The Error Term in the History of Time Series Econometrics
Author(s): Duo Qin and Christopher L. Gilbert
Source: *Econometric Theory*, Vol. 17, No. 2 (Apr., 2001), pp. 424-450
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/3533075

Econometric Theory, **17**, 2001, 424–450. Printed in the United States of America.

THE ERROR TERM IN THE HISTORY OF TIME SERIES ECONOMETRICS

DUO QIN
Queen Mary and Westfield College

CHRISTOPHER L. GILBERT
Vrije Universiteit

We argue that many methodological confusions in time-series econometrics may be seen as arising out of ambivalence or confusion about the error terms. Relationships between macroeconomic time series are inexact, and, inevitably, the early econometricians found that any estimated relationship would only fit with errors. Slutsky interpreted these errors as shocks that constitute the motive force behind business cycles. Frisch tried to dissect the errors further into two parts: stimuli, which are analogous to shocks, and nuisance aberrations. However, he failed to provide a statistical framework to make this distinction operational. Haavelmo, and subsequent researchers at the Cowles Commission, saw errors in equations as providing the statistical foundations for econometric models and required that they conform to a priori distributional assumptions specified in structural models of the general equilibrium type, later known as simultaneous-equations models. Because theoretical models were at that time mostly static, the structural modeling strategy relegated the dynamics in time-series data frequently to nuisance, atheoretical complications. Revival of the shock interpretation in theoretical models came about through the rational expectations movement and development of the vector autoregression modeling approach. The so-called London School of Economics dynamic specification approach decomposes the dynamics of the modeled variable into three parts: short-run shocks, disequilibrium shocks, and innovative residuals, with only the first two of these sustaining an economic interpretation.

1. INTRODUCTION

This paper examines the interpretation of equation errors in time-series econometrics. We contrast the view of errors as what differs trivially from the theoretical model with the view that the errors represent the shocks that are the important driving forces of model dynamics. The history of econometrics may

This is the revised version of a paper given at a session of the American Economic Association meeting held in San Francisco, January 5–7, 1996. We thank John Aldrich, David Hendry, Alessandro Rebucci, Peter C.B. Phillips, and two anonymous referees for helpful comments. Address correspondence to: Duo Qin, Department of Economics, Queen Mary and Westfield College, London E1 4NS, England; e-mail: D.Qin@qmw.ac.uk; Christopher L. Gilbert, Faculteit der Economische Wetenschappen, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands; e-mail: cgilbert@econ.vu.nl.

  0266-4666/01 $9.50

be seen as oscillating between these interpretations of errors, with some econometricians attempting to maintain both simultaneously.

In nascent econometrics, errors were generally disregarded as lacking economic implication and were regarded as simply the estimation by-products of a priori formulated theoretical models (see the empirical studies of American agricultural economists described in Morgan, 1989). This "nonstructural" viewpoint was challenged by the seminal work of Slutsky (1937), who showed, in the form of a difference-equation model, that random shocks could generate cyclical motions in economic variables and postulated shocks of this type as the cause of business cycles. Frisch (1938) tried to reconcile both the nonstructural theme and Slutsky's structural theme by distinguishing between "stimuli" (structural shocks) and "aberrations" (nonstructural disturbances), but he failed to provide a clear formulation of the two in a unified framework.[1] His distinction was lost in the Cowles Commission work. Subsequently, the general perception of errors regressed to the initial view, i.e., that they lack economic significance. This approach was strengthened by the assumption, made for statistical convenience, that the errors followed the serially independent and identical distribution (i.i.d.), on the argument that these errors merely represented the aggregate effects of a large number of individually unimportant omitted variables.[2] However, the i.i.d. assumptions, in particular serial independence, were rarely supported by applied models estimated on time-series data.

A remedy was to allow for autocorrelation of the error term in a largely a priori formulated structural model and to "fix" the statistical properties of coefficient estimators accordingly. This remedy was widely adopted from the 1950's to the early 1970's and remains a component of textbook expositions of mainstream econometrics. However, as theoretical models became more dynamic during the 1970's, encouraged in particular by the rational expectations movement, mainstream methods attracted increasing criticism, and a plethora of rival methodologies emerged from these debates. Two methodologies, the VAR (vector autoregression) approach and the LSE (London School of Economics) approach (see Pagan, 1987, 1995), came into prominence during the 1980's. Both maintain that the presence of residual autocorrelation implies a dynamically misspecified structural model and that empirical modeling should proceed on the basis of innovational errors (i.e., residuals should be unpredictable from the lagged data set). This has resulted in renewed interest in the dynamic structure of models and, with it, a revival in attempts to provide shocks and errors with an economic interpretation. A detailed historical account of these developments is given in the next seven sections. A summary "road map" is provided in the Appendix.

In part, the treatment of error terms in econometrics reflects econometricians' views on the completeness of economic theory. The most prevalent view has been that theories give partial and incomplete descriptions of economic systems, with the implication that the error terms of estimated models complement whatever is left out from the theoretical models. However, there is also a long tradition in econometrics, extending back to Slutsky and Frisch, in which theory relates directly to the stochastic shocks seen as driving the dynamic evolution of the economy. In practice, errors have to fill both of these roles.

The belief that errors solely represent random shocks responsible for the generation of business cycles fails to acknowledge that the properties of regression residuals are determined by the empirical model, sample data, and estimation procedure. On the other hand, the "innovational residuals" model design criterion can result in errors (and therefore possibly also a model) that are not interpretable in relation to any economic theory. The history of time-series econometrics shows how econometricians shifted their positions in attempting to combine these two interpretations. However, attempts at synthesis, which began with Frisch, have not been fully carried through, and there remains a need for a framework that incorporates the two rival interpretations in a logically consistent manner. This defines our agenda.

2. ABERRATIONS AND/OR STIMULI: SLUTSKY, FRISCH, AND TINTNER

Econometric studies prior to and during the 1930's were predominantly based upon deterministically formulated theories. However, deterministic relationships never fitted data with any precision, and so estimation required the use of statistical methods, of which least squares regression became the most prominent. The regression error term took on either of two rationalizations: errors in variables or errors in equations (Morgan, 1989, Ch. 7). On the first rationalization, the error term was regarded as arising purely from imprecise measurement of the theoretical variables and was therefore devoid of theoretical connotation. The second interpretation saw the error term as arising primarily from imprecision associated with possible functional form misspecification and therefore also lacking theoretical interpretation. In either case, the error term was generally omitted from expositions of the theoretical model and regarded as part of the translation from the theoretical model to its empirical counterpart (Morgan, 1991; Gilbert, 1991).

The error term first took on a theoretical role in business cycle studies. A major difficulty in this area was the lack of precisely formulated theoretical models capable of capturing the dynamic cyclical movement observed in many markets and economies. Pioneers of business cycle theory had therefore to start their research by looking for ways to describe cyclicality within a largely univariate context. These variables typically exhibited oscillatory time paths coupled with irregular jumps. In attempting to simultaneously analyze these two features, business cycle theorists were attracted by the time-series literature in statistics, where a standard procedure was to decompose single time series into trend, periodic, and erratic components.

One of the first attempts was due to Slutsky (1937).[3] Inspired by laws of physics and biology describing "how regularities could be derived from a chaos of disconnected elements because of the very disconnectedness," Slutsky postulated that "the summation of random causes may be the source of cyclic, or undulatory processes" observed in economic time series. Slutsky classified "chaotically-random elements" into "coherent and incoherent" series depending on whether the series were serially correlated or not and showed that a coherent (serially correlated) series (e.g., $\{y_t\}$ in (2.1), which follows) could be decomposed into a weighted moving summation, or a weighted moving average (MA), of an incoherent (serially uncorrelated) series (e.g., $\{\varepsilon_t\}$):

$$\begin{cases} y_t &= A_0\varepsilon_t + A_1\varepsilon_{t-1} + \cdots + A_{n-1}\varepsilon_{t-n+1} \\ y_{t-1} &= A_0\varepsilon_{t-1} + \cdots A_{n-1}\varepsilon_{t-n+1} + A_n\varepsilon_{t-n}. \\ \cdots &= \qquad\qquad \cdots \end{cases} \tag{2.1}$$

He does not use the terms *error* or *disturbance*. Slutsky's postulate was immediately recognized as innovative and stimulating by the pioneer econometricians. However, his model left open a number of unanswered issues. In particular, how should the "random causes" $\{\varepsilon_t\}$ be identified from observable economic data? One possible solution was implied by the decomposition theorem of Wold (1938). Wold showed that any stationary time series could be decomposed into a deterministic component together with an infinite MA of serially uncorrelated errors and that any such process may be approximated by a sufficiently high order autoregressive (AR) process (pp. 75–80). He reserved the term *residual* for the approximation error on this process and noted that the autocorrelation of the error terms would die out as the order of the autoregression increases. These important time-series results were to form, nearly 50 years later, the basis both for VAR modeling (see Section 7) and for the so-called LSE approach (see Section 8).

Another route was explored by Frisch, who saw the similarity of Slutsky's model to the time-series methods employed by Yule (1927) and Hotelling (1927). However, there was an additional ingredient required if the model was to be applicable to business cycle models—multivariate extension to causal economic relationships. Inspired also by Wicksell (1907), Frisch (1933) set out to extend the Walrasian tradition of completely deterministic and closed, static models by introducing two new elements. The first was to employ differential equations to capture "theoretical dynamic laws," which he referred to as the "propagation" problem. Taking a simple static model that related the output of capital goods, y, to the output of consumers' goods, x, he extended this to a dynamic model by adding the past stock of the capital goods as an additional explanatory variable,

$$y_t = f(x_t, \textstyle\int g(y_{t-\tau})d\tau; \theta), \tag{2.2}$$

where both f and g are simple linear functions and θ is the parameter set of the two functions. Frisch then supposed that each of the variables in (2.2) could be represented by an oscillatory deterministic time-series path in an exponential form,

$$\begin{cases} x_t & = & a_0 + \sum\limits_k a_k e^{\rho_k t} \\ y_t & = & b_0 + \sum\limits_k b_k e^{\rho_k t} \qquad k = 0,1,2,\ldots, \\ \int g(y_{t-\tau})d\tau & = & c_0 + \sum\limits_k c_k e^{\rho_k t} \end{cases} \tag{2.3}$$

where ρ_k were complex numbers and were found to be determined by the structural parameter set θ but where the coefficient sets $\{a\}$, $\{b\}$, and $\{c\}$ were dependent upon the initial conditions of each variable. Equation (2.3) allowed him to transform (2.2) into a differential equation in terms of y alone (subsequently referred to as the "final-form equation") to focus upon the dynamics of y. The transformation offered models of the Yule–Slutsky class a feasible economic rationale, i.e., a causal economic relationship underpinning an AR or an MA process.

Frisch's second extension was to introduce "erratic shocks," ε_k, to capture the random jumps observed in the actual cyclical movements of economic variables. He referred to this as the "impulse" problem. Because the structural model (2.2) would yield only damped (stable) solutions, Frisch saw these random shocks as "the source of energy which maintains the economic cycles." He maintained that the two parts, when combined, formed "a theoretical setup which seems to furnish a rational interpretation of those movements which we have been accustomed to see in our statistical time data."

The particular differential equation considered by Frisch was

$$y_t = P(t-t_0)y_0 + Q(t-t_0)\dot{y}_0 + \sum_{k=1}^{n} Q(t-t_0)\varepsilon_k. \tag{2.4}$$

The equation is augmented by a set of discrete random impulses $\{\varepsilon_k\}$. Here $P(\cdot)$ and $Q(\cdot)$ are deterministic sinusoidal functions that, in the absence of $\{\varepsilon_k\}$, would relate the current y_t to its initial value. Moreover, the way in which these impulses fed into $\{y_t\}$ was identical to that of the systematic part of its dynamics; i.e., ε_k and $\dot{y}_0$ had the same weight system $Q(\cdot)$ in the equation. Frisch therefore called the model "a linear weight system." Note that Frisch augmented his equations with impulses, whereas Slutsky had decomposed series into what can be regarded as weighted sums of impulses, but that Frisch did not suggest how the impulses $\{\varepsilon_k\}$ could be determined empirically, as with Slutsky.[4]

In parallel to these theoretical innovations, Frisch was also experimenting with estimation methods. This led to his invention of "confluence analysis" (see Hendry and Morgan, 1989). In confluence analysis, regression residuals were seen as resulting from "disturbances of the data" and were seldom either discussed directly or explicitly represented (Frisch, 1934). In a later attempt at synthesis, Frisch categorized the error terms occurring in both his theoretical and empirical models into two types:

> A *disturbance* is a deviation from that situation which should have existed as a consequence of the structure. In other words, it is something *incompatible* with the structure, something new and spontaneous introduced in addition to the structure. Such disturbances may be of two sorts: *aberrations* and *stimuli*. A stimulus is a disturbance that carries on its effects to the subsequent states of the system,—through the structural equations. In other words at any given moment it is the magnitudes of the variates *including* the stimuli that are taken as influencing the further evolution, that is the stimuli act as a sort of permanently changing initial conditions. An aberration is also a departure from the value which a variate should have had according to the structure, but this departure acts only at the actual moment at which it occurs, it is a sort of instantaneous addition—unexplained by the structure—and without any *consequence* for the subsequent states. In other words it is the magnitudes of the variates *exclusive* of the aberrations that act as initial conditions for the subsequent states.
>
> . . .
>
> The existence of aberrations does not necessarily involve any important consequences for the theoretical analysis, it only concerns the statistical technique, but in this respect it is important. The existence of stimuli entails much more far-reaching consequences. The total time shape will now be more or less transformed, for instance damped cycles will become undamped in the long run, but will have a disturbing effect over shorter intervals. The timing between the cycles may be changed from what it is in the stimulus-free system, and entirely new cycles, pure cumulation cycles will emerge. (1938, pp. 408–418, reprinted in Hendry and Morgan, 1995)

Apparently, stimuli correspond to the shocks $\{\varepsilon_k$ of equation (2.4) in Frisch's theoretical studies of business cycles, whereas the aberrations correspond to the measurement error residuals of his confluence analysis. However, despite a remark that it would in principle be possible to include both types of errors in the same structural model, Frisch demonstrated neither how this might be done nor the consequences for estimation.

With hindsight, we can ascribe Frisch's difficulty to ambiguity over what constitutes a structural model. Frisch appears to regard relationships relating static theory as structural (see Frisch, 1934, Part I). In this case, errors are aberrations resulting from lack of full correspondence between observed and theoretical quantities. By contrast, in his business cycle analysis, the equation of interest (2.2) is deterministic, with the impulses appearing as an add-on only in the final form. It is simply unclear whether the add-on impulses imply a randomization of the conditional relationship (2.2) as

$$y_t = f\left(x_t, \int g(y_{t-\tau})d\tau;\theta \right) + u_t \qquad \textbf{(2.5)}$$

or of the $\mathbf{x}_t$ equation in (2.3),

$$x_t = a_0 + \sum_k a_k e^{\rho_k t} + \nu_t. \qquad \textbf{(2.6)}$$

Either way, aberrations were not specified. So whereas aberrations and stimuli are both part of Frisch's equipment, he appears content to invoke whichever was required in a particular context. Confluence analysis never became dynamic, whereas aberrations were ignored in business cycle analysis.

Around the time that Frisch proposed the distinction between aberrations and stimuli, Tintner (1938) proposed an alternative categorization in terms of the agents' economic behavioral rules. He attributed the randomness in economic data to two kinds of error-generating behavior: the first was due to agents' random failures to achieve "the optimum adaptations of the controllable factors," whereas the second resulted from agents' failures to forecast "the uncontrollable factors." Tintner maintained, in agreement with Frisch and other business cycle analysts, that "errors of the second kind are much more fundamental, and have a much deeper influence on economic life and also on the shape of economic time series," and he observed that the residual "which appears as a remainder in most methods of analysis of time series is the result of errors of the first kind." However, Tintner appears to have been aware of the need to allow for both kinds of errors in a dynamic model in acknowledging that the mathematical representations of the errors did "not seem to be unique."

To tackle the problem, Tintner resorted to the variate difference method (1940a). Assuming both kinds of errors to be serially independent and that the second kind follow the normal distribution, he argued that statistical analyses of interdependency among economic time series should be based only upon the random variations in the individual series, which could be obtained via differencing, and that the residuals from such analyses should then correspond to the errors of the second kind (Tintner, 1940b). However, Tintner's ideas and methods were largely overshadowed by the excitement of the profession over Tinbergen's epoch-making model (Tinbergen, 1939).

3. ERRORS IN EQUATIONS: KOOPMANS, HAAVELMO, AND THE COWLES COMMISSION

Tinbergen's famous (1939) model moved business cycle theories from the realm of latent ideals to a world in which hypotheses became measurable and testable against data, and, in so doing, he laid the foundations for postwar macroeconometric model building. In particular, the astructural regression residual–based interpretation of the error terms soon came to dominate the shock-based structural interpretation. The discrete form of equation (2.5) became established as the standard form of structural relation, as in the Tinbergen model where the error terms were interpreted as errors in equations, with ordinary least squares (OLS) regression residuals as their empirical counterparts.

Tinbergen based his choice of the OLS method on the theoretical work of Koopmans entitled *Linear Regression Analysis of Economic Time Series* (1937). In a manner similar to that of Frisch and Tintner, Koopmans categorized the erratic components of economic time series into two types, but unlike them, he reverted to the earlier distinction between errors in variables and errors in equations. The former were described as measurement errors, i.e., "inadequate weighting of an index number" or "inaccuracy in the statistical recording procedure" (Koopmans, 1937, Ch. 1, p. 6); the latter were defined as minor influences on the determined variables by incidental factors omitted from the determining variables of the equations. To simplify the estimation task, he proposed ignoring the first type on the argument that, in most applied cases, their omission should not significantly affect the accuracy of the estimated parameters. His greater worry was the problem of simultaneity, i.e., the problem that the OLS residuals fail to correspond with the equation errors when there is more than one possible economic relation among the regression variables. He proposed to solve the problem by a method based on minimization of a weighted sum of the squared residuals from the regression and the inverse regression.

Koopmans's technical rigor, Frisch's theoretical vision, and Tinbergen's inventive experimentation combined to lay a solid foundation for Haavelmo's *The Probability Approach in Econometrics* (1944). By demonstrating the inconsistency of the OLS estimates of the structural coefficients of a structural econometric model (SEM) (see also Haavelmo, 1943), Haavelmo focused his attention on the link between identification and estimation. He came to the view that simultaneity among the endogenous variables should be made explicit by specifying the joint distribution of the variables, achievable as a mapping from the assumptions on the distribution of error terms. It was therefore crucial that these assumptions be consistent with the a priori postulated theoretical relations. For example, Haavelmo adopted the assumptions of joint normality and serial independence in his illustratory examples (1944, Ch. 4).[5] However, he did not consider serial independence to be essential to an SEM, for he remarked that "if there happens to be a functional relationship between the ε's at two different points of time, the dimensionality of the joint distribution of all the ... ε's can be correspondingly reduced" (1944, p. 89).

Haavelmo's work provided a blueprint for the research program of the Cowles Commission for Research in Economics, when they set out to formalize econometrics in the early 1940's (see, e.g., Qin, 1993). Indeed, their core achievement was in the identification and estimation of the dynamic SEM of a set of variables $x = \{y, z\}$ (y denotes a vector of endogenous variables, z a vector of explanatory variables):

$$Ay_t' + \Gamma w_t' = \varepsilon_t' \qquad \text{with} \quad w_t = [\,y_{t-1} \quad \dots \quad y_{t-\tau} \quad 1 \quad z_t \quad z_{t-1} \quad \dots \quad z_{t-\tau}\,], \tag{3.1}$$

where ε_t denotes errors in equations assumed to have a joint normal distribution.[6] Estimation of (3.1) was recommended via the "reduced form" (see, e.g., Mann and Wald, 1943; Koopmans, 1950):

$$y_t' = -A^{-1}\Gamma w_t' + A^{-1}\varepsilon_t' = \Pi w_t' + u_t'. \tag{3.2}$$

The Cowles work was, for the most part, based on the assumptions of non-autocorrelation and identical distribution of ε_t in addition to the joint normality.

However, there was a strongly held view that the assumption of nonautocorrelation was merely a simplification to ease estimation. An example of possible autocorrelation within a simple SEM was provided by Hurwicz (1944):

$$\begin{cases} Y_t = & bY_{t-1} + a_1 Z_{t-1} + a_2 Z_{t-2} + \varepsilon_t' \\ Z_t = & cY_{t-1} + \varepsilon_t'' \end{cases}. \tag{3.3}$$

He transformed this into a single autoregressive-moving average (ARMA) equation,

$$Y_t = bY_{t-1} + a_1 cY_{t-2} + a_2 cY_{t-3} + \eta_t, \tag{3.4}$$

where $\eta_t = \varepsilon_t' + a_1 \varepsilon_{t-1}'' + a_2 \varepsilon_{t-2}''$, and called it a "reduced equation,"[7] in which η_t formed a "composite disturbance." Obviously, η_t is autocorrelated even if the disturbances ε_t' and ε_t'' in the original structural equations are nonautocorrelated. Hurwicz therefore concluded that "it might well be that some of the distrust with which the 'literary' economists have viewed the 'mathematical' business-cycle theory has arisen from their opposition to the unrealistic postulate of non-autocorrelated 'disturbances.'"

Hurwicz's view was apparently widely shared by his contemporaries. For example, Marschak maintained, in the introductory chapter of the Cowles Commission Monograph 14, that possible serial correlation in the disturbances "must be considered part of the structure" on the ground that there were no economic reasons to rule out the possibility of their forming a "stochastic process, each shock depending on one or more of its predecessors," especially when the observation frequency became high, e.g., "weekly or even quarterly instead of annual time series" (1953, pp. 21–23).

Frisch's difficulty in handling aberrations versus stimuli resurfaces here in Marschak's attempt to provide a structural interpretation of autocorrelated errors but with two additional complications. First, reference to "reduced forms" implies that serial correlation has a direct impact on estimation because reduced forms were regarded merely as a convenient vehicle for estimation, and evidence of residual autocorrelation would normally be taken from the estimated equations. Second, the structuralist tradition of specifying a priori theoretical models in static form provided abundant practical confirmation of the presence of residual serial correlation. Consequently, the prevalent view became that a more sophisticated estimation apparatus was the solution to residual autocorrelation, and this tended to distract attention from interpretation of the error terms in dynamic models.

4. RESIDUAL AUTOCORRELATION: TIME-SERIES ANALYSIS AT THE DAE

The implications of autocorrelated errors for the statistical properties of the estimators of structural coefficients became the next focus of econometric research. Much of this research was undertaken at the Cambridge Department of

Applied Economics (DAE) in a program called "Analysis of Time Series" (see Gilbert, 1988), which resulted in the famous Durbin–Watson test (Durbin and Watson, 1950) and the Cochrane–Orcutt (CORC) estimation procedure (Cochrane and Orcutt, 1949).

Orcutt was a key figure in the Cambridge program. His time-series research started with a correlation analysis in relation to the time series used in Tinbergen's model (1939). From the time-series perspective, he viewed the Tinbergen model as a VAR model and estimated AR(2) (autoregression of order two) models for each of the 52 economic series Tinbergen had used (Orcutt, 1948). These estimates led him to conclude that the series could typically be seen as being generated by the following AR(2) with a unit root

$$y_{i,t} = y_{i,t-1} + 0.3\Delta y_{i,t-1} + \varepsilon_{i,t}. \quad \textbf{(4.1)}$$

Orcutt believed that this explained why many of the residuals of Tinbergen's model exhibited significant degrees of autocorrelation. With Cochrane, he then endeavored to find a general solution for the loss of efficiency in OLS estimates in the presence of such autocorrelation. Orcutt and Cochrane considered two possible ways of solving the problem (Cochrane and Orcutt, 1949). The first was "to change some of the variables, add additional variables, or modify the form of the relation until a relationship involving what appear to be random error terms is found"; the second was to develop more elaborate estimation methods than OLS while leaving the structural specifications fixed. They chose the latter in the belief that economists typically specify structural models in terms of "the most reasonable choice of variables and form of relation" independently of statistical considerations. From this structuralist standpoint, Orcutt and Cochrane asserted that, empirically, there was "strong evidence in favour of the view that the error terms involved in most current formulations of economic relations are highly positively autocorrelated." To overcome the inefficiency of the OLS estimates, $\hat{\beta}$, of a structural model

$$y_t = \beta z_t + u_t \quad \textbf{(4.2)}$$

when $\{\hat{u}_t\}$ turned out to be autocorrelated, Cochrane and Orcutt proposed to append (4.2) by an AR(1) in u_t,

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad \textbf{(4.3)}$$

and to estimate β by an iterative procedure using least squares estimation on the two equations. This is now known as the CORC estimator. Cochrane and Orcutt also observed that the CORC estimator of ρ tended to be biased toward unity when the a priori theoretical models took the simple static form of (4.2), especially when z and y are trending variables, and recommended estimation in first differences, i.e., $\Delta y_t = \beta \Delta z_t + \Delta u_t$, a recommendation very close to that of Tintner nearly 10 years earlier (see Section 2). The recommendation was adopted by Stone (1954) and also by the Netherlands Central Planning Bureau (see Theil, 1958).

The appendage AR(1) of (4.3), which distinguishes between autocorrelated disturbances u_t and white-noise shocks ε_t in a somewhat similar manner to Frisch's distinction between stimuli versus aberrations, suggests the possibility of reviving Frisch's classification. However, this turns out to be illusory because the error terms analyzed by Cochrane and Orcutt lacked theoretical interpretation, either in respect to the structural model (4.2) or in relation to the generation of dynamics, as in Hurwicz (1944). In other words, the addition of (4.3) to (4.2), or the move to first-differenced series, tended to reinforce the prior status of theoretical models and concentrated attention on the statistical issue of efficient estimation. The consequence was that the majority of econometric research was oriented toward developing more complicated estimation tools for fixed, frequently static, theoretical models with relatively little attention devoted to the way more complicated dynamic specifications might explain economic fluctuations.

The independent work of Sargan (1959, 1961) and Durbin (1960) formed a diversion from this orthodox research route. By combining (4.3) with (4.2), they demonstrated that appending an AR(1) error process to a static model is equivalent to dynamic respecification of the original model together with a particular restriction on the coefficient of the lagged exogenous variable:

$$(y_t - \rho y_{t-1}) = \beta(z_t - \rho z_{t-1}) + \varepsilon_t,$$
$$(1 - \rho L)y_t = \beta(1 - \rho L)z_t + \varepsilon_t$$

(L denotes the lag operator, e.g., $Lz_t = z_{t-1}$) or

$$\begin{aligned}(y - \beta z)_t &= \rho(y - \beta z)_{t-1} + \varepsilon_t,\\ &\Rightarrow y_t = \beta z_t + \rho y_{t-1} - \rho z_{t-1} + \varepsilon_t = \beta z_t + \rho y_{t-1} + \alpha z_{t-1} + \varepsilon_t,\\ &\quad \alpha = -\beta\rho. \end{aligned} \tag{4.4}$$

However, neither Sargan nor Durbin pursued this very far. Instead, both focused on how to achieve optimal estimation of combined model (4.2) and (4.3), although Sargan mentioned in passing the possibility of testing the residual autoregressive setting against the more general dynamic autoregressive distributed lag (ADL) model. It was nearly two decades later that the dynamic implications of the common factor restrictions in (4.4) were to be fully explored (see Mizon, 1977; Hendry and Mizon, 1978).

The contributions to time-series estimation of Orcutt, Durbin, and Sargan soon came to be widely adopted as part of mainstream econometrics. These were often presented in textbooks as an extension to a simple-equation model or to an SEM in a purely static form,

$$Ay_t' + \Gamma z_t = \varepsilon_t', \tag{4.5}$$

where the set $\{z_t\}$ no longer explicitly contains any lagged terms, in contrast to the original Cowles model (3.1). This loss of explicit dynamics, hidden now in the appended error process, moved orthodox econometrics even further away from Frisch's program of studying the time-path effects of random shocks on business cycles. In particular, because Frisch's concept of "stimuli" was related directly to dynamic structure, the loss, or at least concealment, of this dynamic structure amounts to loss of the possibility of any structural interpretation of shocks in terms of model dynamics.

5. DISTURBANCES AS MANEUVERABLE UNOBSERVABLES: THEIL AND LEAMER

Although, in principle, it was a prerequisite of the orthodox approach to have a complete a priori formulation of a theoretical model, this prerequisite could not be satisfied in applied contexts. Early applied modelers were familiar with the frequent need for ad hoc adjustments to the structural models that they inherited from theorists (see Qin, 1993, Ch. 6). Theil was among the first econometricians who tried to highlight and tackle the problems associated with such practices. Theil (1958, Ch. 6) argued that such practices amount to changes to the maintained hypotheses and are therefore contrary to standard Neyman–Pearson hypothesis testing methodology. He advocated that changes in the maintained hypothesis be made explicit through what he referred to as "specification analysis." The regression residuals played an essential role in specification analysis. This implied two extensions to the previously existing framework. First, Theil proposed minimization of residual variance as the main criteria for specification choice. He recognized that "there is no law in economics which states that such proportions [of the disturbance] are small or even 'as small as possible,'" but he attempted nevertheless to justify the minimization as the criterion for choosing between different model specifications (Theil, 1957). Second, Theil included both residual autocorrelation and cross-equation dependence as potential problems to be dealt with in specification analysis (Theil, 1958, Ch. 6), thereby reviving the question of how model specification should deal with the error terms, an issue that had been subsumed into estimation in the previous orthodoxy.

Because Theil's specification analysis presumed that modelers "do not know the 'true' specification in general" (Theil, 1958, p. 215), it purported to justify model respecification and the practice of testing between different (often non-nested) theoretical models. This undermined the foundations of the Cowles procedures, which were based on a priori known structural specifications, and denied any independent status to the disturbance term, because equation errors were now contingent on model specification.

Theil's ideas on specification analysis were systematically revived and clarified by Leamer, albeit within a Bayesian framework, 20 years later. Despite the difference of statistical framework, Leamer's stance on specification analysis remained close to that of Theil (see Qin, 1996). Specifically, Leamer took economic models as the given theoretical basis of structural models and focused on the task of filling the gap between the theoretical and empirical models, which he saw as the problem of specification in applied econometrics. He became convinced that the Bayesian approach would provide a systematic albeit honestly subjective way to improve the usual ad hoc practice of specification by many modelers. In his influential 1978 book, Leamer pointed out that modelers' specification choices relate closely to their error specifications and argued that any given model, $y = Z\beta + u$, would give rise to "a tautological definition" of the error term: $u \equiv y - Z\beta$, i.e., u being "all of those things that determine y, excluding $Z\beta$." He therefore interpreted u as the sum of maneuverable "unobservables." This error term could be represented explicitly as $u = R\gamma$ to differentiate it from the explanatory variables Z, which were regarded as fixed. By contrast, R must be supposed "to vary within the confines of some more-or-less well-defined experimental conditions" so that $R\gamma$ would "appear to have been drawn from a particular normal distribution." Leamer inferred that "finding a complete theory" would correspond to eliminating the error term and that attempts to minimize error variance therefore implied searching for theories that would be as complete as possible (Leamer, 1978, pp. 65–66).[8]

The specification of $u = R\gamma$, together with the Bayesian approach of imposing a prior density for β, gave Leamer considerable freedom to experiment with various posterior estimators of β. Here, Leamer and the Bayesian econometricians differ more substantially from Theil and other classical econometricians in that their interest was primarily in the properties of the various posterior estimators of β rather than minimization of the residual variance. Nonuniqueness of $R\gamma$, together with the fact that γ was considered nuisance parameters, allowed the possibility of concealing misspecification in the residuals $\hat{u}$. Consequently, choices of the prior, $p(\beta)$, and the extreme bounds analysis of the fragility of the estimated β dominated revisions of the possibly incorrectly formulated theoretical model $Z\beta$, a move in the opposite direction from that for which Theil had argued. Although, Leamer was conscious of arbitrary features in the formulation of theoretical models, his reliance on Bayesian subjectivity allowed him to maintain a priori theoretical models largely intact in empirical studies. The consequence is that this Bayesian approach to econometrics fails to grapple with the fundamental issues in the structural modeling paradigm but nevertheless subjectively undermines the primacy of theoretical models (see Qin, 1996).

6. SERIALLY DEPENDENT SHOCKS: IMPLICATIONS OF RATIONAL EXPECTATIONS

Frisch's vision of "stimuli" as a means of characterizing the dynamic effects of random shocks on business cycles was revived in the rational expectations (RE) movement, which was centered on developing dynamic, theoretical models along the route initially set by Frisch (1933). In the seminal RE model, Muth (1961) postulates a shock variable u_t in a static SEM of a commodity market,

$$\begin{matrix} Q_t^D = -\beta p_t \\ Q_t^S = \gamma p_t^e + u_t \end{matrix} \qquad Q_t^D = Q_t^S \text{ in market equilibrium,} \tag{6.1}$$

where Q_t^D and Q_t^S denote the amount of the commodity demanded and supplied, respectively, and p_t and p_t^e the market price and its expected value, respectively. To make the expectation variable p_t^e operational, he assumed that "part of the shock variable $[u_t]$ may be predicted on the basis of prior information." He assumed specifically[9]

$$u_t = w(L)\varepsilon_t, \tag{6.2}$$

where $\varepsilon_t \sim i.i.d.(0,\sigma^2)$ and $w(L)$ denotes a matrix polynomial in lags L of possibly infinite order. Substituting this MA representation into (6.1) and taking the expectation of p_t would result in an AR(n) explanation of p_t^e,

$$E(p_t) = p_t^e = \nu(L)p_t, \tag{6.3}$$

where the weights $\nu_j = f(\beta,\gamma,w_i)$ in the matrix polynomial $\nu(L)$ of order n.[10] The dynamics of p_t^e was derived directly from the assumed dynamics of u_t in (6.2), which was in turn derived from the assumption of u_t as exogenous shocks to the model (6.1). These assumptions were accepted without question in many of the subsequent, more elaborate, RE models (see Lucas and Prescott, 1971; Sargent and Wallace, 1975), probably because these models remain primarily theoretical with their focus on the causal relationship between the expectation dynamics and the estimability of the parameters via cross-equation restrictions of the reduced form. These restrictions are important because any exogenous shocks that alter the dynamics of expectation formulation may also alter some parameters of the structural models, a point elaborated by Lucas (1976), who discussed the impact of exogenous policy shocks.

The time-series representation (6.2) of autocorrelated but economically unspecified shocks was soon applied directly to the generation process of a specific explanatory (exogenous) variable. In the case of a single-equation model, the reduced form derived from an RE model becomes an ADL model, as in Sargent and Wallace (1973), whereas in the multiequation case, it becomes a VAR, as in Sargent (1979) and Wallis (1980). Combination of this alternative with a serially dependent shock would lead to the reduced form with, respectively, an ARMA (autoregressive moving average) or VARMA (vector ARMA) model (see Sargent, 1977; Lucas and Sargent, 1979). In each of these cases, the reduced forms of RE models brought back the dynamic SEM of (3.1) by the Cowles Commission.

The dynamics of such RE models presupposes two conditions: weak stationarity of $\{u_t\}$ and autonomy of u_t, be it an exogenous or a shock variable. Furthermore, the representation designated ε_t the "fundamental" role as the driving process of cycles in the endogenous variables (cf. Whiteman, 1983, Ch. 1).[11] However, there is a certain taxonomic ambiguity in assigning to the error term a structural interpretation equivalent to that assigned to specific exogenous variables. For example, Lucas and Sargent (1979) acknowledged that "restrictions . . . governing the behavior of the error terms . . . are harder to motivate theoretically because the errors are by definition movements in the variables which the *economic* theory cannot account for"; and Sargent (1978) wrote, "optimizing, rational-expectations models do not entirely eliminate the need for side assumptions not grounded in economic theory. Some arbitrary assumptions about the nature of the serial-correlation structure of the disturbances and/or about strict econometric exogeneity are necessary in order to proceed with estimation." But the ambiguity was not perceived as problematic.

On the other hand, the status of the error terms in theoretical RE models is clearly contrary to Theil's and Leamer's view that equation errors are maneuverable consequences of model specification and therefore do not have reference outside the context of the model from which they arise. In empirical RE models however, error terms with such maneuverable properties were often added in to reflect perturbation of the exact RE models in sample data. This allows the possibility of coexistence of structural shocks and nonstructural disturbances (see note 6), via decomposition of the error terms of simple, static models into two parts: the structural dynamic shocks and the remaining nonstructural residuals.

Most econometric discussion relating to RE models was within the structuralist paradigm. Technically, the central issues were identification and estimation of the structural parameters from the reduced forms. Empirically, the issue of testing whether significant dynamic feedback effects existed between the endogenous and exogenous variables by means of the Granger-causality tests became central because this would imply lack of invariance of structural parameters in the face of policy or other external shocks (see, e.g., Sargent, 1976). By contrast, relatively little attention was given to verifying the dynamic assumptions of the exogenous shocks.

Subsequently, empirical RE modeling began to expose weaknesses in the dynamic formulation of RE models (see, e.g., Pesaran, 1988). This was the most evident in Sargent's attempts to bridge the RE hypothesis and applied vector time-series models. For example, Sargent (1977) showed, via reduced-form derivation, that Cagan's expectations model (1956) of money demand was a special case of a general vector first-order autoregressive model with moving average errors, and this allowed him to take a time-series approach to testing Cagan's model. In another paper, Sargent (1976) simply started from a time-series model in the Wold moving average form in his discussion of the empirical implications of natural versus unnatural growth rate hypotheses in macroeconomics, though he still referred to the initial time-series model as a reduced-form model. These contributions induced economists and econometricians to adopt a more positive attitude toward the pure time-series modeling approach, which was to become the core of the VAR approach in econometrics.[12]

7. INNOVATIONAL ERRORS: VAR METHODOLOGY

VAR-type models had been used in econometrics from the days of the Cowles Commission (Mann and Wald, 1943). There was initially no suggestion that these VAR representations were nonstructural. When the issues of estimation and identification were discussed by the Cowles Commission authors in the context of the SEM (3.1), the reduced form (3.2) took on the most general form of an open VAR. It was Liu (1960) who first argued that (3.2) was actually the most one could obtain from data information without any arbitrary manipulation of the data, in the sense that (3.1) in fact was a particular form of (3.2) created by a set of a priori restrictions. Nearly two decades later, Liu's viewpoint was fully incorporated by Sims (1980) into the VAR methodology. Influenced by the RE movement and by Box and Jenkins (1970) time-series techniques, the methodology advocated the use of a general VAR form as the time-series representation of the data set.

An early and radical VAR experiment was carried out by Sargent and Sims (1977). On the basis that there was little reliable a priori business cycle theory, they endeavored to analyze the cyclical dynamics reflected in their estimated VAR by factor analysis in the frequency domain. Attention shifted back to the time domain when Sims advocated the use of VAR methodology as a formal alternative to mainstream econometrics in his famous 1980 paper. In particular, he proposed modeling the variables in (3.1) by means of a closed, unrestricted VAR,[13]

$$A(L)x_t = \varepsilon_t, \tag{7.1}$$

where $A(L)$ is a matrix polynomial in L of order n, the magnitude of which should be chosen in such a way to ensure that the model-derived $\{\varepsilon_t\}$ should be an innovation process, i.e., $E(\varepsilon_t | x_{t-1}, x_{t-2}, \ldots) = E(\varepsilon_t | X_{t-1}) = 0$. VAR residuals were therefore serially uncorrelated by construction and not by assumption. As far as the interpretation of ε_t was concerned, Sargent and Sims (1977) recognized that because most macroeconomic theories generated "behavioral equations without residuals, . . . we must tack on residuals to obtain empirically usable models and the theory is silent about the nature of the residuals."[14] This left the status of the errors unclear, but Sims (1980) described ε_t as innovation shocks to certain related modeled variables; e.g., he called the error term in a money-demand equation "money innovation."

Sims's interpretation of the errors as shocks was predicated on the general transformation of (7.1), provided $A(L)$ was invertible (which requires that each component of x is stationary), into the moving average representation (MAR):

$$x_t = A(L)^{-1}\varepsilon_t. \tag{7.2}$$

Notice that equation (7.2) reexpresses in a multivariate context Slutsky's (1927) univariate analysis of "coherent" series (here, the vector x) in terms of a moving average of "incoherent" series (here ε_t). The duality of (7.1) and (7.2) also underlines the problem of nonuniqueness emphasized earlier by Tintner (1938). In particular, the MAR (7.2) restores the possibility of interpreting disturbances in terms of Frisch's "stimuli." Equation (7.2) was in fact used widely by VAR modelers for policy analysis, with the error series $\{\varepsilon_t\}$ being interpreted as shocks. The impact of these shocks through the transmission mechanism $A^{-1}(L)$ was seen as the main generator of business cycles. This type of impulse response analysis requires imposition of a causal ordering (i.e., a recursive structure) on the system. This is equivalent to triangularization of the leading term of $A^{-1}(L)$, but there is no unique way of doing this—alternative triangularization schemes are equivalent to alternative identification assumptions within an SEM framework. VAR analysis simply reverses the traditional sequencing of model identification and estimation. Within simple SVAR (structural VAR) models, alternative identification schemes are available—for example, using a two variable SVAR, Blanchard and Quah (1989) identify through the neutrality requirement that the "monetary" shock should not have any long-run effect on real variables.

Impulse response analysis has been extensively used in the real business cycles (RBC) models, which, interestingly, are often seen as being distinct from VAR models because RBC modelers generally value theory consistency above data consistency and consequently choose to calibrate rather than to estimate the unknown parameters. In the case when the parameters are estimated, RBC models could be viewed as restricted VAR models, quite similar to RE models. RBC modelers would generally specify the error terms of their models as arising from two sources: shocks from exogenous variables, which are commonly treated as evolving from AR processes with random shocks, and measurement errors, or errors of observation, which are actually introduced out of the necessity of model estimation because the number of exogenous shock terms is normally smaller than the total number of equations to be estimated (see Kim and Pagan, 1995). This specification fails to acknowledge the possibility of misspecified models and in particular misspecification arising from omitted variables. When significant discrepancy occurs between the values of their model simulation and the actual data, RBC modelers tend to explain the discrepancy as arising from unimportant or uninteresting aspects of the economy from which their models abstract (see, e.g., Kydland and Prescott, 1991). Explanations of this sort have aroused great distrust of RBC models among econometricians (see, e.g., Quah, 1995; Gregory and Smith, 1995). What is interesting to observe here is how similar interpretation of the error terms, e.g., as exogenous shocks in impulse response analysis, could considerably narrow the gap between seemingly very incompatible methodologies.

The interpretation of errors as shocks in impulse response analysis presupposes that the underlying VAR (7.1) provides an economically valid characterization of the statistical process followed by the x variables. The apparent lack of any restrictions in (7.1) might appear to guarantee that whatever model is valid would be a subset of (7.1). However, this is true only conditional on the (marginalization) decision of which variables to include in the vector x. Although in principle the economist should aspire to be catholic, in practice the limited number of observations forces selection of a very small number of variables. These choices are typically suggested by the same informal and imprecise theory criticized by Sims (1980) as generating "incredible" restrictions. The implication is that the ε will include the innovation component of any marginalization errors in addition to genuine "stimuli."

8. SHORT-RUN DECOMPOSITION: EQUILIBRIUM-CORRECTION MODELS

Although the VAR approach has become increasingly popular over the traditional structuralist approach, many econometricians have resisted total abandonment of structural modeling. In particular, the so-called LSE approach to dynamic modeling (cf. Pagan, 1987) has developed an equilibrium-correction model (ECM) scheme as a middle course between traditional structural modeling and the VAR approach.

The LSE approach is best exemplified in Hendry's formulation (1995), where a structural model based upon a particular theory of interest is considered to be possibly reducible from a VAR. Take, e.g., a simple partial equilibrium theory. The theory of interest is used to decompose the variable set, $x = \{y, z\}$, so that the joint distribution $D(x)$ underlying (7.1) can be factorized into $D(x) = D(y|z)D(z)$ to enable the modeler to discard the marginal distribution $D(z)$ and to search for the structural model within the conditional autoregressive distributed lag (ADL) class

$$y_t = \sum_{i=0}^{n} \Gamma_i z_{t-i} + \sum_{j=1}^{n} \Phi_j y_{t-j} + \varepsilon_t \tag{8.1}$$

based upon the conditional expectation $E(y_t|z_t, X_{t-1}^n)$. As in VAR methodology, the error term in (8.1) is regarded as a model-derived mean-innovation process. In contrast to the VAR methodology, the error term is considered "a compound of many reduction losses" relative to the information available during model specification and therefore "cannot sustain claims to be a 'demand' shock or a 'monetary innovation'" in the sense that it does not correspond to any autonomously identifiable economic factors (Hendry, 1995, p. 359). This insight follows from the explicit focus on the marginalization (variable selection) decision as a crucial component in any modeling exercise.

A crude caricature of the LSE methodology is that the modeler is required to search for, or design, a statistically optimal ADL with the minimum number of lags, n, by minimization of the variance of the derived innovational residuals. Once an optimal ADL has been found, it may be reparameterized into an ECM for the purpose of identifying the structural model corresponding to the theory of interest. For example, the first-order (i.e., $n = 1$) ADL of (8.1) may be transformed into

$$\Delta y_t = \Gamma_0 \Delta z_t + (\Phi_1 - 1)(y - Kz)_{t-1} + \varepsilon_t, \tag{8.2}$$

where $K = \Gamma_0 + \Gamma_1/1 - \Phi_1$. The preceding reparameterization is thought to be especially desirable when $x = \{y, z\}$ exhibit nonstationarity or near nonstationarity. Under such a situation, a necessary condition for y and z to correspond to theory is that they must cointegrate in K to produce a stationary combination: $u_t = y_t - Kz_t$, where K is interpreted as the long-run coefficient of the static, equilibrium condition

$$y = Kz. \tag{8.3}$$

Equation (8.2) may then be rewritten using the disequilibrium series $\{u_t\}$

$$\Delta y_t = \Gamma_0 \Delta z_t + (\Phi_1 - 1) u_{t-1} + \varepsilon_t. \tag{8.2$'$}$$

Notice that (8.2′) decomposes the short run explained variable Δy_t into three types of shocks: the short-run Δz_t, the lagged disequilibrium u_{t-1}, and the innovative error ε_t. It is interesting to ask how these "structural" coefficients may be interpreted. There are two possibilities. One is to confine the structural interpretation to the a priori theoretical relationship, i.e., to the long-run coefficient K alone. Under this interpretation, the estimation residuals will generally be autocorrelated, because (8.2′) reveals

$$u_{t-1} = (\Phi_1 - 1)^{-1} (\Delta y_t - \Gamma_0 \Delta z_t - \varepsilon_t). \tag{8.4}$$

It follows that the $\{\varepsilon_t\}$ but not the $\{u_t\}$ are serially independent by construction. This interpretation rationalizes the conventional practice of estimating a priori structural models in a simple static form allowing for residual autocorrelation. It also implies that it may not always be appropriate to select an estimator for a long-run structural coefficient by requiring that the classical assumptions (in particular, that the errors be white noise) should hold for u_t. Notice however that (8.4) also defines a general function for the autocorrelation derived from the ADL, making any ad hoc imposition of the correlation, e.g., (4.3), testable. The other possibility is to interpret all the estimated coefficients of (8.2) as structural. This interpretation is much broader than the first in that it not only considers the short-run dynamics Δz_t structural but also allows implicitly for K to be data instigated rather than a priori given (this latter case is described as "error correction" instead of "equilibrium correction" in Hendry, 1995). Notice that the innovation property of ε_t becomes one prerequisite of this interpretation. But more interesting, this interpretation provides a possible way of circumventing the difficulties inherent in Frisch's attempt to classify random shocks according to their dynamic roles and to study their impacts on structural models. Following Frisch, we can transform (8.2) into a type of final form,

$$y_t = y_0 + \Gamma_0 \sum_{j=0}^{t-1} \Delta z_{t-j} + (\Phi_1 - 1) \sum_{i=1}^{t} u_{t-i} + \sum_{j=0}^{t-1} \varepsilon_{t-j}. \tag{8.5}$$

Equation (8.5) shows that we can now decompose the input shocks into a set of model-derived, nonstructural, innovational shocks $\{\varepsilon_t\}$ and another set of structural shocks that further divide into the short-run exogenous shocks $\{\Delta z_t\}$ and the long-run disequilibrium shocks $\{u_{t-1}\}$. This decomposition has two advantages over (1938) Frisch's distinction between "aberrations" and "stimuli." From the economic standpoint, it enables us to see the distinct dynamic impacts of the short-run and long-run shocks after filtering out the nonstructural shock designated as innovational and model dependent, and to conceive the long-run equilibrium path as a latent structure imposing a negative feedback on the dependent variable. From the econometric viewpoint on the other hand, we need only require that the structural shocks be weakly stationary and not necessarily free of autocorrelation. Furthermore, the separation of the short-run Δz_t from the long-run cointegrating relation $(y - Kz)$ reduces the likelihood of high collinearity among the coefficient estimates that frequently afflicts VAR model estimation, because the ECM representation of an ADL model resembles the MA representation of a VAR model.

9. PROSPECTS AND CONCLUDING REMARKS

Over recent years, the ECM approach, cointegration techniques, and the associated software have achieved enormous popularity among econometric practitioners. Lack of uniqueness in the structural interpretation of ECM's, which highlights the empirical imprecision of a priori theories, and a desire for statistically better performing models have greatly encouraged data-based model construction. As a result, mean innovation in the residuals has become the benchmark for econometric model search. This has relegated economic theory to the role of a desirable (as reflected in the specification of the cointegrating combination $(y - Kz)$) but inessential accommodating factor, leaving modelers with virtually unlimited latitude with regard to dynamic specification.

Data-oriented approaches of this sort undermine attempts to characterize the resulting models, whether or not structural, as being in any sense "true" or "valid" representations of a "data generation process" (DGP). For example, any claim that an ECM model is indeed structural is likely to meet the objection that supposed structural representations are likely to be chance outcomes of data-based model selection procedures. One possible response is to follow Rissanen (1987, 1989) in entirely abandoning the quest for models with "true" or inherently meaningful parameters[15] and to regard, instead, any statistical models as mathematically linguistic descriptions of certain regularities in data. On this instrumentalist view of modeling, it is futile to predicate model selection on the existence of "true" parameters—rather modelers should search for the most compact models with mean-innovation errors with respect to available information (for recent discussions in econometrics, see, e.g., Phillips, 1996; Phillips and Ploberger, 1996; Phillips and McFarland, 1997; Chao and Chiao, 1998; Reschenhofer, 1999).[16] Methodologically, this development may be seen as blending the ECM approach with a Bayesian perspective to reduce the dangers of overfitting.

In this instrumentalist approach, the error term may be seen as playing the role of demarcating, from given data information, what is unknowable from what is knowable, the demarcation being achieved by requiring the unknowable errors to be mean innovations with respect to the knowable part (i.e., the model itself). Any more extended interpretation of the error terms would appear redundant. But although a completely instrumentalist view of modeling is sustainable so long as the econometrician is concerned only with representation and prediction, intervention requires valid external reference, and this could make thoroughgoing instrumentalist positions unattractive in economics.

Econometricians have come a long way since the days of Slutsky and Frisch in their attempts to build better models. Because theoretical models purporting to describe the economic relationships among macroeconomic time series are often incomplete or otherwise inadequate, error terms became the most convenient expedient to fill the gap. However, this expedient is not without cost, and in particular, the interpretation of these error terms has proved to be problematic and controversial.

Two themes have been evident throughout the history of the subject. The first, which dates from Slutsky but which has been resurrected in RE and VAR modeling, views these errors as shocks that constitute the motive force behind business cycles. The second theme, formalized by Haavelmo, sees these errors as providing the statistical foundations for econometrics. Given an assumed distributional structure for the errors, one can obtain the maximum likelihood estimates of the model coefficients from the parameters of the distribution function.

There has been a tendency for econometricians to assert both of these positions simultaneously. Many methodological confusions in time-series modeling may be viewed as a result of this ambivalence. The interpretation of equation errors as shocks presupposes that they are autonomous and have the same status as other variables in the theoretical economic models. But in that case, there is no general argument that these shocks will exhibit those statistical properties embodied in the classical assumptions. As the second theme became dominant in mainstream econometrics, the remedy for violation of these assumptions became a change in estimation methods. But because the theoretical model remains unmodified, the attempted remedy did not invariably prove effective (poor forecasting performance, unstable coefficients, etc.). The alternative approach is to intermediate theory and data by an empirical model to give a looser representation of the structure. In the time-series literature, the latter route has led to a new class of dynamic models, which includes both VAR models and ECM's. We have argued that it is possible, within this framework, to combine the two positions in a logically coherent manner. VAR or ECM innovational errors map into a representation in which the variable of interest is seen as determined by the history of innovational shocks but also a set of autocorrelated "nuisance" disequilibrium terms. This insight provides the integrating framework within which we may hope to better understand earlier attempts at interpreting errors.

NOTES

1. The term structural takes a variety of meanings in econometrics. We use the term to refer to estimated parameters that are claimed (by the authors of the model) to have reference independent of the estimating model in which they arise. This will generally imply interpretability of the parameters into a theoretical model that will correspond more or less closely to the estimated model. By contrast, nonstructural parameters are parameters that arise out of the empirical specification of a model but that are devoid of any direct theoretical interpretation.

2. This paper covers a long period of econometrics history. External influences from the statistics literature have also been important, but we confine our attention to the econometrics literature.

3. Slutsky's paper was originally published in Russian in 1927.

4. Interestingly, Frisch wrote, "The concrete interpretation of the shock . . . does not interest us for the moment" (1933).

5. Notice that Haavelmo assumed that the error term was "stochastically independent" of its "previous values" instead of being serially uncorrelated (1944, pp. 72, 79).

6. These errors were referred to as "unobserved (latent) variables" and interpreted as "disturbances (or shocks), which represent the aggregate effects of additional unspecified exogenous variables on the economic decisions expressed by each relation" (Koopmans and Hood, 1953, p. 115). A more detailed explanation was recorded in one of the Cowles documents: "*Disturbances* ordinarily include both 'shocks' and 'errors.' Shocks are those factors a given theoretical system does not explicitly take into account or cannot explicitly take into account. They are usually small factors not separately noticeable. Errors are those in observation or measurement, which cannot in any case be made part of the economic theory underlying a system of equations. Unless otherwise stated, disturbances as used later will include only shocks, errors being assumed to be negligible" (Cowles Commission, 1947).

7. Notice that the equation does not relate Y_t to the initial conditions that would qualify it as a final equation.

8. The editor has drawn our attention to a famous remark, attributed to J.W. Tukey but which we have been unable to locate, that "Man makes X, God gives us U." It seems likely that Tukey thought of X as reflecting experimental design, whereas U were uncontrolled factors. More generally from an instrumentalist stance, we could view the "man-made" part, i.e., econometric models, as misspecified models to a greater or lesser extent and the residuals as God's reply to the model misspecification. Notice that such an instrumentalist viewpoint reemerges in the recent development, as discussed briefly in the final section of this paper.

9. Notice the difference of this MA representation from that of Slutsky (1937), who decomposes an economic variable into an MA of shocks rather than decomposing a shock variable into an MA of white-noise errors.

10. Interestingly, Muth chose the special case of $u_t = \sum_0^t \varepsilon_i$ as an example to show the importance of cross-equation restrictions in formulating $p_t^e = f(\beta, \gamma)$.

11. For example, Hansen and Sargent (1980) interpreted the error term in a structural relation as representing a random process observed, and responded to, systematically by those agents whose behavior is described by the relation. They justified the additional imposition of a serially dependent error term in RE models by assuming that the random process observed by the agents was not observed by the econometrician.

12. For instance, Sargent acknowledged several times the influence of the time-series modeling approach on his work (for more details, see Sent, 1998).

13. In the simplest VAR, the only prior restriction imposed on $A(L)$ is $A(0) = I$.

14. The view of the error terms in time-series analysis was also discussed in Granger and Newbold (1977). They noted two basic differences between the structuralist approach and the time-series approach, namely, "the determination of the lag structure of the model and the handling of the residuals." They wrote,

> Residual terms are usually treated by econometricians as being mere nuisances of little real importance. Many econometric texts introduce models as a set of deterministic relationships between variables and then casually add on the "error" terms to equations to account for such things as model misspecification and variable measurement error. The very use of names such as "residuals" and "errors" implies value judgements about the importance of these terms. A better name might be "innovations," although "error" is a proper designation in a forecasting context. (p. 8)

15. Rissanen remarks that "the assumption of a 'true' distribution is not needed in this theory" (1989, p. 4).

16. Rissanen proposes to view models as algorithmic encoding of data, which amounts to the imposition of constraints on the data. Following the theory of algorithmic complexity, he develops a theory of stochastic complexity to capture the essence of modeling as searching for the minimum description length (MDL) or the shortest code length with respect to data information.

REFERENCES

Blanchard, M. & D. Quah (1989) The dynamic effects of aggregate demand and supply disturbances. *American Economic Review* 79, 655–673.

Box G.E.P. & G.M. Jenkins (1970) *Time Series Analysis: Forecasting and Control*. San Francisco: Holden Day.

Cagan, P. (1956) The monetary dynamics of hyperinflation. In M. Friedman (ed.), *Studies in the Quantity Theory of Money*, pp. 25–117. Chicago: University of Chicago Press.

Chao, J.C. & C.S. Chiao (1998) Testing the expectations theory of the term structure of interest rates using model-selection methods. *Studies in Nonlinear Dynamics and Econometrics* 12, 95–108.

Cochrane, D. & G. Orcutt (1949) Application of least squares regression to relationships containing autocorrelated error terms. *Journal of American Statistical Association* 44, 32–61.

Cowles Commission (1947) *Five-Year Report 1942–46*. University of Chicago.

Durbin, J. (1960) Estimation of parameters in time-series regression models. *Journal of the Royal Statistical Society, Series B* 22, 139–153.

Durbin, J. & G.S. Watson (1950) Testing for serial correlation in least squares regression: I. *Biometrica* 37, 409–428.

Frisch, R. (1933) Propagation problems and impulse problems in dynamic economics. In *Economic Essays in Honour of Gustav Cassel*, pp. 171–205. London: Allen and Unwin.

Frisch, R. (1934) *Statistical Confluence Analysis by Means of Complete Regression Systems*. Oslo: Universitets Økonomiske Institutt.

Frisch, R. (1938) Statistical versus Theoretical Relations in Economic Macrodynamics (reproduced by University of Oslo in 1948 with Tinbergen's comments). Memorandum, Oslo. Also in D.F. Hendry & M.S. Morgan (eds.) (1995) *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press.

Gilbert, C.L. (1988) *The Development of Econometrics in Britain since 1945*. Ph.D. thesis, Oxford University.

Gilbert, C.L. (1991) Do economists test theories? Demand analysis and consumption analysis as tests of theories of economic methodology. In N. de Marchi & M. Blaug (eds.), *Appraising Economic Theories: Studies in the Methodology of Research Programs*, pp. 137–168, Cheltenham: Edward Elgar.

Granger, C.W.J. & P. Newbold (1977) *Forecasting Economic Time Series*. New York: Academic Press.

Gregory, A.W. & G.W. Smith (1995) Business cycle theory and econometrics. *Economic Journal* 105, 1597–1608.

Haavelmo, T. (1943) The statistical implications of a system of simultaneous equations. *Econometrica* 11, 1–12.

Haavelmo, T. (1944) *The Probability Approach in Econometrics*. Supplement to *Econometrica* 12.

Hansen, L.P. & T.J. Sargent (1980) Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control* 2, 7–46.

Hendry, D.F. (1995) *Dynamic Econometrics*. Oxford: Oxford University Press.

Hendry, D.F. & G.E. Mizon (1978) Serial correlation as a convenient simplification, not a nuisance: A comment on a study on the demand for money by the Bank of England. *Economic Journal* 88, 549–563.

Hendry, D.F. & M.S. Morgan (1989) A re-analysis of confluence analysis. *Oxford Economic Papers* 41, 35–52.

Hotelling, H. (1927) Differential equations subject to error. *Journal of the American Statistical Association* 22, 283–314.

Hurwicz, L. (1944) Stochastic models of economic fluctuations. *Econometrica* 12, 114–124.

Kim, K. & A.R. Pagan (1995) The econometric analysis of calibrated models. In M.H. Pesaran & M. Wickens (eds.), *Handbook of Applied Econometrics*. Oxford: Blackwell.

Koopmans, T.C. (1937) *Linear Regression Analysis of Economic Time Series*. Haarlem: Netherlands Economic Institute.

Koopmans, T.C. (ed.) (1950) *Statistical Inference in Dynamic Economic Models*. Cowles Commission Monograph 10. New York: Wiley.

Koopmans, T.C. & W.C. Hood (1953) The estimation of simultaneous linear economic relationships. In W.C. Hood & T. Koopmans (eds.), *Studies in Econometric Method*, Cowles Commission Monograph 14, pp. 112–199, New Haven, Connecticut: Yale University Press.

Kydland, F. & E. Prescott (1991) The econometrics of the general equilibrium approach to business cycles. *Scandinavian Journal of Economics* 93, 161–178.

Leamer, E.E. (1978) *Specification Searches: Ad Hoc Inference with Non-Experimental Data*. New York: Wiley.

Liu, T.-C. (1960) Underidentification, structural estimation and forecasting. *Econometrica* 28, 855–865.

Lucas, R.E. (1976) Econometric policy evaluation: A critique. In K. Brunner & A. Meltzer (eds.), *Stabilization of the Domestic and International Economy*, pp. 7–29. Amsterdam: North-Holland.

Lucas, R.E. & E.C. Prescott (1971) Investment under uncertainty. *Econometrica* 39, 659–681.

Lucas, R.E. & T.J. Sargent (1979) After Keynesian macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review* 3(1), 1–16.

Mann, H.B. & A. Wald (1943) On the statistical treatment of linear stochastic difference equations. *Econometrica* 11, 173–220.

Marschak, J. (1953) Economic measurements for policy and prediction. In W.C. Hood & T. Koopmans (eds.). *Studies In Econometric Method*, Cowles Commission Monograph 14, pp. 1–26. New Haven, Connecticut: Yale University Press.

Mizon, G.E. (1977) Inferential procedures in nonlinear models: An application in a UK industrial cross section study of factor substitution and returns to scale. *Econometrica* 45, 1221–1242.

Morgan, M.S. (1989) *The History of Econometric Ideas*. Cambridge: Cambridge University Press.

Morgan, M.S. (1991) The stamping out of process analysis in econometrics. In N. de Marchi & M. Blaug (eds.), *Appraising Economic Theories: Studies in the Methodology of Research Programmes*, pp. 237–272. Cheltenham: Edward Elgar.

Muth, J.F. (1961) Rational expectations and the theory of price movements. *Econometrica* 29, 315–335.

Orcutt, G. (1948) A study of the autoregressive nature of the time series used for Tinbergen's model of the economic system of the United States 1919–1932. *Journal of the Royal Statistical Society, Series B* 10, 1–45.

Pagan, A.R. (1987) Three econometric methodologies: A critical appraisal. *Journal of Economic Surveys* 1, 3–24.

Pagan, A.R. (1995) Three econometric methodologies: An update. In L. Oxley, D.A.R. George, C.L. Roberts, & S. Sayer (eds.), *Surveys in Econometrics*, pp. 30–41. Oxford: Blackwell.

Pesaran, M.H. (1988) *The Limits to Rational Expectations*. Oxford: Blackwell.

Phillips, P.C.B. (1996) Econometric model determination. *Econometrica* 64, 763–812.

Phillips, P.C.B. & J.W. McFarland (1997) Forward exchange market unbiasedness: The case of the Australian dollar since 1984. *Journal of International Money and Finance* 16, 885–907.

Phillips, P.C.B. & W. Ploberger (1996) An asymptotic theory of Bayesian inference for time series. *Econometrica* 64, 381–412.

Qin, D. (1993) *The Formation of Econometrics: A Historical Perspective*. Oxford: Oxford University Press.

Qin, D. (1996) Bayesian econometrics: The first twenty years. *Econometric Theory* 12, 500–516.

Quah, D.T. (1995) Business cycle empirics: Calibration and estimation. *Economic Journal* 105, 1594–1596.

Reschenhofer, E. (1999) Improved estimation of the expected Kullback–Leibler discrepancy in case of misspecification. *Econometric Theory* 15, 377–387.

Rissanen, J. (1987) Stochastic complexity. *Journal of the Royal Statistical Society, Series B*, 223–239 and 252–265.

Rissanen, J. (1989) *Stochastic Complexity in Statistical Inquiry*. Singapore: World Scientific Publishing.

Sargan, J.D. (1959) The estimation of relationships with autocorrelated residuals by the use of instrumental variables. *Journal of the Royal Statistical Society, Series B* 21, 91–105.

Sargan, J.D. (1961) The maximum likelihood estimation of economic relationships with autoregressive residuals. *Econometrica* 29, 414–426.

Sargent, T.J. (1976) A classical macroeconometric model for the United States. *Journal of Political Economy* 84, 207–237.

Sargent, T.J. (1977) The demand for money during hyperinflations under rational expectations. *International Economic Review* 18, 59–82.

Sargent, T.J. (1978) Estimation of dynamic labor demand schedules under rational expectations. *Journal of Political Economy* 86, 1009–1044.

Sargent, T.J. (1979) Estimating vector autoregressions using methods not based on explicit economic theories. *Federal Reserve Bank of Minneapolis Quarterly Review* 3(3), 8–15.

Sargent, T.J. & C.A. Sims (1977) Business cycle modeling without pretending to have too much a priori economic theory. In C.A. Sims (ed.) *New Methods in Business Cycle Research: Proceedings from a Conference*, pp. 45–109. Federal Reserve Bank of Minneapolis.

Sargent, T.J. & N. Wallace (1973) Rational expectations and the dynamics of hyperinflation. *International Economic Review* 14, 328–350.

Sargent, T.J. & N. Wallace (1975) "Rational" expectations, the optimal monetary instrument, and the optimal money supply rule. *Journal of Political Economy* 83, 241–254.

Sent, E.-M. (1998) *The Evolving Rationality of Rational Expectations: An Assessment of Thomas Sargent's Achievements*. Cambridge: Cambridge University Press.

Sims, C.A. (1980) Macroeconomics and reality. *Econometrica* 48, 1–48.

Slutsky, E. (1937) The summation of random causes as the source of cyclic processes. *Econometrica* 5, 105–146 (originally published in Russian in 1927).

Stone, R.H. (1954) *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom 1920–1938*. Cambridge: Cambridge University Press.

Theil, H. (1957) Specification errors and the estimation of economic relationships. *Review of International Statistical Institute* 25, 41–51.

Theil, H. (1958) *Economic Forecasts and Policy*. Amsterdam: North-Holland.

Tinbergen, J. (1939) *Statistical Testing of Business Cycle Theories*. Geneva: League of Nations.

Tintner, G. (1938) A note on economic aspects of the theory of errors in time series. *Quarterly Journal of Economics* 53, 141–149.

Tintner, G. (1940a) *The Variate Difference Method*. Cowles Commission Monograph 5. Bloomington, Indiana.

Tintner, G. (1940b) The analysis of economic time series. *Journal of the American Statistical Association* 35, 93–100.

Wallis, K.F. (1980) Econometric implications of the rational expectations hypothesis. *Econometrica* 48, 49–73.

Whiteman, C.H. (1983) *Linear Rational Expectations Models: A User's Guide*. Minneapolis: University of Minnesota Press.

Wicksell, K. (1907) Krisernas gåta. *Statsøkonomisk Tidsskrift*, 255–286, Oslo.

Wold, H. (1938) *A Study in the Analysis of Stationary Time Series*. Uppsala: Almqvist and Wiksells.

Yule, G. (1927) On a method of investigating periodicities in disturbed series, with special application to Wolfert's sun spot numbers. *Philosophical Transactions of the Royal Society, Series A* 226, 267–298.

Appendix: A Road Map

Authors	Key viewpoints
Slutsky (1927)	Represented cyclical variables as moving averages of independent shocks that are the source of economic dynamics.
Frisch (1933)	Saw economic variables as having a damped deterministic sinusoidal component augmented by a moving average of random "impulses."
Frisch (1934)	Distinguished empirically between "aberrations" and "stimuli"—"stimuli" correspond to "impulses" and are transmitted into the variable dynamics, whereas "aberrations" are transient and play the role of measurement errors.
Wold (1938)	Decomposed any stationary univariate time series into a deterministic component plus an infinite moving average, which, under invertibility, may be expressed as an autoregression with finite but sufficiently long lags plus a serially uncorrelated residual.
Tintner (1938)	Distinguished between small, random optimization errors, reflected in estimation residuals, and often large forecasting errors, which generate economic dynamics.
Tinbergen (1939)	Saw error terms as errors in equations, as distinct from measurement errors; errors are not interpreted.
Haavelmo (1944)	Modeled distribution of variables of interest on the basis of assumed joint normal distribution of equation errors; recognized serial independence as a simplification of the normal distribution.
Mann and Wald (1943) and Koopmans (1950)	Recommended reduced-form estimation, so that reduced-form errors become a linear transformation of the original structural errors.
Hurwicz (1944)	Noted that errors on the "reduced" equation may be autocorrelated.
Orcutt (1948)	Viewed macroeconomic variables as typically following an AR(2) with a unit root.
Cochrane and Orcutt (1949)	Saw error terms as generally being autocorrelated and recommended estimation of differenced equations.
Theil (1958)	Argued that true structure is unknown and that residual and cross-equation correlation suggest respecification—errors atheoretical.
Sargan (1959, 1961) and Durbin (1960)	Argued that autoregressive models should be seen as restricted versions of more general distributed lag models.

(continued)

Appendix: Continued

Authors	Key viewpoints
Muth (1961)	Economic dynamics arises from autocorrelated shocks; agents forecast the noninnovational component of these shocks.
Leamer (1978)	Saw errors as the sum of maneuverable unobservables dependent on model specification and so atheoretical.
Lucas and Sargent (1979)	Saw reduced forms of RE models as having ADL or VAR form—errors are constructed and hence in principle atheoretical; they are nevertheless used to examine impulse responses.
Sims (1980)	Argued that VAR residuals should be innovation by construction and could be viewed as shocks to modeled variables as a result of equivalence between the VAR and MA representations.
Hendry (1995)	Viewed errors as model-derived mean-innovation processes resulting from statistical reduction and thus totally atheoretical.