assumption that its disturbance has zero mean is a real restriction.

Let us suppose that equations (1–2) describe the behavior of two sectors of the economy, perhaps (1) for the suppliers and (2) for the demanders of a product. Such equations were called *structural* by Haavelmo (1944). The idea here is that a change in the form or parameters of one structural equation does not affect other structural equations: in this example a change in the demand equation does not affect the supply equation, and vice versa. Haavelmo used the name *structure* to denote such a system of structural equations when numerical values are specified for all its parameters, including the parameters of the joint distribution of the disturbances as well as the coefficients of the equations.

Marschak (1950) characterized a *model* by means of a set of a priori information (true or false) about the structure, information obtained from economic theory, not from observed values of the variables. Ideally this information tells how many equations there are, what variables appear in each one, which variables are endogenous and which are exogenous, what lagged endogenous variables (if any) appear in the model, what is the form of each equation, and what is the form of the joint distribution of the disturbances (normality is often assumed, for example). Leonid Hurwicz (1950) defined a model as a set of structures that are admissible according to this a priori information. Thus the problem in the Cowles approach is first to construct the model, and second to estimate the structure.

Often one wishes to forecast the future value of an endogenous variable. In general, a structural equation cannot do this, because each structural equation typically includes two or more endogenous variables. To obtain such forecasts, one must solve the system of structural equations algebraically for the endogenous variables. The result is called the *reduced form*. Each of its equations gives the value of one endogenous variable in terms of the predetermined variables, disturbances, and structural parameters.

To solve the system (1–2) for its reduced form, we compute its determinant, which we denote by δ:

$$\delta = 1 - \beta_1\beta_2 \tag{3}$$

Using this expression, the reduced form of the system (1–2) is

$$y_{1t} = (1/\delta)\,[(\gamma_1 - \beta_1\gamma_3)x_{1t} + \gamma_2 x_{2t} + (\epsilon_{1t} - \beta_1\epsilon_{2t})] \tag{4}$$

$$y_{2t} = (1/\delta)\,[(\gamma_3 - \beta_2\gamma_1)x_{1t} - \beta_2\gamma_2 x_{2t} + (\epsilon_{2t} - \beta_2\epsilon_{1t})]. \tag{5}$$

Sometimes it is convenient to give new names to the parameters and disturbances of the reduced form, and rewrite it thus:

$$y_{1t} = \pi_{11}x_{1t} + \pi_{12}x_{2t} + v_{1t} \qquad (6)$$

$$y_{2t} = \pi_{21}x_{1t} + \pi_{22}x_{2t} + v_{2t}. \qquad (7)$$

Notice that if there is a change in any of the parameters of either one of the structural equation (1) or (2), then the parameters of both of the reduced form equations change, because every one of the β's and γ's appears in both reduced form equations (remember that δ depends on all of the β's). Thus the structural equations can be thought of as the elementary building blocks of the system, while each reduced form equation is a mixture of the structural equations.

To use the reduced form for forecasting future endogenous variables, one obviously needs estimates of its parameters, and forecasts of future values of the predetermined variables and disturbances. Forecasts of exogenous variables must be obtained by some method external to the model. Future values of lagged endogenous variables are obtained either from past data, or from the past operation of the system. The best forecast of a random disturbance is its expected value, typically zero.

The system of equations (1–2) appears to be static, because all the variables are dated at the same time, t. It is indeed static if the y's and x's at time t are four different variables, such as price, quantity sold, income, and production cost, all pertaining to the same period of time. Then we think of the y's as endogenous and the x's as exogenous.

b. *Model 2: A Dynamic Version of Model 1*

However, if either or both of the x's stands for a previous period's value of one of the y's, then the system is dynamic. For example, suppose that by definition

$$x_{2t} = y_{1,t-1}. \qquad (8)$$

This could be because in equation (1) y_1 depends on adaptive expectations about the future value of y_2. Alternatively, it could be because in (1) there is a lag in the response of y_1 to changes in y_2.

The dynamic case of Model 1 obtained by specifying that x_2 is last period's value of y_1, as in (8), will be called *Model 2*. In Model 2 x_2 cannot be exogenous, because it is determined by the past operation of the system. As noted earlier, x_{2t} (that is, $y_{1,t-1}$) is said to be *predetermined at time t*, and y_{1t} and y_{2t} are said to be *jointly dependent at time t.*

The preceding explanations of exogenous and predetermined variables are plausible, but are not suitable for rigorous statistical analysis. Koopmans (1950b) gave rigorous definitions as follows. (i) A variable x (meaning the whole series of values of x for all t) is *exogenous* if and only if the value of x at every time period is statistically independent of all disturbances in the system, past, present, and future. Note that an exogenous variable can be stochastic, or not. (ii) A variable is *endogenous* if it is not exogenous. (iii) A variable x_t is *predetermined at time t* if and only if x_t is statistically independent of all present and future disturbances in the system, but not necessarily of past disturbances. Thus in Model 2, where (8) holds, x_{2t} (that is, $y_{1,t-1}$) will be predetermined at time t if ϵ_1 and ϵ_2 are serially independent, but it cannot be exogenous: it cannot be independent of the previous disturbance $\epsilon_{1,t-1}$ because it is influenced by $\epsilon_{1,t-1}$ through the lagged version of equation (1). Exogenous variables are always predetermined, but the converse is not true. (iv) A variable dated at time t is *jointly dependent at time t* if it is not predetermined at time t. In Model 2, note that $y_{1,t-1}$

is jointly dependent at time $t - 1$, but predetermined at time t or later.

A dynamic system determines the time paths of the endogenous variables from time t onwards, if it is given the values of the parameters, the time paths of the exogenous variables and disturbances, and the initial values of the lagged endogenous variables as of time t. To study the behavior of a dynamic system we assume that the exogenous variables and disturbances are held constant for the future. The system may generate cycles, or not. It may be stable (that is, the endogenous variables may approach equilibrium values), or not. In order to discover the dynamic behavior of a system, we need to find its *final form*, that is, a set of equations similar to the reduced form except that each endogenous variable is expressed in terms of *its own* lagged values (not those of others) and exogenous variables and disturbances (this term is due to Tinbergen 1939).

Let us return to the dynamic Model 2. We can still use the structural system (1–2) and its reduced form (4–5), remembering that x_{2t} is $y_{1,t-1}$ as in (8). In this case the final form equation for y_1 is obtained simply by substituting $y_{1,t-1}$ for x_{2t} in the reduced form equation (4). It expresses y_{1t} as a function of x_{1t} and $y_{1,t-1}$, thus

$$y_{1t} = (1/\delta)\,[(\gamma_1 - \beta_1\gamma_3)x_{1t} + \gamma_2 y_{1,t-1} + (\epsilon_{1t} - \beta_1\epsilon_{2t})]. \quad (9)$$

The final form equation for y_2 is found as follows: first write equation (2) for time $t - 1$; solve the result for $\beta_2 y_{1,t-1}$; then substitute the resulting expression in (5) to get

$$y_{2t} = (1/\delta)\,[(\gamma_3 - \beta_2\gamma_1)x_{1t} + \gamma_2 y_{2,t-1} - \gamma_2\gamma_3 x_{1,t-1} + (\epsilon_{1t} - \beta_2\epsilon_{1t} - \gamma_2\epsilon_{2,t-1})]. \quad (10)$$

Notice that when the exogenous variable x_1 and the disturbances ϵ_1 and ϵ_2 are held constant, the final form equations (9) and (10) are alike except for their constant terms. Each can be written as

$$y_{it} = (\gamma_2/\delta)y_{i,t-1} + \theta_i, \qquad i = 1 \text{ or } 2 \quad (11)$$

where θ_i is a constant depending on the values of x_1 and the parameters and ϵ's. These final form equations are first order linear difference equations. Each has the same characteristic root, γ_2/δ. The system is stable for both y_1 and y_2 if the root is less than one in absolute value, and unstable for both otherwise. Further, it has oscillations for both if the root is negative, and not otherwise. The fact that all endogenous variables in the system have the same dynamic behavior is not an accident; it is typical of linear systems.

The equilibrium values of y_1 and y_2 can be found from (11) by setting the lagged and current values equal; the result is

$$y_{it} = \theta_i/[1 - (\gamma_2/\delta)]. \quad (12)$$

c. *The Recursive Form*

An interesting special case of a structural model occurs when two conditions are met: (i) one of the equations contains only a single jointly dependent variable, another equation contains only two (that one and one other), another contains only three (those two and one other), and so on, until the last equation which contains them all; (ii) the disturbances to all the equations at any time t are statistically independent of each other. Such a model is said to be *recursive*. Model 1 above would be recursive if either β_1 or β_2 were zero (or both) and the disturbances ϵ_{1t} and ϵ_{2t} were independent of each other for all t. In that case, if β_1 were zero, y_{1t} would be predetermined in equation (2); similarly, if β_2 were zero, y_{2t} would be predetermined in equation (1); and if both β_1 and β_2 were zero, the two equations would determine y_{1t} and y_{2t} independently of each other.

The Cowles theoretical results were of course not confined to a two-equation model. They dealt with arbitrary numbers of equations, endogenous variables, exogenous variables, and lags of the endogenous variables. In more general systems, where the number of endogenous variables appearing with a lag is greater than one, and/or where the longest lag of any endogenous variable is greater than one, each final equation is of second or higher order, that is, contains two or more lagged values of its endogenous variable. In that case the characteristic roots can be complex numbers and hence the dynamic path of the system can involve cycles of any length, not merely oscillations up and down in alternate periods as in the case of a first-order final equation such as (9) or (10).

Thus far we have seen how the system of equations (1–2) in Model 1 can handle the presence of constant terms, and how it can describe either a static or a dynamic model. In the next section, by assigning different possible values (sometimes zero) to its parameters, we will obtain and analyze several interesting special cases.

4. *The Identification Problem and Its Theoretical Solution*

a. *Model 3: An Example of the Problem*

As noted in Section 1, econometric equations cannot be estimated unless they are identified. Those who first tried to estimate simultaneous equations models dealt with supply and demand. They stumbled when they came to the identification problem. They used a special case of Model 1 in which equations (1) and (2) are respectively the supply and demand curves for a good, y_1 is quantity, y_2 is price, x_1 is a constant equal to one so that γ_1 and γ_3 are the equations' constant terms, and γ_2 is zero, so that the only systematic variables in the model are price and quantity. We shall call this special case *Model 3*. It is as follows:

$$y_{1t} + \beta_1 y_{2t} = \gamma_1 + \epsilon_{1t} \quad \text{supply} \qquad (13)$$

$$\beta_2 y_{1t} + y_{2t} = \gamma_3 + \epsilon_{2t} \quad \text{demand.} \qquad (14)$$

As noted above, they understood that if the supply curve alone shifts (which (13) will if $\gamma_1 + \epsilon_1$ changes but $\gamma_3 + \epsilon_2$ and β_2 do not), the observed price and quantity data trace out the demand curve, and if demand alone shifts, the data trace out the supply curve, but if both curves shift, the data trace out neither curve. Hence if one does not observe when and how the curves shift, one cannot tell whether a curve traced out by the price and quantity data is the supply curve, the demand curve, or neither. They did not know what to do about this.

The Cowles workers called it the problem of the *identification* of structural equations. It is not a problem of uncertainty associated with sampling variation: it does not go away as the sample size becomes infinitely large.

b. *Conditions for Identification*

Koopmans (1949) defined an *identified* equation as one whose parameters can be deduced from observed data, using the a priori information that is incorporated in the model (with no uncertainty except the sampling error that occurs in small samples). He gave necessary and sufficient conditions for the identification of the parameters of an equation in a linear system, based on the knowledge or belief that certain parameters in the equation are zero. His *order condition* is this: if an equation in such a system is to have all its parameters identified, at least $G - 1$ of its parameters must be believed to be zero, where G is the number of equations in the system. The converse is almost always true as well, that is, if an equation satisfies this condition, its parameters are almost certain to be identified. It is called the order condition because it can be expressed by saying that the order (i.e., the number of columns) of a certain matrix[9] must be at least $G - 1$. When we count the number of parameters believed to be zero, it makes no difference whether a parameter is a constant term, or a coefficient of a jointly dependent or a predetermined variable.

There is no identification problem with reduced form equations,[10] because each contains only one dependent variable. The order condition holds automatically for reduced form equations because each is a complete one-equation system, $G = 1$, and each excludes $G - 1$ = none of its own variables. Because reduced form equations are identified, any structural equation whose parameters can be deduced from a knowledge of the reduced form parameters is also identified, and conversely.

The identifiability of structural equations when some parameters are believed to be zero will now be illustrated with several special cases of Model 1. Because there are two equations, $G - 1 = 1$. Hence to be identified a structural equation must have at least one parameter that is believed to be zero, that is, must exclude at least one variable.

c. *Model 3 Again: An Unidentified Model*

Consider Model 3, the one that baffled the early investigators. Its reduced form is as follows, where as before δ is given by (3):

$$\begin{aligned} y_{1t} &= (1/\delta)[(\gamma_1 - \beta_1\gamma_3) + (\epsilon_{1t} - \beta_1\epsilon_{2t})] \\ &= \pi_{11} + v_{1t} \qquad (15) \end{aligned}$$

$$\begin{aligned} y_{2t} &= (1/\delta)[(\gamma_3 - \beta_2\gamma_1) + (\epsilon_{2t} - \beta_2\epsilon_{1t})] \\ &= \pi_{21} + v_{2t}. \qquad (16) \end{aligned}$$

It is a special case of (4–5) and of (6–7). Here the reduced form parameters π_{11} and π_{21} are automatically identified. The sample mean of the data for y_1 (quantity) is an estimate of the parameter π_{11} in (15), and the sample mean of y_2 (price) is an estimate of π_{21} in (16).

Now consider the structural parameters of Model 3, β_1, β_2, γ_1, and γ_3. The reduced form parameters are functions of them, thus (remember that δ is a function of β_1 and β_2):

$$\pi_{11} = (\gamma_1 - \beta_1\gamma_3)/\delta \qquad (17)$$

$$\pi_{21} = (\gamma_3 - \beta_2\gamma_1)/\delta. \qquad (18)$$

It is not possible to determine the values of the four unknown structural parameters β_1, β_2,γ_1, and γ_3 from π_{11} and π_{21} and these two equations. Thus in Model 3 the structural parameters are not identified. The order condition confirms this: neither of the equations (13–14) in Model 3 excludes any of the model's variables, whereas identification requires the exclusion of at least one.

---

[9] The matrix in question is obtained by omitting from the matrix of the model's structural coefficients all columns in which the equation does not have a prescribed zero. The so-called rank condition, also set forth by Koopmans (1949), says that, when the a priori information takes the form of such zero restrictions, an equation is identified if and only if the rank of that same matrix is $G - 1$. The rank condition is almost certain to be satisfied if the order condition is, because the parameters in this matrix are very unlikely to have a set of values that will make its columns linearly dependent. However, if a model is segmentable, in the sense that a subset of its equations determine a subset of its dependent variables, the rank condition can fail even if the order condition is met (see Christ 1966, p. 345, ex. 3.10 for an example). Note also that there are other kinds of identifying information besides the requirement that a parameter in an equation be zero. For example, there may be a restriction relating two or more parameters in the same or different equations, or there may be restrictions on the parameters of the probability distribution of disturbances. Identifiability conditions for such information are beyond the scope of this paper (Koopmans, Rubin, and Leipnik 1950; Leon Wegge 1965; Franklin Fisher 1966).

[10] If some of the predetermined variables in a reduced form equation are linearly dependent, then their coefficients are not identified, and cannot be estimated. One remedy for this, if the sample size is large enough, is to omit one (or if necessary more than one) of the predetermined variables, until the remaining ones are linearly independent. Then the least squares estimates can be computed.

d. *Model 4: An Identified Model*

Now consider a case in which all structural parameters are identified. Such a case is obtained from Model 1 by restricting γ_1 to be zero and requiring x_1 to be constant and equal to one. It is similar to Model 3 except that now γ_1 is required to be zero and γ_2 is not. It will be called *Model 4.* It is as follows:

$$y_{1t} + \beta_1 y_{2t} = \gamma_2 x_{2t} + \epsilon_{1t} \quad \text{supply} \qquad (19)$$

$$\beta_2 y_{1t} + \quad y_{2t} = \gamma_3 \quad + \epsilon_{2t} \quad \text{demand.} \qquad (20)$$

Note that each of the equations of Model 4 has one parameter specified to be zero: in (19) it is the constant term (that is, the coefficient of x_1), and in (20) it is the coefficient of x_2. The reduced form of Model 4 is as follows:

$$y_{1t} = (1/\delta)\,[-\beta_1\gamma_3 + \gamma_2 x_{2t} + (\epsilon_{1t} - \beta_1\epsilon_{2t})] = \pi_{11} + \pi_{12}x_{2t} + v_{1t} \qquad (21)$$

$$y_{2t} = (1/\delta)\,[\gamma_3 - \beta_2\gamma_2 x_{2t} + (\epsilon_{2t} - \beta_2\epsilon_{1t})] = \pi_{21} + \pi_{22}x_{2t} + v_{2t}. \qquad (22)$$

(Here the symbols π_{11} and π_{21} are different from those in equations (15–18).) Again the reduced form parameters are automatically identified, and hence can be estimated. They are functions of the structural parameters, thus:

$$\pi_{11} = -\beta_1\gamma_3/\delta \qquad (23)$$

$$\pi_{21} = \gamma_3/\delta \qquad (24)$$

$$\pi_{12} = \gamma_2/\delta \qquad (25)$$

$$\pi_{22} = -\beta_2\gamma_2/\delta. \qquad (26)$$

Now the four structural parameters can be deduced from the four equations (23–26) and the values of the reduced form parameters: β_1 is estimated by the ratio of the estimates of $-\pi_{11}$ and π_{21}, β_2 is estimated by the ratio of the estimates of $-\pi_{22}$ and π_{12}, δ is estimated from the estimates of the β's using its definition $\delta = 1 - \beta_1\beta_2$, γ_2 is estimated by the product of the estimates of δ and π_{12}, and γ_3 is estimated by the product of the estimates δ of and π_{21}. Thus the structural parameters are identified. The order condition is consistent with this, for the number of variables excluded from each of the structural equations (19–20) is $G - 1 = 1$, as required.

e. *Other Cases*

Now consider Model 1. Equation (1) excludes no variables, and hence is not identified. Equation (2) excludes one variable, x_2, and hence is identified. This result can be verified by comparing the reduced form equations (4–5), and noting that β_2 can be estimated because it is the ratio of $-\pi_{22}$ to π_{12}, as in Model 4, but β_1 cannot be estimated. In Model 2 the situation is the same as in Model 1.

If demand depends only on current price, and supply depends only on last period's price, then the identification problem is solved, because any mixture of the supply and demand curves will contain both current and lagged prices, and hence cannot be mistaken for either demand or supply (Mordecai Ezekiel 1928). This is exemplified by Model 1 if we assume that $\beta_1 = 0$, x_{2t} = last period's price $y_{2,t-1}$, and x_1 is constant and equal to one. Such a model is called the cobweb model, because the dynamic behavior of a graph of price vs. quantity resembles a spider's web. If the disturbances in the two equations are independent at each time t, such a system is recursive.

The order condition says that the number of the model's variables excluded from an identified equation must be *at least* $G - 1$. If it is exactly $G - 1$, the equation is said to be *just identified.* If it is more than $G - 1$, the equation is said to be *overidentified.* Model 4 is just identified. In Model 1, if γ_3 were specified to be zero but γ_1 and γ_2 were not, then equation (2) would be overidentified, because it would exclude two of the model's variables.

When an equation is overidentified,

there are restrictions on the parameters of the reduced form. For example, in the case just mentioned, where equation (2) is overidentified because $\gamma_3 = 0$, a comparison of equations (4) and (5) shows that the ratios π_{21}/π_{11} and π_{22}/π_{12} must be the same, because both must be equal to $-\beta_2$.

5. *Simultaneous Equations Estimation and Testing at Cowles*

a. *Estimation Methods*

Under suitable statistical assumptions, least squares yields consistent and asymptotically unbiased estimators of reduced form equations like (6–7), because each of them contains only one dependent variable.[11] However, for simultaneous equations like (1–2), least squares yields biased and inconsistent estimators of structural parameters except in special cases.[12] The problem is that in such a system, even if all structural equations are identified, each equation contains more than one jointly dependent variable.

Hence for simultaneous equations, new estimation methods were needed. These methods, coupled with the solution of the identification problem, constitute the heart of the Cowles revolution in econometric methods.

Using models similar to the just-identified Model 4 in the previous section, Haavelmo (1943) derived estimators of the structural parameters equivalent to those described above for Model 4 and used by Philip Wright (1928) and Tinbergen (1929–30)—but he did not refer to their work. These estimators are now called *indirect least squares* estimators, because the first step is to estimate the reduced form by least squares, and the second step is to deduce estimated values of the structural parameters from the estimated reduced form. He showed that his estimators were consistent.

Haavelmo (1947) used the indirect least squares method to estimate the marginal propensity to consume from annual U.S. data for 1922–1941. His consumption equation was

$$c_t = \alpha y_t + \beta + u_t \qquad (27)$$

where c, y, and u are respectively real per capita consumption, disposable income, and a disturbance. This equation is just identified in his two-equation model, treating consumer saving as the exogenous variable, and also in his three-equation model, treating gross private investment plus the government deficit as the exogenous variable (the estimate in this case was based on data for only 1929–1941 because the Commerce Department investment data did not go back before 1929). The two indirect least squares estimates of α were respectively 0.67 and 0.71. The corresponding least squares estimates for the same two periods were respectively 0.73 and 0.72, larger than the indirect least squares estimates, as predicted by Haavelmo's analysis. These results illustrate a frequent finding: even though least squares estimators are not consistent, they usually turn out in practice to be quite close to consistent estimators. Haavelmo constructed confidence interval estimates for the marginal propensity to consume, based on estimates of the standard deviation of its estimator. He also constructed joint confidence intervals for it and other parameters of the model. When he included lagged income in the equation,

---

[11] See Section 6. If in addition all predetermined variables are exogenous, least squares estimators of the reduced form become unbiased. If disturbances are normally distributed, least squares estimation of the reduced form is a maximum likelihood method.

[12] This was shown by Haavelmo (1943, 1944). Jean Bronfenbrenner (1953) gives a clear and simple analysis of least squares bias. A recursive model is one special case in which least squares estimation can be unbiased and consistent.

he obtained slightly lower estimates for 1922–1941 than without it.

For estimating the parameters of a just-identified structural equation the *instrumental variables* method[13] consists of the following steps: First, using the sample data, form the moment of the equation with each of the predetermined variables in the model (not, as in least squares, with each of the equation's right-hand-side variables), and assign a zero value to the moment of the disturbance with each predetermined variable. The result is a system of equations in which the unknowns are the instrumental variables estimates, and the known quantities are the sample moments of the variables in the equation with the predetermined variables in the model. Then solve this system algebraically for the estimates of the parameters.

For a just-identified equation the instrumental variables method is equivalent to indirect least squares. In this case both methods work because the number of structural parameters to be estimated is just equal to the number of equations to be solved for them: in the case of instrumental variables it is the number of equations relating the structural estimates to moments, and in the case of indirect least squares it is the number of equations relating the structural estimates to the estimated reduced form parameters. In both cases it is the number of parameters that are specified to be zero in the equation. The indirect least squares method is exemplified by the discussion following equations (23–26) for Model 4 above.

For an unidentified equation both methods fail, as they should, because the number of equations to be solved for the estimates of the structural parameters is too few. For indirect least squares this is exemplified by equation (1) of Model 1: its parameters cannot be got from the estimates of the reduced form (4–5).

For an overidentified equation, the number of equations to be solved is more than needed, and they will contradict each other with probability one if the sample is finite (because the estimated values of the parameters will differ randomly from the true values which satisfy the a priori restrictions exactly). In such a case, if there are n more a priori restrictions than needed, one can ignore n of them (any n of them will do), thus pretending that the equation is just identified. One can then use either indirect least squares or instrumental variables. The disadvantage of this approach is that it wastes some of the a priori information, and gives different results depending on which n restrictions are ignored.

When the method of maximum likelihood is applied to the problem of estimating all the structural parameters of the system at once, using all the a priori restrictions for the entire model, on the assumption that the disturbances are normally distributed, the resulting estimators are called *full information maximum likelihood* estimators, FIML for short, because they use all the available a priori information. Under suitable assumptions FIML estimators of structural parameters are consistent and asymptotically efficient, approximate confidence-region estimators of these parameters can be obtained, and approximate t-tests and F-tests of their statistical significance can be performed (Koopmans, Rubin, and Leipnik 1950).[14]

If the maximum likelihood method is applied to one identified structural equation at a time (whether it is overidentified or not), using only the a priori restrictions

---

[13] The instrumental variables method goes back to Sewall Wright (1925), Olav Reiersøl (1941, 1945), and Robert C. Geary (1949). Reiersøl was a visitor at the Cowles Commission in the summer of 1949.

[14] This is a difficult technical paper. The "suitable conditions" are described in Section 6 below.

that concern that equation, the result is called the *limited information maximum likelihood* method, LIML for short, because it uses only some of the a priori information (Girshick and Haavelmo 1947; Koopmans and Hood 1953). LIML yields the same estimates for an overidentified equation as would be obtained if the reduced form were estimated by least squares subject to the restrictions on the reduced form that are implied by overidentification of the structural equation, and the indirect least squares method were then applied. LIML remains consistent even if the equations are nonlinear, so long as they are linear in unknown parameters and certain moment matrices have appropriate probability limits (Anderson and Rubin 1949, 1950).

For a just-identified equation, LIML reduces to the indirect least squares and instrumental variables methods. For an unidentified equation it fails, as it should. For an overidentified equation, it gives a unique result.[15]

b. *Applications*

Empirical work in the 1940s and early '50s was done with mechanical desk calculators which, unlike many of today's cheap electronic pocket calculators, couldn't compute square roots or logarithms or exponentials with a single command. The iterative computations for FIML are so complex that the method was never applied by the Cowles workers, except for an illustrative computation for a simplified model containing only three equations. For most of their empirical work, the Cowles workers used LIML, which is also iterative, but much simpler than FIML.[16] Later, with the advent of electronic computers, FIML became a practical method.

The first application of LIML was the estimation of a five-equation model of food demand, using annual U.S. data for 1922–1941 (Girshick and Haavelmo 1947). Some of the equations were just identified, and of course for them the LIML estimates were the same as the indirect least squares estimates. In 1947 it was not yet possible to give confidence interval estimates for the overidentified equations because the standard deviation of the LIML estimator for this case had not yet been derived.

The first macroeconometric model using the new methods was presented in Klein's monograph, *Economic Fluctuations in the United States 1921–1941* (1950). The model has 16 equations, including demand equations for active money balances, idle money balances, consumer goods, owner-occupied hous-

---

[15] The idea behind LIML is this: Each structural equation in a linear system must be a linear combination of just those reduced form equations that concern the jointly dependent variables that are not excluded a priori from the structural equation. The coefficients of the linear combination are the structural equation's coefficients of those jointly dependent variables. An estimate of a *just*-identified equation can be expressed exactly as a linear combination of those estimated reduced form equations; the indirect least squares method finds that linear combination. But this cannot be done with an *over*identified equation, because no linear combination of the least squares estimates of the reduced form equations can exactly satisfy all the a priori restrictions (except in infinite samples). What the LIML method does is to find the linear combination of the estimated reduced form equations that comes closest to satisfying the a priori restrictions, in the sense of making the so-called variance ratio as close to one as possible. The variance ratio is the ratio between two variances, each of which is the variance of the residual of a regression of the jointly dependent component of the equation being estimated (for equation (1) or (2) this is the expression on the left side of the equality sign) on a set of predetermined variables. One of these regressions uses *all* of the model's predetermined vaiables, and the other uses *only* those that are not excluded a priori from the equation. Since the excluded variables are supposed to have zero coefficients, the two regressions should be the same except for sampling variation, and so should their residual variances. Hence the LIML estimation procedure was designed to make the ratio of their residual variances as close to one as possible. See also note 25 below.

[16] Chernoff and Nathan Divinsky (1953) described the iterative computation methods for FIML and LIML, and reproduced the actual computations made with desk calculators for one of Klein's simple models.

ing, rental housing, plant and equipment, inventories, and labor; an equilibrium equation for dwelling space; adjustment equations for output, interest rate, and rent; and four definitional equations. Klein checked the identifiability of the equations, and estimated them by the LIML method using U.S. data for 1921–1941. He gave no estimates of the standard deviations of the LIML estimates, because the necessary formula still had not been derived. This model had a clearly Keynesian flavor.

The necessity for testing an econometric model against new data, after it has been constructed and estimated, was not emphasized in the theoretical econometric publications of the Cowles Commission. A simple procedure is to use the estimated model to forecast data that were not used to estimate it, or, even better, were not available when the model was specified, and conclude that something is wrong if the forecast errors are too large. Holbrook Working and Harold Hotelling (1929) and Hotelling (1942) had derived the standard deviation of the error of a post-sample forecast, made from a single equation estimated by least squares, under the hypothesis that the estimated equation still holds during the forecast period, and showed how to use it to test that hypothesis: if the forecast error is too large relative to this standard deviation, the hypothesis is rejected. Rubin (1948, unpublished) showed how to do the same thing with the post-sample residual from a structural equation that has been estimated by LIML. Variants of this test were applied by Andrew Marshall (1949), Christ (1951), Stephen G. Allen (1954), and Hildreth and Francis G. Jarrett (1955) to decide which LIML-estimated structural equations continued to describe data in the post-sample period, and which ones did not.

Andrew Marshall's M.A. thesis at the University of Chicago, entitled "A Test of Klein's Model III for Changes of Structure" (1949, unpublished), used Klein's model to compute residuals for each equation for 1946 and 1947 (skipping the war years 1942–45 on the ground that normal economic behavior was superseded by wartime regulations),[17] to see which equations held up well and which didn't. Five didn't. This type of test had been proposed, and carried out for one equation, by Friedman in his review (1940) of Tinbergen's model of the U.S. There Friedman pointed out correctly that Tinbergen's procedure (which has become common among model builders) of trying many different forms for each equation, and choosing one with a high correlation coefficient, is appropriate as a way of deriving tentative hypotheses, but does not constitute an empirical test of the equations so chosen. We will expand on this point at the end of this section.

"A Test of an Econometric Model for the United States, 1921–1947" (Christ 1951) built upon Andrew Marshall's results. The five Klein equations that did not hold up well for 1946 and 1947 in Marshall's tests were revised, and the revised model was reestimated by LIML using data for 1921–1941 and 1946–1947. The reduced form of the revised model was then used to compute forecasts for 1948. For six of 13 variables, least squares estimates of the revised model's reduced form made better forecasts for 1948 than the naive methods of assuming no change from the previous year, or no change in the trend from the previous two years, and for seven of 13, worse.[18]

---

[17] An unusual example of a consumption equation successfully fitted right through World War II is Richard E. Brumberg (1956). For consumer durables, he defined and measured consumption not in the usual way as purchases of goods, but as the flow of services represented by interest and depreciation.

[18] Andrew Marshall also used naive model tests, but for structural equations rather than reduced forms.

When a structural equation is overidentified, restrictions on the reduced form are implied (recall the discussion of this point at the end of Section 4e above). A chi-square test of whether the data reject these restrictions was proposed by Anderson and Rubin (1950). The Anderson-Rubin test of overidentifying restrictions was performed by Christ (1951); they were rejected for four of his ten stochastic equations at the 95 percent level, and seven of ten at the 90 percent level. Robert Basmann (1960) proposed an F test of overidentifying restrictions on a structural equation, equivalent to the Anderson-Rubin chi square test in infinite samples but more accurate in small samples.

Friedman in his discussion of Christ's model endorsed the testing of such models against post-sample data, but ventured the

> hunch . . . that attempts to proceed now to the construction of additional models along the same general lines will, in due time, be judged failures. (1951, p. 112)

Instead, he wrote,

> I believe our chief hope is to study the sections covered by individual structural equations separately and independently of the rest of the economy. These remarks obviously have a rather direct bearing on the desultory skirmishing between what have been loosely designated as the National Bureau and the Cowles Commission techniques of investigating business cycles. . . . As the National Bureau succeeds in finding some order, some system, in the separate parts it has isolated for study its investigations will increasingly have to be concerned with combining the parts—putting together the structural equations. As the Cowles Commission finds that its general models for the economy as a whole are unsuccessful, its investigators will increasingly become concerned with studying the individual structural equations, with trying to find some order and system in component parts of the economy. Thus, I predict the actual work of the two groups of investigators will become more and more alike. (p. 114)

This prediction has been confirmed only partly. On the one hand, complete macroeconometric models have proliferated, and many now make useful forecasts of the U.S. economy several quarters ahead (Stephen McNees 1988). On the other hand, it is true that macroeconometric models have not revealed an invariant fundamental structure of the economy, and in recent years the emphasis in mainstream econometric work has shifted toward individual equations or incomplete models.[19]

The last of the Cowles applied econometric studies at Chicago was *A Statistical Study of Livestock Production and Marketing* (Hildreth and Jarrett 1955).[20] It is based on a model of seven equations, describing the livestock production function, the farm demands for feed grain and for protein feed, the farm supply of livestock, the demand for livestock, and supply equations for feed grains and protein feeds. The first five equations were estimated by LIML using annual U.S. data for 1920–1949. Estimated standard deviations were given. For the produc-

---

[19] "The recurrent theme [of Epstein's book] is the persistent gap between the theoretical and empirical achievements of structural estimation" (Epstein 1987, p. 3). Friedman remains skeptical of mainstream econometrics' treatment of single equations such as the demand for money (Friedman and Anna Schwartz 1982, the attack by David Hendry and Neil Ericsson 1991, and the Friedman and Schwartz reply 1991).

[20] Among other applied pieces by Cowles staff in Chicago were the following. Marschak and William H. Andrews, Jr. (1944) estimated the relation between inputs and outputs in the productive process. George Borts (1952) estimated the railroad production function. His study differed perforce from most production function studies because the common carrier law implies that output is a predetermined variable for the railroads, not a decision variable. Allen (1954) estimated a six-equation model of the market for linseed and flaxseed oil. All equations were identified (some were overidentified). They were estimated by LIML using quarterly data for 1926–1939. Estimated standard deviations were given. Forecasts were made for the four quarters of 1940. For most variables in most quarters, naive forecasts assuming no change from 1939 were better than the model's forecasts.

tion function the Anderson-Rubin test rejected the hypothesis that the overidentifying restrictions are correct, but for four other equations it accepted the overidentifying restrictions.

Particularly praiseworthy is Hildreth and Jarrett's procedure for testing whether the estimated equations predict future data well or not: the tests were done with 1950 data, which were not only outside of the sample used for estimation, they were not even available when the model was specified. This means that when the authors chose their model, they could not have been influenced by knowledge of the data that they were going to try to predict. Hence their test is even more stringent than one based on a model's ability to predict data that were already known to the model-builders when they built the model.[21] Most of the model's equations described 1950 as well as they did the sample data, and most of them had smaller errors in 1950 than did the naive models mentioned above.

Many of the Cowles applied studies presented their data, so that other researchers could verify their computations.

6. *The Cowles Assumptions*

The assumptions under which the Cowles workers derived the asymptotic distributions and properties of their estimators were quite strong. There has been much discussion and criticism of these assumptions in the subsequent literature, especially with regard to their realism, and many modifications have been proposed. (Friedman's criticisms of the Cowles approach were not based on its assumptions, for in his well-known piece on the methodology of economics (1953) he argued that a theory should be judged by its results, rather than by whether its assumptions are realistic or not.) In most of the Cowles theoretical work, the assumptions included the following:

(a) Economic behavior is governed by simultaneous equations. This assumption has been questioned by Herman Wold (1953). He contends that economic relations cannot be simultaneous, but must be recursive, on the ground that every economic action is a response to some previous action (however short the response time may be). Robert Strotz and Wold (1960) present an interesting explanation of how simultaneous equations can be appropriate tools even if Wold is correct in principle: Suppose that economic interactions are always sequential, not simultaneous, but that there is always a very short delay, say one day, between any agent's action and the next agent's subsequent response; suppose also that all our data correspond to aggregates or averages over a time-interval that is much longer than the one-day response delay, say one calendar year. Then it is not possible to fit a model to the sequence of decisions that occur at one-day intervals: the required daily data are missing. However, if we fit a simultaneous equations model to annual data, we will be approximately correct. For example, if demand decisions follow price-setting decisions by one day, then our demand equation should say that total demand during a calendar year (January 1 through December 31) depends on the average price during the 365-day period starting and ending one day earlier (from December 31 of the preceding year through December 30 of the current year), but we do not make a large error by saying that it depends on the average price during the calendar year. Wold's plea for recursive systems did not convince many economists, and the simultaneous equations framework is widely accepted. Of course, lagged variables can

---

[21] A defense of this more stringent test is given by Christ (1966, pp. 546–48, and 1993).

be (and often are) included in simultaneous equations to reflect dynamic effects.

(b) The equations of the model are linear in systematic variables and disturbances. This assumption is strong, but may be approximately correct for short periods of time. As noted above, Anderson and Rubin weakened it to require only linearity in unknown parameters, without destroying most of the theoretical econometric results. This weakening permits the handling of the types of nonlinearity that can be removed by a suitable transformation, such as in Cobb-Douglas equations where the appropriate transformation is to take logarithms. In both its forms it is testable, but the Cowles group did not devise tests of it. A procedure for deciding whether a relationship is linear or logarithmic is provided by George E. P. Box and D. R. Cox (1964). Empirical studies have increasingly used nonlinear models in recent years. This represents an advance in technique (not in principle) that has been made possible by the advent of electronic computers and by the derivation of nonlinear estimators (Stephen Goldfeld and Richard Quandt 1972; Takeshi Amemiya 1983).

(c) The systematic variables are observable, without error. The first part of this assumption is strong. Because agents' expectations about the future are not directly observable, they are ruled out of consideration unless one can represent them by observable variables. More about this later. Errors of measurement of variables are real, but most Cowles workers ignored them in favor of disturbances to equations. Exceptions were Koopmans (1937), as noted above, and Chernoff and Rubin (1953).

(d) Variables change at discrete time intervals, such as yearly, or quarterly, or monthly, rather than continuously. This assumption is made essentially for convenience, to allow the use of difference equations to explain dynamic relations among annual, quarterly, or monthly data. It has been adopted in most econometric work.[22]

(e) It is known which variables are exogenous, and which are predetermined at time t. This assumption is strong. If variables are erroneously specified to be predetermined or exogenous, estimators become biased, just as least squares estimators are biased in simultaneous equations systems. More about this later.

(f) The determinant of the coefficients of the jointly dependent variables is not zero, so that the reduced form exists. This assumption has to be made: if it were not true, the model would be useless, because it could not determine the values of the jointly dependent variables.

(g) The predetermined variables are linearly independent, so that their moment matrix can be inverted. This assumption requires that the sample size be greater than or equal to the number of predetermined variables in the model. If that is so, it is easy to satisfy, for if it were false, one could make it true without loss simply by omitting one of the linearly dependent predetermined variables (or more than one, if necessary).

(h) The structural equations are identified by a priori restrictions on their parameters (possibly including restrictions on the parameters of the distribution of disturbances), and these restrictions are correct. This assumption has been severely criticized long ago by Ta-Chung Liu (1960), and more recently by Christopher Sims (1980; see below). As noted earlier, overidentifying restrictions have often been rejected.

(i) The disturbances are normally dis-

---

[22] Clifford Wymer (1972) and Giancarlo Gandolfo (1981) have begun to use continuous-time models and differential equations instead.

tributed random variables with zero means and finite and constant variances and covariances, and their covariance matrix is nonsingular. This assumption is strong. However, the Cowles group showed that normality is not necessary for most of their results. The assumption of zero means has been discussed above, in Section 3a. The assumption of a nonsingular covariance matrix is plausible unless the system includes definitional equations that have no disturbances, but this can be handled by eliminating those equations by substitution (this step should be taken prior to estimation of the complete system as a whole).

(j) The disturbances are serially independent. This assumption is testable, but the Cowles group did not devise tests for it. Instead, they used available tests. At first they used the test of B. I. Hart and John Von Neumann (1942). Hildreth and Jarrett (1955) used the more appropriate test of James Durbin and Geoffrey S. Watson (1950, 1951). Serial correlation of disturbances constituted a serious difficulty for the Cowles workers, because most of their theorems assumed it away. (Serial correlation of observed variables is very common, and is no problem.) They did not devote much effort to finding estimators that were appropriate in its presence. Subsequent work has provided solutions to this problem for many cases (John Denis Sargan 1959; Ray C. Fair 1972).

(k) The equation system is dynamically stable. This is assured by assuming that the exogenous variables are considered fixed in repeated samples, and their moments have finite limits as the sample size goes to infinity, and all characteristic roots of the final equations of the system are less than one in absolute value. Some such condition is needed to prove that estimators are consistent and to derive their asymptotic variances and covariances.

The Cowles derivations of the properties of their estimators were almost entirely asymptotic, that is, for infinitely large samples, which we never have in practice. Of course, there is some comfort in knowing that one's estimators are consistent, and asymptotically unbiased and efficient, but this tells nothing about how large the bias or variance may be in samples of size 20 or 30 years, or 60 quarters, which were the sample sizes used in their applied work. Small sample properties are difficult to derive, and the Cowles group made little progress with this problem. The first successes were for very special cases of two or three equations (Basmann 1961, 1963) but progress has now been made with more general cases (Anderson 1982; Peter C. B. Phillips 1983).

Most of the Cowles group dealt with time-series data, because of their interest in the amelioration of business cycles. Their methods are applicable to pure cross section data, where t = 1,2,3, . . . stands not for the first, second, and third time periods, etc., but for the first, second, and third members of the cross section sample, etc. Note, however, that the concept of serial correlation makes little sense when the data are for cities or counties or families or firms, unless they are arranged in some meaningful order, such as size.

Panel data, consisting of a time series of cross section data, were rare in the 1940s, and the Cowles group did essentially nothing with the problems of estimation from such data.

7. *An Evaluation of the Cowles Econometric Contributions*

a. *A Recapitulation*

Let us begin this section with a reminder of what econometrics is about, and a look back at the main features of the Cowles program. Econometrics has been aptly defined by a committee of the Econometric Society as

> the quantitative analysis of actual economic phonomena based on the concurrent development of theory and observation, related by appropriate methods of inference. (Paul Samuelson, Koopmans, and Stone 1954, p. 142)

We have seen that the two mainstays of the Cowles revolution in econometric methods were the probability approach and the concept of a simultaneous equations model.

The probability approach is well established in econometrics, and is no longer in question. Experience has shown that (except for definitional identities) no quantitative theoretical economic relationship fits observed data exactly. Hence each econometric equation is formulated so as to include not only observed variables, but also an unobservable catch-all variable representing the discrepancy between the computed and observed values of the equation. This catch-all variable is treated as stochastic. If plausible assumptions (sometimes testable) are made about the probability distributions of such stochastic variables, then plausible conclusions follow about the distributions of estimators of parameters of the equations, and plausible statistical inferences can be drawn. With the increased attention given to behavior under uncertainty, stochastic elements are also becoming important in economic theory. This is especially evident in the theory of rational expectations. Estimation methods have been developed to take account of rational expectations (Kenneth F. Wallis 1980; Frederic Mishkin 1983). Hence I believe that the obituary in the last sentence of Morgan's otherwise excellent book (1990, p. 264) is premature: "By the 1950s the founding ideal of econometrics, the union of mathematical and statistical economics into a truly synthetic economics, had collapsed."

Likewise, simultaneous equations models are well established in econometrics—and in economics in general. Economic phenomena are the result of interactions among many agents. Any single economic relationship in which we may be interested interacts with other relationships. Hence we are led from single equations to systems of several equations. The most general case is that of simultaneous equations. Some economists believe that in principle all economic equation systems are recursive, but even if this is correct, as noted earlier, a simultaneous system can be regarded as approximately correct. Hence most economists think in terms of simultaneous equations systems, even when investigating a single equation at a time.

We have seen that the main Cowles contributions in theoretical econometrics were the solution of the identification problem and the development of methods of estimation that have desirable properties for simultaneous equations. These were impressive intellectual achievements, and they have profoundly affected econometrics. But there is controversy even today about their value for applied work. The controversy relates not so much to the probability approach (which is almost universally accepted in econometrics), but rather to the nature and implications of the concept of simultaneous equations. In what follows we shall first consider in several subsections the process of proposing and provisionally adopting the specifications of an econometric model, then consider the present position of the Cowles simultaneous equations estimation methods, and finally sum up.

b. *Why Is Identification Interesting?*

Consider first the identification problem. Remember that structural equations cannot be estimated unless they can be identified. Why is it important to be able to identify and estimate structural equations? After all, the reduced form of the model is what one needs in order to make predictions about the effects of policy changes upon endogenous variables. So why do we care about the structural equations that underlie the reduced form? There are three reasons.

One is that if a structural model is correctly specified and is overidentified, then more efficient estimates of the reduced form can be obtained by solving the estimated structure than by estimating the reduced form by least squares directly. This is implied by the discussion of overidentified equations at the end of Section 4e.

A second is that if the results of using a model are unsatisfactory, it is important to find out what part or parts of the model are at fault. One can test each estimated structural equation separately in order to see which, if any, are performing well, and which one or ones are not. Then one knows which parts of the model to revise.

A third reason is that the economy's structure may undergo a change after it has been estimated, but before the period that one wants to predict. The possibility of a change in the economy's structure was of deep concern to the Cowles group. Marschak (1947, 1953) correctly pointed out that if there is a change in the structure of the economy, that is, if the value of a structural parameter changes (or worse, if the form of a structural equation changes), then in general every reduced form equation changes, and therefore the old reduced form is no longer relevant for forecasting. If one has not estimated the old structure, one has no way of knowing how the reduced form will change. But if one has estimated the old structure, and if one knows the time and the nature of the structural change, one can then find the new structure, and can solve it to get the new reduced form, and hence can use it to make forecasts for the period after the change. This argument foreshadows the critique of Robert Lucas (1976; see below). Marschak would have agreed with Lucas that it is important to know whether the structure of the economy has changed, and if so, when and how, if one is to make good forecasts from parameter estimates that are based on past data.

c. *The Attacks of Liu and Sims*

The information that would enable one to identify a model comes from the theoretical restrictions that are imposed on the model. There is no controversy over the correctness of the theorems that tell whether a model can be identified. The controversy is over whether information sufficient to identify econometric models can be obtained in practice.

The main attacks have come from Liu (1960) and Sims (1980). Both are based on the view that the simultaneous interactions of economic variables are so pervasive that most structural relationships contain all or nearly all of the variables in the economy, and are therefore not in fact identified. If so, it is pointless to try to estimate them. Liu concluded that in such cases we are forced to retreat to the estimation of reduced form equations, each of which expresses an endogenous variable as a function of exogenous variables and possibly of lagged endogenous variables. Of course this would mean that we are unable to use the Marschak strategy to deduce the new values of reduced form parameters that will prevail if a structural change occurs after we have estimated the old reduced form.

Sims also regards the restrictions that would be needed to identify structural equations as "incredible." But his recommendation is different from Liu's. He proposes the use of vector autoregression (VAR). With this method, one chooses a set of variables thought to be relevant to the problem at hand, and estimates the regression of each one on lagged values of itself and all the others. No exogenous variables are involved (but some could be included, thus making Sims' recommendation more like Liu's). No economic theory is needed to specify a priori restrictions on structural relationships. The only role for economic theory is to assist in the choice of variables and the length of lags (though the latter is usually decided empirically—it cannot be so large that the number of parameters in any single VAR equation exceeds the size of the available sample). Of course a VAR can be regarded as the reduced form of some structural model, one that contains the same variables as the VAR. Epstein (1987, pp. 219–20) properly faults VAR for providing no way to distinguish between competing theories of the way the economy responds to policy changes.

In my view we have sufficient information to identify many structural equations. The usefulness of supply and demand analysis lies in the fact that many of the factors that affect the demand curve of a good have little or no effect on the supply curve, and vice versa. The same is true of saving and investment analysis, and many other types of economic analysis. And we have the Basmann (1960) test for overidentifying restrictions. It may be the case that some of the restrictions used for identification purposes are only approximately correct. In an important paper Franklin Fisher (1961) explains why it makes sense to use such restrictions.

Related to VAR is another approach which also downplays economic theory, autoregressive integrated moving average (ARIMA) time series analysis (Box and Gwilym M. Jenkins 1970). Here a function of current and lagged values of an observable variable is set equal to a function of current and lagged values of a white noise disturbance (random, serially independent, with zero mean). The autoregressive (AR) part gives the variable's current value in terms of its own past values; the integrated (I) part is needed to undo the difference operator if differences (of first or higher order) rather than levels of the variable are involved; and the moving average (MA) part describes the way the current and lagged white noise disturbances are combined. An ARIMA model implies that the current value of a variable can be explained as a combination of current and (perhaps infinitely many) past values of a white noise disturbance. ARIMA models can be univariate or multivariate. VAR is a special case of a multivariate ARIMA.

Zellner and Franz Palm (1974) propose an ingenious combination of Cowles-type structural modeling and ARIMA modeling: they suggest using a structural model to explain the behavior of endogenous variables in terms of exogenous variables, and an ARIMA model to explain the the behavior of exogenous variables. They show that if this is done, then the behavior of all variables, endogenous as well as exogenous, is modeled by ARIMA processes. But it strikes me as most unsatisfactory to explain everything as the outcome of combinations of white noise disturbances.

d. *The Lucas Critique*

Lucas (1976) issued a serious warning regarding the use of estimated econometric models to predict the effects of future changes in economic policy. He pointed out that if some of the parameters of a model reflect the adaptation of private behavior to a previously maintained policy reaction function, then, if the policy reaction function is changed, private behavior will re-adapt to the new function, and as a result the previously estimated parameters will no longer describe the response of the economy. He and others have given likely examples, based on the hypothesis of rational expectations. This argument is a more sophisticated treatment of structural change than Marschak's. Marschak did not display an explicit understanding of the Lucas critique, though in my view he came close to it.

The Lucas critique implies that if a change in a policy reaction function occurs which changes private behavior parameters, then the reduced form of the system will change, as will any VAR or ARIMA that previously described the system. Hence, in my view, the Lucas critique is just as damaging to Liu's and Sims' procedure, and to ARIMA, as it is to the Cowles procedure. The proper response to the Lucas critique is to specify econometric models in such a way that their parameters are invariant to changes in public policy response functions. Admittedly this is easier said than done.

Thomas Sargent (1976) expressed Lucas's point in a different way. He noted that if the monetary policy regime is to be changed from a constant-money-growth rule to a feedback rule, it is possible to write the reduced form of the system in two versions that are observationally equivalent during a sample period under the constant-money-growth rule, in the sense that both versions describe the sample data with identical residuals, but one of the reduced forms suffers a parameter change when the regime changes, and the other does not. The difference between the two reduced forms lies in the manner in which expectations are formed, and the manner in which they affect behavior. Old-regime estimates of the reduced form, in either version, or in both versions, will not reveal which version will change when the regime changes, and which will remain invariant. Hence reduced-form estimation alone will not settle the question of whether a feedback rule is better than a constant-money-growth rule.

e. *Model Specifications: Where Do They Come From?*

The Cowles Commisssion approach to econometrics was built on the premise that correct a priori specifications were already available for the models that were to be estimated by its methods. The Cowles theoretical econometric work did not have much to say about the process of specifying models, rather taking it for granted that economic theory would do that, or had already done it.

Ideas for econometric model-building can certainly come from economic theory, by way of the postulate that consumers and producers seek to optimize subject to the constraints they face. But ideas can also come from examining data and searching for patterns. Economic theory is not powerful enough today, and may never be, to tell us everything we want to know about the specification of our models, that is, the variables to be included in each equation, the classification of variables as endogenous or exogenous, the mathematical form of equations (linear, logarithmic, quadratic, or the like), the nature of the random disturbances, and so on. And although economic theory is rather good at deriving equilibrium relationships, it is not very good at deriving the adjustment behavior of a system that is out of equilibrium. Thus there is much room for empirical investigation in the process of model building, even before one arrives at the problem of parameter estimation.[23]

Both VAR and ARIMA can be characterized, with some exaggeration, as mindless data-mining, because both

---

[23] Heckman in his review (1992) of Morgan (1990) takes a similar view, and objects to Morgan's narrow implicit definition of econometrics.

share the attribute that no use is made of economic knowledge except to choose the variable(s) to be studied and the length of the lags to be used. One can apply these techniques to economic data without knowing much about economics, or to biological data without knowing much about biology. However, data mining can certainly be justified as a method of searching for regularities in the data. And when regularities are found that persist from decade to decade, or from place to place, then there is something for economic theory to try to understand.

In the celebrated exchange between Koopmans and Rutledge Vining concerning Burns and Mitchell's *Measuring Business Cycles* (1946), the magnum opus of the National Bureau of Economic Research as of that time, Koopmans (1947) began by attacking Burns and Mitchell for having no theory about business cycles, and hence having no clear idea about what to measure or what hypotheses to test. Rutledge Vining (1949) replied that Koopmans appeared to reject any style of research that differed from his own, and defended the attempt to increase our understanding of business cycles by careful observation and recording of many economic variables that behave cyclically. In his rejoinder Koopmans (1949) conceded that he and Vining were not so far apart after all, but he reiterated his defense of the Cowles approach. Perhaps as a result of this exchange, the Cowles Commission's original motto, adopted in 1932 and emblazoned on the covers and title pages of early monographs, "Science is Measurement," was changed in 1952, at the suggestion of Hildreth, to "Theory and Measurement" (Christ 1952, p. 62).

f. *Model Specifications: Are They Correct?*

Econometric model specifications must be tested before they can be accepted as correct, or as approximately correct. The Cowles econometric theory workers did not have much to say about how to decide whether a model had been correctly specified or not.

In any applied econometric study, the distinction between variables whose behavior is to be explained and variables whose behavior is to be taken as given is fundamental. This corresponds to the distinction between endogenous and exogenous variables. Although as noted earlier Koopmans (1950b) had given a rigorous statistical definition of exogeneity, the Cowles group had no method of testing whether the designation of variables as exogenous had been done correctly or not. Several such tests are now available (Sims 1972; De-Min Wu 1973; John Geweke 1978, 1984; J. A. Hausman 1978, Robert F. Engle, David Hendry, and Jean-Francois Richard 1983; and Engle 1984).

When post-sample data demand the rejection of the hypothesis that an estimated equation has not changed, there are several possibilities: (1) The form of the equation was correctly specified, and did not change, but its parameters changed after the sample period. (2) The form of the equation was correct during the sample period, but changed after the sample period. (3) Both the form and parameters of the true equation are unchanged, but the form specified by the model builder was incorrect in the first place, so that the estimated equation is only an approximation to the true one, and the approximation no longer held in the post-sample period. For example, suppose that the true relation between two variables is a curve that rises until 1993 and falls thereafter, but a linear equation is fitted to the data before 1993. The fitted equation does not turn down after 1993, but the true relation does, so that post-1993 forecast errors will be large, and will reject the hypothesis that the fitted straight line describes the new data.

The appropriate research strategy differs among these cases: if the form of the equation is correctly specified and only its parameters have changed, one should re-estimate it with post-change data, but if the form of the equation is incorrect, one should try a new form. The Cowles group threw little light on how to distinguish these situations. Subsequent work has made progress with this important problem; among the pioneers are Hausman (1978) and Edward Leamer (1978).

g. *The Cowles Simultaneous Equations Estimation Methods*

For simultaneous equations models the Cowles theorems about least squares bias and about the consistency of limited and full information maximum likelihood estimators are correct, but these methods are not the most commonly used methods today.

The limited information maximum likelihood method (LIML) has become a curiosum, displaced by the two-stage least squares (2SLS) method of Henri Theil (1953, 1958) and Basmann (1957). The reasons are that 2SLS is much simpler to compute, not requiring iterative computations as does LIML; LIML occasionally gives highly implausible estimates in small samples, which 2SLS does not,[24] and the two methods are asymptotically equivalent. For just identified equations, 2SLS gives the same results as LIML, indirect least squares, and instrumental variables. For overidentified equations, it has a better method than LIML for coming as close as possible to satisfying the overidentifying restrictions.[25]

The full information maximum likelihood method (FIML) has been largely displaced by three-stage least squares (3SLS), because the two are asymptotically equivalent, and 3SLS is simpler to compute, not requiring iterative computations as does FIML. For a system in which every equation is just identified, 3SLS gives the same result as 2SLS (Zellner and Theil 1962).

Experience suggests that the problem of least squares bias is not very serious. Indeed, it is unusual for least squares estimates to differ much from 2SLS. Hence the least squares method is still quite commonly used even for equations that are part of a simultaneous equations sytem. When a method free of least squares bias is desired, the instrumental variables method is often used because it is so simple to work with, but many popular econometrics programs for personal computers now include 2SLS, and some even include 3SLS and FIML.

8. *A Summing Up*

Although the Cowles Commission had an indelible effect on the field of econometrics, its lasting contributions are less than what its econometric pioneers hoped for. Those lasting contributions are principally in theoretical rather than applied econometrics: the probability approach, simultaneous equations models, the distinction between structural and reduced-form equations, the distinction between endogenous and exogenous variables, the identification problem and its theoretical solution, and the initiation of research on the construction and statis-

---

[24] Theil (1958, 1961) shows a series of graphs that explain why LIML occasionally yields weird estimates while 2SLS does not. Klein and Mitsugu Nakamura (1962) also throw light on this question.

[25] 2SLS makes the variance difference as close to zero as possible. The variance difference is the difference between the two variances that are described in footnote 15.

tical properties of estimators. It has been left to others to fill the lacunae in the Cowles theoretical econometric effort, such as model specification, specification tests, endogenizing of policy variables, time-varying parameters, serial correlation and heteroskedasticity of disturbances, errors of measurement, and the development of more useful estimators such as two-stage and three-stage least squares and estimators for nonlinear models.

As for applied results, macroeconometric models have established a short-term forecasting record good enough to keep them in business. But Alfred Cowles' dream of predicting the stock market, and Marschak's dream of predicting the effects of economic policy variables so as to control business cycles, are not yet realized.

REFERENCES

ALLEN, STEPHEN G. "Inventory Fluctuations in Flaxseed and Linseed Oil, 1926–1939," *Econometrica*, July 1954, 22(3), pp. 310–27.

AMEMIYA, TAKESHI. "Non-Linear Regression Models," in GRILICHES AND INTRILIGATOR 1983, pp. 333–89.

ANDERSON, THEODORE W. "Small-Sample Distribution Theory," in *Advances in econometrics*. Econometric Society Monographs in Quantitative Economics, Econometric Society Publication No. 2. Ed.: WERNER HILDENBRAND. Cambridge, Eng.: Cambridge U. Press, 1982, pp. 109–22.

______. "Trygve Haavelmo and Simultaneous Equation Models," *Scand. J. Statist.*, Jan. 1991, *18*(1), pp. 1–19

ANDERSON, THEODORE W. AND RUBIN, HERMAN. "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Ann. Math. Statist.*, Mar. 1949, *20*(1), pp. 46–63.

______. "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Ann. Math. Statist.*, Dec. 1950, *21*(4), pp. 570–82.

BASMANN, ROBERT L. "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," *Econometrica*, Jan. 1957, 25(1), pp. 77–83

______. "On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics," *J. Amer. Statist. Assoc.*, Dec. 1960, *55*(292), pp. 650–59.

______. "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in Two Leading Overidentified Cases," *J. Amer. Statist. Assoc.*, Sept. 1961, *56*(295), pp. 619–36.

______. "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in a Leading Three Equation Case," *J. Amer. Statist. Assoc.*, Mar. 1963, *58*(301), pp. 161–71.

BORTS, GEORGE H. "Production Relations in the Railway Industry," *Econometrica*, Jan. 1952, *20*(1), pp. 71–79.

BOX, GEORGE E. P. AND COX, D. R. "An Analysis of Transformations," *J. Roy. Statist. Soc.*, *Series B*, 1964, *26*(2), pp. 211–52.

BOX, GEORGE E. P. AND JENKINS, GWILYM M. *Time series analysis: Forecasting and control*. San Francisco: Holden-Day, 1970.

BRONFENBRENNER, JEAN. "Sources and Size of Least-Squares Bias in a Two-Equation Model," Chapter IX in HOOD AND KOOPMANS 1953, pp. 221–35.

BRUMBERG, RICHARD E. "An Approximation to the Aggregate Saving Function," *Econ. J.*, Mar. 1956, *66*(2), pp. 66–72.

BURNS, ARTHUR F. AND MITCHELL, WESLEY C. *Measuring business cycles*. New York: National Bureau of Economic Research, 1946.

CHERNOFF, HERMAN AND DIVINSKY, NATHAN. "The Computation of Maximum-Likelihood Estimates of Linear Structural Equations," Chapter X in HOOD AND KOOPMANS 1953, pp. 236–302.

CHERNOFF, HERMAN AND RUBIN, HERMAN. "Asymptotic Properties of Limited-Information Estimates Under Generalized Conditions," Chapter VII in HOOD AND KOOPMANS 1953, pp. 200–12.

CHRIST, CARL F. "A Test of an Econometric Model for the United States, 1921–1947," in *Conference on business cycles*. Universities-National Bureau Committee for Economic Research, New York: National Bureau of Economic Research, 1951, pp. 35–107.

______. "History of the Cowles Commission, 1932–1952," in *Economic theory and measurement: A twenty year research report 1932–1952*, Chicago: Cowles Commission for Research in Economics, 1952, pp. 3–65.

______. *Econometric models and methods*. New York: Wiley, 1966.

______. "Early Progress in Estimating Quantitative Economic Relationships in America," *Amer. Econ. Rev.*, Dec. 1985, 75(6), pp. 39–52.

______."Assessing Applied Econometric Results," *Fed. Res. Bank St. Louis Rev.*, Mar./Apr. 1993, *75*(2), pp. 71–94.

COBB, CHARLES W. AND DOUGLAS, PAUL H. "A Theory of Production," *Amer. Econ. Rev.*, Mar. 1928, *18*(1), pp. 139–65.

COWLES, ALFRED. "Can Stock Market Forecasters Forecast?"*Econometrica*, July 1933, *1*(3), pp. 309–24.

DURBIN, JAMES AND WATSON, GEOFFREY S. "Testing for Serial Correlation in Least Squares Regression. I and II," *Biometrika*, Dec. 1950, *37*(4), pp. 409–28; June 1951, *38*(2), pp. 159–78.

ENGLE, ROBERT F. "Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics," Chapter 13 in GRILICHES AND INTRILIGATOR, Vol. II, 1984, pp. 776–826.

ENGLE, ROBERT F.; HENDRY, DAVID F. AND RICHARD, JEAN-FRANCOIS. "Exogeneity," *Econometrica*, Mar. 1983, *51*(2), pp. 277–304.

EPSTEIN, ROY J. *A history of econometrics*. Contributions to Economic Analysis No. 165. Amsterdam: North-Holland, 1987.

EZEKIEL, MORDECAI. "Statistical Analyses and the 'Laws' of Price," *Quart. J. Econ.*, Feb. 1928, *42*(1), pp. 199–227.

FAIR, RAY C. "Efficient Estimation of Simultaneous Equations with Auto-Regressive Errors by Instrumental Variables," *Rev. Econ. Statist.*, Nov. 1972, *54*(4), pp. 444–49.

FISHER, FRANKLIN M. "On the Cost of Approximate Specification in Simultaneous Equation Estimation," *Econometrica*, Apr. 1961, *29*(2), pp. 139–70.

______. *The identification problem in econometrics*. New York: McGraw-Hill, 1966.

FRIEDMAN, MILTON. Review of Tinbergen (1939), *Amer. Econ.Rev.*, Sept. 1940, *30*(3), pp. 657–61.

______. "Comment" (on Christ 1951), in *Conference on business cycles*. Universities-National Bureau Committee for Economic Research. New York: National Bureau of Economic Research, 1951, pp. 107–14.

______. "The Methodology of Positive Economics," in *Essays in positive economics*. Chicago: U. of Chicago Press, 1953, pp. pp. 3–43.

______. *A theory of the consumption function*. Princeton, NJ: Princeton U. Press for the National Bureau of Economic Research, 1957.

FRIEDMAN, MILTON AND BECKER, GARY. "A Statistical Illusion in Judging Keynesian Models," *J. Polit. Econ.*, Feb. 1957, *65*(1), pp. 64–75.

______."Reply" to Klein (1958), *J. Polit. Econ.*, Dec. 1958, *66*(6), pp. 545–49

FRIEDMAN, MILTON AND SCHWARTZ, ANNA J. *Monetary trends in the United States and the United Kingdom, Their relation to income, prices, and interest rates, 1867–1975*. Chicago: U. of Chicago Press for the National Bureau of Economic Research, 1982.

______. "Alternative Approaches to Analyzing Economic Data," *Amer. Econ. Rev.*, Mar. 1991, *81*(1), pp. 39–49.

FRISCH, RAGNAR. "Pitfalls in the Statistical Construction of Demand and Supply Curves," *Veröffentlichungen der Frankfurter Gesellschaft für Konjunkturforschung*, Neue Folge, Heft 5. Leipzig: Hans Buske, 1933.

______. "Statistical Versus Theoretical Relations in Economic Macrodynamics." A memorandum prepared for the Business Cycle Conference at Cambridge, England, July 18th-20th, 1938, to discuss Professor J. Tinbergen's publications of 1938 for the League of Nations. Mimeographed, 19 pp.

GANDOLFO, GIANCARLO. *Qualitative analysis and econometric estimation of continuous time dynamic models*. Amsterdam: North-Holland, 1981.

GEARY, ROBERT C. "Determination of Linear Relations Between Systematic Parts of Variables with Errors of Observation the Variances of Which Are Unknown," *Econometrica*, Jan. 1949, *17*(1), pp. 30–58.

GEWEKE, JOHN. "Testing the Exogeneity Specification in the Complete Dynamic Simultaneous Equation Model," *J. Econometrics*, Apr. 1978, *7*(2), pp. 163–85.

______. "Inference and Causality in Economic Time Series Models," in GRILICHES AND INTRILIGATOR, Vol. II, 1984, pp. 1102–44.

GIRSHICK, MEYER A. AND HAAVELMO, TRYGVE. "Statistical Analysis of the Demand for Food: Examples of Simultaneous Estimation of Structural Equations," *Econometrica*, Apr. 1947, *15*(2), pp. 79–110; Abridged as Chapter V in HOOD AND KOOPMANS 1953, pp. 92–111.

GOLDFELD, STEPHEN M. AND QUANDT, RICHARD E. *Nonlinear methods in econometrics*. Amsterdam: North-Holland, 1972.

GRANGER, CLIVE W. J. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, July 1969, *37*(3), pp. 424–38.

GRILICHES, ZVI AND INTRILIGATOR, MICHAEL D., eds. *Handbook of econometrics*, Vols. I, II, and III. Amsterdam: North-Holland, 1983, 1984, 1986.

HAAVELMO, TRYGVE. "The Statistical Implications of a System of Simultaneous Equations," *Econometrica*, Jan. 1943, *11*(1), pp. 1–12.

______. "The Probability Approach in Econometrics," *Econometrica*, July 1944, *12*(Supplement), pp. 1–115; a preliminary mimeographed version entitled "On the Theory and Measurement of Economic Relations" was issued in Cambridge, MA in 1941.

______. "Methods of Measuring the Marginal Propensity to Consume," *J. Amer. Statist. Assoc.*, Mar. 1947, *42*(237), pp. 105–22; reprinted as Chapter IV in HOOD AND KOOPMANS 1953, pp. 75–91.

HART, B. I. AND VON NEUMANN, JOHN. "Tabulation of the Probabilites for the Ratio of the Mean Square Successive Difference to the Variance," *Ann. Math. Statist.*, June 1942, *13*, pp. 207–14.

HAUSMAN, J. A. "Specification Tests in Econometrics," *Econometrica*, Nov. 1978, *46*(6), pp. 1251–71.

HECKMAN, JAMES. "Haavelmo and the Birth of Modern Econometrics: A Review of *The History of Econometric Ideas* by Mary Morgan," *J. Econ. Lit.*, June 1992, *30*(2), pp. 876–86.

HENDRY, DAVID F. AND ERICSSON, NEIL R. "An Econometric Analysis of U. K. Money Demand in *Monetary Trends in the United States and the United Kingdom* by Milton Friedman and Anna J. Schwartz," *Amer. Econ. Rev.*, Mar. 1991, *81*(1), pp. 8–38.

HILDRETH, CLIFFORD. *The Cowles Commission in Chicago, 1939–1955. Lecture notes in economics and mathematical systems*. Vol. 271. Berlin: Springer, 1986.

HILDRETH, CLIFFORD AND JARRETT, FRANCIS G. *A statistical study of livestock production and market-*

ing. Cowles Commission Monograph 15. New York: Wiley, 1955.

Hood, William C. and Koopmans, Tjalling C., eds. *Studies in econometric method.* Cowles Commission Monograph 14. New York: Wiley, 1953.

Hotelling, Harold. "Problems of Prediction," *Amer. J. Sociol.*, July 1942, *48*(1), pp. 61–76.

Hurwicz, Leonid. Five papers in Koopmans 1950a.

Klein, Lawrence R. *Economic fluctuations in the United States, 1921–1941.* Cowles Commission Monograph 11. New York: Wiley, 1950.

———. "The Friedman-Becker Illusion," *J. Polit. Econ.*, Dec. 1958, *66*(6), pp. 539–45.

Klein, Lawrence R. and Nakamura, Mitsugu. "Singularity in the Equation Systems of Econometrics: Some Aspects of the Problem of Multicollinearity," *Int. Econ. Rev.*, Sept. 1962, 3(3), pp. 274–99.

Koopmans, Tjalling C. *Linear regression analysis of economic time series.* Haarlem: de Erven F. Bohn, 1937.

———. "Measurement without Theory," *Rev. Econ. Statist.*, Aug. 1947, 29(3), pp. 161–72.

———. "Identification Problems in Economic Model Construction," *Econometrica*, Apr. 1949a, *17*(2), pp. 125–44; reprinted with minor revisions as Chapter II in Hood and Koopmans 1953, pp. 27–48.

———. "A Reply" (to Vining 1949), *Rev. Econ. Statist.*, May 1949b, *31*(2), pp. 86–91.

———, ed. *Statistical inference in dynamic economic models.* Cowles Commission Monograph 10. New York: Wiley, 1950a.

———. "When Is an Equation System Complete for Statistical Purposes?" 1950b; Chapter XVII in Koopmans 1950a, pp. 393–409.

Koopmans, Tjalling C. and Hood, William C. "The Estimation of Simultaneous Linear Economic Relationships," Chapter VI in Hood and Koopmans 1953, pp. 112–99.

Koopmans, Tjalling C.; Rubin, Herman and Leipnik, Roy B. "Measuring the Equation Systems of Dynamic Economics," Chapter II in Koopmans 1950a, pp. 53–237.

Leamer, Edward E. *Specification searches.* New York: Wiley, 1978.

Lenoir, Marcel. *Etudes sur la formation et le mouvement des prix.* Paris: Giard et Briere, 1913.

Liu, Ta-Chung. "Underidentification, Structural Estimation, and Forecasting," *Econometrica*, Oct. 1960, *28*(4), pp. 855–65.

Lucas, Robert E., Jr. "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conf. Ser. Public Policy*, 1976, *1*, pp. 19–46.

Mann, Henry B. and Wald, Abraham. "On the Statistical Treatment of Linear Stochastic Difference Equations," *Econometrica*, July-Oct. 1943, *11*(3&4), pp. 173–220.

Marschak, Jacob. "Economic Structure, Path, Policy, and Prediction," *Amer. Econ. Rev.*, May 1947, *37*(2), pp. 81–84.

———. "Statistical Inference in Economics: An Introduction," Chapter I in Koopmans 1950a, pp. 1–50.

———. "Economic Measurements for Policy and Prediction," Chapter I in Hood and Koopmans 1953, pp. 1–26.

Marschak, Jacob and Andrews, William H., Jr. "Random Simultaneous Equations and the Theory of Production," *Econometrica*, July-Oct. 1944, *12*(3&4), pp. 143–205.

Marshall, Andrew W. "A Test of Klein's Model III for Changes of Structure." Unpub. M.A. thesis, U. of Chicago, 1949; a summary appears in Christ 1951.

McNees, Stephen K. "How Accurate Are Macroeconomic Forecasts?" *New Eng. Econ. Rev.*, July-Aug. 1988, pp. 15–36.

Mishkin, Frederic S. *A rational expectations approach to macroeconometrics.* Chicago: U. of Chicago Press for the National Bureau of Economic Research, 1983.

Mitchell, Wesley C. *Business cycles.* Berkeley: U. of California Press, 1913.

———. *Business cycles: The problem and its setting.* New York: National Bureau of Economic Research, 1927.

Morgan, Mary S. *The history of econometric ideas.* Cambridge, Eng.: Cambridge U. Press, 1990.

Phillips, Peter C. B. "Exact Small Sample Theory in the Simultaneous Equations Model," in Griliches and Intriligator 1983, pp. 449–516.

Reder, Melvin W. "Chicago Economics: Permanence and Change," *J. Econ. Lit.*, Mar. 1982, *20*(1), pp. 1–38.

Reiersøl, Olav. "Confluence Analysis by Means of Lag Moments and Other Methods of Confluence Analysis," *Econometrica*, Jan. 1941, 9(1), pp. 1–24.

———. "Confluence Analysis by Means of Instrumental Sets of Variables," *Arkiv for Mathematik, Astronomi och Fysik*, Aug. 1945, *32*(1), 4, 119 pp.

Rubin, Herman. "The Approximate Distribution of Calculated Disturbances." Cowles Commission Discussion Paper, Statistics, No. 318, duplicated, 1948.

Samuelson, Paul A.; Koopmans, Tjalling C. and Stone, J. Richard N. "Report of the Evaluative Committee for Econometrica," *Econometrica*, Apr. 1954, *22*(2), pp. 141–46.

Sargan, John Denis. "The Estimation of Relationships with Autocorrelated Residuals by the Use of Instrumental Variables," *J. Roy. Statist. Soc.*, Series B, 1959, 21(1), pp. 91–105.

Sargent, Thomas J. "The Observational Equivalence of Natural and Unnatural Rate Theories of Macroeconomics," *J. Polit. Econ.* , June 1976, *84*(3), pp. 631–40.

Schultz, Henry. *The theory and measurement of demand.* Chicago: U. of Chicago Press, 1938.

Sims, Christopher A. "Money, Income, and Causality," *Amer. Econ. Rev.*, Sept. 1972, *62*(4), 540–52.

———. "Macroeconomics and Reality," *Econometrica*, Jan. 1980, *48*(1), pp. 1–48.

Strotz, Robert H. and Wold, Herman O. A. "A

Triptych on Causal Chain Systems," *Econometrica*, Apr. 1960, *28*(2), pp. 417–63.

THEIL, HENRI. "Estimation and Simultaneous Correlation in Complete Equation Systems." The Hague: Central Planning Bureau, 1953; reprinted *Henri Theil's Contributions to economics and econometrics*. Eds.: BALDEV RAJ AND J. KOERTS. Dordrecht: Kluwer Academic, 1991, ch. 6.

______. *Economic forecasts and policy*. Contributions to Economic Analysis No. 15. Amsterdam: North-Holland, first and second editions, 1958 and 1961.

TINBERGEN, JAN. "Bestimmung und Deutung von Angebotskurven: Ein Beispiel," Z. *Nationalokon.*, 1929–30, *1*, pp. 669–79.

______. *An econometric approach to business cycle problems*. Paris: Hermann et Cie., 1937

______. *Statistical testing of business-cycle theories, Vol. II: Business cycles in the U.S.A. 1919–1932*. Geneva: League of Nations, 1939.

______. *Selected papers*. Eds.: L. H. KLAASSEN, L. M. KOYCK, AND H. J. WITTEVEEN. Amsterdam: North-Holland, 1959.

VINING, RUTLEDGE. "Koopmans on the Choice of Variables to Be Studied and of Methods of Measurement," *Rev. Econ. Statist.*, May 1949, *31*(2), pp. 77–86 (including a "A Rejoinder" to KOOPMANS 1949b, pp. 91–94).

WALLIS, KENNETH F. "Econometric Implications of the Rational Expectations Hypothesis," *Econometrica*, Jan. 1980, *48*(1), pp. 49–73.

WEGGE, LEON L. "Identifiability Criteria for a System of Equations as a Whole," *Australian J. Statist.*, Nov. 1965, *7*(3), pp. 67–77.

WOLD, HERMAN O. A. in association with LARS JUREEN. *Demand analysis: A study in econometrics*. New York: Wiley, 1953.

WORKING, ELMER J. "What Do Statistical 'Demand Curves' Show?" *Quart. J. Econ.*, Feb. 1927, *41*(1), pp. 212–35.

WORKING, HOLBROOK AND HOTELLING, HAROLD. "Applications of the Theory of Error to the Interpretation of Trends," *Amer. Statist. Assoc. J.*, Mar. 1929, *24*(165A), pp. 73–85.

WRIGHT, PHILIP G. "Review of *Economic Cycles* by Henry Moore," *Quart. J. Econ.*, May 1915, *29*(2), pp. 631–41.

______. *The tariff on animal and vegetable oils*. New York: Macmillan for the Insitutute of Economics, 1928.

WRIGHT, SEWALL. *Corn and hog correlations*. U.S. Department of Agriculture Bulletin No. 1300. Washington, DC: U.S. GPO, 1925.

WRIGHT, SEWALL with the collaboration of PHILIP G. WRIGHT. "The Method of Path Coefficients," *Ann. Math. Statist.*, 1934, *5*, pp. 161–215.

WU, DE-MIN. "Alternative Tests of Independence Between Stochastic Regressors and Disturbances," *Econometrica*, July 1973, *41*(4), pp. 733–50.

WYMER, CLIFFORD R. "Econometric Estimation of Stochastic Differential Equation Systems," *Econometrica*, May 1972, *40*(3), pp. 565–77.

ZELLNER, ARNOLD AND PALM, FRANZ. "Time Series Analysis and Simultaneous Equation Econometric Models," *J. Econometrics*, May 1974, *2*(1), pp. 17–54.

ZELLNER, ARNOLD AND THEIL, HENRI. "Three-stage Least Squares: Simultaneous Estimation of Simultaneous Equations," *Econometrica*, Jan. 1962, *30*(1), pp. 54–78.