Journal of Monetary Economics 27 (1991) 381–423. North-Holland

The causal direction between money and prices

An alternative approach*

Kevin D. Hoover
University of California, Davis, Davis, CA 94916, USA

Received December 1988, final version received March 1991

Causality is viewed as a matter of control. Controllability is captured in Simon's analysis of causality as an asymmetrical relation of recursion between variables in the unobservable data-generating process. Tests of the stability of marginal and conditional distributions for these variables can provide evidence of causal ordering. The causal direction between prices and money in the United States 1950–1985 is assessed. The balance of evidence supports the view that money does not cause prices, and that prices do cause money.

1. Introduction

The debate in economics over the causal direction between money and prices is an old one. Although in modern times it had seemed largely settled in favor of the view that money causes prices, the recent interest in real

*This paper is a revision of part of my earlier working paper, 'The Logic of Causal Inference: With an Application to Money and Prices', No. 55 of the series of working papers in the Research Program in Applied Macroeconomics and Macro Policy, Institute of Governmental Affairs, University of California, Davis. I am grateful to Peter Oppenheimer, Peter Sinclair, Charles Goodhart, Thomas Mayer, Edward Learner, Thomas Cooley, Stephen LeRoy, Leon Wegge, Nancy Wulwick, Diran Bodenhorn, Clive Granger, Paul Holland, the editors of this journal, an anonymous referee, and the participants in macroeconomics seminars at the University of California, Berkeley, Fall 1986, the University of California, Davis, Fall 1988, and the University of California, Irvine, Spring 1989, for comments on this paper in its various previous incarnations. I am especially grateful to Steven Sheffrin for many fruitful discussions, much encouragement, and sound advice, and to David Hendry, Neil Ericsson, and Arthur Havenner for useful discussions of econometric issues. Nicholas Ramsing provided excellent research assistance.

business cycle models has in fact reopened the question. Much of the research into this question has applied tests of Granger-causality [Granger (1969)]. If we take the view that the question of interest is: supposing the Federal Reserve can control the stock of money, can it thereby also control the level of prices (or the rate of inflation)?, then another technique is needed. Tests of Granger-causality neither ask nor answer this question [Granger (1980, p. 1)].[1] The approach which will be applied in this paper is based on the analysis of causality due to Simon (1953); see also Simon and Iwasaki (1988). Simon's paper is an attempt to characterize the causal relation. The methodological and philosophical foundations of this approach are developed in Hoover (1990). A central conclusion of that paper is that causal relations cannot be judged from any statistical test alone. Rather once the causal relation has been logically characterized, it becomes clear that prior knowledge about institutions and the economic environment must be combined with (available) statistical techniques to gather evidence about the actual causal direction in the economy. The problem of the causal direction between money and prices is taken to be a testbed for this approach.

2. Does money cause prices?: The current debate

Perhaps the most famous explicit debate about causal order in empirical economics is over the monetarist claim that money causes prices and income [e.g., Fisher (1911/1931), Friedman and Schwartz (1963a)]. Despite its central place in monetarist thinking, surprisingly little evidence has been brought to bear directly on this claim [cf. Friedman and Schwartz (1963a, b), Cagan (1965, ch. 6)]. This is an old debate. Yet, far from being moribund, it has recently been revived in a surprising quarter. New classical economists are often seen as the successors to monetarism; nevertheless, some new classical economists now doubt the causal efficacy of money over income and, in some respects, over prices. Black (1970, 1972) and Fama (1980) have argued that a large portion of the money stock, bank deposits ('inside money'), responds passively to independent changes in economic activity, and that prices are determined almost completely by the amount of currency (and central bank reserves) in circulation ('outside money').[2] Black and Fama's view favors real business cycle models [e.g., Kydland and Prescott (1982), Long and Plosser (1983)], in which all real activity is determined without reference to monetary variables.[3] If real business cycle models are correct the correlation between money and nominal income arises because real activity causes money and not

[1]See Hoover (1988a, ch. 8) for a general discussion of Granger-causality and Zellner (1979) for a discussion of its relationship to other notions of causality.

[2]Hoover (1988a, ch. 5) and (1988c) present and criticize Fama's argument in detail.

[3]For a discussion of these models, see Hoover (1988a, ch. 3).

because money causes real activity. King and Plosser (1984) present empirical evidence to bolster this claim. The new classical resuscitation of the notion of an endogenous money stock – a notion that had been peculiarly Keynesian [e.g., Gurley and Shaw (1960), Kaldor (1982)] – reopens an issue that many monetarists thought was dormant, if not finally laid to rest.

One conclusion of the famous debate between Tobin (1970) and Friedman (1970) is that both seem to have an implicit notion that causality is not about temporal ordering but about controllability. In contrast to many other studies using different concepts of causality, we shall attempt to develop an approach which is germane to this specific question.

3. A causal concept and its empirical relevance

3.1. A probabilistic approach

The central problem of causal inference is that, although we observe data, the enduring relationships between these data cannot be observed directly but must be inferred indirectly. Were we willing to assume the truth of sufficiently strong identifying assumptions, the problem of causal inference would disappear. Such assumptions, however, beg the question, and are too often, to use Christopher Sims's (1980, p. 1) term, 'incredible'. We should like to have empirical support for causal orderings and not simply assume that we know them *a priori*.[4]

Let us suppose that the observable data are generated by an unobservable process (data-generating process) which can be described by their joint probability distribution.[5] For two variables (think of them as money and prices), the definition of a conditional distribution implies that this joint distribution may be partitioned two different ways into the product of conditional and marginal distributions:

$$D(M,P) = D(M|P)D(P) = D(P|M)D(M). \tag{1}$$

If the joint probability distribution is multivariate normal, then the four distributions that constitute the two alternative partitions can be interpreted as regression equations. What is more, in a sense to be made clear presently,

[4] A fuller discussion of the general methodological issues arising from the relationship between causality and identifiability is found in Hoover (1989).

[5] The use of 'data-generating process' seems close to its use in the econometrics literature. E.g., Spanos (1986, pp. 661–662), writes: 'The concept of the *actual data generating process* (DGP) is used to designate the phenomenon of interest which a theory purports to explain. The concept is used in order to emphasize the intended scope of the theory as well as the source of the observable data.... For example, ...if the intended scope of the theory is to explain observed quantity or/and price changes over time, then the actual DGP should refer to the actual market process giving rise to the observed data.'

each partition represents a different causal order: the first can be read as P causes M, and the second as M causes P.

If we are dealing with a set of data drawn from a single policy regime, the joint probability distribution will remain stable. Furthermore, both partitions will remain stable as well. There will be no basis for discriminating between the two causal orderings implicit in the partitions. The two partitions are *observationally equivalent*.[6] Sargent (1976) makes this same point when he shows that natural and nonnatural rate models generate empirically equivalent forms for estimation. A single set of data cannot discriminate between them.

What we will show presently is that if there are interventions in the data-generating process (i.e., changes of regime), and *crucially* if we can attribute the interventions either to the money-determination process or to the price-determination process, then the patterns of stability and instability of the conditional and marginal distributions in the alternative partitions give important information about the underlying causal order in the data-generating process.[7]

In particular, if *money causes prices* and if the intervention was in the money-determination process, say a change in Federal Reserve policy, one would not expect either $D(M|P)$ or $D(M)$ to remain stable; but one would expect $D(P|M)$ to be stable, although not $D(P)$. If *money causes prices* and the intervention was in the price-determination process, say a period of price controls, one would not expect $D(P|M)$ or $D(P)$ to remain stable; but, one would expect $D(M)$ to be stable, although not $D(M|P)$. That is, if *money causes prices*, the partition $D(P|M)D(M)$ is preferred to the partition $D(M|P)D(P)$. Of course the problem is symmetrical. If prices cause money, we simply interchange M and P in our expectations of stability or instability. 'Stability', as we use it here, is closely related to Engle, Hendry, and Richard's (1983, esp. pp. 281–285) notion of superexogeneity. *Superexogeneity* is defined to be weak exogeneity together with parameter invariance with respect to some class of admissible interventions. *Weak exogeneity* obtains when the joint probability distribution of a group of variables can be factored into a conditional distribution and a marginal distribution in such a way that efficient estimates of the parameters of interest may be recovered from the conditional distribution without knowledge of the marginal distribution. The conditioning variables are then said to be weakly exogenous with respect to the parameters of interest. Both weak exogeneity and superexogeneity are relative to the choice of parameterization. Causal order, as we shall define it,

[6]On the problem of observational equivalence in causal systems, see Basmann (1965, 1988) and Granger (1969, p. 374).

[7]Neftci and Sargent (1978) pioneered the use of regime changes as a source of information to help discriminate between otherwise observationally equivalent theories. Also see Miller (1983) and Blanchard (1984). None of these authors, however, casts the problem as one of causal inference.

is a property not of a particular parameterization, but of the underlying data-generating process [cf. Basmann (1988, p. 74)]. Going beyond Engle et al. (1983, p. 284), who seek superexogeneity '[t]o sustain conditional inference in processes subject to interventions...', we find both the presence and the absence of superexogeneity of alternative conditional distributions in the face of particular classes of interventions to be crucial parts of the evidence needed to infer causal direction. As will become clear presently, evidence on the invariance or noninvariance of the alternative marginal distributions is also crucial.

3.2. *The structure of empirical evidence*

To make the causal interpretation of these stability conditions persuasive, consider a simple example of a data-generating process in which money causes prices:[8]

$$P = aM + \varepsilon, \quad \varepsilon \sim \mathrm{N}(0, \sigma_\varepsilon^2), \tag{2}$$

$$M = b + \nu, \quad \nu \sim \mathrm{N}(0, \sigma_\nu^2), \tag{3}$$

where $\operatorname{cov}(\varepsilon, \nu) = 0$, $\mathrm{E}(\varepsilon_t \varepsilon_{t-j}) = 0$, and $\mathrm{E}(\nu_t \nu_{t-j}) = 0$, $j = 1, 2, \ldots, \infty$.

M causes P according to Simon's (1953) analysis because M is (block-) recursively ordered ahead of P: M is determined independently of P but is itself an essential determinant of P.[9]

The reduced forms of eqs. (2) and (3) are

$$P = ab + a\nu + \varepsilon, \tag{4}$$

$$M = b + \nu. \tag{5}$$

Now consider the two possible partitions of the joint probability distribution $D(M, P)$:

$$D(P|M) = \mathrm{N}(aM, \sigma_\varepsilon^2), \tag{6}$$

$$D(M) = \mathrm{N}(b, \sigma_\nu^2), \tag{7}$$

$$D(M|P) = \mathrm{N}\left(\left[a\sigma_\nu^2 P + b\sigma_\varepsilon^2\right]/\left[a^2\sigma_\nu^2 + \sigma_\varepsilon^2\right], \left[\sigma_\varepsilon^2\sigma_\nu^2\right]/\left[a^2\sigma_\nu^2 + \sigma_\varepsilon^2\right]\right), \tag{8}$$

$$D(P) = \mathrm{N}(ab, a^2\sigma_\nu^2 + \sigma_\varepsilon^2). \tag{9}$$

[8] Hoover (1988b) contains two more complex examples, one involving rational expectations.

[9] Hoover (1990) provides a full discussion of Simon's analysis of causality and generalizes it [following Mesarovic (1969) and Katzner (1983)] to stochastic, nonlinear systems, possibly involving cross-equation restrictions of the sort that are sometimes generated by rational expectations. It is also shown that Simon's analysis is fully consistent with Mackie's (1980) philosophical analysis of causality based on conditional propositions.

The parameters of the price-determination process are a and σ_ε^2, and the parameters of the money-determination process are b and σ_v^2. Now suppose that we have some way of assigning interventions not to particular parameters, for we assume that the actual data-generating process cannot be observed, but to one or other of these two processes. For example, suppose that the Federal Reserve changes the conduct of monetary policy, then either b or σ_v^2 change. In either case, $D(M|P)$ and $D(M)$ will change, as is to be expected; but notice that $D(P)$ will also change and, crucially, that $D(P|M)$ will be invariant. Suppose on the other side that a price control regime is introduced which alters either a or σ_ε^2. In either case, $D(P|M)$ and $D(P)$ will change; but notice that $D(M|P)$ will also change and, crucially, that $D(M)$ will be invariant. Because money causes prices in the true underlying data-generating process, the partition $D(P|M)D(M)$ is clearly more stable to well-defined interventions than the partition $D(M|P)D(P)$. If prices had caused money in the data-generating process, these results would of course been reversed.

Applications may prove less straightforward than this simple example suggests. First, rational expectations can complicate things. As shown in Hoover (1988b), if money causes prices in a model in which agents base expected prices on a rational anticipation of the actions of the monetary authority, what Hendry (1988a) calls a 'feedforward mechanism', then $D(P|M)$ is no longer invariant with respect to interventions in the money-determination process. Causal direction may nevertheless still be unambiguously defined as running from money to prices [Hoover (1990, app.)]. The evidence for this causal ordering is restricted, however, to the invariance of $D(M)$ with respect to interventions in the price-determination process.[10]

Second, the errors in the data-generating process [(2) and (3)] are assumed to have zero covariance. If both are driven by an omitted common third factor, then this covariance may not be zero. If the third factor is observable, the obvious solution is to condition on it, in which case analogous stability results with respect to interventions in M and P go through. This strategy is employed in the empirical application in section 4 below. If the third factor is unobservable, the causal order in the data-generating process may nevertheless by precisely characterized as described in Hoover (1990, app.). Just as when rational expectations are introduced, there will be cross-equation restrictions. The conditional distribution, $D(P|M)$, will no longer be invari-

[10]Hendry (1988a) uses this difference between the case of rational expectations and the case presented in the text as a basis for discriminating between 'feedforward' and 'feedback' or rule-of-thumb mechanisms. He implicitly assumes that prices, income, and interest rates cause money. Then, the failure of the conditional model for money to be superexogenous would be compatible with cross-equation restrictions implied by an expectations mechanism as in the Lucas critique; whereas superexogeneity in the face of the noninvariance of marginal models for prices, income, and interest rates would not. He presupposes, but never tests directly for, causal direction.

ant, but the marginal distribution corresponding to the true causal order [in this case, $D(M)$] will remain invariant to interventions in the price-determination process.

With due account for these complications, the analysis suggests a general strategy for identifying causal orderings. Each of the conditional and marginal distributions in eqs. (6)–(9) can be interpreted as regression equations. It should be possible to use institutional and historical knowledge to identify periods in which there are probably no important interventions in either the money-determination or the price-determination processes. During such periods the regression equations should all show stable estimated coefficients. If we could then identify periods in which there are interventions clearly associated with the money-determination process and ones clearly associated with the price-determination process, we could check the patterns of relative stability of the alternative partitions and thereby determine with which causal ordering (if either) the data are consistent.

Our method resembles Neftci and Sargent's (1978) attempt to discriminate empirically between the natural-rate and the nonnatural-rate hypotheses. Sargent (1976) had shown that the two hypotheses could be observationally equivalent within a single regime.[11] Neftci and Sargent recognized that data from different regimes could nonetheless help to discriminate between them, even in the absence of precise identification. Our method differs from Neftci and Sargent in that it is more self-consciously interested in causal order and it aims at more precisely identifying the source, nature and timing of the regime changes.

4. An application to money and prices

The strategy outlined in the last section is conceptually simple, although not easy to implement. To demonstrate how it might be put into practice, we shall now consider the problem of causally ordering money and prices in the United States over the period 1950–1985. We proceed in three steps.

The first step is to construct a chronology of potentially important interventions in the money-determination and price-determination processes. In particular, we look for significant changes in monetary and credit policy and regulation, wage and price controls, and the structure of relative prices. This chronology is used to identify a tranquil, baseline period – a period in which there are no obvious important interventions in either the money-determination or the price-determination processes.

The second step is to estimate regressions corresponding to the appropriate marginal and conditional distribution functions over the baseline period.

[11]See Hoover (1988a, ch. 8, esp. pp. 182–185) for a discussion of observational equivalence.

If this period has been correctly identified as tranquil, all of these regressions should be invariant – i.e., show stable estimated coefficients.

The third step is to use recursive regression techniques to check for the relative stability of the various conditional and marginal distributions. Two types of related evidence shall be considered. The first is the general relative stability of the coefficients of the different regressions. The second coordinates the institutional record with the econometric evidence. Thus, one would expect, for example, the Nixon price controls to affect the regressions that have prices as dependent variables. Suppose there were no evidence of coefficient instability during the control period in regressions with money as a dependent variable *and* with prices among the independent variables. And suppose there were evidence of instability in a money regression omitting prices. There would, then, be some support for the proposition 'prices cause money'.

While we wholeheartedly agree with the spirit of Romer and Romer (1989) that only a 'narrative' approach that uses the historical record can make statistical results persuasive about causal questions, our method is crucially different from the Romers' method. The Romers look for policy decisions that may well be realizations of a stable rule, and look for the effects of those decisions in the errors of a presumptively stable vector autoregression. Here we look for interventions that fundamentally change policy rules, institutions or the economic environment, and look for the effects of those interventions in the breakdown of stable statistical relationships. These sorts of interventions are less likely to be endogenous than those highlighted by the Romers.[12]

4.1. Stable and unstable periods

Our first step was to examine recent economic history to construct a chronology of potentially important interventions in the money-determina-

[12] Cf. Romer and Romer (1989, pp. 182–183). The exogeneity of interventions can of course still be questioned. Sims (1982, 1986) argues that regime changes are rare, that most putative regime changes are really just observations drawn from a superregime. Sims's argument is considered extensively in Hoover (1988a, pp. 192–202). To some extent, any study of policy, such as ours, accepts Sims's point: we regard resetting of the discount rate or the Federal funds rate not as regime changes but as implementations of a single regime. Sims would, however, have us ramify the single regime to a higher level, so that even changes in which variables the Federal Reserve is targetting or in its operating procedures are just draw from one regime. The essential question is, how far can this process of ramification reasonably be taken? Hoover (1988a) argues that Sims carries this process too far: that there are policy actions that, because they are sufficiently rare, economic actors cannot easily situate as drawings from distributions with well-defined probabilities; and that, if carried to the utmost, Sims's view eliminates a substantial role for the normative analysis of policy [cf. Sargent (1984)]. As a practical matter, there is not, and could not be, a simple empirical test of whether our view or Sims's view is correct. But it should be noted that Sims's view implies that the correct model is a complicated nonlinear one. Were such a model developed and shown to be an adequate, stable characterization of the data, our position would have to be reconsidered. Until then, the assumption that regimes change is a reasonable one.

tion and price-determination processes. Although this chronology is not reproduced here, several points should be noted.[13] First, 1954:I–1966:II is a tranquil period with no obviously important interventions in either the money-determination or the price-determination processes. This period begins after the end of the Korean War and its economic disruptions, and ends the quarter before the Federal Reserve's 'credit crunch' of 1966.

Second, both the period before this tranquil period, 1950:I–1953:IV, and the period after, 1966:III–1985:IV, contain important interventions. Some of the important ones in the earlier period are: wage and price controls (1951:I–1953:I), credit controls (1950:III–1952:II), the Fed/Treasury Accord (1951:I) leading to the withdrawal of price supports for government securities and the adoption of the 'bills-only' doctrine (finalized 1953:IV). In the later period, some important interventions are: the credit crunch and associated policy measures (1966:III), the introduction of lagged reserve accounting (1968:III) and the reintroduction of contemporaneous reserve accounting (1984:I), the Nixon wage/price controls (1971:III–1974:II), the progressive adoption of explicit monetary targets (from 1972:I on), the introduction of interest-bearing checking accounts (1972:II–1980:I), the experiment with reserve base control (1979:IV–1982:III), the oil price shocks (1974:I and 1979:IV), and credit controls (1980:I).

Third, the structure of relative prices changed significantly throughout the 1950:I–1985:IV period. Fig. 1 plots the ratios of several important price indices. During the tranquil period, these ratios show steady trends. But in the periods before 1954:I and after 1966:II, the changes are far more erratic, larger, and more rapid.

4.2. Specification methodology

The second step is to estimate regressions corresponding to the appropriate marginal and conditional distribution functions over the period 1954:I–1966:II.

Noninvariance of estimated regressions might result from two different sources: a policy intervention or functional misspecification. The first is crucial to revealing a causal ordering; the second is a nuisance and must be guarded against. To obtain well-specified regressions, we use the general-to-specific modeling technique of Hendry and Richard (1982).

A number of authors have explained and defended Hendry and Richard's modeling methodology.[14] P.C.B. Phillips (1988, pp. 352–353) provides a

[13] For a detailed chronology 1950–1979 see Hoover (1985, ch. 8).

[14] See Hendry (1983, 1987), McAleer, Pagan, and Volcker (1985), and Gilbert (1986). Pagan (1987) surveys and compares it with the approaches of Leamer and of Sims. For applied examples, see Rose (1985) and Hoover (1988d).

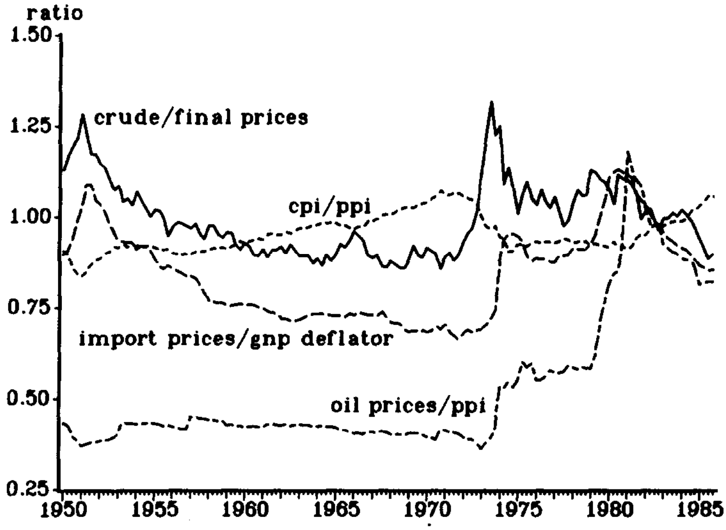

Fig. 1. Ratios of selected price indices. Crude prices are Producer Price Index: Crude Materials for Further Processing and final prices are Producer Price Index: Finished Goods from Producer Price Indexes by Stage of Processing; PPI is the Producer Price Index: All Commodities; oil price is the Producer Price Index: Crude Petroleum; 1982 = 100. CPI is the Consumer Price Index: All Urban Consumers (All Items); 1982–84 = 100. Data are for the last month of the quarter; not seasonally adjusted. Original source: Department of Labor. Import prices and the GNP deflator are the Implicit Price Deflators for Gross National Product and for Imports; 1982 = 100; quarterly; seasonally adjusted. Original source: Department of Commerce. All data are from CITIBASE: Citibank economic database (July 1990)

succinct description of the main points:

> 'In the Hendry methodology [model selection] involves working back from a general unrestricted dynamic specification towards a parsimoniously reparameterized model whose regressors are temporal transformations that are interpretable in some economic sense and nearly orthogonal... [T]he objective of the Hendry approach is to seek out a single-equation model that is a tentatively adequate, conditional data characterization. Such a model satisfies the following criteria...
>
> (a) The model is coherent with the data (i.e., fits the data up to an innovation that is white noise and, further, a martingale difference sequence relative to the selected data base);
> (b) It validly conditions on variables that are weakly exogenous with respect to the parameter of interest;
> (c) It encompasses rival models;
> (d) Its formulation is consistent with economic theory;
> (e) It has parsimoniously chosen and orthogonal decision variables.'

Phillips describes a specification search procedure precisely like the one used in section 4.3 below. On the basis of econometric theory, he shows that such a search is a practical procedure for reducing a cointegrated system to a single error-correction equation. Comparing Hendry and Richard's procedure to the theoretically optimal method for achieving the reduction, Phillips (1988, p. 357) concludes

> 'that the Hendry methodology comes remarkably close to achieving [this] optimal inference procedure. In some cases it actually does so and in other cases it can be further modified to achieve it... These findings can be taken to provide strong support for the Hendry approach.'

It remains true, of course, that since high t-statistics and other desirable characteristics such as white-noise errors are used to design the regression, these characteristics cannot themselves be used within sample as valid tests. The regressions reported in the second step are *designed* to have desirable properties. To guard against *insidious* dating mining, some observations of the baseline period are retained for an out-of-sample stability check. Similarly, the only test statistics from which inferences are drawn directly are either against alternative models or, more importantly, out-of-sample stability tests in the third step. There is no specification search over these data; hence the test statistics bear their usual interpretation.

Some would argue that it would be better to use vector autoregressions [e.g., Sims (1980)] than to follow Henry and Richard's methodology. But this poses serious difficulties. The large number of coefficients to be estimated in a typical VAR would eat up the limited degrees of freedom in our dataset. In addition, such overparameterization usually implies that coefficient estimates are very imprecise, which would reduce the power of tests of structural change. This would pose particular difficulty in this study since detection of changing coefficients is the desideratum of the underlying analysis of causal inference.[15]

4.3. Baseline regressions

To apply the general-to-specific modeling technique, we begin with an unrestricted distributed lag regression [UDL(5)] of the (log-) levels of money

[15]Some might prefer an automatic procedure to reduce the VARs to a more parsimonious form; but, given that there are many plausible criteria [e.g., see Box and Pierce (1970), Ljung and Box (1978), Akaike (1973), Schwarz (1978)], all the same issues of judgment that an automatic procedure aims to avoid would be reintroduced into the design of the reduction procedure. Relentless application of encompassing tests against seriously maintained rivals provides an *objective* specification test even when an informal reduction procedure is followed [see Mizon and Richard (1986), Hendry and Richard (1987), and Hendry (1988b) for details of encompassing and its relation to nonnested hypothesis testing].

Table 1

Baseline price regressions.[a]

Model	Equation[b]
(1.1) Complete: $D(P\|M,r,Z)$	$\begin{array}{lllll} \Delta CPI = & 0.18\,\Delta_4 W_{-1} - & 0.0033\,\Delta SR + & 0.0022\,\Delta_2 SR_{-4} - & 0.0046\,\Delta_2 LR_{-4} \\ (0.0037) & (0.044) & (0.0013) & (0.00079) & (0.0026) \\ {[0.0037]} & {[0.039]} & {[0.0014]} & {[0.00067]} & {[0.0027]} \end{array}$ $\begin{array}{llll} -\,0.066(CPI-W)_{-5} - & 0.071CPI_{-5} + & 0.16\,\Delta_2 M_{-1} + & 0.56 \\ (0.017) & (0.028) & (0.062) & (0.19) \\ {[0.013]} & {[0.023]} & {[0.057]} & {[0.15]} \end{array}$ + 3 seasonal dummies
(1.2) Marginal of money: $D(P\|r,Z)$	$\begin{array}{llll} \Delta CPI = & 0.11\,\Delta_2 W_{-1} + & 0.21\,\Delta_2 W_{-3} - & 0.0029\,\Delta(LR-SR)_{-4} \\ (0.0037) & (0.060) & (0.055) & (0.0011) \\ {[0.0037]} & {[0.057]} & {[0.052]} & {[0.0008]} \end{array}$ $\begin{array}{llll} -0.0038LR_{-4} - & 0.030CPI_{-5} - & 0.071(CPI-W)_5 + & 0.41 \\ (0.0016) & (0.026) & (0.017) & (0.17) \\ {[0.0016]} & {[0.027]} & {[0.016]} & {[0.17]} \end{array}$ + 3 seasonal dummies
(1.3) Marginal of interest rates: $D(P\|M,Z)$	$\begin{array}{llll} \Delta CPI = & 0.14\,\Delta_4 W_{-1} - & 0.033\,\Delta_2\Delta_2 M - & 0.076(CPI-M)_{-5} \\ (0.0037) & (0.032) & (0.012) & (0.013) \\ {[0.0037]} & {[0.038]} & {[0.010]} & {[0.011]} \end{array}$ $\begin{array}{l} -\,0.036 + 2 \text{ seasonal dummies} \\ (0.0062) \\ (0.0052) \end{array}$
(1.4) Marginal of interest rates and money: $D(P\|Z)$	$\begin{array}{llll} \Delta CPI = & 0.020\,\Delta_4 W_{-1} - & 0.055(CPI-W)_{-5} - & 0.0040CPI_{-5} \\ (0.0037) & (0.046) & (0.018) & (0.026) \\ {[0.0037]} & {[0.045]} & {[0.017]} & {[0.028]} \end{array}$ $\begin{array}{l} +\,0.38 + 3 \text{ seasonal dummies} \\ (0.18) \\ {[0.19]} \end{array}$

[a]All estimates by ordinary least squares using PC-GIVE, version 5.0.

[b]50 observations, 1954:I–1966:II. Mean in parentheses, standard deviation in square brackets beneath the dependent variable. Standard errors in parentheses and heteroscedasticity-corrected standard errors in square brackets beneath independent variables.

Table 1 (continued)

Summary statistics[c]									
R^2	*SER*	*RSS*	AR(·) $F(\cdot,\cdot)$	Normality $\chi^2(2)$	ARCH(·) $F(\cdot,\cdot)$	Chow − 1 $F(\cdot,\cdot)$	Chow − 2 $F(\cdot,\cdot)$	Nested in UDL(5) $F(\cdot,\cdot)$
0.86	0.0022	0.00019	(4) 0.46 (4,35)	1.19	(4) 0.19 (4,31)	1.29 (4,35)	1.42 (11,28)	1.33 (22,15)
0.84	0.0023	0.00021	(4) 0.30 (4,36)	0.01	(4) 0.69 (4,32)	1.48 (4,36)	0.77 (10,30)	1.01 (23,17)
0.81	0.0024	0.00026	(4) 0.09 (4,40)	0.47	(4) 0.63 (4,36)	0.76 (4,40)	0.87 (6,38)	0.40 (23,29)
0.80	0.0025	0.00027	(4) 0.71 (4,39)	1.22	(4) 0.54 (4,35)	2.10 (4,39)	1.04 (7,36)	0.42 (35,8)

[c]*SER* = standard error of regression. *RSS* = residual sum of squares. *AR*(·) = test for residual autocorrelation up to order (·); statistic distributed $F(\cdot,\cdot)$. Normality = Jarque and Berra's (1980) test for normal residuals; statistic distributed $\chi^2(2)$. ARCH(·) = Engle's (1982) test for autoregressive conditional heteroscedasticity up to order (·); statistic distributed $F(\cdot,\cdot)$. Chow − 1 = Chow test of whole sample against a sample shorter by four observations; statistic distributed $F(\cdot,\cdot)$. Chow − 2 = Chow test of whole sample against both half samples; statistic distributed $F(\cdot,\cdot)$. Nested in UDL(5) = test of reported regression against a vector autoregression including the level and five lags of each basic variable (i.e., without differences or lags) in the reported regression; statistic distributed $F(\cdot,\cdot)$.

or prices on five lags of themselves and on the current and five lagged values of each of the independent variables (which vary depending upon which conditional or marginal distribution we are estimating) and on a constant and three seasonal dummies (all data are seasonally unadjusted).

Money, interest rates, and prices are represented by quarterly observations of M1, the three-month Treasury bill rate, and the consumer price index (CPI).[16] These variables are chosen partly for their appropriateness to the problem at hand (M1 was after all the Federal Reserve's premier monetary aggregate, and the CPI is the most publicly visible price index) and partly in order to get consistent data back to 1950 (especially in the case of the T-bill rate). The other variables were chosen on the basis of encompassing tests from those suggested by the relevant economic theory.

Currency is within the direct control of the Federal Reserve.[17] Therefore, currency could be a cause of prices but could not be caused by prices (see sections 4.4 and 4.7 below). Hence, the money variable (M) in price regressions is M1, which includes currency, and the money variable in money regressions is demand deposits (DD), defined as M1 excluding currency.

The complete price regression [(1.1) in table 1] is

$$\begin{aligned}\Delta CPI = {} & \underset{[0.039]}{0.18\,\Delta_4 W_{-1}} - \underset{[0.0014]}{0.0033\,\Delta SR} + \underset{[0.00067]}{0.0022\,\Delta_2 SR_{-4}} \\ & - \underset{[0.0027]}{0.0046\,\Delta_2 LR_{-4}} - \underset{[0.013]}{0.066\,(CPI-W)_{-5}} \\ & - \underset{[0.023]}{0.071 CPI_{-5}} + \underset{[0.057]}{0.16\,\Delta_2 M_{-1}} + \underset{[0.015]}{0.56} \\ & + 3 \text{ seasonal dummies,}\end{aligned}$$

$R^2 = 0.86$, AR(4): $F(4,35) = 0.46$,
$SER = 0.0022$, ARCH(4): $F(4,31) = 0.19$,
Normality: $\chi^2(2) = 1.19$, Nested in UDL(5): $F(22,15) = 1.33$,
Chow − 1: $F(4,35) = 1.29$, Chow − 2: $F(11,28) = 1.42$.

[16] Natural logarithms of all variables except interest rates are used. Interest rates are pre-tax in price regressions and after-tax in money regressions. Exact sources and definitions of all the data used in this paper are given in Hoover (1988b, app.), and are available separately from the author.

[17] It is often argued that currency is in practice in the control of the public. This is, however, simply the result of the Federal Reserve's having adopted a fully accommodating reaction function and so does not reverse causality.

W is the wage rate, SR a short rate of interest, and LR a long rate. Heteroscedasticity-corrected standard errors are shown in square brackets.

This regression is in mixed 'levels-and-differences' or 'error-correction' form, which permits it to capture short-run behavior while remaining consistent with the degree of cointegration between variables (i.e., with their long-run equilibrium behavior) [see Harvey (1981, pp. 290–292)]. The long-run solution is

$$\overline{CPI} = 0.48\overline{W} + 4.1\,.$$

Neither money nor interest rates appear in the long-run solution. This is because only their differences enter regression (1.1). This form appeared to encompass every rival that included money or interest rates in levels. Originally several additional regressors were considered. Time deposits, labor productivity, the unemployment rate, and unit labor costs were omitted from the specification search because the null hypothesis of no effect of their levels and five lags in UDL(5) could not be rejected at the 5 percent confidence level.

The specification search was conducted retaining four observations for out-of-sample tests of stability. Thus the statistic labelled 'Chow – 1' reports the test of the constancy of the coefficients of the regression over the period 1954:I–1965:II against the full sample (1954:I–1966:II).[18] The null hypothesis of constant coefficients cannot be rejected at the 5 percent confidence level. We take this as evidence in favor of the specification, and report only the regression using the full sample.

In addition to this out-of-sample test, the statistic labelled 'Chow – 2' also reports a within-sample test to buttress the claim that the baseline period is in fact tranquil in the sense that there is no structural change in the regression during it. A standard Chow test is run with the sample divided in half (1954:I–1960:I and 1960:II–1966:II).[19] Again, the null hypothesis of constant parameters cannot be rejected.

Although the functional form of the regression was determined through an extensive search, only the final form is reported here. The null hypothesis that this form is a valid restriction of the corresponding UDL(5) cannot be rejected at the 5 percent confidence level [see the statistic labelled 'Nested in UDL(5)'].

[18]The test is Chow's first test which can be computed even when the number of observations in one subsample is less than the number of regressors [Chow (1960, pp. 594, 595, 598)].

[19]The standard Chow test is Chow's second test [Chow (1960, pp. 595–599)].

Table 2

Baseline money regressions.[a]

Model	Equation[b]
(2.1) Complete: $D(DD\mid P,r,Z)$	$\Delta_4(DD-CPI) = 0.91\,\Delta_4(DD-CPI)_{-1} - 0.26\,\Delta_4 CPI + 1.56\,\Delta CPI_{-4}$ (0.010) (0.038) (0.070) (0.21) [0.023] [0.033] [0.075] [0.21] $-\ 0.012\,\Delta_4 SR_{-1} - 0.018\,\Delta(LR-SR) - 0.010\,\Delta(LR-SR)_{-3}$ (0.0015) (0.0037) (0.0044) [0.0018] [0.0034] [0.0045] $+\ 0.078\,\Delta_2\,\Delta GNP_{-2} - 0.0079(DD-GNP-CPI)_{-5}$ (0.015) (0.0065) [0.015] [0.0060]
(2.2) Marginal of prices: $D(DD\mid r,Z)$	$\Delta_4 DD = 1.00\,\Delta_4 DD_{-1} - 0.68\,\Delta DD_{-4} + 0.13\,\Delta_3(\tfrac{1}{2}GNP + \tfrac{1}{2}GNP_{-1})$ (0.024) (0.076) (0.16) (0.041) [0.017] [0.065] [0.16] [0.036] $-\ 0.0088\,\Delta_4 SR_{-1} - 0.011\,\Delta(LR-SR)_{-3} - 0.0057(LR-SR)_{-5}$ (0.0016) (0.0049) (0.0025) [0.0013] [0.0045] [0.0024] $+\ 0.0046 DD_{-5}$ + 3 seasonal dummies (0.0010) [0.0010]
(2.3) Marginal of interest rates: $D(DD\mid P,Z)$	$\Delta DD = -\ 0.55\,\Delta DD_{-3} - 0.098 CPI_{-5} + 0.12\,\Delta_4 GNP$ (0.0060) (0.15) (0.033) (0.029) [0.0170] [0.15] [0.032] [0.032] $-\ 0.35[DD - \tfrac{1}{2}(GNP+CPI)]_{-5} + 1.28$ (0.083) (0.33) [0.071] [0.30] + 3 seasonal dummies
(2.4) Marginal of prices and interest rates: $D(DD\mid Z)$	$\Delta DD = 0.21\,\Delta DD_{-1} + 0.0078 DD_{-1} + 0.082\,\Delta GNP$ (0.0060) (0.13) (0.0011) (0.053) [0.0170] [0.14] [0.0010] [0.055] $-\ 0.19\,\Delta GNP_{-4}$ + 3 seasonal dummies (0.051) [0.044]

[a,b,c]See notes to table 1.

Table 2 (continued)

Summary statistics[c]									
R^2	SER	RSS	AR(·) $F(\cdot,\cdot)$	Normality $\chi^2(2)$	ARCH(·) $F(\cdot,\cdot)$	Chow − 1 $F(\cdot,\cdot)$	Chow − 2 $F(\cdot,\cdot)$	Nested in UDL(5) $F(\cdot,\cdot)$
0.097	0.0047	0.00093	(4) 0.66 (4,38)	0.24	(4) 0.38 (4,34)	0.96 (4,38)	2.13 (8,34)	0.91 (25,17)
0.98	0.0050	0.00099	(4) 0.37 (4,36)	1.00	(4) 1.08 (4,32)	0.90 (4,36)	1.21 (10,30)	0.80 (17,23)
0.93	0.0054	0.0012	(4) 1.14 (4,38)	0.27	(4) 2.34 (4,34)	1.48 (4,38)	0.87 (8,34)	0.53 (13,21)
0.91	0.0057	0.0014	(4) 0.95 (4,39)	2.37	(4) 0.93 (4,35)	1.40 (4,39)	1.74 (7,36)	1.11 (8,35)

In addition, the regression passes a battery of standard diagnostic tests. We cannot reject at the 5 percent level the null of:

(i) autocorrelation up to fourth order (see the statistic labelled 'AR'),
(ii) autoregressive conditional heteroskedasticity up to fourth order (see the statistic labelled ARCH), or
(iii) normality of the estimated residuals (see the statistic labelled 'Normality').

Our conclusion must be that the complete price regression is tentatively acceptable as a parsimonious representation of the corresponding distribution function. Other, better representations may exists – using additional independent variables or alternative functional forms. If they are truly better, they would have to pass the same battery of tests *and encompass* the reported regression.

The complete money regression [(2.1) in table 2] is

$$\begin{aligned}\Delta_4(DD-CPI) = {} & \underset{[0.033]}{0.91}\,\Delta_4(DD-CPI)_{-1} \underset{[0.075]}{-0.26}\,\Delta_4 CPI \\ & \underset{[0.21]}{+1.56}\,\Delta CPI_{-4} \underset{[0.0018]}{-0.012}\,\Delta_4 SR_{-1} \\ & \underset{[0.0034]}{-0.018}\,\Delta(LR-SR) \underset{[0.0045]}{-0.010}\,\Delta(LR-SR)_{-3} \\ & \underset{[0.015]}{+0.078}\,\Delta_2\Delta GNP_{-2} \underset{[0.0060]}{-0.0079}\,(DD-GNP-CPI)_{-5},\end{aligned}$$

$$\begin{array}{ll} R^2 = 0.97, & \mathrm{AR}(4)\colon F(4,38) = 0.66, \\ SER = 0.0047, & \mathrm{ARCH}(4)\colon F(4,34) = 0.38, \\ \text{Normality: } \chi^2(2) = 0.24, & \text{Nested in UDL}(5)\colon F(25,17) = 0.91, \\ \text{Chow}-1\text{: } F(4,38) = 0.96, & \text{Chow}-2\text{: } F(8,34) = 2.13. \end{array}$$

As with the complete price regression, we cannot reject the null hypotheses that the complete money regression has stable coefficients, white noise errors, and is a valid restriction of a more general dynamic form.

The long-run solution is

$$\overline{DD} = \overline{GNP} + \overline{CPI}.$$

In other words, money and nominal income are homogeneous of degree one as many theories suggest.

The complete price regression is an estimate of the distribution of prices conditional on money, interest rates, and other variables $[D(P|M, r, Z)]$. The complete money regression is an estimate of demand deposits conditional on prices, interest rates, and other variables $[D(DD|P, r, Z)]$.

To obtain the various marginal distributions, we omitted the variable with respect to which the distribution was to be marginalized and its lags from UDL(5) and conducted a completely new specification search. Table 1 reports the complete price regression as well as marginalizations with respect to interest rates, money, and interest rates and money. Table 2 reports the complete money regression as well as marginalizations with respect to prices, interest rates, and prices and interest rates. Each of the reported regressions passes the entire battery of specification tests described in detail for the complete price regression.

Among the additional variables, interest rates are singled out for special treatment because of the possibility, especially given the *modus operandi* of monetary policy, of substantial interaction between money and interest rates. So, even though no full investigation of the causal direction between interest rates and either money or prices will be undertaken, regressions of both these variables marginal of interest rates are reported.

4.4. Digression: Reaction function or causal relation?

Regressions (2.1) and (2.3) in table 2 have prices as independent variables. The existence of policy reaction functions that change from time to time does not itself pose a problem for our method, but may in fact supply the very sorts of interventions in the money-determination process that yield useful information for causal inference. Still, we must rule out the possibility that the only reason prices ever appeared as independent variables in money regressions was merely because the Federal Reserve had an inflation target and adjusted the money supply in response to price information.[20] Further examination of the baseline regressions suggests that this is not the case.

A Federal Reserve reaction function would have to run from prices through its directly controlled instruments – reserves, the discount rate, and the Federal funds rate – and, finally, through the effects of these instruments on the banking system and the public to the stock of deposits. In that case, if prices were to appear in the money regressions, it would be because prices were proxying for the Federal Reserve's policy instruments. If that were so, then the relationship between demand deposits and these instruments should

[20]The possibility that it would pose a problem has been suggested by several commentators on earlier drafts.

Table 3

Encompassing tests for policy reaction functions.

Regression[a] (2.1)	Regression[a] (2.3)	plus	Additional variables	over	Sample period
1.49 $F(6,36)$	1.60 $F(6,36)$		Total reserves		1954:I–1966:II
1.70 $F(18,15)$	2.12 $F(18,15)$		Total reserves Discount rate, Federal funds rate		1956:II–1966:II

[a]Each regression adds the current level plus five lags of the noted variables to one of the regressions in table 2 and tests their joint significance.

Critical values: $F_{0.95}(6,36) = 2.37$, $F_{0.95}(18,15) = 2.35$.

be closer and more direct than that between demand deposits and prices. If this proves not to be so, we can fairly rule out the reaction function story.[21]

Table 3 presents tests of whether adding Federal Reserve policy instruments to those baseline regressions that include prices as regressors makes any difference. The policy instruments are total reserves, the discount rate, and the Federal funds rate. Because the Federal funds rate is available only after 1955:I, all three instruments could be included only if the regression began in 1956:II (allowing for five lags), two years after the beginning of the baseline period. Therefore, one set of regressions is run over 1956:II–1966:II. The other set uses only total reserves, omitting both interest rates, because it is the interaction of the Federal funds rate and the discount rate which is relevant to bank borrowing decisions. It runs the entire baseline period, 1954:I–1966:II.

In no case can we reject the hypothesis at the 5 percent confidence level that the policy instruments may be excluded from the baseline regressions [(2.1) and (2.3)] plus policy instruments. The baseline regressions therefore *parsimoniously encompass* the less restricted regressions in which they are nested. According to a theorem reported by Hendry and Richard (1987, sects. 6–8) they in turn encompass any regressions *minimally nested* in the less restricted regression.[22] Hence, regressions (2.1) and (2.3) encompass every regression that adds any linear combination of the policy variables to

[21]Cagan (1965, pp. 235–239) employs a similar argument in his study of the money supply.

[22]One regression *parsimoniously encompasses* another if it encompasses the other and is nested within the other [Hendry and Richard (1987)]. Two regressions are *minimally nested* in a third if no variable can be omitted from the third without one or both of the first two failing to be nested in the resulting regression.

them. The policy variables therefore appear to be redundant, and it is unlikely that prices in (2.1) and (2.3) are related to deposits simply because of a policy reaction function.[23]

4.5. Techniques for identifying structural breaks

The third step in our investigation is to check for the relative stability of the various conditional and marginal distributions.

Two related types of evidence shall be considered. The first is the general relative stability of the coefficients of the different regressions. The second coordinates the institutional record with the econometric evidence. Thus, one would expect, for example, the Nixon price controls to affect the regressions that have prices as dependent variables. Suppose there were no evidence of coefficient instability during the control period in regressions with money as a dependent variable *and* with prices among the independent variables. And suppose there were evidence of instability in a money regression omitting prices. There would, then, be some support for the proposition 'prices cause money'.

There are two questions relating to relative stability: do structural breaks occur at all? And, given that they occur, exactly when? A test which accurately answers the first question may still not give very good information about the second. While both questions interest us, the second is more important to the issues at hand. Although the literature is burgeoning, most recent work on detecting structural breaks when break points are unknown [e.g., Andrews (1989), Kim and Siegmund (1989), Hansen (1990), and Chu (1990)] aims to answer the first question. [An exception is Banerjee, Lumsdaine, and Stock (1990), who address the second question in attempting to discriminate between unit roots and breaks in trends for GNP.] The powers of alternative tests have been worked out by Monte Carlo methods, usually against the simple alternative hypothesis of discrete one-time shifts in coefficients. Unfortunately, neither the precise nature of the structural breaks (discrete or gradual) nor their number or timing can be known in advance. What is more, the exact effects of pretest bias due to our specification search or out-of-sample test statistics are unknown. Working them out using Monte Carlo methods would again be sensitive to the maintained assumption about the true data-generating process in a situation in which we are not really justified in adopting the 'Axiom of Correct Specification' [Leamer (1978)].

[23]An even better test would compare (2.1) and (2.3) not to themselves augmented by the policy variables but to the appropriate UDL(5) augmented by the policy variables. Unfortunately, given only 50 observations in the baseline period, degrees of freedom become a problem.

Recognizing that these problems remain open issues for methodological research, we nevertheless employ several tests based on recursive estimation.[24] Recursive estimation reestimates each regression adding one period at a time to the beginning or to the end of the sample period. Each of the regressions in tables 1 and 2 are projected separately forward and backward. For each specification, we report the four tests covering the periods before and after the baseline period.

(i) *A constant-base sequential Chow test.* The test divides the sample into two parts, the first is always the baseline period (hence 'constant-base'), the second comprises the observations from the end point of the baseline period to the last observation used in the recursive estimate.[25] Thus, if the baseline period is N observations and the total sample is K observations, the second part starts out as one observation, and increases one observation for each recursive reestimation, until finally there are $K-N$ observations. A sequence of $K-N$ Chow statistics is then computed.[26]

The sequence of Chow statistics from this test are not independent. Their true distribution has not been worked out. One way of interpreting them, however, as a means of *localizing* a structural break is counterfactually: suppose that someone challenges us claiming that our baseline regression breaks down at some particular point out of sample, the single Chow statistic out of our sequence for that point is the correct test with a critical value of the size reported in conventional F-tables. Thus, as a summary statistic for this test, we report the 'break point' – the first time the Chow statistic exceeds its 5 percent critical value.

(ii) *A constant-horizon sequential Chow test.* The test reports the Chow statistics from splitting the sample at each observation between the end point of the baseline period and the end point of the sample. All $N+K$ observations are always used (hence 'constant-horizon').

As with the constant-base test, the Chow statistics in this sequence are not independent. Using the maximum value of these statistics as the indicator of a single, discontinuous change in coefficients is equivalent to Quandt's (1960) maximum likelihood method [see Hansen (1990, p. 2)]. The distribution of this 'max-Chow' statistic has only recently been worked out [Andrews (1989), Chu (1990), Kim and Siegmund (1989)]. As a summary statistic for this test, we report the value and the date of the maximum Chow statistic. Since the natural alternative hypothesis to the null of stable coefficients with the

[24]See Harvey (1981, pp. 54–57) on the technique of recursive regression.

[25]The endpoint may be either the first or last observation in the baseline period depending on whether we are projecting backwards or forwards.

[26]All three sequential Chow tests use Chow's first test, see footnote 17.

max-Chow test is a single, discontinuous change in the true coefficients, and we have no reason to rule out gradual change or multiple discontinuities, other characteristics of the constant-horizon sequence (such as local maxima) may also be reported [cf. Hansen (1990, p. 3)].

(iii) *A one-step-ahead sequential Chow test.* The test reports the sequence of Chow statistics computed from comparing each recursive estimate to the one that uses just one more observation.

Because the one-step-ahead prediction errors of recursive regressions are independent, each Chow statistic in this sequence is itself independent [Harvey (1981, pp. 55–56)]. Thus, in large sample using a 5 percent critical value, about 5 percent of the statistics should exceed the critical value. We therefore report the ratio of the number of violations to the total number of statistics in the sequence. To get a further handle on the significance, we also report the value and the date of the maximum one-step-ahead Chow statistic as well as its maximum scaled by its 5 percent critical value. When it is particularly informative we reproduce the plot of the entire sequence of statistics.

(iv) *Plots of recursively estimated regression coefficient against their standard errors.* Visual inspection of these plots appears in simulation studies often to be a reliable way to determine the date of a structural break.[27] Because of the large numbers of plots generated, only the most informative ones are reproduced.

The number of tests examined of course means that we use no simple, mechanical decision procedure to determine if and when a break occurs. This is appropriate since the nature, timing, and possible multiplicity of interventions in the underlying data-generating process cannot be known *a priori*. This is in the spirit of the pioneering work of Brown, Durbin, and Evans (1975, pp. 149–150):

> '...the techniques are designed to bring out departures from constancy in a graphic way instead of parameterizing particular types of departure in advance and then developing formal significance tests intended to have high power against these particular alternatives. From this point of view the significance tests suggested should be regarded as yardsticks for

[27]Based on Cai (1990) and work in progress by Yongxin Cai. The standard errors reported are the same as for the corresponding OLS regression, and are used as *informal* guides to sampling variability. Formal tests would have to be constructed along similar lines to Banerjee et al. (1990), who derive correct asymptotic Dickey–Fuller t-statistics for tests of the unit root hypothesis using recursive regressions.

the interpretation of data rather than leading to hard and fast decisions.'[28]

Such an attitude is clearly in keeping with the spirit of the 'narrative' approach in which multiple sources of information are combined to construct a persuasive case.

There are, no doubt, possible pitfalls. First, it might be argued that the results are not robust to alternative, plausible choices of baseline periods. While this is possible, it would be difficult to demonstrate. The baseline period was not chosen arbitrarily, but because prior historical/institutional evidence suggested that it was likely to be stable. Unfortunately, there are not enough degrees of freedom within the baseline period chosen according to this prior evidence to attempt specification searches across alternative subsamples. One check was made, however. Whenever out-of-sample stability tests showed a longer period of stability than the initial baseline, this longer period was taken to be the baseline and within-sample and out-of-sample tests were rerun to confirm the stability of the baseline period and the location of the structural break. These results are not reported formally, because in every case they confirm our initial results.

Second, it might be argued that the techniques of identifying and dating structural breaks or comparative stability of alternative marginalizations are inappropriate or inefficient. Improved techniques are always welcome; but they must be produced and employed in the flesh, so to speak, if the objection is to be effective.

4.6. *Out-of-sample projections*

4.6.1. *Backward recursions, 1953:IV–1950:I*

Table 4 presents some summary statistics from the sequential Chow tests projecting the regressions in table 1 backward to 1950:I.

The striking thing about the projections of price regressions [(4.1)–(4.4)] is their similarity. It is clear, particularly from the one-step-ahead Chow tests, that stability can be rejected in this period for all the price regressions. The break points for the constant-base Chow tests occur at 1952:III in every case. The maxima for the constant-horizon Chow tests occur at 1951:II in every case. Fig. 2 is a typical coefficient plot from the price regressions. The behavior of the constant in the conditional price regression [projection (4.1)] is clearly different before 1951:II than after 1952:II. Plots of the remaining

[28]We in fact also examined, but do not report, the cusum and cumsum-squared plots developed in Brown et al. (1975). In practice they seemed to be less useful at disciminating between alternative regressions than those tests we do report. Hansen (1990, p. 3 passim) gives theoretical reasons why this should be so.

Table 4

Chow tests of structural stability, projections backwards, 1953:IV–1950:I.

	Chow tests				
	Constant-base	Constant-horizon	One-step-ahead		
Projection of[a]	Break point[b]	Maximum (Date)	Ratio[c]	Maximum (Date)	Scaled maximum[d] (Date)
(4.1) Prices complete [Regression (1.1)]	1952:III	24.0 (1951:II)	0.31	92.5 (1951:I)	22.9 (1951:I)
(4.2) Prices marginal of money [Regression (1.2)]	1952:III	28.8 (1951:II)	0.31	104.5 (1951:I)	25.9 (1951:I)
(4.3) Prices marginal of interest rates [Regression (1.3)]	1952:III	44.2 (1951:II)	0.44	108.5 (1051:I)	27.0 (1951:I)
(4.4) Prices marginal of money & interest rates [Regression (1.4)]	1952:III	31.3 (1951:II)	0.31	97.2 (1951:I)	24.2 (1951:I)
(4.5) Money complete [Regression (2.1)]	1950:II	13.6 (1950:II)	0.25	13.6 (1950:I)	3.4 (1950:I)
(4.6) Money marginal of prices [Regression (2.2)]	1950:III	4.9 (1950:IV)	0.19	9.4 (1950:III)	2.3 (1950:III)
(4.7) Money marginal of interest rates [Regression (2.3)]	—[e]	2.9 (1950:II)	0.00	3.5 (1953:I)	0.9 (1953:I)
(4.8) Money marginal of prices & interest rates [Regression (2.4)]	1951:III	2.6 (1952:I)	0.13	6.5 (1951:IV)	1.6 (1951:IV)

[a]Each projection begins with the baseline regression indicated in parentheses and extends the sample starting date backwards one period at a time, reestimating using recursive least squares from 1953:IV to 1950:I (12 additional observations beyond the baseline sample). The three indicated types of Chow tests are calculated at each step.

[b]'Break point' is the date at which the Chow statistic first exceeded its 5 percent critical value.

[c]'Ratio' is the number of violations of the 5 percent critical value divided by the number of observations in the projection period.

[d]'Scaled maximum' is the maximum of the Chow statistics when computed as ratios to their 5 percent critical values.

[e]Never rejects at the 5 percent critical value.

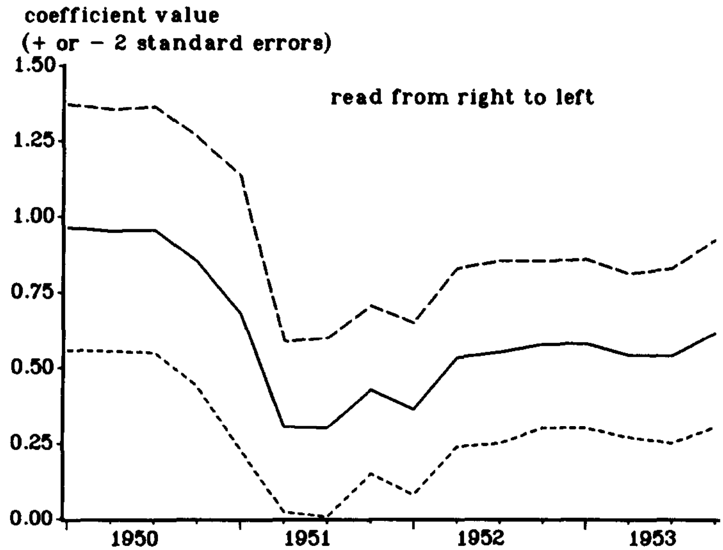

Fig. 2. Constant in regression (1.1) recursively estimated backward from 1953:IV to 1950:I [projection (4.1)].

coefficients (not shown) in the four price regressions tell the same story. A few show little change relative to their standard errors, and a few pick up the change only at 1951:II, but most show structural breaks precisely parallel to those in fig. 2.

None of the regressions is obviously more stable than the others, and none shows any obvious structural break not common to all of the others. This suggests that we have identified interventions in the price-determination process – a conclusion supported by the institutional record. The structural break at 1951:I corresponds exactly to the imposition of wage and price controls during the Korean War, and the break at 1952:II corresponds to the dismantling of those controls.[29]

The projections of the money regressions [(4.5)–(4.8)] tell a different story. The one-step-ahead Chow tests reject parameter stability for every money regression *except* projection (4.7), money marginal of interest rates. While on the constant-base Chow tests the other projections reach break points between 1950:II and 1951:III, parameter constancy is never rejected for projection (4.7). Similarly, while the constant-horizon Chow statistics for the other

[29]The process of dismantling wage and price controls began in 1952:I. Although it did not officially end until 1953:I, it effectively ended in 1952:IV with the resignation of the industrial members of the Wage Board in protest over the board's failure to prevent an increase in wages in the steel industry. See Council of Economic Advisers (1951–53).

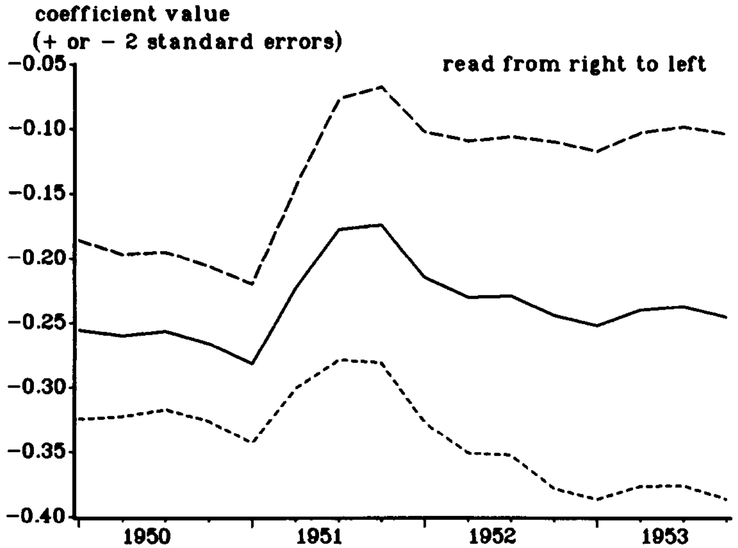

Fig. 3. Coefficient on $\Delta_4 CPI$ in regression (2.1) recursively estimated backward from 1953:IV to 1950:I [projection (4.5)].

regressions reach maxima 1.2 to 3.4 times their 5 percent critical values between 1950:II and 1952:I, projection (4.7) shows only a borderline rejection at 1953:III at exactly the 5 percent critical value; parameter stability can be rejected nowhere else.

Most of the coefficients in projection (4.5), the conditional money regression, show some fluctuations but would not be judged unstable given their standard errors. The coefficient on $\Delta_4 CPI$ (fig. 3), however, clearly changes between 1951:III and 1951:I. The coefficient on ΔCPI_{-4} (not shown) is similar.

In contrast, the coefficients in projection (4.6), money marginal of prices, appear to be stable within relatively large standard-error bands. The coefficients on $\Delta(LR - SR)_{-3}$ and $(LR - SR)_{-5}$ become insignificant near the beginning of the sample. Thus, even though stability is clearly rejected on the Chow tests, it shows up as a general increase in the variance not in a well-defined shift in the means of the coefficients.

Fig. 4 shows that the coefficient on DD_{-1} in projection (4.8), money marginal of prices and interest rates, shifts gradually but significantly over the projection period. There is no single date that can be localized as the point of structural breakdown. The coefficient on ΔGNP_{-4} (not shown) behaves similarly. The remaining coefficients are stable given their standard errors.

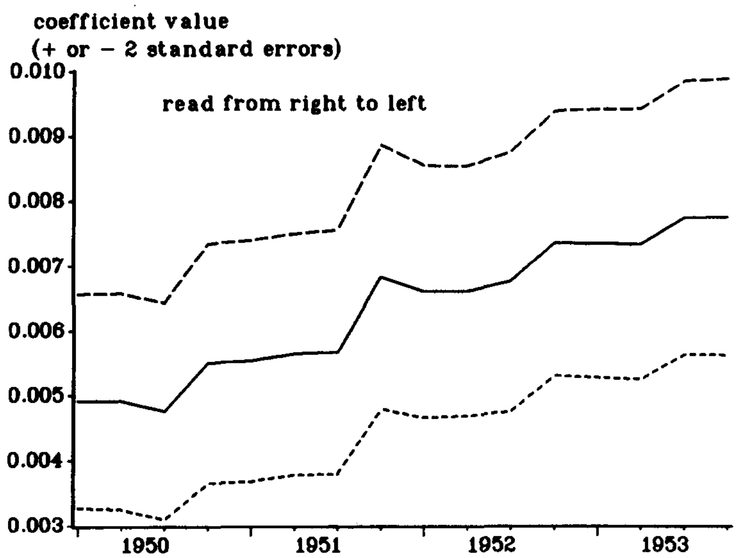

Fig. 4. Coefficient on DD_{-1} in regression (2.4) recursively estimated backward from 1953:IV to 1950:I [projection (4.8)].

Fig. 5 shows the coefficient on CPI_{-5}, the *most* unstable coefficient from projection (4.7), money marginal of interest rates. Except for the borderline significant shift between 1953:III and 1953:I, the coefficient is stable and well determined statistically. The remaining coefficients (not shown) are even more stable.

Credit controls were imposed in 1950:III, tightened in 1951:I, and removed gradually over 1951:IV to 1952:II [Council of Economic Advisers (1950–53)]. When added to the fact that marginalizing with respect to interest rates is stabilizing, these interventions in the interest-rate-determination process suggest that regression (2.3) should be regarded as the conditional model. In that case, the fact that its projection (4.7) is stable, while marginalizing further with respect of prices [projection (4.8)] reintroduces instability, appears to provide precisely the sort of discrimination between a conditional and marginal model needed to indicate causal direction. The timing of the structural breaks in projection (4.8) is unclear. If, however, we judge by the constant-base Chow test, then the break at 1951:III corresponds approximately to the introduction of wage and price controls. Precise timing may be unclear in this case because the structure of relative prices changed significantly and continuously over this period – especially with major boom and bust in commodity prices and different degrees of stringency in wage and price controls. Fig. 1 shows dramatic shifts in relative prices just at the points

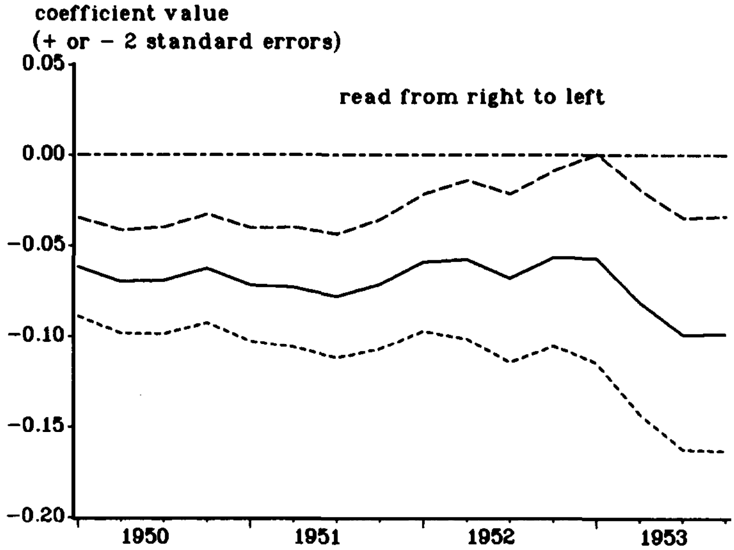

Fig. 5. Coefficient on CPI_{-5} in regression (2.3) recursively estimated backward from 1953:IV to 1950:I [projection (4.7)].

at which the price regressions show structural instability. The rise in the *PPI* relative to the *CPI* and oil prices in 1951:I and its recovery in 1952:IV is compatible with a wage and price regime which controlled retail prices more stringently than wholesale prices and which became gradually less stringent over time [Council of Economic Advisers (1951, pp. 144–146)]. The rise in import and crude prices relative to the prices of final goods may reflect controls as well, but mostly reflects an exogenous boom in world commodity prices. Not all of these changes would necessarily show up as dramatic breakdowns of *CPI* regressions. Yet they may still have had major differential effects on the alternative money regressions. The weight of evidence on the basis of general stability is in favor of the view that prices cause money and interest rates do not: money regressions become more stable if interest rates are omitted and less stable if prices are omitted.

4.6.2. Forward recursions, 1966:III–1985:IV

Table 5 presents summary statistics from the sequential Chow tests based on projecting the regressions in table 1 forward over 1966:III to 1985:IV.

Consider first the price regressions [projections (5.1)–(5.4)]. The one-step-ahead Chow statistics indicate that none of the regressions is stable across the entire projection period. Fig. 6 is a plot of the one-step-ahead Chow statistics for projection (5.1), the complete price regression. It shows the

Table 5

Chow tests of structural stability, projections forwards, 1966:III–1985 IV.

	Chow tests				
	Constant-base	Constant-horizon	One-step-ahead		
Projection of[a]	Break point[b]	Maximum (Date)	Ratio[c]	Maximum (Date)	Scaled maximum[d] (Date)
(5.1) Prices complete [Regression (1.1)]	1970:III	10.0 (1969:IV)	0.20	24.8 (1982:I)	6.3 (1982:I)
(5.2) Prices marginal of money [Regression (1.2)]	1973:II	8.4 (1972:IV)	0.22	15.5 (1974:I)	3.9 (1974:I)
(5.3) Prices marginal of interest rates [Regression (1.3)]	1967:I	13.5 (1971:II)	0.24	48.6 (1983:IV)	12.4 (1983:IV)
(5.4) Prices marginal of money & interest rates [Regression (1.4)]	1973:III	10.4 (1972:IV)	0.22	21.9 (1983:IV)	5.6 (1983:IV)
(5.5) Money complete [Regression (2.1)]	1966:III	6.1 (1966:III)	0.20	32.3 (1981:II)	8.2 (1981:II)
(5.6) Money marginal of prices [Regression (2.2)]	1966:III	5.1 (1978:I)	0.17	31.0 (1981:II)	7.9 (1981:II)
(5.7) Money marginal of interest rates [Regression (2.3)]	1967:III	6.1 (1967:I)	0.19	11.9 (1975:IV)	3.0 (1975:IV)
(5.8) Money marginal of prices & interest rates [Regression (2.4)]	1966:III	5.3 (1974:IV)	0.24	18.2 (1979:II)	4.6 (1979:II)

[a]Each projection begins with the baseline regression indicated in parentheses and extends the sample starting date forwards one period at a time, reestimating using recursive least squares from 1966.III to 1985.IV (78 additional observations beyond the baseline sample). The three indicated types of Chow tests are calculated at each step.

[b]'Break point' is the date at which the Chow statistic first exceeded its 5 percent critical value.

[c]'Ratio' is the number of violations of the 5 percent critical value divided by the number of observations in the projection period.

[d]'Scaled maximum' is the maximum of the Chow statistics when computed as ratios to their 5 percent critical values.

particular difficulty of fitting the periods 1973–1975 and 1979–1984. Inspection of the analogous plots for the other price regressions (not shown) reveals that this is a common feature.

The fact that all price regressions show a similar pattern suggests a break in the price-determination process. One might think that Nixon's price controls could explain the earlier break; but the timing is not quite right. Phase 3 of the price controls, which *relaxed* controls somewhat, was adopted in 1973:I and phase 2, which was stricter, was reimposed in 1973:II. If the common structural breaks in the price equations are to be attributed to price controls, why is there no indication of a common break corresponding to the introduction of the stricter phase 1 (1971:III) or phase 2 (1971:IV)?

A more likely source for these common breaks, which seems to correspond more closely to the apparent timing, is the massive change in the structure of relative prices associated with the oil price increase starting 1974:I and the somewhat earlier commodity price boom. Fig. 1 shows massive changes in the ratios of import prices to the *GNP* deflator and in crude to final goods prices that correspond closely to the indicated structural breaks and can be construed to be violations of the *ceteris paribus* conditions implicit in the *CPI* equations estimated here.

A similar explanation would easily account for the later break. Oil prices rose massively beginning in mid-1979, peaked in 1981:III, and fell substantially until the end of our sample. There were, to be sure, events in the monetary sector – particularly the Federal Reserve's changes in operating procedures in 1979 and 1982 – that might also be thought of as sources of such structural breaks, but for the fact that the breaks are common across all four price regressions.

Although it seems clear that there are interventions in the price-determination process that are common across price regressions, it is still possible to discriminate among them. Fig. 6 shows that the complete price regression probably breaks down before 1973:I as well. Inspection of the plots of the one-step-ahead Chow statistics for the other price regressions (not shown) reveals that the regression marginal of interest rates [projection (5.3)] shows signs of structural break not only before the first common break (1973:I–1975:II) but after it as well. The regressions marginal of money [projection (5.2)] and of money and interest rates [projection (5.4)] shows no signs of breaks other than the two common breaks.

Using the length of time beyond the baseline period before a constant-base Chow test indicates rejection as one measure of general stability suggests that the price regression marginal of money and interest rates [projection (5.4)] is the most stable, followed by that marginal of money [projection (5.2)] and the complete price regression [projection (5.1)], with that marginal of interest rates the least stable. The maxima of the constant-horizon Chow statistics can be ranked by date identically.

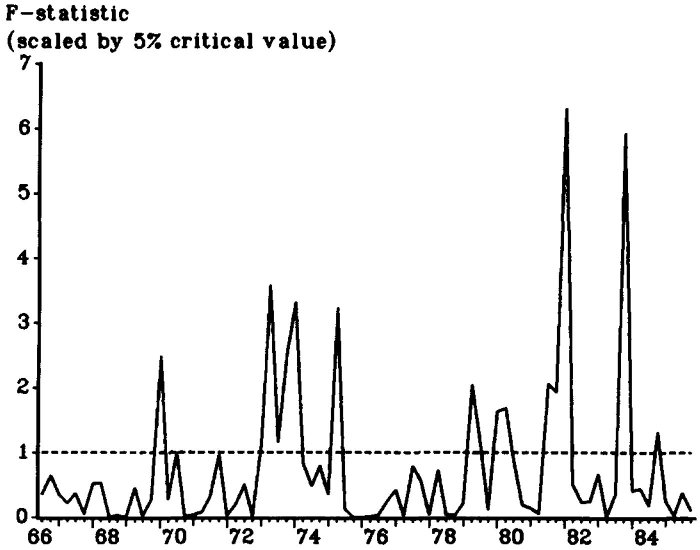

Fig. 6. One-step-ahead Chow statistics from the recursive estimation of regression (1.1) forward from 1966:III to 1985:IV [projection (5.1)].

Fig. 7 plots the constant from the complete price regression [projection (5.1)]. It shows a pattern common to most of the coefficients in all of the price regressions – a hump that peaks about 1972:IV. Judging by the standard error bands, the constant is clearly different after 1975:II than before 1973:I, and clearly different at 1985:IV than at 1979:I. This is typical of other coefficients in other projections and confirms our general conclusions about interventions in the price-determination process. Fig. 7 also shows, however, that the behavior of the constant in the complete-price regression is different before 1969:I than after 1970:I. The coefficient on $\Delta_2 M_{-1}$ in projection (5.1), not shown, is similar. The coefficient on the error-correction term $(CPI-M)_{-5}$ in projection (5.3), fig. 8, shows a somewhat earlier but similar shift in behavior. The constant in projection (5.3) is similar. Inspection of the coefficient plots for the remaining regressions indicates nothing similar.

The evidence therefore suggests that there are two common interventions in the price-determination process, and that regressions which include money show structural break even before the first intervention. Both examination of coefficient plots and more general measures of structural stability confirm that the omission of money from price regressions is stabilizing.

The forward projection of the money regressions (5.5)–(5.8) are somewhat trickier to interpret than those for the price regressions. The one-step-ahead Chow statistics indicate that none of the four is stable. The maxima for the

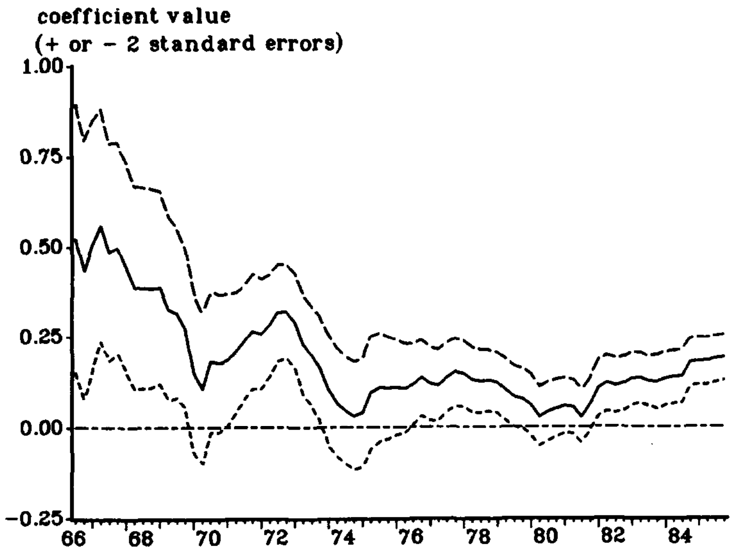

Fig. 7. Constant in regression (1.1) recursively estimated forward from 1966:III to 1985:IV [projection (5.1)].

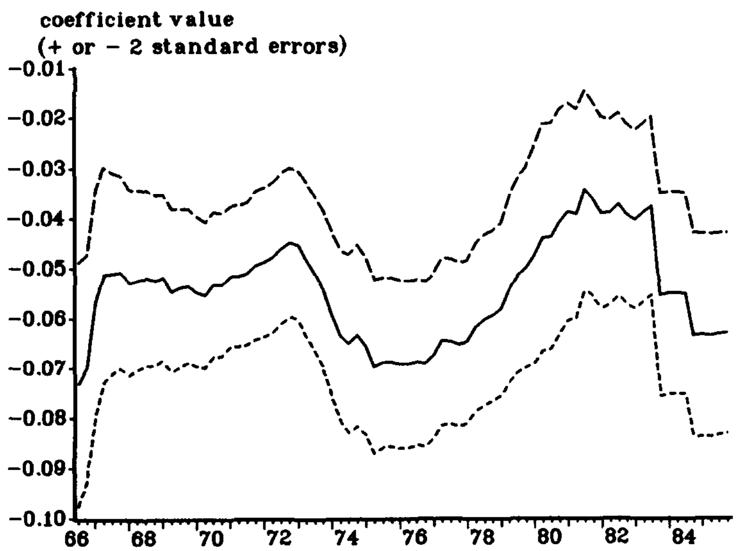

Fig. 8. Coefficient on $(CPI - M)_{-5}$ in regression (1.3) recursively estimated forward from 1966:III to 1985:IV [projection (5.3)].

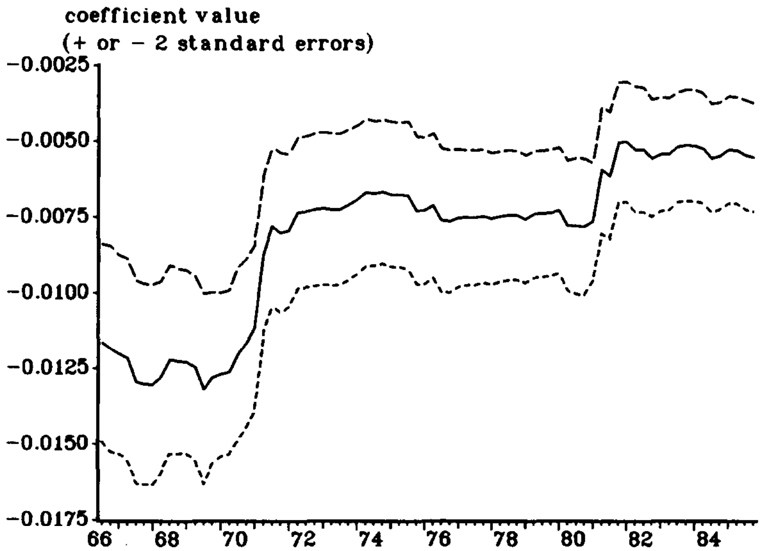

Fig. 9. Coefficient on $\Delta_4 SR_{-1}$ in regression (2.1) recursively estimated forward from 1966:III to 1985:IV [projection (5.5)].

statistic for both the complete money stock [projection (5.5)] and the regression marginal of prices [projection (5.6)] occurs at 1981:II and is twice as large as the maxima for the other two regressions and six to fifteen times as large as the statistics for the other two regressions at 1981:II. This is suggestive of instability associated with the interest-rate-determination process, since only projections (5.5) and (5.6) include interest rates. Fig. 9, which plots the coefficient on $\Delta_4 SR_{-1}$ in projection (5.5), makes the point dramatically. The coefficient shows two distinct breaks, and the second one coincides with maximum for the one-step-ahead statistic. The plot of the coefficient on $\Delta_4 SR_{-1}$ for projection (5.6), not shown, is nearly identical. The other coefficients on interest rate variables in those regressions show similar though less dramatic behavior. Such instability, common to money regressions both with and without prices as a conditioning variable, suggests that, as with the backward projections, we should regard the regression of money marginal of interest rates as the conditional distribution.

The constant-base Chow tests place a break point for the regression of money marginal of interest rates [projection (5.7)] at 1967:III. The Chow tests reject constancy at increasingly high levels for every period thereafter. For the regression of money marginal of prices and interest rates [projection (5.8)], constancy is rejected immediately at 1966:III and at increasingly high levels for every period but one thereafter. The constant-base Chow tests for the other two money regressors also reject stability in 1966:III and thereafter.

Clearly the omission of interest rates is a stabilizing factor. The structural break in three of the regressions at 1966:III suggests that a shock to the interest-rate-determination process is the source. This would clearly need further confirmation from a study of interest-rate regressions, although this is outside the scope of out current study. The institutional record suggests that this is a plausible hypothesis. The Federal Reserve instituted a 'credit crunch' in 1966:III driving interest rates up to the point that regulation Q ceilings on savings and time deposits were binding. The scope of regulation Q ceilings was widened to cover savings and loan deposits. Banks were urged to restrain loan supply. And special rules were adopted for the administration of the discount window. By 1967:II the credit crunch was over, and special arrangements were dismantled [Wojnilower (1980), Board of Governors (1967, 1968)].

The fact that all regressions indicate a structural break by 1967:III suggests that the source of the instability is the money-determination process.[30] It is more difficult to find a likely source for this break in the institutional record.

[30]It will surely occur to many readers that the period of the 'missing money' in which U.S. money stock equations break down is not until 1973, so that the breakdown of the regressions reported here might be simply misspecification. Furthermore, some authors [e.g., Baba et al. (1987) and Rose (1985)] report parameter constancy even across the 1973 period. It should be recalled that the original Goldfeld (1976) missing-money paper as well as Baba et al., Rose, and the survey by Gordon (1984) estimate up through 1973. The most compelling evidence for misspecification rather than structural break would be to show that there exists an encompassing regression that remains stable across the periods that we have identified as structural breaks. Direct comparison of any of these studies with ours is not easy since they all take M1 rather than its demand deposit component to be the dependent variable and they all use information outside of our baseline period in formulating their models. Because of the difficulties of obtaining the precise data and other relevant information, we do not attempt to reproduce all of the studies. But to illustrate that the existence of these 'stable' regressions is not necessarily evidence against our results, consider Rose's (1985, p. 446) eq. 7 and his version of Goldfeld's equation [Rose (1985, p. 441, eq. 2)]. I was unable to obtain the original data from the author and, instead, had to follow the indications in his data appendix and use published sources. The Goldfeld equation shows roughly the right magnitudes on the coefficients and all the right signs, but has a standard error of regression substantially less than that reported by Rose. Rose's own equation shows roughly the correct magnitudes and signs for most regressors, but the cointegration term is positive rather than negative and its standard error of regression is somewhat higher than reported by Rose. [On the difficulties of replication of econometric results, see Dewald, Thursby, and Anderson (1986).] When the Goldfeld model is reestimated over the baseline period, it can be rejected against the corresponding UDL(5) with $F(20,27) = 2.42$. Rose's model cannot be rejected against UDL(5) with $F(17,23) = 1.85$. Direct comparison with out regressions in table 2 is not possible since the dependent variables are different. It is worth noting, however, that both the Goldfeld and Rose regressions have higher percentage standard errors than any of our money regressions, and that both reject stability on the Chow test at 1966:III, like all of the regressions reported in table 4, except money marginal of interest rates. When Chow tests are run for later periods, Goldfeld's equation shows stability up to 1973:IV – i.e., until the period of the missing money; while Rose's is stable up to 1979:II, which is compatible with what Rose himself reports. It therefore appears that both equations provide further evidence of a structural break at 1966:III in equations like our complete money stock, and this despite the fact that Rose's equation appears to be stable across the period of the missing money when its initial estimation period ends in 1973:IV.

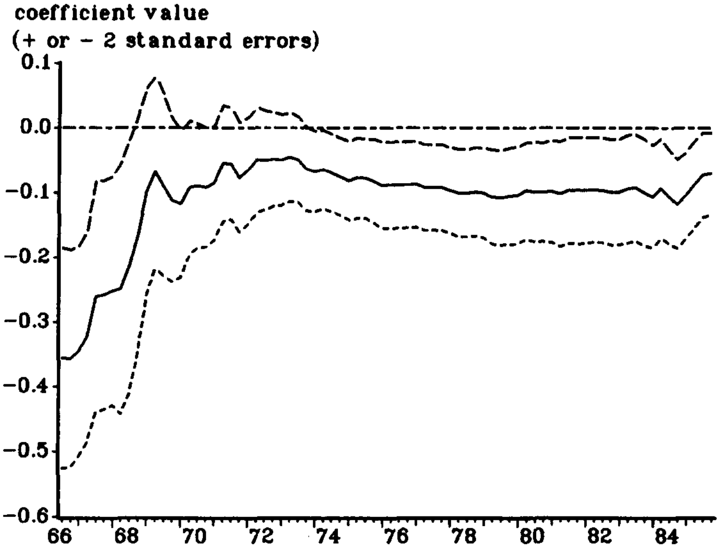

Fig. 10. Coefficient on $(DD - \frac{1}{2}(GNP + CPI))_{-5}$ in regression (2.3) recursively estimated forward from 1966:III to 1985:IV [projection (5.7)].

However, it should be noted that the credit crunch itself was a watershed. Before the 1966:III, regulation Q had been allowed to bind only sporadically, when monetary policy was tight. Afterwards, regulation Q was binding most of the time until it was finally eliminated in the mid-1980's.

The different performance of projections (5.7) and (5.8) before 1967:III favors the view that omitting prices from a money regression is destabilizing. But this is a weak reed. The coefficient plots provide additional evidence. Fig. 10 plots the coefficient on the error-correction term, $(DD - \frac{1}{2}(GNP + CPI))_{-5}$, from the regression of money marginal of interest rates [projection (5.7)]. The plot clearly reflects the structural break common to all the money regressions early on. After 1969:II, however, the coefficient is very stable. Every coefficient but one in projection (5.7) shows analogous behavior. The exception is the coefficient on ΔDD_{-3} (not shown), which again shows the early common structural break but which then drifts toward statistical insignificance over the sample.

In contrast, fig. 11 shows that the coefficient on ΔDD_{-1} from the regression marginal of prices and interest rates [projection (5.3)] drifts around over sample. Casual visual inspection suggests that its most dramatic change occurs in early 1974, about the time of the oil crisis. The plots of the coefficients in projection (5.8) are again, with a single exception, analogous. The coefficient on ΔGNP (not shown) is fairly stable albeit with large

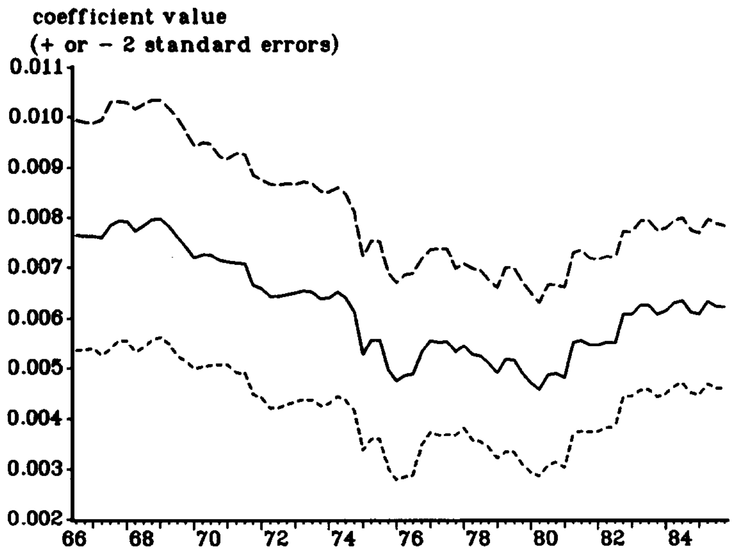

Fig. 11. Coefficient on DD_{-1} in regression (2.4) recursively estimated forward from 1966:III to 1985:IV [projection (5.8)].

standard errors. While, far from dramatic, the evidence favors the view that the omission of prices is destabilizing.

4.6.3. Conclusions

Taking all the evidence together from the two sets of recursions, a consistent and surprising pattern emerges: money regressions marginal of prices are less stable than money regressions conditional on prices; price regressions marginal of money are more stable than price regressions conditional on money. On balance this supports the view that prices cause money, and that money does not cause prices.

The projections of price regressions in both periods before and after the baseline seem clearly to reflect known interventions in the price-determination process. While there appears to be little evidence to discriminate among the price regressions in the earlier period, there is some evidence in the later period to suggest that marginalizing with respect to money is stabilizing at precisely the points at which the money regressions indicate a common intervention. The interaction of interest rates with the other variables in the price regressions appears complicated and potentially confusing. It is certainly worthy of further detailed investigation.

In contrast, the evidence of the money regressions appears more clear-cut: marginalizing with respect to interest rates stabilizes the projections, and further marginalizing with respect to prices destabilizes the projections. This is most clear-cut in the earlier period in which there do not appear to be any interventions in the money-determination process to muddy the waters. In the later period, the balance of evidence clearly supports this view, although the fact that there appears to be a common monetary intervention early in the projection period makes interpretation somewhat more difficult.

4.7. *Does currency cause prices?*

The money variable in out price regressions is the Federal Reserve's M1. Theoretical work by Fama (1980) and empirical work by King and Plosser (1984) suggests that although demand deposits do not influence prices, currency does. If they are correct, using M1 as the money variable may mask the causal role of currency. To check whether we have been misled, the complete price equation was reestimated over the original baseline period (1954:I–1966:II) with the current and five lags of currency as additional regressors. The test of whether or not they are statistically significant is a test of the restriction that currency and demand deposits enter the complete price equation with the same coefficients. With $F(6,33)=0.54$, the null of common coefficients cannot be rejected at any conventional level of significance. It would appear, therefore, that we are correct to treat M1 as a whole and not to separate out the influences of demand deposits and currency. The evidence against money causing prices is thus also evidence against currency causing prices.

5. Summary and conclusions

What economists, statisticians or, for that matter, laymen mean when they ask, does A cause B?, is fraught with difficulty. One thing that seems often to be meant, however, is, were one to make A happen would B also happen? Does control of the supply of money give us control of prices as well? Taking this general approach to causality, we saw how Simon's analysis allows us to characterize causality in a representation of the unobserved data-generating process. We also saw that within a single regime such a characterization is not unique; it suffers from observational equivalence. Fortunately, we were also able to show that if one looks beyond a single regime, if one looks to interventions in the data-generating process, then it is possible to use information on the stability or lack of stability of the different conditional and marginal distributions into which a joint probability distribution can be factored to discriminate between alternative causal orderings.

The relation of stability of regressions to causal ordering provides a general strategy for gathering evidence on causal direction. This strategy was applied to money, prices, and interest rates in the United States from 1950 to

1985 with a somewhat surprising result: the evidence supports the view that prices cause money, and not that money causes prices.

The limited scope of the investigation may be questioned. If only the range of possible causal interactions had been expanded to include more fully the role of interest rates or the role of real variables and so forth, the results might have been different. To this criticism there is no answer except that every study must be circumscribed. In the case of interest rates, for example, to have investigated their role as fully as we have money and prices would have more than doubled the size of the investigation. And while the precise role of interest rates remains unclear, the evidence is that, whether or not interest rates are included, money regressions are more stable with prices as regressors that without, and price regressions are more stable without money as regressors than with. Our main conclusions seem to be robust. The limited nature of the present investigation simply calls for further research along these same lines – a highly desirable thing in itself.

Our results are clearly anti-monetarist. But they are also limited. Monetarism can be thought of as a set of beliefs: the money supply is exogenous and in the control of the central bank; the demand for money is stable; money is neutral in the long run; and the short-run channels connecting money to prices and output are direct and not mediated through financial markets [see Friedman (1974), Friedman and Schwartz (1982, ch. 2)].

This study does not directly address controllability of the money supply; that would require examination of the causal links between the Federal Reserve's instruments – reserves, the discount rate, and the federal funds rate – and the monetary aggregates, such as M1.[31]

Many studies, not just ours, have given reason to doubt that money-demand functions are stable [e.g., Judd and Scadding (1982)]. Our preferred partition of the joint probability distribution of money and prices involves a price regression with money not appearing as a regressor. The preferred money regression includes prices but not interest rates as regressors. Regression (2.3), which omits interest rates, shows a long-run price elasticity with respect to nominal income of only 1/2, rejecting neutrality.

Finally, and most directly, our results point to an absence of a *direct* causal linkage between money and prices, which clearly contradicts strict monetarism. Notice, however, that we have not thoroughly examined the linkages between real GNP and money and prices, and we have only partly examined the linkages between interest rates and money and prices. It is therefore possible to maintain that our results are correct but that money affects prices through indirect channels. Such a position might be called monetarist by some, but it is certainly is not the classic monetarism of Milton Friedman.

[31]Hoover (1985) attempts to address this question with mixed success.

Details aside, the central claim of monetarism is often painted with a broad brush: 'the central fact is that inflation is always and everywhere a monetary phenomenon. Historically, substantial changes in prices have always occurred together with substantial changes in the quantity of money relative to output' [Friedman (1968, pp. 105–106); cf. Lucas (1977, pp. 232–233)]. Although our study covers a limited period and lacks the historical sweep of Friedman and Schwartz's (1963a, 1982) research, it is consistent with the monetarist observation, although not with the monetarist interpretation of that observation. Our preferred causal ordering (prices cause money, money does not cause prices) would generate a consistently high correlation between money and prices in the long run. Everyone, including Friedman, recognizes that one cannot infer causal direction from such a correlation. Friedman and Schwartz examined the consistently high correlation of money and prices across large institutional changes. But, hitherto no one has examined the *relative* stability of alternative conditional and marginal distributions, which provides a basis for discriminating between a high correlation due to money causing prices and one due to prices causing money.

Our results partly, although not completely, support the views of some advocates of real business cycles. Like them, we find that prices (nominal income) do not cause demand deposits (inside money). Unlike them, we do not find evidence that currency (outside money) causes prices.

The empirical results presented here are substantive. They are also, and perhaps more importantly, an illustration of a method. Clearly, causal linkages among economic variables require further investigation. The evidence of this research is that there is a fruitful way to proceed.

References

Akaike, H., 1973, Information theory and an extension of the maximum likelihood principle, in: B.N. Petrov and F. Csaki, eds., Second international symposium on information theory (Akademiai Kiade, Budapest).

Andrews, D.W.K., 1989, Tests of parameter instability and structural change with an unknown change point, Unpublished typescript (Yale University, New Haven, CT).

Baba, Yoshihisa, David F. Hendry, and Ross M. Starr, 1987, U.S. money demand, 1960–1984, Discussion papers in economics no. 27, Dec. (Nuffield College, Oxford).

Banerjee, Anindya, Robin L. Lumsdaine, and James H. Stock, 1990, Recursive and sequential tests of the unit root hypothesis: Theory and international evidence, NBER working paper no. 3510, Nov.

Basmann, R.L., 1965, A note on the statistical testability of 'explicit causal chains' against the class of 'interdependent' models, Journal of the American Statistical Society 60, 1080–1093.

Basmann, R.L., 1988, Causality tests and observationally equivalent representations of econometric models, Journal of Econometrics 39, 69–104.

Black, Fischer, 1970, Banking and interest rates in a world without money, Ch. 1 in: Fischer Black, 1987, Business cycles and equilibrium (Basil Blackwell, Oxford).

Black, Fischer, 1972, Active and passive monetary policy in a neoclassical model, Ch. 2 in: Fischer Black, 1987, Business cycles and equilibrium (Basil Blackwell, Oxford).

Blanchard, Olivier J., 1984, The Lucas critique and the Volcker deflation, American Economic Review 74, 211–215.

Board of Governors of the Federal Reserve System, 1967–1968, Annual report (Washington, DC).
Box, G.E.P. and D.A. Pierce, 1970, Distributions of residual autocorrelations in autoregressive integrated moving average models, Journal of the American Statistical Association 65, 1509–1526.
Brown, R.L., J. Durbin, and J.M. Evans, 1975, Techniques for testing the constancy of regression relationships over time, Journal of the Royal Statistical Society B 37, 149–192 (with discussion).
Cagan, Philip, 1965, The determinants of the money stock in the United States 1875–1963 (National Bureau of Economic Research, New York, NY).
Cai, Yongxin, 1990, The practical importance of the Lucas critique: An empirical study of the invariance of structural parameters under policy regime changes, Unpublished Ph.D. prospectus, Sept. (University of California, Davis, CA).
Chow, Gregory C., 1960, Tests of equality between sets of coefficients in two linear regressions, Econometrica 28, 591–605.
Chu, Chia-Shang James, 1990, Test for parameter constancy in stationary and nonstationary regression models, Unpublished typescript, Revised Jan. (University of California, San Diego, CA).
Council of Economic Advisers, 1950–1953, The economic situation at mid-year (Government Printing Office, Washington, DC).
Dewald, William G., Jerry G. Thursby, and Richard G. Anderson, 1986, Replication in empirical economics: The Journal of Money, Credit and Banking project, American Economic Review 76, 587–603.
Engle, Robert, E., 1982, Autoregressive conditional heteroskedasticity, with estimates of the variance of United Kingdom inflations, Econometrica 50, 987–1008.
Engle, Robert F., David F. Hendry, and Jean-Francois Richard, 1983, Exogeneity, Econometrica 51, 277–303.
Fama, Eugene, 1980, Banking in the theory of finance, Journal of Monetary Economics 6, 39–57.
Fisher, Irving, 1911/1931, The purchasing power of money (Macmillan, New York, NY).
Friedman, Milton, 1968, Dollars and deficits: Inflation, monetary policy and the balance of payments (Prentice Hall, Englewood Cliffs, NJ).
Friedman, Milton, 1970, Reply to Tobin, Quarterly Journal of Economics 84, 318–327.
Friedman, Milton, 1974, A theoretical framework for monetary analysis, in: Robert J. Gordon, ed., Milton Friedman's monetary framework (University of Chicago Press, Chicago, IL).
Friedman, Milton and Anna J. Schwartz, 1963a, A monetary history of the United States (Princeton University Press, Princeton, NJ).
Friedman, Milton and Anna J. Schwartz, 1963b, Money and business cycles, Reprinted in: Milton Friedman, 1969, The optimum quantity of money and other essays (Macmillan, London).
Friedman, Milton and Anna J. Schwartz, 1982, Monetary trends in the United States and the United Kingdom: Their relation to income, prices and interest rates 1867–1975 (University of Chicago Press, Chicago, IL).
Gilbert, Christopher L., 1986, Professor Hendry's econometric methodology, Oxford Bulletin of Economics and Statistics 48, 283–307.
Goldfeld, Stephen M., 1976, The case of the missing money, Brookings Papers on Economic Activity 3, 683–730.
Gordon, Robert J., 1984, The short-run demand for money: A reconsideration, Journal of Money, Credit and Banking 16, 403–434.
Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, reprinted in: Robert E. Lucas, Jr. and Thomas J. Sargent, eds., 1981, Rational expectations and econometric practice (George Allen and Unwin, London).
Granger, C.W.J., 1980, Testing for causality: A personal viewpoint, Journal of Economic Dynamics and Control 2, 329–352.
Gurley, John and Edwin Shaw, 1960, Money in a theory of finance (Brookings, Washington, DC).
Hansen, Bruce I., 1990, Testing for structural change of unknown form in models with non-stationary regressors, Unpublished typescript, Revised March (University of Rochester, Rochester, NY).

Harvey, A.C., 1981, The econometric analysis of time series (Philip Allan, Deddington, Oxford).
Hendry, David F., 1983, Econometric modelling: The 'consumption function' in retrospect, Scottish Journal of Political Economy 30, 193–220.
Hendry, David F., 1987, Econometric methodology: A personal perspective, Ch. 10 in: Truman F. Bewley, ed., Advances in econometrics, Vol. 2 (Cambridge University Press, Cambridge).
Hendry, David F., 1988a, The encompassing implications of feedback versus feedforward mechanisms in econometrics, Oxford Economic Papers N.S. 40, 132–149.
Hendry, David F., 1988b, Encompassing, National Institute of Economic Review, Aug., 88–92.
Hendry, David F. and Jean-Francois Richard, 1982, On the formulation of empirical models in dynamic econometrics, Journal of Econometrics 20, 3–33.
Hendry, David F. and Jean-Francois Richard, 1987, Recent developments in the theory of encompassing, in: B. Cornet and H. Tulkens, eds., Contributions to operation research and econometrics, The XXth anniversary of CORE (MIT Press, Cambridge, MA).
Hoover, Kevin D., 1985, Causality and invariance in the money supply process, Unpublished D. Phil dissertation, Trinity term (Oxford University, Oxford).
Hoover, Kevin D., 1988a, The new classical macroeconomics: A sceptical inquiry (Basil Blackwell, Oxford).
Hoover, Kevin D., 1988b, The logic of causal inference: With an application to money and prices, Research program in applied macroeconomics and macro policy, Working paper no. 55, Aug. (University of California, Davis, CA).
Hoover, Kevin D., 1988c, Money, prices and finance in the new monetary economics, Oxford Economic Papers N.S. 40, 150–167.
Hoover, Kevin D., 1988d, On the pitfalls of untested common-factor restrictions: The case of the inverted Fisher hypothesis, Oxford Bulletin of Economics and Statistics 50, 125–138.
Hoover, Kevin D., 1989, Econometrics as observation: The Lucas critique and the nature of econometric inference, Forthcoming in: Ingrid Rima, ed., Measurement, quantification and economic analysis.
Hoover, Kevin D., 1990, The logic of causal inference: Econometrics and the conditional analysis of causation, Economics and Philosophy 6, 207–234.
Jarque, C.M. and A.K. Berra, 1980, Efficient tests for normality, homoscedasticty and serial independence of regression residuals, Economics Letters 6, 255–259.
Judd, John P. and John L. Scadding, 1982, The search for a stable money demand function: A survey of the post-1973 literature, Journal of Economic Literature 20, 993–1024.
Kaldor, Nicholas, 1982, The scourge of monetarism (Oxford University Press, Oxford).
Katzner, Donald W., 1983, Analysis without measurement (Cambridge University Press, Cambridge).
Kim, H.J. and D. Siegmund, 1989, The likelihood ratio test for a change point in simple linear regression, Biometrika 76, 409–423.
King, Robert G. and Charles I. Plosser, 1984, Money, credit and prices in a real business cycle, Americal Economic Review 74, 363–380.
Kydland, Finn E. and Edward C. Prescott, 1982, Time to build and aggregate fluctuations, Econometrica 50, 1345–1369.
Leamer, Edward E., 1978, Specification searches: Ad-hoc inference with nonexperimental data (Wiley, New York, NY).
Ljung, G.M. and G.E.P. Box, 1978, On a measure of lack of fit in time series models, Biometrika 65, 297–303.
Long, John B., Jr. and Charles I. Plosser, 1983, Real business cycles, Journal of Political Economy 91, 39–69.
Lucas, Robert E., Jr., 1977, Understanding business cycles, in: Robert E. Lucas, Jr., 1981, Studies in business-cycle theory (Basil Blackwell, Oxford) 215–239.
McAleer, Michael, Adrian R. Pagan, and Paul A. Volker, 1985, What will take the con out of econometrics?, Americal Economic Review 75, 292–307.
Mackie, J.L., 1980, The cement of the university: A study in causation (Clarendon Press, Oxford).
Mesarovic, Mihajlo D., 1969, Mathematical theory of general systems and some economic problems, in: H.W. Kuhn and G.P. Szegö, eds., Mathematical systems and economics I (Springer-Verlag, Berlin) 93–116.

Miller, Preston J., 1983, Higher deficit policies lead to inflation, Federal Reserve Bank of Minneapolis Quarterly Review 7, 8–19.
Mizon, Grayham E. and Jean-Francois Richard, 1986, The encompassing principle and its application to testing non-nested hypotheses, Econometrica 54, 657–678.
Neftci, Salih and Thomas J. Sargent, 1978, A little bit of evidence on the natural rate hypothesis from the U.S., Journal of Monetary Economics 4, 315–320.
Pagan, Adrian, 1987, Three econometric methodologies: A critical appraisal, Journal of Economic Surveys 1, 3–24.
Phillips, Peter C.B., 1988, Reflections on econometric method, Economic Record 64, 344–359.
Quandt, Richard, 1960, Tests of the hypothesis that a linear regression system obeys two separate regimes, Journal of the American Statistical Association 55, 324–330.
Romer, Christina D. and David H. Romer, 1989, Does monetary policy matter? A new test in the spirit of Friedman and Schwartz (with discussion), in: Olivier Jean Blanchard and Stanley Fischer, eds., NBER macroeconomics annual 1989 (MIT Press, Cambridge, MA).
Rose, Andrew K., 1985, An alternative approach to the American demand for money, Journal of Money, Credit and Banking 17, 439–455.
Sargent, Thomas J., 1976, The observational equivalence of natural and unnatural rate theories of macroeconomics, Reprinted in: Robert E. Lucas, Jr. and Thomas J. Sargent, 1981, Rational expectations and econometric practice (George Allen and Unwin, London) 553–562.
Sargent, Thomas J., 1984, Autoregressions, expectations and advice, American Economic Review 74, 408–415.
Schwarz, G. 1978, Estimating the dimension of a model, Annals of Statistics 6, 461–464.
Simon, Herbert A., 1953, Causal ordering and identifiability, Reprinted in: Herbert A. Simon, 1957, Models of man (Wiley, New York, NY).
Simon, Herbert A. and Yumi Iwasaki, 1988, Causal ordering, comparative statics and near decomposability, Journal of Econometrics 39, 149–173.
Sims, Christopher A., 1980, Macroeconomics and reality, Econometrica 48, 1–48.
Sims, Christopher A., 1982, Policy analysis with econometric models, Brookings Papers on Economic Activity 1, 107–152.
Sims, Christopher A., 1986, Are forecasting models usable for policy analysis?, Federal Reserve Bank of Minneapolis Quarterly Review 10, 2–15.
Spanos, Aris, 1986, Statistical foundations of econometric modelling (Cambridge University Press, Cambridge).
Tobin, James, 1970, Money and income: Post hoc ergo proper hoc?, Quarterly Journal of Economics 84, 301–317.
Wojnilower, Albert M., 1980, The central role of credit crunches in recent financial history, Brookings Papers on Economic Activity 2, 277–326.
Zellner, Arnold A., 1979, Causality and econometrics, in: Karl Brunner and Allan H. Meltzer, eds., 1979, Three aspects of policy making: Knowledge, data and institutions, Carnegie–Rochester Conference Series on Public Policy 10, 9–54.