Journal of Monetary Economics 16 (1985) 283–308. North-Holland

ATHEORETICAL MACROECONOMETRICS
A Critique

Thomas F. COOLEY and Stephen F. LeROY *

University of California, Santa Barbara, CA 93106, USA

The spurious nature of the restrictions used to identify many macroeconometric models has led some researchers to advocate a style of econometric inquiry that is less dependent on prior theoretical restrictions of the sort that were central to the approach of the Cowles Commission. This development, which we call atheoretical macroeconometrics, is summarized and evaluated in the current paper. It is contrasted with an updated version of the Cowles Commission approach. We conclude that while some of the exercises of atheoretical macroeconometrics are valid, those that have attracted the most attention and appear the most innovative – exogeneity testing, impulse response analysis and policy analysis using estimated vector autoregressions – are based on incorrect analysis.

1. Introduction

In the 1940s economists at the Cowles Commission and the National Bureau of Economic Research debated the relative roles of induction and deduction in economics, particularly macroeconomics. Koopmans (1947) likened the situation to that in astronomy where Brahe and Kepler carefully documented the stylized facts of planetary motion, which facts remained unexplained until Newton articulated the underlying theory that was capable of reconciling them. In Koopmans' view the business cycle research of the NBER represented the 'Kepler style' of inquiry, while the program of the Cowles Commission advocated a unification of the Kepler style with the 'Newton style'. The succeeding two decades were characterized by growing optimism about the utility and relevance of Keynesian macroeconometric models based on the Cowles approach. More recently, however, the ascendence of neoclassical macroeconomics, with its emphasis on rational expectations and general equilibrium modeling, has engendered much skepticism about the validity of the *a priori* restrictions used to identify Keynesian econometric models. The reasons for this skepticism, being related to the well-known Lucas (1976) critique, are sufficiently familiar not to require review here.

* We are indebted to David Bessler, Edward Leamer, Jack Marshall, Charles Nelson, William Parke, Richard D. Porter, John W. Pratt, Neil Raymon, Andrew Rose, and John Seater for helpful comments on an earlier version of this paper. The suggestions of Robert King and Charles Plosser were particularly helpful in preparing the current version.

One response to this disenchantment with the received macroeconometrics has been to return to a Kepler style of inquiry that is less dependent on prior theoretical restrictions but that uses modern time series methods. This development, which is due primarily to Christopher Sims, together with his coworkers and students, is summarized and evaluated in the present paper. Briefly, our conclusion is that while some of the interrelated analytical exercises which we collectively term 'atheoretical macroeconometrics' are valid – forecasting, data description, hypothesis-searching, and certain types of theory-testing – those applications which have attracted the most attention and which appear most original – exogeneity testing, innovation accounting, impulse response analysis, policy appraisal – are based on incorrect analysis. All the criticisms of atheoretical macroeconometrics set out in this paper have been made before, some many times. The present paper is motivated by our conviction that, despite this, many practicing macroeconomists either are unaware of these criticisms or minimize their importance.

We begin in section 2 with an introduction to multiple time series methods consisting of a non-technical description of vector autoregressions (VARs). Section 3 consists of an itemization and brief explanation of some of the applications of VARs: forecasting, causality testing, tests of theories, hypothesis-seeking, data characterization, innovation accounting, impulse response analysis and policy analysis. We then review in section 4 the received Cowles Commission approach to econometric modeling. Included here is our interpretation of the distinctions between deep and shallow parameters, and between exogenous and endogenous variables. In section 5 the use of the terms 'causality' and 'exogeneity' in atheoretical macroeconometrics is contrasted with that of 'causality' by economists using the older methods. The burden of this discussion is that the concept of causality or exogeneity that is tested by the Granger and Sims tests is not closely related to the causality of the Cowles Commission economists, contrary to the presumption implicit in the usage of atheoretical macroeconometrics. Section 6 extends the critique of atheoretical macroeconometrics to innovation accounting, impulse response analysis, and policy analysis, which, like exogeneity testing, will not sustain the interpretations frequently placed upon them. In section 7 we formulate at a more abstract level the distinction between two alternative modeling strategies: structural and non-structural. Our view is that here we are doing no more than making explicit a dichotomy that was implicit in the use of the term 'structural' by the Cowles economists, and in their dictum that correlation does not imply causation. Section 8 offers our conclusions.

2. Introduction to multiple time series models

Underlying all applications of atheoretical macroeconometrics is the vector autoregression, or VAR model. Now, a scalar (as distinguished from vector) autoregression is just a regression of a variable on its own past values:

$$x_t = \sum_{i=1}^{n} \pi_i x_{t-i} + u_t, \tag{1}$$

where u_t is a mean-zero, serially uncorrelated unobservable scalar random variable, x_t is an observable scalar random variable, and the π are scalar parameters. A vector autoregression, as its name implies, differs from the above in that x_t and u_t are vectors and the coefficients are square matrices:

$$x_t = \sum_{i=1}^{n} \Pi_i x_{t-i} + u_t, \qquad \mathrm{E}(u_t u_t') = \Sigma. \tag{2}$$

In estimating VAR models, the econometrician makes no attempt to use theory to distribute zeros in the coefficient matrices, so that prior information guides only the selection of the variables to enter x_t and the lag length n. In Sims' (1980a) paper, around which we structure our discussion, the variables comprising x_t are money, GNP, the unemployment rate, the wage rate, the price level and an import price index, and the lag length is four quarters. Thus the Π_i which are to be estimated consist of four six-by-six matrices.

The errors in (2) are not assumed to be contemporaneously uncorrelated. To facilitate interpretation it is customary to premultiply (2) by the unique triangular matrix with units on the main diagonal that diagonalizes the error covariance matrix:

$$Tx_t = T\sum_{i=1}^{n} \Pi_i x_{t-i} + \eta_t, \qquad \mathrm{E}(\eta_t \eta_t') = D, \tag{3}$$

where $\eta_t = Tu_t$ and $D = T\Sigma T'$, a diagonal matrix. The η_t are termed the *orthogonalized innovations.*

All this is easier to interpret in a simple two-variable example. We assume that the maximum lag length is one period and, following Sims (1972), that the two variables are the money stock m and income y.[1] Then the reduced form corresponding to (2) would be estimated as

$$m_t = \pi_{11} m_{t-1} + \pi_{12} y_{t-1} + u_{mt}, \tag{4}$$

$$y_t = \pi_{21} m_{t-1} + \pi_{22} y_{t-1} + u_{yt}, \tag{5}$$

where $\mathrm{E}(u_{mt}^2) = \sigma_{mm}$, $\mathrm{E}(u_{yt}^2) = \sigma_{yy}$, and $\mathrm{E}(u_{mt} u_{yt}) = \sigma_{my}$. If (4) is multiplied by

[1]Alternatively, we may think of y as a vector of quantities and relative prices consisting of all the elements of x except m, as in Sims (1980a, p. 27). Only a minor change in notation is required to accommodate this interpretation.

σ_{my}/σ_{mm} and the result subtracted from (5), the transformed system is

$$m_t = \rho_{11} m_{t-1} + \rho_{12} y_{t-1} + \eta_{mt}, \tag{6}$$

$$y_t = \delta m_t + \rho_{21} m_{t-1} + \rho_{22} y_{t-1} + \eta_{yt}, \tag{7}$$

where η_{mt} and η_{yt} are uncorrelated both contemporaneously and serially, and δ and the ρ_{ij} are defined from the π_{ij} and σ_{ij} in an obvious way. Eqs. (6) and (7) correspond to (3) above.

3. Applications of vector autoregressions

VAR models, as described in the preceding section, underlie all applied work in atheoretical macroeconometrics. In this section we describe some of the uses to which VAR models are put:

(a) *Forecasting*. Since VAR models allow complete flexibility and generality (except for the linearity assumption) in specifying the correlations between future, present and past realizations of the system variables, they have a natural application to forecasting. Examples of the use of VAR models in forecasting are Sims (1982) and Doan, Litterman and Sims (1984).

(b) *Causality tests*. Suppose we have two time series $\{m_t\}$ and $\{y_t\}$. The series $\{y_t\}$ fails to Granger-cause $\{m_t\}$ according to the Granger (1969) test if, in a regression of m on lagged m and lagged y, the latter takes on a zero coefficient. In terms of the VAR model presented above, the regression of interest is just eq. (4), and the term that must equal zero is the coefficient π_{12}. Similarly, $\{y_t\}$ fails to Granger-cause $\{m_t\}$ according to the Sims (1972) test if, in a regression of y on lagged y and future m, the latter takes on a zero coefficient.[2] Although there are econometric differences between these tests [see Hosoya (1977), Pierce (1977), Feige and Pearce (1979), Chamberlain (1982), Florens and Mouchart (1982) and Kohn (1982)], it is known that the Granger and Sims tests are implications of the same null hypothesis [Jacobs, Leamer and Ward (1979)]. If y fails to Granger-cause m, it is said that m is *exogenous* with respect to y. If in addition m does Granger-cause y, m is said to be *causally prior* to y.

To understand the Granger and Sims tests, suppose that the Federal Reserve determines the money stock m by spinning a roulette wheel. The money stock can depend on past as well as present spins of the wheel, but assume that the

[2]Alternatively, one can prewhiten the data by using Box–Jenkins methods to obtain the univariate innovations of each series, and then ascertain whether the innovations in $\{x_t\}$ can be predicted from those in $\{y_t\}$. This latter procedure is advocated by Pierce and Haugh (1977).

Federal Reserve pays no attention whatever to y in setting the money stock. Now, if in this environment one regresses the money stock on its own past values and past y, in large samples the latter will take on a zero coefficient (Granger test). Similarly, because the Federal Reserve pays no attention to y in setting m, there is no causal link either way between y and future values of m. Hence the prediction of y implied by past y will not be changed if future m is included among the predictors (Sims test).

(c) *Tests of theories.* Many theories, especially in macroeconomics and finance, have the implication that one variable of the system should fail to Granger-cause another. The Granger and Sims procedures therefore constitute tests of these theories which are valid in the usual sense: if Granger-causality is absent, the theory is supported, while if it is present the theory is not supported. Examples are:

(1) *Martingale models.* In a wide class of finance models, rates of return have the 'efficient capital markets' property

$$E(r_t|I_t) = \rho, \tag{8}$$

where r_t is the rate of return on some security or durable good from t to $t+1$, I_t is the information set, assumed identical over individuals, and ρ is a positive constant [see LeRoy (1982) and the papers cited there for discussion of the theoretical basis for (8)]. It follows from (8) that no variable in I_t Granger-causes r_t, so that the Granger and Sims procedures provide tests of capital market efficiency.

(2) *Optimal control.* If y_t is a target variable and x_t a control variable, then under certain restrictive conditions (linear structure, quadratic criterion function, no adjustment costs) optimal control will reduce y_t to white noise, so that the failure of x_t to Granger-cause y_t is an implication of the model.

(3) *Rational expectations monetarism.* According to classical monetarism, variations in the money stock are a consequence of the animal spirits of central bankers, and are not a response to events in the economy. If so, GNP should fail to Granger-cause money. Further, one rational expectations version of monetarism implies that only the unexpected portion of monetary changes affects GNP, so that money should fail to Granger-cause GNP. This rational expectations model implies that neither GNP nor money should Granger-cause the other [Sargent (1976, 1979), Nelson (1979)].

(4) *Measurement error.* If, in a supply–demand model, either the supply or demand equation is without error, and if price is measured subject to error,

quantity will appear to be causally prior to observed price. More generally, '[w]henever there is a very large measurement error in one variable, not in the other, the other variable will tend to appear exogenous in an equation with the error-ridden variable on the left' [Sims (1977), p. 35)].

(5) *Permanent-income consumption models.* If individuals maximize $\mathrm{E}\Sigma_t(1+\rho)^{-t}u(c_t)$ subject to a random income stream and constant real interest rate r, the first-order condition is

$$\mathrm{E}_t(u'_{t+1}) = \frac{1+\rho}{1+r}u'_t, \tag{9}$$

Given certain additional assumptions, (9) says that the marginal utility of consumption follows a martingale with drift $(1+\rho)/(1+r)$ [Hall (1978)]. If utility is quadratic, the marginal utility of consumption is a linear function of its level, implying that income fails to Granger-cause consumption.

(d) *Hypothesis-seeking.* One can, of course, reverse the sequence of theorizing and empirical testing. That is, econometricians can use VAR models to generate stylized facts about causal orderings of macroeconomic variables that seem to be robust empirically. Then theorists would try to explain these patterns. For example, Sims (1972) found that causality was unidirectional from money to income. However, when a nominal interest rate was added to the $m-y$ system, it turned out that interest rate innovations explain most of the comovement in m and y [Sims (1980b)]. Litterman and Weiss (1985) constructed a model that explains this change in the causal ordering. Another example is Ashenfelter and Card's (1982) use of a VAR model to characterize the time series behavior of wages, prices, interest rates and unemployment. They then compared these stylized facts to predictions of alternative theoretical models of the labor market.

(e) *Data characterization.* VAR models can be used to provide summary characterizations of the cyclical behavior of a system of macroeconomic variables. To see this, note that under weak restrictions on the coefficient matrices, the VAR model (3) can be inverted and written in moving average form

$$x_t = \sum_{i=0}^{\infty} \Lambda_i \eta_{t-i}, \tag{10}$$

where the η_t are contemporaneously and serially uncorrelated. Whether the system will be explosive or damped, with or without cycles, depends on the roots of the matrices Λ_i, which can be readily calculated from an estimated VAR model [Sims (1980b)].

(f) *Innovation accounting and impulse response analysis.* The purpose of innovation accounting (or, synonymously, variance decompositions) is to determine the proportion of each variable's forecast error that is attributable to each of the orthogonalized innovations in the VAR model. From the moving average representation (10) of a VAR model, each variable can be written as a function of the innovations, so that the response of the ith element of x_{t+k} to the innovation in the jth variable at date t is just the i, j element of the matrix Λ. A tabulation of those responses for $k = 0, 1, \ldots$ is called an impulse response function. Because the covariances among the innovations are zero by definition, the variance of each variable will be a weighted sum of the variances of each variable, with the weights being determined by the elements of the Λ_k. Innovation accounting is the exercise of determining which innovations contribute to the forecast error of each variable.

(f) *Policy analysis.* In recent papers, Sims (1980a, 1982) argued that VAR models are useful for analysis of the effects of alternative monetary or fiscal policies. Assuming that a Granger or Sims test indicates that the policy instrument is exogenous, one can view VAR model forecasts conditional on different hypothetical values of the instrument as capturing the effect of alternative instrument settings on the endogenous variables. In an extension of this argument, Sims (1982) contended further that existing Keynesian models may also be serviceable for policy analysis, again subject to the proviso that the hypothesized exogeneity of the instrument be verified via a Granger or Sims test, even though the theoretical restrictions incorporated in Keynesian models are 'incredible'. This is so, Sims argued, because these restrictions do not greatly affect the estimated reduced form, so that even if the structural form of the model is grossly misspecified, the forecasts the model generates are in practice not too different from those of a VAR model. Here, of course, Sims explicitly took issue with Lucas and Sargent's (1979) opinion that 'the difficulties [with Keynesian macroeconomic models] are *fatal*: that [they] are of *no* value in guiding policy and this condition will not be remedied by modifications along any line which is currently being pursued' (p. 2, emphasis in original).

4. The Cowles Commission program

In the 1940s and 1950s economic theorists and econometricians, particularly several associated with the Cowles Commission, developed an interrelated set of ideas and analytical tools for the estimation of macroeconometric models. The names of Marschak, Koopmans and Simon are prominent. The Cowles

Commission program constituted, at least in principle, the methodological underpinning of the Keynesian macroeconometric models against which both neoclassical macroeconomics and atheoretical macroeconometrics were a reaction. In this section we review the Cowles Commission program before going on to the main task of the paper: determining the relation between the Cowles program and that of atheoretical macroeconometrics, and undertaking a critique of the latter. Like others, we have found that econometricians working in the Cowles tradition have given only vague and incomplete definitions even of such central notions as 'exogeneity' and 'structure'. Greater precision is needed if comparison is to be made with atheoretical macroeconometrics, and we therefore fill in some of the gaps in the Cowles exposition. Our discussion seems to us to be consistent with the spirit of the Cowles program, but we concede the possibility that we may be reading our own ideas onto those of the Cowles economists.

Cowles economists distinguished between parameters and variables: the latter vary over time while the former do not.[3] A further distinction is that between deep and shallow parameters (the terms 'structural' and 'non-structural' are sometimes used in place of deep and shallow; we prefer to avoid ambiguity by reserving the former terms to the use defined in section 7). A parameter is considered to be a deep parameter if the model-builder is prepared to assume that it is functionally independent of other deep parameters. For example, parameters representing preferences or production possibilities are deep: a change in production possibilities entails no change in preferences, or vice versa. The idea of invariance under intervention is implicit in that of deep parameters – if a parameter is deep, it makes sense to compare the *cet. par.* behavior of endogenous variables for different values of that parameter. The reason this procedure makes sense is that '*cet. par.*' has unambiguous meaning: all other deep parameters are assumed invariant under the assumed change in the parameter of interest. Thus, for example, it makes good sense to compare the behavior of asset prices under risk-neutrality and risk-aversion, assuming implicitly that there is no difference in production possibilities.

On the other hand, for a shallow parameter such a conceptual experiment is ill-defined. For example, suppose that in a particular model it could be proved that the money stock m and the nominal interest rate i covary negatively, and also that m and y covary positively. The question 'would m and y still covary positively if instead m and i covaried positively?' is, in general settings, ambiguous. This is so because these covariances are not deep parameters, but rather are functions of deep parameters. Except perhaps in special cases, it will

[3] This distinction seems entirely elementary; we remind the reader of it only because it has been thoroughly blurred in the neoclassical treatment of policy regimes [see Cooley, LeRoy and Raymon (1984a, b)].

be unclear what intervention in the deep parameters is conceived to have led to the assumed change in the covariance of m and i, and the effect on the covariance of m and y will vary in different cases. Accordingly, the answer to questions like that just posed can only be 'it depends'.

Turning from parameters to variables, the Cowles economists distinguished between exogenous and endogenous variables. As the name implies, an exogenous variable is one determined outside the model, while an endogenous variable is determined by the model. Since an exogenous variable is determined outside the model, it is legitimate to perform hypothetical experiments consisting of varying any of the exogenous variables, *cet. par.* (where '*cet. par.*' means that all other exogenous variables are assumed unchanged), and determining the effect of this intervention on the endogenous variables. Thus for the Cowles economists causation is bound up with exogeneity: x can be causally prior to y if and only if x is exogenous and y endogenous.

To clarify subsequent analysis it is convenient to write whatever model is under discussion so that all the exogenous variables are mutually uncorrelated. This assumption is no more than a normalization: any system can be written so as to have the desired property, although, as we will see, it matters how the reformulation is done. The point of the exercise is to enforce an explicit statement of what exogeneity properties are and are not assumed, rather than allow these assumptions to remain buried in uninterpreted correlations. For example, suppose that we start with the model

$$m = \varepsilon_1, \tag{11}$$

$$y = \gamma m + \varepsilon_2, \tag{12}$$

where ε_1 and ε_2 are correlated. If the analyst were willing to assume that the correlation between ε_1 and ε_2 occurred because ε_1 determines a component of ε_2 – i.e., $\varepsilon_2 = \lambda\varepsilon_1 + \varepsilon_3$ – then (11)–(12) could be rewritten as

$$m = \varepsilon_1, \tag{13}$$

$$y = \gamma m + \lambda\varepsilon_1 + \varepsilon_3, \tag{14}$$

with ε_1 and ε_3 uncorrelated. Since m is exogenous in this setup, the effect of a change in m on y is well-defined: $\mathrm{d}y = (\gamma + \lambda)\mathrm{d}\varepsilon_1$. Here $\gamma + \lambda$ could be estimated by regression; γ and λ, of course, are not separately identified.

If, on the other hand, the analyst were willing to specify that the correlation between ε_1 and ε_2 of (11)–(12) owes to a causal link in the reverse direction, the system would be rewritten

$$m = \varepsilon_4 + \delta\varepsilon_2, \tag{15}$$

$$y = \gamma m + \varepsilon_2, \tag{16}$$

with ε_2 and ε_4 uncorrelated. Now the question 'What is the effect of m on y?' is not well-posed, since the answer depends on whether the assumed shift in m is due to an underlying change in ε_4 (in which case the answer is $\mathrm{d}y = \gamma\,\mathrm{d}\varepsilon_4$) or in ε_2 (in which case the answer is $\mathrm{d}y = (\gamma\delta + 1)\,\mathrm{d}\varepsilon_2$). If the analyst were willing to assume neither of these one-way causal links, the model could be rewritten as

$$m = \varepsilon_1 + \delta\varepsilon_2, \tag{17}$$

$$y = \gamma m + \lambda\varepsilon_1 + \varepsilon_2, \tag{18}$$

which contains the two examples just discussed as special cases, or

$$m = \varepsilon_1 + \varepsilon_3, \tag{19}$$

$$y = \gamma m + \varepsilon_2 + \delta\varepsilon_3, \tag{20}$$

which attributes the link between ε_1 and ε_2 of (11)–(12) to the common effect of some third unobserved variable ε_3. In both cases the question 'What is the effect of m on y?' is again not well-posed.

The same point may be put in terms of the model

$$m = \theta y + \varepsilon_1, \tag{21}$$

$$y = \gamma m + \varepsilon_2, \tag{22}$$

which differs from (11)–(12) in that the error in the first equation has been redefined.[4] Suppose ε_1 and ε_2 are uncorrelated. It is obvious from the reduced form

$$m = (\varepsilon_1 + \theta\varepsilon_2)/(1 - \gamma\theta), \tag{23}$$

$$y = (\gamma\varepsilon_1 + \varepsilon_2)/(1 - \gamma\theta), \tag{24}$$

that the question 'what is the effect of m on y, *cet. par.*?' is ill-defined because, as above, it is unclear whether the change in m is due to an intervention in ε_1, in ε_2, or both. The ambiguity disappears, of course, if $\theta = 0$, so that (21)–(22) becomes equivalent to (11)–(12).

We have seen that the notion of exogeneity involves the idea of intervention: a change in an exogenous variable is envisaged, and the effect of this intervention on the endogenous variables is calculated. Exogeneity also involves the idea of invariance under intervention: a *cet. par.* assumption is made, and this

[4] In terms of the ε_1 and ε_2 of eqs. (11) and (12), ε_1 $(= m - \theta y)$ of eq. (21) equals $(1 - \theta\gamma)\varepsilon_1 - \theta\varepsilon_2$.

restriction must have unambiguous meaning so that the hypothesized intervention is clearly defined. In order that the required invariance under intervention have unambiguous meaning in all contexts, particularly large systems, the assumption that all exogenous variables be uncorrelated is required. Fortunately, however, in particular applications it is frequently not necessary that all conceivable hypothetical experiments have meaning in all possible contexts. By limiting the set of interventions one seeks to analyze and by exploiting prior structural restrictions, it is often possible to demonstrate that the effects of a particular intervention do not depend on whether a particular correlation between exogenous variables is viewed as causal. In such cases there is no reason to bother with the exercise just outlined of rewriting the model so that all the exogenous variables are mutually uncorrelated.

To see this, consider the model

$$m = \varepsilon_1, \tag{25}$$

$$z = \tau m + \varepsilon_2, \tag{26}$$

$$y = \gamma m + \beta z + \varepsilon_3, \tag{27}$$

where the analyst is for some reason willing to assume that ε_3 is uncorrelated with either ε_1 or ε_2, but not that ε_1 and ε_2 are uncorrelated. In terms of the definition of exogeneity outlined above, only ε_3 is exogenous: because of the assumed correlation between ε_1 and ε_2, the effect of an intervention in m (or, equivalently, ε_1) on z, *cet. par.*, is ambiguous. But note that in this case, the effect of m on y, *cet. par.*, does not depend on whether the correlation between ε_1 and ε_2 comes about because ε_1 causes ε_2, ε_2 causes ε_1, or both respond to a third variable. This is so because in this context the '*cet. par.*' proviso ensures that z is held constant as m is varied: if the intervention in ε_1 is $d\varepsilon_1$, that in ε_2 must be $d\varepsilon_2 = -\tau d\varepsilon_1$. Thus even though m is not exogenous as defined above – so that, accordingly, the effect of an intervention in m on z is not well-defined – still the effect of m on y, *cet. par.*, is unambiguous. In this case we would say that m is exogenous with respect to y, but not with respect to z. Evidently m is exogenous with respect to y if and only if m is uncorrelated with the unobserved determinant(s) of y; this, accordingly, may be taken as a definition of the meaning of exogeneity of one variable with respect to another.[5]

[5] This definition is not completely explicit in that it contains the notion of a subset of variables which 'determines' a particular endogenous variable. This in turn involves the Cowles economists' concept of a causal ordering [see Simon (1953)]. Rather than prolong our detour from atheoretical macroeconometrics by explicating this idea, we will simply hope that most readers will find the definition in the text sufficiently clear.

Or consider another model:

$$z_1 = \varepsilon_1, \tag{28}$$

$$z_2 = \varepsilon_2, \tag{29}$$

$$m = \tau z_1 + \varepsilon_3, \tag{30}$$

$$y = \gamma z_2 + \varepsilon_4, \tag{31}$$

where it is assumed only that neither ε_1 nor ε_2 is correlated with either ε_3 or ε_4. Since ε_1 can be correlated with ε_2, and ε_3 with ε_4, there are no exogenous variables. But by the above definition, z_1 is exogenous with respect to m, and z_2 is exogenous with respect to y. Hence τ captures the effect of an intervention in z_1 on m, and γ that of z_2 on y. Further, these equations could be estimated consistently by least squares.

The Cowles economists were careful to distinguish between correlation and causation. If the data were generated by

$$m = \theta y + \varepsilon_1, \tag{32}$$

$$y = \gamma m + \varepsilon_2, \tag{33}$$

with ε_1 and ε_2 uncorrelated, it is evident that neither m nor y is exogenous with respect to the other. Hence neither can be viewed as causing changes in the other. But none of this prevents us from defining from (32) and (33) the model

$$y = \delta m + \eta, \tag{34}$$

where $\delta = \operatorname{cov}(m, y)/\operatorname{var}(m)$, and the unobservable η, which is uncorrelated with m, is defined as $y - \delta m$. Here δ is a shallow parameter, assuming that γ, θ, σ_1^2 and σ_2^2 are deep parameters. The usefulness of eq. (34) is that it allows direct calculation of the conditional expectation $\mathrm{E}(y|m) = \delta m$ and the conditional variance $\mathrm{v}(y|m) = \sigma_\eta^2$, which are useful in many applications. But the Cowles economists were scrupulously careful to distinguish conditional correlation from causation. The statement 'on average, when m is high y is also high' is very different from 'an increase in m causes an increase in y'. As we have just seen, causation has to do with interventions, which in turn involve an invariance property that is satisfied only under an exogeneity assumption. The random term η, being correlated with the true exogenous variables ε_1 and ε_2, cannot itself be treated as exogenous. The fact that η is (by construction) orthogonal to m does not justify its treatment as an exogenous variable. Under the prior restriction $\theta = 0$, of course, m becomes an exogenous variable, and the conditional correlation of m and y may then be identified with a causal relation.

As just observed, the Cowles distinction between correlation and causation implies that the presence or absence of exogeneity cannot be inferred from the data. The presence of a correlation between money and income is consistent either with money being exogenous with respect to income, income being exogenous with respect to money, or neither variable directly determining the other, both instead being functions of some third variable which is exogenous. The Cowles economists viewed exogeneity as one type of restriction on the parameter space required for identification. In the absence of other prior restrictions derived from theory, the assumption of exogeneity is under-identifying or just-identifying, and hence is not testable. A set of over-identifying restrictions on the parameter space can, of course, be tested, and therefore if the hypothesis of exogeneity is included with a sufficient number of other restrictions (i.e., sufficient to over-identify the model), a joint test of exogeneity and these other restrictions can be conducted [Wu (1973), Hausman (1978)].

The exogenous/endogenous dichotomy among variables just discussed closely parallels that between deep and shallow parameters treated above: the idea of invariance under intervention is central in both cases. Because exogenous variables are conceived to be determined outside the model, hypothetical experiments consist of varying one exogenous variable. *cet. par.*, and determining the effect on the endogenous variables. In order that such experiments be well-defined it is essential to specify precisely what is invariant under the hypothesized intervention; this is the role of the uncorrelatedness assumption. It is, of course, generally inadmissible to inquire as to the effect of a change in one endogenous variable on another, again because the underlying experiment that led to the assumed variation in the endogenous variable is ambiguous. In the case of interventions among parameters, the situation is precisely analogous: only interventions involving deep parameters are generally well-defined, since only then does the assumption that other deep parameters are invariant to the proposed intervention have unambiguous meaning. Similarly, as we have seen, it is inadmissible to hypothesize an intervention involving a shallow parameter, again due to the ambiguity in the nature of the intervention.

The idea of invariance under intervention provides a link between the concepts of deep parameters and exogenous variables. The difference between the two has to do with the nature of the intervention. In the case of parameters, one is comparing different models by doing an exercise in comparative statics or dynamics. For example, one compares the behavior of agents in an economy in which monetary policy is, and has always been, conducted according to a constant growth rate rule with that in an economy in which some other monetary control rule is, and has always been, in place. In the case of variables, on the other hand, one is considering causal relations within a given model by assuming different time paths for variables which are inputs to the model. For example, imagine that at some date the monetary authority decides henceforth to stabilize the monetary growth rate at some predetermined level; what effect will this have on the behavior of agents? The two experiments are very different: the latter experiment, not the former, is the relevant one if the economist wishes to determine the effects of switching to a constant growth rate rule at some given date. This point has been missed in much analysis of policy regimes [again, see Cooley, LeRoy and Raymon (1984a,b)].

Let us now extend these ideas to dynamic models. Consider the model

$$m_t = \theta y_t + \beta_{11} m_{t-1} + \beta_{12} y_{t-1} + \varepsilon_{1t}, \tag{35}$$

$$y_t = \gamma m_t + \beta_{21} m_{t-1} + \beta_{22} y_{t-1} + \varepsilon_{2t}, \tag{36}$$

with ε_{1t} and ε_{2t} contemporaneously and serially uncorrelated. Here the ε_{it} are exogenous, with the m_t and y_t being endogenous. These equations contain the restriction that future-dated variables are excluded from the determination of current-dated variables. Because of this restriction, it is immediately evident that m_{t-j} and y_{t-j}, $j = 1,2,\ldots$, are exogenous with respect to m_t and y_t. Further, if $\theta = 0$, then m_t is also exogenous with respect to y_t. In the context of time-series analysis, it is customary to say that in this case m_t, m_{t-j} and y_{t-j}, $j = 1,2,\ldots$, are *predetermined*. If m_t is predetermined, an intervention consisting of varying m_t is unambiguously associated with a change in ε_{1t}, and in no other exogenous variables. Hence the effect on y_t is $\gamma\, \mathrm{d}m_t$.

Sometimes, however, one wishes to consider an intervention consisting of altering an entire time path of m_t, not just a single realization of m_t. In such cases it is necessary to distinguish between two types of experiment: those involving intercept shifts in the equation determining m_t [with predeterminedness, this is unambiguously eq. (35)], or interventions involving directly specifying a time path for m_t itself. The difference is that the former intervention allows the effect of the intercept shift on the m to cumulate via the $\beta_{12} y_{t-1}$ term, whereas in the latter intervention intercept adjustments are chosen to wash out this indirect effect. If $\beta_{12} = 0$, the two experiments produce the same answer. In this case – $\theta = \beta_{12} = 0$ – m is said to be *strictly exogenous* as well as predetermined.[6] Note that strict exogeneity is not needed to render unambiguous either of these interventions; for this predeterminedness is sufficient. Strict exogeneity only assures that the two types of intervention are in fact equivalent.

[6]A recent paper by Engle, Hendry and Richard (1983) suggests that the notions of predeterminedness and strict exogeneity are not adequately clear, particularly in nonlinear models. They propose in addition the concepts of weak exogeneity, strong exogeneity and super-exogeneity. As we discuss only linear models the concepts of predeterminedness and strict exogeneity correspond to their notions of weak and strong exogeneity respectively. Our analysis, however, stresses that exogeneity is a property of a model, while theirs describes it as a property of the likelihood function that summarizes a model. A discussion more closely related to ours is Leamer (1985).

5. Granger causality and Cowles causality

As things stand, there exists at least the appearance of conflict between the exogeneity tests of atheoretical macroeconometrics and the dictum of the Cowles economists that, in the absence of prior restrictions, empirical testing of exogeneity restrictions is impossible. To resolve the apparent conflict, it is necessary to clarify the relation between Granger non-causality, predeterminedness and strict exogeneity. Consider again the money–income model

$$m_t = \theta y_t + \beta_{11} m_{t-1} + \beta_{12} y_{t-1} + \varepsilon_{1t}, \tag{37}$$

$$y_t = \gamma m_t + \beta_{21} m_{t-1} + \beta_{22} y_{t-1} + \varepsilon_{2t}, \tag{38}$$

where the errors are contemporaneously and serially uncorrelated. As we saw in the preceding section, m is predetermined for y if $\theta = 0$, while m is strictly exogenous for y if $\theta = \beta_{12} = 0$. Finally, if (37)–(38) are solved for the reduced form

$$m_t = \pi_{11} m_{t-1} + \pi_{12} y_{t-1} + u_{mt}, \tag{39}$$

$$y_t = \pi_{21} m_{t-1} + \pi_{22} y_{t-1} + u_{yt}, \tag{40}$$

y fails to Granger-cause m if $\pi_{12} = 0$. Now, π_{12} is given by

$$\pi_{12} = \frac{\theta\beta_{22} + \beta_{12}}{1 - \theta\gamma}.$$

Plainly Granger non-causality is neither necessary nor sufficient for predeterminedness: $\theta = 0$ neither implies nor is implied by $\pi_{12} = 0$. Since predeterminedness is the exogeneity concept relevant for the analysis of interventions, it follows that the Granger and Sims tests are irrelevant to whether a causal interpretation of a conditional correlation is justified.

Further, predeterminedness is also the exogeneity concept relevant for econometric estimation, implying that the Granger and Sims tests are equally irrelevant to the question of whether a model is consistently estimated. Hence Lucas and Sargent (1979, p. 6) are incorrect in taking builders of large econometric models to task for failing to perform exogeneity tests:

> 'Sims showed that the hypothesis that x_t is strictly econometrically exogenous... necessarily implies certain restrictions that can be tested given time series on the y's and x's. Tests along the lines of Sims's ought to be used routinely to check classifications into exogenous and endoge-

nous sets of variables. To date they have not been. Prominent builders of large econometric models have even denied the usefulness of such tests.'

The error here is in assuming that strict exogeneity rather than predeterminedness is relevant.

With regard to strict exogeneity, matters are more complicated. Strict exogeneity does indeed imply Granger non-causality, so that failure of a Granger or Sims test is evidence against strict exogeneity, subject to the usual significance criteria.[7] But the converse is not true: acceptance of Granger non-causality does not imply (although it is consistent with) strict exogeneity: $\pi_{12} = 0$ implies $\theta\beta_{22} + \beta_{12} = 0$, not $\theta = \beta_{12} = 0$.

We observed in the preceding section that predeterminedness is the relevant exogeneity concept for the analysis of interventions. The stronger restriction of strict exogeneity guarantees that the two types of interventions that can be performed in dynamic models – intercept shifts and altered time paths for one of the variables – are in fact equivalent. Given predeterminedness, the strict exogeneity of m is equivalent to the Granger non-causality of m. It therefore can be established either by a Granger or Sims test, or simply by performing the relevant simulations and comparing the results. But for predeterminedness itself, which is needed to justify any of the interventions under discussion, Granger non-causality is irrelevant, as are the Granger and Sims tests.

Many expositions of atheoretical macroeconometrics diminish the importance, of, or ignore altogether, the ambiguity in the interpretation of a finding of Granger non-causality. For example, in his well-known paper testing the monetarist tenet that money is exogenous against the Keynesian view that money may have a substantial endogenous component, Sims (1972) wrote that Granger-causality

> 'is easily testable: if *and only if* causality runs one way from current and past values of some list of exogenous variables to a given endogenous variable, then in a regression of the endogenous variable on past, current and future values of the exogenous variables, the future values of the exogenous variables should have zero coefficients.' (p. 541, emphasis supplied)

[7]Jacobs, Leamer and Ward (1979) stated that even though $\pi_{12} \neq 0$ means that we can reject $\theta = \beta_{12} = 0$, we cannot reject the neighboring hypothesis that θ and β_{12} are almost zero. Thus '... the usefulness of the test is restricted to the unlikely circumstances when only the sharp hypothesis is of interest' (p. 405). Their reasoning appears to be that if the parameters are such that $1 - \theta\gamma$ is close to zero, it is possible for π_{12} to be very large even though θ and β_{12} are near zero. Thus, in loose terms, even if $\theta \cong \beta_{12} \cong 0$, rejection of $\pi_{12} = 0$ is probable if $1 - \theta\gamma \cong 0$. This argument is incorrect: even though $1 - \theta\gamma \cong 0$ creates a presumption that π_{12} will be large in absolute value, it does not create a presumption that the hypothesis that $\pi_{12} = 0$ will be rejected. Jacobs, Leamer and Ward overlooked the fact that if $1 - \theta\gamma \cong 0$, the reduced-form errors have high variances, making rejection of $\pi_{12} = 0$ improbable. The two effects cancel, so that if $\gamma \cong \beta_{12} \cong 0$, the hypothesis that $\pi_{12} = 0$ will probably be accepted regardless of the magnitude of $1 - \theta\gamma$.

Deletion of 'and only if' is required if this statement is to be rendered correct. Further, we are not informed why causality is to be identified with strict exogeneity, rather than predeterminedness as argued above.

Sims's empirical finding was that income does not Granger-cause money, but money does Granger-cause income. For him a 'natural' and 'appropriate' interpretation of the former results is that the data verify the strict exogeneity of money:

> '(t)he most conservative way to state the results for money and income is that they show it to be unreasonable to interpret a least-squares lag distribution for money on GNP as a causal relation, and that they provide no grounds for asserting that distributed lag regressions of GNP on money do not yield estimates of a causal relation. It is natural, and I believe appropriate, to phrase the result more positively: the data verify the null hypothesis that distributed lag regressions of GNP on money have a causal interpretation.' (p. 542)

Characteristically in the literature under review, the statement that exogeneity is testable is accompanied neither by any showing as to why the relevant exogeneity concept is strict exogeneity rather than predeterminedness, nor by the proviso that the interpretation of the test is unambiguous only in the case of rejection of Grange non-causality. For example, Geweke (1978) wrote:

> 'The specification of exogeneity is usually made *a priori*. If the specification is incorrect the otherwise identifying restrictions imposed on structural equations may not be sufficient to identify those equations.... It is therefore desirable to test the exogeneity specification rather than let it remain a mere assertion.' (p. 163)

In places [e.g. Sims (1977), Geweke (1984)] acknowledgment is made of the ambiguity in the interpretation of Granger non-causality outlined above. When this is done, it is under the rubric of 'spurious exogeneity'. It is known that in some cases the structure of the model generates a spurious conclusion of exogeneity according to the Granger–Sims tests. Indeed, the various theories – such as efficient markets tests – that generate Granger non-causality as testable implications re-emerge here as instances of spurious exogeneity. We see now that what appeared earlier as a virtue of Granger–Sims tests – that they can be used to test theories – is a fault in the present context: in precisely these cases Granger noncausality cannot be identified with strict exogeneity. But by and large, spurious exogeneity is not seen as a problem worth losing sleep over, except in cases where there is some specific reason to suspect its existence. In the following passage, Sims (1977) seems to suggest that in any particular instance we need only run through the list of cases which are known to produce spurious exogeneity. If none of these seems directly similar to the case at hand, we can dismiss the possibility of spurious exogeneity and identify

Granger–Sims non-causality with strict exogeneity:

> 'Do the examples...suggest a likely mechanism for a spurious M-to-GNP causal ordering: that is, is there an alternative explanation for an M-to-GNP ordering which would imply the GNP-on-M regression does not have the postulated behavioral interpretation? The answer, in my opinion, is no. Money is not the price of a durable good. It is hard to see why there should be much greater measurement error in GNP than in M. For an optimal control model to produce an M-to-GNP ordering would require that GNP be subject to control while M was the target variable, which seems implausible. There appears to be no good reason to believe money supply or demand should be an exact function of current and past one-step-ahead prediction errors in some other variable. While the list of mechanisms to generate causal orderings given in section 3 is certainly not exhaustive, none of the possibilities listed there applies naturally to the money and GNP case.
>
> Thus, it appears that economists who do not believe that GNP on money distributed lag regressions are structural ought to be basing their argument either on rational expectations or on the straightforward, old-fashioned possibility of Type II error.' (pp. 42–43)

It is one thing to say that acceptance of Granger non-causality is consistent with strict exogeneity. But Sims here goes further and argues that, failing some specific reason to suspect the existence of spurious exogeneity, Granger non-causality verifies strict exogeneity. Based on classical statistical theory, such a conclusion is indefensible: since that element of the parameter space associated with the null hypothesis is observationally equivalent to elements of the parameter space in the rejection region, no observational evidence can verify the null hypothesis.[8] Sims' position only makes sense if it is interpreted as an implicitly Bayesian argument: if one had a prior density that assigned negligible weight to regions in the parameter space observationally equivalent to the null hypothesis compared to the weight assigned to the null hypothesis itself, then a finding of Granger non-causality would lead to a posterior distribution that strongly favors strict exogeneity. We do not know whether the many econometricians who have viewed Granger non-causality as evidence in favor of strict exogeneity would acknowledge having such priors. We surely would not – the many examples (cited above) of models in which theoretical restric-

[8]See Breusch (1980) for the interpretation of tests performed on underidentified models. Breusch emphasized that if the maintained model is underidentified only compound null hypotheses have any observable implications at all. Further, such compound hypotheses can be rejected but not confirmed since every element of the parameter space in the acceptance region is observationally equivalent to elements in the rejection region. In the present case, the Granger–Sims tests can falsify the hypothesis of strict exogeneity, but cannot verify it. It is therefore unclear why Sargent (1979) regards a proof of the necessity of Granger non-causality for strict exogeneity as controverting Nelson's (1979) statement that 'exogeneity cannot be verified from non-experimental data...'.

tions generate Granger non-causality but not strict exogeneity seem to us to argue strongly against such a story.

6. Impulse response analysis, innovation accounting and policy analysis

Our criticism of impulse response analysis and innovation accounting is an extension of – and is implied by – the criticism outlined in the preceding section of causality as treated in atheoretical macroeconometrics. Both of these exercises consist of identifying conditional correlations with causal orderings. Such identification is justified only under a predeterminedness assumption which is untestable in the absence of prior restrictions derived from theory. In particular, it is not tested by the Granger or Sims tests. Failing predeterminedness, it makes no more sense in a dynamic setting to interpret impulse response functions as capturing the effect of an intervention in m on y than it did in a static setting to regard the regression of y on m as capturing the effect of an intervention in the latter on the former [see the discussion following eq. (34) above].

In the literature of atheoretical macroeconometrics, on the other hand, impulse response functions and the like are routinely interpreted as reflecting causal orderings:

> 'Price innovations are of negligible importance in the U.S. system. In the German system price innovations are a major source of disturbance...(they) produce a large sustained drop in real GNP and a persistent decline in the real wage...' [Sims (1980a, p. 22)]

> 'It is likely for the stock of money to grow less and unemployment to be higher following an unexpected increase in interest rates. Investment, change in business inventories and real GNP growth show small drops following such a shock while inflation shows very little response.' [Litterman (1981, p. 32)]

> 'Innovations in both the full employment surplus and the growth rate of money increase the variability of relative prices as does an increase in the inflation rate.' [Fischer (1981, p. 408)]

In the absence of prior predeterminedness restrictions, such interpretations are completely unjustified; the reasoning is exactly the same here as in the static models discussed in section 4.[9]

There is really no more that need be said. But it may be helpful to make the same point in a different way. Thus let us assume that the data are generated

[9] The point being made here – that VAR models sustain the interpretations made of them only under a prior predeterminedness restriction – is well expressed in Sachs' (1982) comments on Sims (1982).

by (37)–(38), which is parameterized so that predeterminedness and strict exogeneity appear as special cases ($\theta = 0$ and $\theta = \beta_{12} = 0$). Then we will mimic the estimation and renormalization of a VAR model so as to see what it means to postulate an intervention in the money stock innovations. The reduced form corresponding to (37)–(38) is

$$m_t = \pi_{11} m_{t-1} + \pi_{12} y_{t-1} + u_{mt}, \tag{39}$$

$$y_t = \pi_{21} m_{t-1} + \pi_{22} y_{t-1} + u_{yt}. \tag{40}$$

Consider that this model has been estimated by least squares. It is easily verified that the reduced form errors have variances $\sigma_{mm} = (\sigma_1^2 + \theta^2\sigma_2^2)/(1-\theta\gamma)^2$, $\sigma_{yy} = (\gamma^2\sigma_1^2 + \sigma_2^2)/(1-\theta\gamma)^2$ and covariance $\sigma_{my} = (\gamma\sigma_1^2 + \theta\sigma_2^2)/(1-\theta\gamma)^2$. Now, we can replace (40) by itself less (39) multiplied by $\delta = \sigma_{my}/\sigma_{mm}$ to obtain the VAR model

$$m_t = \rho_{11} m_{t-1} + \rho_{12} y_{t-1} + \eta_{mt}, \tag{41}$$

$$y_t = \delta m_t + \rho_{21} m_{t-1} + \rho_{22} y_{t-1} + \eta_{yt}. \tag{42}$$

Here η_{mt} and η_{yt}, the orthogonalized innovations in m and y, are given by $\eta_{mt} \equiv u_{mt}$ and $\eta_{yt} \equiv u_{yt} - \gamma u_{mt}$. Expressed in terms of the original errors and parameters, the orthogonalized innovations are just

$$\eta_{mt} = (\varepsilon_{1t} + \theta\varepsilon_{2t})/(1-\theta\gamma), \tag{43}$$

$$\eta_{yt} = \left\{ (\gamma\varepsilon_{1t} + \varepsilon_{2t}) - \left(\frac{\gamma\sigma_1^2 + \theta\sigma_2^2}{\sigma_1^2 + \theta^2\sigma_2^2} \right) (\varepsilon_{1t} + \theta\varepsilon_{2t}) \right\} \Big/ (1-\theta\gamma). \tag{44}$$

Now, what does it mean to hypothesize an intervention in η_{mt}? Plainly the effect on η_{yt}, and therefore on all future values of m and y, depends on whether ε_{1t} or ε_{2t} caused the change in η_{mt}. Here we have no more than a repetition of the simple static analysis of section 4. Hence the question 'What is the effect of an innovation in m on y?', to which impulse response functions and innovation accounting purportedly give the answer, is in fact ambiguous.

But suppose that on prior grounds we know that $\theta = 0$. Then $\delta = \gamma$, $\eta_{mt} = \varepsilon_{1t}$, $\eta_{yt} = \varepsilon_{2t}$, and the estimated VAR model is just the original system (37)–(38), with the predeterminedness restriction $\theta = 0$ imposed. Thus we are back to the point made above that a prior predeterminedness assumption is required if the orthogonalized innovations are to be treated as exogenous variables.

A potential misunderstanding must be pointed out. Sometimes one encounters the impression that analysts believe the interpretation of VAR models to be ambiguous only when alterations in the ordering of the variables produce major changes in its properties [Friedman (1981), Fischer (1981), Goldfeld (1982)]. To see that the ambiguity is not restricted to this case (and in fact has nothing to do with the degree to which the model's properties are sensitive to the ordering of the variables), note first that the ordering of variables matters only if, and to the extent that, the reduced-form errors are contemporaneously correlated. The ordering is immaterial if the reduced-form errors are uncorrelated, so that no renormalization is needed to diagonalize the error covariance matrix. But it is easy to see that the fact that the results are insensitive to the ordering of variables does not justify causal interpretations. The relevant assumption that must be made to justify such interpretations, that these errors are true exogenous variables, implies but is certainly not implied by their uncorrelatedness.

Turning to policy analysis, we must distinguish between questions like 'What would be the behavior of y if m had always been and will always be generated by a constant growth-rate rule rather that a variable growth-rate rule?' and 'What would be the effect of switching to a constant growth-rate rule in the future, given that the past was whatever it was?' The former question is answered by determining the effect of an intervention in parameters, while the latter involves an intervention in variables. VAR models are entirely unsuited to the first type of question (and their proponents do not assert the contrary). This is so because interventions in parameters have unambiguous meaning only if the parameters are deep, so that the *cet. par.* assumption has clear meaning. VAR models, of course, do not satisfy this property. In the case of interventions involving variables, the fact that the parameters of VAR models are shallow causes no problem since no intervention in parameters is contemplated. However, for an intervention in policy variables to have clear meaning, an exogeneity assumption is needed. Such an assumption is, as we have noted, untestable in the absence of prior restrictions.

We conclude that VAR models are not useful for analyzing interventions either in parameters or in variables.

7. Structural and non-structural models

In discussing the Cowles Commission program, we pointed out the essential role of the distinction between deep and shallow parameters. Here we distinguish between structural and non-structural models and model-building. As before, we believe that our treatment is in spirit with that of the Cowles group, as with that of such discussions as Hurwicz (1962) and Pratt and Schlaifer (1984), but we would not insist on the point. We will see that with the dichotomy between structural and non-structural models in hand, a much more concise formulation of our critique of atheoretical macroeconometrics becomes possible.

The distinction between structural and non-structural modeling has to do with the difference between the two sets of game rules that apply in each case. Therefore to define structural and non-structural modeling, and thereby distinguish the two, it suffices to identify the differences between the two sets of game rules. This we do in the following paragraphs. As with any question pertaining to modeling strategy, the choice between structural and non-structural modeling depends on the nature of the questions to be answered: if these can be answered with non-structural methods, those are to be used since fewer assumptions are required; if not, it is necessary to engage in structural modeling.

Non-structural models. The exclusive purpose of non-structural modeling is to capture the probabilistic characteristics of the data under examination and to answer questions that can be resolved with that information. In linear–normal models, for example, these characteristics can be subsumed in the means and covariances of the data. Sometimes it is easiest to summarize these means and covariances simply by specifying that, for example, the data are distributed as multivariate normal with given mean vector and covariance matrix. At other times, however, it is convenient to construct new theoretical entities–errors – with given probability distribution, and to define the observable variables as functions of these errors, where the functions incorporate unobserved parameters. Since in non-structural modeling these errors are arbitrary theoretical constructs, their probabilistic characteristics are of no direct interest; they matter only insofar as they implicitly describe the probabilistic characteristics of the observed variables. Because non-structural modeling is concerned exclusively with characterizing the data, it makes sense to define two non-structural models as observationally equivalent if and only if they generate the same probability distribution for the observed variables. A *renormalization* can be defined as substitution of one model for another, where the two are observationally equivalent. In non-structural modeling the choice among equivalent models – or, to say the same thing, the choice of normalizations – is purely a matter of convenience.

We have seen that in non-structural modeling the analyst permits himself to substitute one observationally equivalent model for another at his convenience. Because he does not attempt to distinguish among observationally equivalent models, it follows that *the only questions he can ask of the model are those which have the same answers for all observationally equivalent versions of the model.*

Structural models. Let us now consider structural models. These are not distinguished from non-structural models by virtue of any formal property possessed by one and not the other, but by the fact that different rules of the game apply, as indicated below. Structural models are used, not to characterize the data, but to answer questions about the effects of interventions, either in terms of parameters or exogenous variables.

Accordingly, it is necessary that structural models correctly specify the deep parameters or the exogenous variables, or both, depending on the intervention to be analyzed. But this means that observationally equivalent models cannot be substituted one for another arbitrarily, since if one model correctly specifies deep parameters and exogenous variables, then even in some other model which is observationally equivalent the parameters will be shallow and the unobserved variables will not be exogenous. For example, suppose that m is exogenous, and is related to y according to

$$m = \varepsilon_m, \tag{45}$$

$$y = \gamma m + \varepsilon_y, \tag{46}$$

where ε_m and ε_y are normally and independently distributed. Obviously the model

$$y = \eta_y, \tag{47}$$

$$m = \tau y + \eta_m, \tag{48}$$

(where η_y and η_m are exogenous) tells a very different story about the effect of interventions even if γ and the variances of η_y and η_m are such that the bivariate density of m and y is the same in both models. Further, if γ is a deep parameter and ε_m and ε_y are exogenous variables, then τ is shallow (being a function of γ and the variances of ε_m and ε_y) and η_y and η_m are not exogenous (each being correlated with the true exogenous variables ε_m and ε_y).

At a more formal level, the preceding discussion implies that structural models differ from non-structural models in that the former do not have the equivalence relation possessed by the latter (except for such trivial renormalizations as units changes): models which are equivalent if interpreted non-structurally are generally different if interpreted as structural models. Thus, treating a particular model as structural constitutes a much stronger restriction than if the same model is accorded only a non-structural interpretation: in the former case one is, in effect, rejecting as false an entire class of models which are admitted as equivalent representations in the latter case. However, under the stronger restriction of structuralness a broader set of interpretations is generated, as would be expected, because one can ask questions that have different answers in different observationally equivalent models.

Our complaint about atheoretical macroeconometrics can now be summarized. It is that the distinction between structural and non-structural models is not observed. If VAR models are treated as structural, explicit justification is necessary for whatever triangularity or orthogonality assumption is made – these are not arbitrary normalizations, but substantive restrictions on the parameter space that must be justified from theory. If, on the other hand, VAR models are interpreted as nonstructural, then triangularization and orthogonalization are in fact arbitrary normalizations not requiring theoretical justification. But then the model cannot be interpreted as doing more than summarizing the correlations in the data – no statement dealing with causation or the effect of interventions is admissible.

In large measure the appeal of VAR models – particularly to those more interested in econometrics than economics – is that they appear to offer a way to generate the same kind of output as structural models, but without the input of explicit economic theory. This combination is indeed attractive, but if our analysis is correct it must be rejected as illusory.

8. Conclusions

Assuming that VAR models are interpreted as non-structural, the only conclusions that can properly be generated from them are those which are invariant across observationally equivalent versions of the same model. This excludes any kind of causal statement since, as we have stressed, different but observationally equivalent versions of a given model have different causal interpretations. Now, there certainly exist important applications of VAR models that in fact have this invariance property. Forecasts, being constructed directly from the reduced form, will be invariant across observationally equivalent versions of a given model. The same is true of Granger causality. Therefore theories which have implications for the outcome of causality tests – we summarized several such theories in section 3 – can be tested using the methods of atheoretical macroeconometrics. Finally, VAR models can be used to determine the existence of Granger-causal orderings even in the absence of any theoretical reason to expect them, the idea being that theorists will regard the outcomes of such exercises as stylized facts requiring subsequent explanation in terms of structural models. Whether or not such hypothesis-seeking is likely to be fruitful is an open question, but we do not object to it in principle.

But these valid applications of VAR models, while important, are hardly original. It comes as no surprise that econometric forecasts can be constructed from reduced forms. Similarly, the work of econometricians testing capital market efficiency is not advanced by observing that they are in effect determining whether any publicly available data Granger-causes financial rates of return, these being merely new words for old ideas. Atheoretical macroeconometrics has received widespread attention, not from such uncontroversial applications, but rather from its use in analyzing causal orderings and policy interventions. As we have seen, the nature of the criticism of such exercises depends on whether VAR models are interpreted as structural or non-structural, which, given the looseness of exposition, is largely a matter of the personal preference of the reader. If the models are interpreted as non-structural, we view the conclusions as unsupportable, being structural in nature. If the models are interpreted as structural, on the other hand, the restrictions on error distributions adopted in atheoretical macroeconometrics are not arbitrary renormalizations, but prior identifying restrictions. As such, they require justification from theory. Failing such justification (and it is seldom offered), the conclusions are equally without support.

References

Ashenfelter, O. and David Card, 1982, Time series representations of economic variables and alternative models of the labour market, Review of Economic Studies, 761–782.

Breusch, Trevor, 1980, Testing hypotheses in unidentified models, Reproduced (University of Southampton).

Chamberlain, Gary, 1982, The general equivalence of Granger and Sims causality, Econometrica 50, 569–582.

Cooley, Thomas F., Stephen F. LeRoy and Neil Raymon, 1984a, Modeling policy interventions, Reproduced (University of California, Santa Barbara, CA).

Cooley, Thomas F., Stephen F. LeRoy and Neil Raymon, 1984b, Econometric policy evaluation: Note, American Economic Review 74, 467–470.

Doan, Thomas, Robert Litterman and Christopher Sims, 1984, Forecasting and conditional projection using realistic prior distributions, Econometric Reviews, 1–100.

Engle, Robert E., David F. Hendry and Jean-Francois Richard, 1983, Exogeneity, Econometrica 51, 277–304.

Feige, Edgar L. and Douglas K. Pearce, 1979, The casual causal relationship between money and income, Review of Economics and Statistics 61, 521–533.

Fischer, Stanley, 1981, Relative shocks, relative price variability, and inflation, Brookings Papers on Economic Activity, 381–431.

Florens, J.P. and M. Mouchart, 1982, A note on noncausality, Econometrica 50, 583–592.

Friedman, Benjamin, 1981, The roles of money and credit in macroeconomic analysis, Working paper no. 831 (NBER, Cambridge, MA).

Geweke, J., 1978, Testing the exogeneity specification in the complete dynamic simultaneous equations model, Journal of Econometrics 7, 163–185.

Geweke, J., 1984, Inference and causality in economic time series models, in: Z. Griliches and M. Intriligator, eds., Handbook of econometrics, Vol. II (North-Holland, Amsterdam).

Goldfeld, Stephen M., 1982, Comments, Brookings Papers on Economic Activity, 153–157.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37, 24–36.

Hall, Robert E., 1978, Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence, Journal of Political Economy 86, 971–1007.

Hausman, J.A., 1978, Specification tests in econometrics, Econometrica 49, 1251–1271.

Hosoya, Y., 1977, On the Granger condition for non-causality, Econometrica 45, 1735–1736.

Hurwicz, Leonid, 1962, On the structural form of interdependent systems, in: Ernest Nagel et al., eds., Logic, methodology and philosophy of science (Stanford University Press, Stanford, CA) 232–239.

Jacobs, Rodney L., Edward E. Leamer and Michael P. Ward, 1979, Difficulties with testing for causation, Economic Inquiry, 401–413.

Kohn, R., 1981, A characterization of Granger–Sims exogeneity, Economics Letters 8, 129–133.

Leamer, Edward E., 1985, Vector autoregressions for causal inference?, in: K. Brunner and A. Meltzer, eds., Understanding monetary regimes, Carnegie-Rochester conference series on public policy, Vol. 22 (North-Holland, Amsterdam) 255–303.

LeRoy, Stephen F., 1982, Expectation models of asset prices: A survey of theory, Journal of Finance 37, 185–217.

Litterman, Robert, 1981, Multivariate autoregressions with prior information, Reproduced (MIT, Cambridge, MA).

Litterman, Robert and Laurence Weiss, 1985, Money, real interest rates and output: A reinterpretation of postwar U.S. data, Econometrica 53, 129–156.

Lucas, Robert E., Jr., 1976, Econometric policy evaluation: A critique, in: K. Brunner and A. Meltzer, eds., The Phillips curve and labor markets, Carnegie-Rochester series on public policy, Vol. 1 (North-Holland, Amsterdam) 19–46.

Lucas, Robert E., Jr., and Thomas J. Sargent, 1979, After Keynesian macroeconomics, Federal Reserve Bank of Minneapolis Quarterly Review, 1–16.

Nelson, Charles R., 1979, Granger causality and the natural rate hypothesis, Journal of Political Economy 87, 390–394.

Pierce, David A. and Lawrence D. Haugh, 1977, Causality in temporal systems: Characterization and a survey, Journal of Econometrics 5, 265–293.

Pratt, John W. and Robert Schlaifer, 1984, On the nature and discovery of structure, Journal of the American Statistical Association 79, 9–21.

Sachs, Jeffrey D., 1982, Comments, Brookings Papers on Economic Activity, 157–162.

Sargent, Thomas J., 1976, A classical macroeconomic model for the United States, Journal of Political Economy 84, 207–238.

Sargent, Thomas J., 1979, Causality, exogeneity, and natural rate models: Reply to C.R. Nelson and B.T. McCallum, Journal of Political Economy 87, 403–409.

Simon, Herbert A., 1953, Causal ordering and identifiability, Ch. III in: William C. Hood and Tjalling C. Koopmans, eds., Studies in econometric method (Cowles Foundation).

Sims, Christopher A., 1972, Money income and causality, American Economic Review 62, 540–552.

Sims, Christopher A., 1977, Exogeneity and causal ordering in macroeconomic models, in: C.A. Sims, ed., New methods of business cycle research (Federal Reserve Bank of Minneapolis, MN).

Sims, Christopher A., 1980a, Macroeconomics and reality, Econometrica 48, 1–47.

Sims, Christopher A., 1980b, Comparison of interwar and postwar business cycles: Monetarism reconsidered, American Economic Review 70, 250–259.

Sims, Christopher A., 1982, Policy analysis with econometric models, Brookings Papers on Economic Activity, 107–164.

Wu, De-Min, 1973, Alternative tests of independence between stochastic regressions and disturbances, Econometrica 41, 733–750.

Zellner, Arnold, 1979, Causality and econometrics, in: K. Brunner and A. Meltzer, eds., Three aspects of policy and policymaking: Knowledge, data and institution, Carnegie-Rochester conference series on public policy, Vol. 10 (North-Holland, Amsterdam) 9–54.