Journal of Econometrics 40 (1989) 161–181. North-Holland

INTERPRETING THE EVIDENCE ON MONEY-INCOME CAUSALITY*

James H. STOCK

Harvard University, Cambridge, MA 02138, USA

Mark W. WATSON

Northwestern University, Evanston, IL 60208, USA

Previous researchers have reached strikingly different conclusions about the usefulness of money for forecasting real output based on closely related regression-based tests. An examination of this and additional new evidence reveals that innovations in M1 have statistically significant marginal predictive value for industrial production, both in a bivariate model and in a multivariate setting including a price index and an interest rate. This conclusion follows from focusing on the trend properties of the data, both stochastic and deterministic, and from drawing inferences using asymptotic theory that explicitly addresses the implications of these trends for the distributions of the various test statistics.

1. Introduction

Whether movements in money help to predict movements in income has been an enduring topic in empirical macroeconomics. This question has traditionally been associated with the monetarist–Keynesian debate over the effectiveness of monetary policy. Recently, however, some authors have suggested that equilibrium models in which aggregate movements have real rather than monetary origins imply that money should have limited value for forecasting income or output [King and Plosser (1984), Bernanke (1986), and Eichenbaum and Singleton (1986)]. Since Sims' (1972, 1980a) seminal articles, empirical investigations into this relation have focused on whether money 'Granger causes' output in either a bivariate or a multivariate framework.

Unfortunately, researchers using only slightly different specifications have reached disconcertingly different conclusions using postwar U.S. data. Focusing on systems with money, output, prices and interest rates, these different findings can be summarized as three puzzles. First, in a three-variable system with log output, log money and inflation and in a four-variable system in

*The authors are grateful to Ben Bernanke for encouragement in undertaking this project and to L. Christiano, L. Hansen, L. Ljungqvist, C. Sims and two anonymous referees for helpful discussions and comments on an earlier draft. This research was supported in part by NSF Grants SES-84-08797 and SES-85-10289.

which an interest rate is added, Eichenbaum and Singleton (1986) found sharp reductions in the importance of money when the tests were performed using log differences of the variables rather than log levels with a time trend.

Second, Sims (1980b) found that adding an interest rate to the three-variable system specified in levels (log output, log money and log price) reduced the fraction of the variation in output explained by innovations in money. As Bernanke (1986) emphasized, however, when a time trend is added to this regression, King (1984) and Runkle (1987) (using a linear time trend) and Litterman and Weiss (1985) (using a quadratic time trend as well) report a substantial increase in this fraction, indicating an increased predictive role for money in the four-variable system.

Third, the conclusions are sensitive to the sample period. For example, Eichenbaum and Singleton (1986) emphasize that the evidence supporting a causal role for money in a detrended levels specification is much weaker in a sample that excludes the 1980's than if the 1980's are included [for a further discussion see King (1984, app. B), Christiano and Eichenbaum (1987) and Stock (1987a)].

At the heart of the first two puzzles is the use of different techniques to 'detrend' time series that arguably contain unit roots. Christiano and Ljungqvist (1987) and Ohanian (1986) have recently provided Monte Carlo evidence that non-standard 'unit root' distributions might play a role in resolving these puzzles. Specifically, Christiano and Ljungqvist (1987) compute the distribution of a levels causality test statistic in a bivariate model of money and output under the assumption that money and output are not cointegrated, but have unit roots and drifts; they find a substantial rightward shift in the distribution of the F-statistic under the null of no causality. Ohanian (1986) reports similar 'unexpected' distributional shifts in a Monte Carlo experiment using a vector autoregression (VAR) including money, output, prices, an interest rate and an artificially generated independent random walk. While these authors offer no theoretical explanation for their Monte Carlo results, their findings suggest that non-standard distribution theory may prove important in reconciling the empirical puzzles.

Our objective is to develop an explicit empirical characterization of money, output, prices and interest rates that resolves these three puzzles and permits us to draw conclusions concerning the money-output relation in the postwar U.S. To facilitate the application of existing asymptotic distribution theory, we focus on Granger causality tests rather than variance decompositions.

Our main finding is that although the growth of M1 itself does not Granger cause growth in the industrial production index, the deviation of money growth from a linear time trend does. This result obtains in a bivariate system with money and output, in a trivariate system in which we add prices and in a four-variable system in which we add interest rates. Moreover, the key restriction that appears to be rejected is the long-run sensitivity of output to changes in *detrended* money growth. This conclusion is robust to changes in the sample period, to estimation using two-, three- and four-variable specifications, and to changes in the number of lags in the relevant regressions.

The analysis focuses on the effects of different detrending techniques when the regressors contain some unit roots. Our first step, reported in section 2, is to ascertain the orders of integration and cointegration of the variables, as well as whether the series appear to have polynomial time trends. We use monthly data from January 1959 to December 1985, with industrial production as the output measure.[1] Consistent with the findings of other researchers, over the postwar period in the U.S. each of these variables appears to contain one unit root and possibly a deterministic drift, so that (for example) output growth is stationary. In addition, money growth (Δm_t, where m_t denotes the log of nominal M1) seems to be well described as being stationary around a small but statistically significant time trend. That is,

$$\Delta m_t = \gamma_1 + \gamma_2 t + \Delta\mu_t, \tag{1}$$

where $\Delta\mu_t$ is a stationary, mean zero process. Thus $\Delta\mu_t$ can be thought of as 'detrended nominal money growth'; alternatively, μ_t is the *stochastically* trending component of money. Using the results of this analysis, in section 3 we provide direct evidence that $\Delta\mu_t$ Granger causes Δy_t. Moreover, the growth in detrended money is found to be non-neutral, in the sense that the sum of the coefficients on lagged $\Delta\mu_t$ in the Δy_t equations are statistically significant and positive.

In section 4, the characterization of section 2 and the theoretical results of Sims, Stock and Watson (1986) are used to address the first two puzzles. They show that the usual Gaussian asymptotic distribution theory can be used to interpret F-tests in two cases: (i) the restrictions being tested can be expressed as restrictions on mean zero, stationary variables, or (ii) the restrictions involve some variables that are dominated by a polynomial in time (and perhaps some additional variables that are mean zero and stationary), but there are no other linear combinations of regressors that are dominated by *stochastic* trends. Otherwise, the F-statistics will have non-standard distributions. In particular, these results imply that causality tests computed from regressions involving levels of the data typically will have non-standard distributions.

These theoretical results are applied to the different specifications in the literature. When they are warranted, we use the usual p-values, predicated upon asymptotic normality; when they are not, the p-values are computed by numerical integration of the non-standard asymptotic distributions. The pro-

[1]1959 was chosen as the start of the sample because of the change in definitions of the monetary aggregates that year. While most other researchers have used data from the 1950's and late 1940's, these earlier series involve unofficial retrospectively constructed data that presumably contains additional measurement error.

posed resolution of the first two puzzles involves both these distributional issues and the finding of sections 2 and 3 that detrended money growth is non-neutral with respect to output growth. Thus the differences in conclusions drawn from the various specifications arises partly because of previously incorrect asymptotic inference in the levels regressions, and partly because including time as a regressor allows the dominant component of money growth (its time trend) to be stripped away, revealing the stochastic component ($\Delta\mu_t$) that has predictive value for industrial production.

The final puzzle is addressed in section 5 by repeating the analysis of sections 2–4 for a shorter sample ending in September 1979. In this period inflation is better characterized as being stationary around a time trend, accounting for many of the observed changes in the causality F-statistics. While the neutrality proposition is rejected for this shorter sample, upon controlling for the predictive value of the Treasury bill rate $\Delta\mu_t$ appears to have less predictive content for industrial production in the shorter than in the longer sample. Our conclusions are summarized in section 6.

Before proceeding, we emphasize that our use of the terms 'causality' and 'neutrality' is narrow. By causality we refer only to certain zero restrictions in a VAR: specifically, that lagged values of money do not enter the income equation. Similarly, by neutrality we mean that the *sum* of the coefficients on lagged money – or (detrended) money growth, depending on the specification – equals zero in the output regressions. This mechanical (and conventional) definition of neutrality means that the innovations in the money equation in the VAR will have no long-run effect on output. There are at least three reasons why non-neutrality in a VAR need not imply the non-neutrality of monetary policy. First, as discussed (for example) by Blanchard and Watson (1986), VAR innovations are one-step-ahead forecast errors, which in general will be combinations of underlying structural innovations. Second, in keeping with the efforts of earlier researchers, our examination focuses on M1 – which includes both inside and outside money – thereby weakening the link to monetary policy. Finally, we consider systems with restricted information sets, which arguably will obscure the link between structural disturbances and VAR residuals.[2]

2. Trend properties of the data

Sims, Stock and Watson (1986) show that the asymptotic distributions of neutrality and causality tests are sensitive to unit roots and time trends in the series. This suggests that the finite-sample distributions of these tests will depend on these trending and 'unit root' characteristics as well. Since this

[2]See Geweke (1984, 1986a) for further discussions of causality, predictability, exogeneity and neutrality.

asymptotic theory will be used to approximate these sampling distributions, we begin by characterizing the time trend and unit root properties of the data. The data – representative of those used by earlier researchers – consist of monthly observations on the log of seasonally adjusted nominal M1 (m), the log of industrial production (y), the log of the wholesale price index (p) (not seasonally adjusted), and the secondary market rate on 90-day U.S. Treasury bills (r), obtained from the Citibase data base. Stated sample lengths refer to the estimation period, with earlier observations being used for initial conditions as necessary.

It is now widely recognized that many macroeconomic series appear to contain unit roots [e.g., Nelson and Plosser (1982), Perron (1986), Perron and Phillips (1986), Stock and Watson (1986a, b)]. Table 1 investigates the possibility that these series might have up to two unit roots and time trends up to second order. The first column presents a modification of the Stock–Watson (1986a) q_f-test (described in appendix A) for a single unit root when there might be a quadratic time trend; in no case is there significant evidence against the unit root hypothesis. The next two columns present the results of Stock–Watson (1986a) and Dickey–Fuller (1979) tests for a second unit root, i.e., for a unit root in the first difference of the series, allowing for the alternative that the series is stationary in first differences around a linear time trend. These tests suggest that no series contains two unit roots.[3]

To ascertain the order of the deterministic components in the specification, the first difference of each series was regressed against a constant, time, and six of its own lags (the latter to obtain correct standard errors); the t-statistics on the time trend are reported in the sixth column of table 1. Money growth exhibits clear evidence of a deterministic trend. Performing the same test without a time trend (column 7) indicates that, while output and prices have significant drifts, the interest rate does not.

Since output growth, inflation and the first difference of the interest rate appear not to exhibit time trends, the fourth and fifth columns of table 1 reports tests for a unit root in levels, allowing only for the process to be stationary around a linear time trend under the alternative. Again, all tests fail to reject the null of a unit root. We conclude that, over the 1960:2–1985:12 sample, y and p are well described as processes having a single unit root with

[3]The tests in table 1 were repeated using twelve rather than six lags. Similar results obtained, although with twelve lags it was more difficult to reject the null that inflation has a second unit root. One potential problem with inferences concerning inflation is that wholesale prices exhibit substantial heteroskedasticity during the 1973–1975 oil price rise and the lifting of price controls. We thus computed modified, weighted Dickey–Fuller and Stock–Watson test statistics in which this heteroskedasticity was modeled using a dummy variable for 1973:1–1975:1. With both six and twelve lags, the weighted tests more strongly rejected the hypothesis that p has a second unit root: the largest root in inflation, after adjusting for heteroskedasticity and short-run serial correlation, was estimated to be 0.83. Thus inflation appears not to be integrated over the sample, although it exhibits substantially greater dependence than (say) output growth.

Table 1

Tests for integration, cointegration and time trends, 1960:2–1985:12.[a]

(A) *Univariate tests*

	Unit root tests					t-statistics for a regression of Δx on:	
Series	$q_f^{T^2}[z]$	$q_f^T[\Delta z]$	D-F $\hat{\tau}_\tau[\Delta z]$	$q_f^T[z]$	D-F $\hat{\tau}_\tau[z]$	time	constant
y	−16.10	−166.0**	−5.36**	−10.50	−2.16	−0.52	2.27*
m	−23.52	−242.3**	−6.54**	−3.14	−1.40	4.25**	4.57**
p	−2.28	−232.6**	−3.37*	−6.27	−2.17	0.83	2.19*
r	−11.47	−215.5**	−8.67**	−11.44	−1.97	−0.38	0.36

B) *Multivariate tests*

System	Unit roots under null and alternative	$q_f^{T^2}$
m, y	2 vs. 1	−22.4
m, y, p	3 vs. 2	−16.7
m, y, p, r	4 vs. 3	−15.4

[a]Significant at the 1% (**), 5% (*) and 10% ($^+$) level. All statistics are based on regressions with six lags. $q_f^{T^2}[z]$ denotes the Stock–Watson (1986a) $q_f^{T^2}(1,0)$-statistic computed using the level of each variable; $q_f^T[\Delta z]$ denotes the $q_f^T(1,0)$-statistic computed using the first difference of each variable; $\hat{\tau}_\tau[\Delta z]$ denotes the Dickey–Fuller (1979) t-statistic computed using the first difference of x; and similarly for $q_f^T[z]$ and $\hat{\tau}_\tau[z]$. Critical values for the $\hat{\tau}_\tau$-statistic are from Fuller (1976, p. 373); for the $q_f^T(1,1)$-statistic from Stock and Watson (1986a); and for the $q_f^{T^2}(k, k-1)$-statistics from appendix A.

drift; r has a single unit root without drift; and m has the univariate representation (1).

Multivariate representations of these data will thus be characterized by a linear time trend and at most as many unit roots as variables. However, if the series have stochastic trends in common, their multivariate representations will be cointegrated as defined by Engle and Granger (1987) and so will contain a reduced number of unit roots. However, tests for cointegration among these data (reported in panel B of table 1) fail to reject the null that the series are not cointegrated, suggesting that the multivariate specifications have as many unit roots as variables.[4]

Summarizing, while the power of unit roots and cointegration tests can be low, the foregoing discussion suggests the specification:

$$\begin{aligned}
\Delta y_t &= \alpha_{y0} + \phantom{\alpha_{m1}t +{}} \Delta\eta_t,\\
\Delta m_t &= \alpha_{m0} + \alpha_{m1}t + \Delta\mu_t,\\
\Delta p_t &= \alpha_{p0} + \phantom{\alpha_{m1}t +{}} \Delta\pi_t,\\
\Delta r_t &= \phantom{\alpha_{m0} + \alpha_{m1}t +{}} \Delta\rho_t,
\end{aligned}$$

[4]Similar results obtain using twelve rather than six lags to construct the cointegration tests.

where $\Delta\eta_t$, $\Delta\mu_t$, $\Delta\pi_t$, and $\Delta\rho_t$ are mean zero stationary processes. Letting x_t denote (y_t, m_t) in the bivariate case, (y_t, m_t, p_t) in the trivariate case, and (y_t, m_t, p_t, r_t) in the four-variable case, letting ξ_t denote the corresponding subvector of $(\eta_t, \mu_t, \pi_t, \rho_t)$, and assuming that $\Delta\xi_t$ has a VAR(p) representation, this system can be rewritten as

$$\Delta x_t = \alpha_0 + \alpha_1 t + \Delta\xi_t, \qquad A(L)\Delta\xi_t = \varepsilon_t, \tag{2}$$

where ε_t is a vector of innovations. Thus

$$A(L)\Delta x_t = \gamma_0 + \gamma_1 t + \varepsilon_t \quad \text{or} \quad \tilde{A}(L)x_t = \gamma_0 + \gamma_1 t + \varepsilon_t, \tag{3}$$

where $\gamma_0 + \gamma_1 t = A(L)(\alpha_0 + \alpha_1 t)$ and $\tilde{A}(L) = (1 - L)A(L)$. Thus the VAR representations in (3) summarize the unit root and time trend results of table 1. Clearly the restriction that $\Delta\mu_t$ not enter the $\Delta\eta_t$ equation in (2) is the same as Δm_t not entering the output equations in (3).

3. Evidence of non-neutrality

This section provides empirical evidence for the proposition that industrial production is not neutral to changes in the growth rate of detrended money. To do so, we examine three alternative regressions of output growth on lagged output and money growth, each differing in their treatment of the apparent trend in money growth. Consider first the output equation in a bivariate VAR(p) of money growth and output growth without a time trend:

$$\Delta y_t = \beta_{y1} + \beta_{ym}(L)\Delta m_{t-1} + \beta_{yy}(L)\Delta y_{t-1} + \varepsilon_t. \tag{4}$$

The usual test for Granger causality tests the p restrictions that the coefficients on lagged Δm_t are zero, i.e., that $\beta_{ym}(L) = 0$. One of the restrictions being tested is that money is 'neutral' in the sense that $\beta_{ym}(1) \equiv \sum_{i=1}^{p}\beta_{ymi} = 0$. To focus on this restriction, note that $\beta_{ym}(L) = \beta_{ym}(1) + \beta^*_{ym}(L)(1 - L)$ (where $\beta^*_{ymj} = -\sum_{i=j+1}^{p}\beta_{ymi}$); thus (4) can be rewritten as

$$\text{(I)} \qquad \Delta y_t = \beta_{y1} + \beta_{ym}(1)\Delta m_{t-1} + \beta^*_{ym}(L)\Delta^2 m_{t-1} + \beta_{yy}(L)\Delta y_{t-1} + \varepsilon_t,$$

where $\Delta^2 \equiv (1 - L)^2$. Thus a t-test on the coefficient on Δm_{t-1} in (I) provides a simple test of the hypothesis that $\beta_{ym}(1) = 0$.

Since money growth contains a deterministic time trend, a test of $\beta_{ym}(1) = 0$ in (I) does not provide a direct test of the neutrality of *detrended* money growth. This alternative hypothesis can be examined in a regression in which Δm_t is replaced by the residual from regressing Δm_t on a constant and time; in large samples, this should reflect more closely the effect of $\Delta\mu_t$ on Δy_t.

Table 2

Neutrality and causality tests, 1960:2–1985:12.[a]

(A) *Neutrality tests*: $\hat{\beta}_{ym}(1)$ *and t-statistics*

Specification	Six lags of money			Twelve lags of money		
	y, m	y, m, p	y, m, p, r	y, m	y, m, p	y, m, p, r
(I) Const; Δm_t	0.63**	0.72**	0.69**	0.46+	0.73**	0.67*
	(3.29)	(3.71)	(3.07)	(1.93)	(2.90)	(2.48)
(II) Const; $\Delta\hat{\mu}_t$	1.21**	1.00**	1.12*	1.53**	1.62**	1.48**
	(4.77)	(4.28)	(3.66)	(3.80)	(3.93)	(3.32)
(III) Const, t; Δm_t	1.22**	1.09**	1.12**	1.55**	1.61**	1.48**
	(4.78)	(4.22)	(3.64)	(3.82)	(3.92)	(3.31)

(B) *Causality tests*: *F-statistics and p-values*

Specification	Six lags of money			Twelve lags of money		
	y, m	y, m, p	y, m, p, r	y, m	y, m, p	y, m, p, r
(I) Const; Δm_t	2.65	2.83	2.39	1.63	1.52	1.46
	(0.016)	(0.011)	(0.029)	(0.082)	(0.115)	(0.141)
(II) Const; $\Delta\hat{\mu}_t$	4.66	3.59	3.06	2.57	2.13	1.88
	(0.000)	(0.002)	(0.007)	(0.003)	(0.015)	(0.037)
(III) Const, t; Δm_t	4.67	3.50	3.04	2.58	2.09	1.87
	(0.000)	(0.002)	(0.007)	(0.003)	(0.018)	(0.038)

[a] t-statistics in panel A are significant at the 1% (**), 5% (*) and 10% (+) level. All statistics are based on regressions with p lags of money growth rearranged as in (4) in the text, and with twelve lags of Δy, Δp and Δr as appropriate, where $p = 6$ in the first panel and $p = 12$ in the second panel. In panel A, t-statistics are in parentheses, and in panel B p-values are in parentheses. Significance levels and p-values are based on the usual asymptotic normal and F approximations.

Denoting detrended money growth by $\Delta\hat{\mu}_t$, one could consider

$$\text{(II)} \qquad \Delta y_t = \beta_{y1} + \beta_{ym}(1)\Delta\hat{\mu}_{t-1} + \beta^*_{ym}(L)\Delta^2 m_{t-1} + \beta_{yy}(L)\Delta y_{t-1} + \varepsilon_t.$$

Finally, a third approach to testing whether $\Delta\mu_t$ Granger causes Δy_t would be to detrend all the regressors in (I). This could be done either in a series of preliminary regressions or by estimating

$$\text{(III)} \qquad \Delta y_t = \beta_{y1} + \beta_{y2}t + \beta_{ym}(1)\Delta m_{t-1} + \beta^*_{ym}(L)\Delta^2 m_{t-1} + \beta_{yy}(L)\Delta y_{t-1} + \varepsilon_t.$$

The results of estimating (I)–(III) are presented in table 2. According to the analysis of section 2, none of the regressors contain unit roots, although in (I) and (III) the regressors include variables that are dominated by linear time trends. Thus the usual standard errors and procedures of asymptotic inference are valid and are used in the table.

Focusing first on the bivariate relation, the neutrality proposition is rejected in all three specifications. However, in the regressions with detrended money the point estimate of $\beta_{ym}(1)$ more than doubles, and the t-statistic rises sharply. When the Granger causality proposition is tested using the specifications (I)–(III), the predictive role of money is substantially enhanced using the detrended data. The increase in the F-statistic from (I) to (II) in panel B is consistent with the proposition that much of the power of the causality test arises from the non-neutrality of $\Delta\mu_t$, which is tested only in specifications (II) and (III) (at least asymptotically).

The results for the three- and four-variable systems are similar: detrended money growth exhibits stronger non-neutrality, and thus more evidence of causality, than does money growth itself.[5]

In all systems, $\hat{\beta}_{ym}(1)$ rises somewhat upon increasing the number of lags of money growth from six to twelve, suggesting a small but positive sum of the coefficients on $\Delta\mu_t$ for lags 7–12. The corresponding t-statistics are slightly reduced, but for the detrended specifications they typically remain significant at the 1% level. In contrast, increasing the money lag length sharply reduces the Granger causality F-tests (although all detrended specifications still reject at the 5% level). This is consistent with the proposition that increasing the number of lags of money vitiates the influence of the non-neutrality on the test statistic.[6]

4. The puzzling causality tests

We return to the first two puzzles stated in the introduction by examining the effects of different treatments of time trends and unit roots on F-tests of the predictive content of money. Table 3 contains F-tests for the two-, three- and four-variable systems for specifications using growth rates and log levels, using different number of lags and different orders of deterministic time trends.

The Sims–Stock–Watson (1986) guidelines can be used to ascertain which statistics in table 3 have the usual asymptotic distribution. The test results in

[5]Similar point estimates for $\beta_{ym}(1)$ (in the range 0.9 to 1.4) obtain when these equations are re-estimated with the data aggregated to the quarterly level. The t-statistics are smaller, although significant at the 10% level.

[6]As an additional specification check, these regressions were re-estimated using 24 lags of all regressors; similar estimates of $\beta_{ym}(1)$ obtained. With 36 lags the point estimates dropped to near zero, even though the hypothesis that these additional lags were zero could not be rejected at the 50% level. This low estimate of $\beta_{ym}(1)$ is consistent with Geweke's (1986a) finding of neutrality (in our sense) in comparable two-, three- and four-variable systems using a VAR(36) spectral estimator. While Berk (1974) proved the pointwise consistency and asymptotic normality of univariate AR(p_T) spectral estimators for $p_T \to \infty$, we interpret his condition that $\lim_{T\to\infty} p_T^3/T = 0$ as suggesting that spectral estimators at frequency zero will tend to be more precise using short (e.g., $p = 6$–12) rather than long (e.g., $p = 36$) autoregressions.

Table 3

Money–income causality tests, 1960:2–1985:12.[a]

(A) *Levels on levels*: $y_t = f(m_{t-j}, y_{t-j}, x_{t-j})$
Six lags of m and twelve lags of y, p and r

Included polynomials in time	System: y, m	y, m, p	y, m, p, r
1	5.36 (0.000, < 0.01)	3.74 (0.001, 0.02)	2.52 (0.022, 0.07)
$1, t$	6.55 (0.000, < 0.01)	3.46 (0.003, < 0.01)	3.65 (0.002, < 0.01)
$1, t, t^2$	4.70 (0.000, < 0.01)	2.73 (0.014, 0.11)	2.64 (0.017, 0.14)

(B) *Levels on differences and levels*: $y_t = f(\Delta m_{t-j}, y_{t-j}, x_{t-j})$
Six lags of m and twelve lags of y, p and r

Included polynomials in time	System: y, m	y, m, p	y, m, p, r
1	6.07 (0.000, < 0.01)	4.43 (0.000, < 0.01)	3.21 (0.005, 0.03)
$1, t$	5.79 (0.000)	3.25 (0.004)	2.96 (0.008)
$1, t, t^2$	5.36 (0.000)	3.31 (0.004)	3.07 (0.006)

(C) *Differences on differences*: $\Delta y_t = f(\Delta m_{t-j}, \Delta y_{t-j}, \Delta x_{t-j})$
Six lags of m and twelve lags of y, p and r

Included polynomials in time	System: y, m	y, m, p	y, m, p, r
1	2.65 (0.016)	2.83 (0.011)	2.39 (0.029)
$1, t$	4.67 (0.000)	3.50 (0.002)	3.04 (0.007)
$1, t, t^2$	4.72 (0.000)	2.87 (0.010)	2.50 (0.023)

(D) *Differences on differences*: $\Delta y_t = f(\Delta m_{t-j}, \Delta y_{t-j}, \Delta x_{t-j})$
Twelve lags of all variables

Included polynomials in time	System: y, m	y, m, p	y, m, p, r
1	1.63 (0.082)	1.52 (0.115)	1.46 (0.141)
$1, t$	2.58 (0.003)	2.09 (0.018)	1.87 (0.038)
$1, t, t^2$	2.67 (0.002)	1.72 (0.062)	1.56 (0.104)

[a]The F-statistics test the hypothesis that all of the coefficients on the lags of m_t (or Δm_t, as appropriate) in the respective output equations are zero; p-values appear in parentheses. When just one p-value appears, it is based on the usual asymptotic approximation using the F-distribution. When two p-values appear, the first is based on the usual F-distribution, and the second is based on the non-standard asymptotic distribution, computed numerically as described in appendix B.

table 1 are consistent with all the regressors in table 3, panel A having unit roots and with none being cointegrated. Since not all the restrictions being tested can be represented as restrictions on mean zero, stationary variables, the F-statistics will have non-standard distributions. In panel B the regressors other than money growth contain unit roots. When time is included as a regressor (panel B, rows 2 and 3), the restrictions being tested are effectively on detrended money growth, which is mean zero and stationary, so the usual distributions apply. However, when time is excluded (row 1), money growth is dominated by a time trend and a linear combination of money growth and output is dominated by a *stochastic* trend, with which money growth is in turn correlated. Thus one of the restrictions being tested – in particular, the neutrality proposition that the sum of the coefficients on money growth is zero – involves a coefficient on a variable that is dominated by a time trend, resulting in a non-standard distribution. Finally, since none of the regressors in panels C or D are dominated by stochastic trends, the usual p-values apply. In the non-standard cases, both the usual (incorrect) p-values and the 'unit roots' distribution p-values (computed as described in appendix B) are reported.

The first puzzle is that substantially larger F-statistics obtain using detrended levels rather than first differences of the data. The gap is particularly wide between the F-statistic based on a VAR(12) in differences (panel D, row 1) – the specification examined by Eichenbaum and Singleton (1986) – and the F-statistic based on a linearly detrended levels regression with six lags of m and twelve lags of the other variables (panel A, rows 2 and 3): in the four-variable system, the twelve-lag differences regression has a p-value of 0.141, while the detrended levels regression has an (incorrect) p-value of 0.002. However, we will now argue that the findings reported in section 2, along with the non-standard asymptotic distribution theory, provide a consistent interpretation of these results.

Considering first the bivariate results, comparing panels C and D suggests that the non-zero coefficients are those on the first lags: testing the coefficients on later lags serves only to reduce the F-statistic, holding constant the number of other variables in the regressions. In addition, including time in the regression increases the F-statistic. These observations are consistent with the implications of table 2 that non-neutrality of $\Delta\mu_t$ is key to the large F-statistics. If the unit root assumption is correct, then the regressions with time trends in panels B and C test the same causality restrictions, although the test statistics will differ because of sampling variability: in fact, the reported tests under these alternative specifications suggest a predictive role for money. Thus much of the puzzling gap between the differences and levels results is largely accounted for by using a F-test that focuses more directly on the neutrality proposition, and by recognizing that it is *detrended* money growth that appears to Granger cause output.

The analysis in the three- and four-variable case is similar. Those specifications with six lags of money and twelve lags of the other variables which test the neutrality of $\Delta\mu_t$ (panel B, rows 2 and 3; panel C, rows 2 and 3) yield qualitatively similar results: all reject non-causality at the 5% level, although the specific p-values vary across specifications.

Comparing the usual and the non-standard p-values for the detrended levels regression suggests two general observations. First, the usual (incorrect) p-values typically are 'too small', i.e., the significance is overstated when the unit roots nature of the distribution is ignored.[7] Second, the (correct) p-values in the levels regressions tend to be larger than the (correct) p-values in the mixed levels/differences specification in panel B. Since one additional restriction is being tested in the panel A regressions – that the *level* of money (or detrended money) Granger causes output growth – a possible explanation is that this restriction is not violated, reducing the power of the tests in panel A. However, the (correct) p-values in panel A, row 3 are consistent with the p-values in the differences regressions in panel C in which the restrictions on detrended money are tested (panel C, rows 2 and 3).[8]

Summarizing, in a given system the non-standard asymptotics and the proposition that $\Delta\mu_t$ rather than Δm_t itself Granger causes output growth provides a consistent explanation for the larger F-statistics in the detrended levels model.

The second puzzle concerns the importance of m_t when r_t is added in levels specifications with and without linear and quadratic trends. In the literature, this puzzle arises from comparing variance decompositions, but it is reflected in the F-statistics in table 3 as well: the gap between the three- and four-variable F-statistics is less when trends are added to the levels specification (panel A, row 1 vs. row 2). Indeed, the F-statistic increases from 3.46 to 3.65 when interest rates are added in the specification with a trend, although the corrected p-values change only slightly.[9] This drop is consistent with the

[7] This is consistent with Christiano and Ljungqvist's (1987) Monte Carlo evidence on the bivariate money–output F-test in the levels regression without a time trend.

[8] The F-statistics of 2.65 in table 3, panel C, row 1 and 6.07 in panel B, row 1 for the bivariate models may at first seem contradictory, since both correspond to coefficients on money growth excluding time trends from the regression. However, these statistics in fact test different restrictions. In the pure differences specification, money growth is dominated by its time trend, and the neutrality of detrended money growth is not tested. In the mixed differences/levels specification, money growth is 'detrended' by industrial production, which is dominated by a time trend. Thus this statistic can be rewritten to test (among other restrictions) whether a linear combination of detrended money growth and the stochastic part of output Granger causes output. While this linear combination does appear to Granger cause output, this still is not the hypothesis of interest (concerning $\Delta\mu_t$ alone), which is untested until a time trend is included in the regression.

[9] In finite samples, these F-statistics are only loosely linked to the variance decompositions. In addition, as Runkle (1987) emphasizes, the variance decompositions can be very imprecisely estimated, leading to spurious inferences based on modest shifts in the variance decomposition point estimates. Thus associating the second puzzle with the shifts in panel A of table 3 is only approximate.

previous argument that detrended money growth rather than Δm_t Granger causes output, since in this case both of the levels regressions without a trend are misspecified, having omitted time as a regressor. The relative direction of the 'bias' resulting from omitting time in the three- and four-variable systems is uncertain on *a priori* grounds, although this bias would be expected to have different effects in the two systems. Thus the 'explanation' of the second puzzle is simply that the regressions in panel A, row 1 have omitted a variable (time) that the analysis of sections 2 and 3 suggests is important for isolating the permanent role of money innovations in forecasting output growth. According to this view, the specifications including time (and the non-standard p-values) constitute the correct basis for inference, in which the marginal predictive content of money changes little when the interest rate is added.

The preceding analysis assumes that inflation and detrended money growth are stationary. While these specifications were suggested by the tests of section 2, an alternative characterization of these series would be that both have unit roots, perhaps with drifts. Are the previous conclusions sensitive to this modification? Supposing first that inflation were integrated, the non-standard p-values for the levels regressions reported above would have been computed incorrectly (since the incorrect number of unit roots was included in the multivariate specification). However, the F-tests in the detrended regressions in panels B, C and D would still have the usual asymptotic distributions, since the restrictions could be written in terms of mean zero, stationary variables (detrended money growth). This suggests that our main conclusion – that detrended money growth helps to predict industrial production – would be unchanged were inflation an integrated process. In contrast, if money growth has a unit root, then all the reported tests would involve restrictions on integrated variables, and none of the reported p-values would apply. However, the discussion of section 2 suggests that the argument for this particular alternative characterization is weak.

5. Subsample stability

This section addresses the third puzzle, the apparent instability of the causality tests when the 1980's are eliminated from the sample. The calculations reported in sections 2–4 were repeated using data from January 1959 to September 1979. The results are reported in tables 4–6.

The integration, cointegration, and time trend tests (table 4) yield the same inferences as when the full sample is used, with one important exception: during the shorter sample, inflation seems to be well characterized as containing a linear time trend. While modeling inflation as having a linear time trend in the shorter sample but as having no time trend in the larger sample is logically inconsistent, the purpose of this initial 'identification' phase of the analysis is to obtain an empirical representation of the variables that will serve as a framework for the asymptotic approximations to the sampling distribu-

Table 4

Tests for integration, cointegration and time trends, 1960:2–1979:9.[a]

(A) *Univariate tests*

Series	Unit root tests					t-statistics for a regression of Δx on:	
	$q_f^{\tau^2}[z]$	$q_f^{\tau}[\Delta z]$	D-F $\hat{\tau}_\tau[\Delta z]$	$q_f^{\tau}[z]$	D-F $\hat{\tau}_\tau[z]$	time	constant
y	−12.28	−132.9**	−4.85**	−8.30	−2.26	−0.28	2.26**
m	−15.00	−205.9**	−5.36**	−5.14	−1.87	3.26**	3.97**
p	−9.56	−212.5**	−4.09**	−2.28	−0.91	2.98**	2.19*
r	−9.34	−180.4**	−5.97**	−9.60	−1.80	1.22	1.06

(B) *Multivariate tests*

System	Unit roots under null and alternative	$q_f^{\tau^2}$
m, y	2 vs. 1	−11.1
m, y, p	3 vs. 2	−9.1
m, y, p, r	4 vs. 3	−7.7

[a]See the notes to table 1.

tions. Since a linear time trend appears to be an important empirical feature of inflation in the shorter sample, we adopt a characterization of the four variables during 1959–1979 in which all contain unit roots, there is no cointegration, and inflation – as well as money growth – has a linear time trend:

$$\Delta y_t = \alpha_{y0} + \Delta\eta_t,$$

$$\Delta m_t = \alpha_{m0} + \alpha_{m1}t + \Delta\mu_t,$$

$$\Delta p_t = \alpha_{p0} + \alpha_{p1}t + \Delta\pi_t,$$

$$\Delta r_t = \Delta\rho_t.$$

Comparing the neutrality tests using the full sample (table 2, panel A) with those using the shorter sample (table 5, panel A), in both cases the point estimate of $\beta_{ym}(1)$ in the bivariate case doubles when a time trend is added to the regression. In contrast, in the shorter sample the point estimate changes little when time is added to the three- and four-variable specifications. This apparent discrepancy has a simple explanation in terms of the identified trend in inflation in the shorter sample: when inflation has a trend component, including inflation as a regressor is essentially including time, plus some stationary errors, as a regressor. Thus inflation detrends money growth,

Table 5

Neutrality and causality tests, 1960:2–1979:9.[a]

(A) *Neutrality tests*: $\hat{\beta}_{ym}(1)$ *and t-statistics*

Specification	Six lags of money			Twelve lags of money		
	y, m	y, m, p	y, m, p, r	y, m	y, m, p	y, m, p, r
(I) Const; Δm_t	0.65*	1.09**	1.03**	0.46	1.52**	1.25**
	(2.52)	(3.97)	(3.55)	(1.51)	(4.02)	(3.13)
(II) Const; $\Delta\hat{\mu}_t$	1.26**	1.07**	0.99**	1.41**	1.50**	1.13*
	(3.57)	(2.99)	(2.66)	(2.65)	(2.83)	(1.98)
(III) Const, t; Δm_t	1.26**	0.87*	0.84*	1.43**	1.42**	1.06$^+$
	(3.57)	(2.41)	(2.26)	(2.66)	(2.71)	(1.90)

(B) *Causality tests*: *F-statistics and p-values*

Specification	Six lags of money			Twelve lags of money		
	y, m	y, m, p	y, m, p, r	y, m	y, m, p	y, m, p, r
(I) Const; Δm_t	1.44	3.14	2.68	1.31	2.11	1.60
	(0.201)	(0.006)	(0.016)	(0.215)	(0.018)	(0.095)
(II) Const; $\Delta\hat{\mu}_t$	2.52	1.99	1.75	1.73	1.40	1.09
	(0.022)	(0.069)	(0.112)	(0.062)	(0.166)	(0.375)
(III) Const, t; Δm_t	2.52	1.50	1.49	1.73	1.21	0.98
	(0.022)	(0.179)	(0.182)	(0.063)	(0.277)	(0.472)

[a]See the notes to table 2.

permitting the estimation of the sum of the coefficients on a linear combination of $\Delta\mu_t$ and $\Delta\pi_t$, which is evidently close to the sum of the coefficients on $\Delta\mu_t$ alone.

The causality tests in panel B of table 5 exhibit two major differences from the corresponding tests in the longer sample. First, these results generally indicate less influence of money on output during the shorter period, particularly using twelve lags. Second, given this generally decreased influence, the major discrepancy between the F-tests in tables 2 and 5 is the significant causation evident in specification (I) in the shorter sample for the systems including prices. A possible explanation for this is that detrended inflation ($\Delta\pi_t$) helps to predict output growth. When time is omitted from the regression, money growth is 'detrended' by inflation. Thus one of the restrictions being tested is that a linear combination of $\Delta\mu_t$ and $\Delta\pi_t$ – rather than $\Delta\mu_t$ itself – Granger causes industrial production growth; if $\Delta\pi_t$ Granger causes Δy_t, omitting the time trend would spuriously indicate a greater role for money.

The full set of causality results (table 6) also indicate a diminished role for money in predicting output in this shorter sample. The first puzzle of the

Table 6

Money–income causality tests, 1960:2–1979:9.[a]

(A) *Levels on levels*: $y_t = f(m_{t-j}, y_{t-j}, x_{t-j})$

Six lags of m and twelve lags of y, p and r

Included polynomials in time	System		
	y, m	y, m, p	y, m, p, r
1	2.93 (0.009, 0.04)	2.04 (0.062, 0.52)	1.09 (0.372, 0.88)
$1, t$	4.17 (0.001, < 0.01)	1.22 (0.295, 0.85)	1.64 (0.138, 0.68)
$1, t, t^2$	2.56 (0.020, 0.07)	1.14 (0.338, 0.59)	1.17 (0.322, 0.51)

(B) *Levels on differences and levels*: $y_t = f(\Delta m_{t-j}, y_{t-j}, x_{t-j})$

Six lags of m and twelve lags of y, p and r

Included polynomials in time	System		
	y, m	y, m, p	y, m, p, r
1	3.39 (0.003, 0.01)	2.18 (0.047)	1.20 (0.306)
$1, t$	3.40 (0.003)	1.47 (0.189)	1.17 (0.322)
$1, t, t^2$	2.76 (0.013)	1.56 (0.161)	1.43 (0.204)

(C) *Differences on differences*: $\Delta y_t = f(\Delta m_{t-j}, \Delta y_{t-j}, \Delta x_{t-j})$

Six lags of m and twelve lags of y, p and r

Included polynomials in time	System		
	y, m	y, m, p	y, m, p, r
1	1.44 (0.201)	3.14 (0.006)	2.68 (0.016)
$1, t$	2.52 (0.022)	1.50 (0.179)	1.49 (0.182)
$1, t, t^2$	2.50 (0.023)	1.42 (0.207)	1.43 (0.206)

(D) *Differences on differences*: $\Delta y_t = f(\Delta m_{t-j}, \Delta y_{t-j}, \Delta x_{t-j})$

Twelve lags of all variables

Included polynomials in time	System		
	y, m	y, m, p	y, m, p, r
1	1.31 (0.215)	2.11 (0.018)	1.60 (0.095)
$1, t$	1.73 (0.063)	1.21 (0.277)	0.98 (0.472)
$1, t, t^2$	1.75 (0.059)	1.22 (0.272)	0.96 (0.490)

[a]See the notes to table 3.

introduction – the difference in p-values between the differences and detrended levels specifications – is still present in the bivariate results (panel C, row 1 vs. panel A, row 2). The explanation is the same as before: focusing on fewer restrictions on lagged money, detrending money growth, and using the non-standard critical values accounts for the apparent discrepancy. In the multivariate models, however, the F-tests in the differences specifications are comparable to or exceed those in the levels specifications. One explanation for this reversal is that given in the discussion of table 5, panel B; the similarity of the (correct) p-values in the detrended differences and detrended levels regressions with six lags of m support this view.

The second puzzle (adding a trend to the levels specification alters the effect of money when interest rates are included) is summarized in panel A, rows 1 and 2. The preceding analysis suggests that the anomalous result here is the three-variable levels specification without a trend. Thus the 'explanation' of this puzzle is the same as in the longer sample, appealing to the bias induced by omitting time as a regressor and to the overstatement of the p-values arising from using the usual asymptotic theory.

Summarizing, when the time trend in inflation is recognized and when the correct critical values are used, money growth appears to help forecast industrial production in a bivariate model in both samples. However, its marginal predictive content drops in multivariate settings in the shorter sample. Since the point estimates of $\beta_{ym}(1)$ are similar across samples and across the two-, three- and four-variable systems, this decrease appears to be attributable to a combination of (i) relatively more of the *short-run* movements in industrial production being explained by innovations in money during the 1980's than before; (ii) increased variability in the regressors in the 1980's, and therefore greater power in the F-tests; and (iii) simply having more observations in the longer sample. In particular, using corrected p-values, in the shorter sample the role of money drops substantially when price is added to the specification, while it does not in the longer sample. In contrast to Sims' (1980b) findings, the predictive value of money does not change appreciably when an interest rate is added to the *detrended* specifications in either sample.

6. Conclusions

The asymptotic p-values reported in tables 3 and 6 have been computed under specific assumptions about the unit root and trend behavior of the various series. If these assumptions are incorrect, some or all of these asymptotic p-values would be incorrectly calculated, depending upon the specific assumption in question. This can be interpreted as a drawback of our two-step procedure of first testing for trends, then calculating p-values conditional on the trend specification. An alternative approach – at least in theory – would be to specify a prior distribution over different trend specifications, to assume

Gaussian innovations in the VAR's, and to compute posterior odds ratios, integrating over the various cases. However, the computational burden in this alternative approach is substantially greater than that entailed here; for an application and discussion in a comparatively simple three-lag univariate autoregressive model, see Geweke (1986b). More importantly, we suggest that the major economic import of these results – the rejection of the 'neutrality' and Granger causality propositions – is largely unaffected by this criticism. One message of Sims, Stock and Watson (1986) is that certain linear combinations can be tested using the usual asymptotics even if there are unit roots. Our key results – those in tables 2 and 4 and in tables 3 and 5, panel C, row 2 – will have normal distributions even if many of our stochastic trend identifications are incorrect.

Subject to this caveat, these results suggest three conclusions. First, innovations in the growth of nominal M1 are useful in forecasting industrial production in the bivariate money–output relation. Second, over the entire sample period monetary innovations have predictive value for industrial production beyond information contained in prices and interest rates, although the marginal significance levels rise somewhat in the multivariate systems. Third, the additional role of money over prices and interest rates is substantially less if attention is restricted to a shorter sample ending in 1979. However, the difference between the two samples might simply arise from the greater variation in output, money, prices and interest rates over the 1980 and 1982 recessions, so that the tests in the longer sample would have substantially more power.

This analysis suggests two reasons that past researchers have found conflicting results. First, many of the earlier specifications failed to focus on innovations in money, i.e., on the non-deterministic component of money growth that we find to have predictive content for long-run movements in industrial production. Second, incorrect p-values were used to analyze the levels regressions, typically resulting in overstating the predictive content of money. These observations provide consistent explanations of the puzzles stated in the introduction.

Appendix A

Testing for a unit root with a quadratic time trend

The Stock–Watson (1986a) $q_f^\tau(k, m)$-statistic tests the null hypothesis that a n-vector of time series variables has $k \leq n$ unit roots, perhaps with drift, against the alternative that the variables have among them only $m < k$ unit roots, perhaps with drift. In its simplest form, $q_f^\tau(k, m) = T[\mathrm{Re}(\hat{\lambda}_{m+1}) - 1]$, where $\hat{\lambda}_{m+1}$ is the $(m+1)$th largest root of the first-order sample autocorrelation matrix of the linearly detrended vector process. The drift in the stochastic

Table 7

Critical values for the $q_f^{\tau^2}(k, k-1)$-statistic.[a]

Critical value	$k=1$	$k=2$	$k=3$	$k=4$
1%	−35.5	−46.0	−55.7	−64.5
5%	−27.9	−36.9	−45.3	−53.5
10%	−24.1	−32.5	−41.0	−48.6

[a] *Source*: Authors' calculations.

trends results in a linear time trend in the series themselves plus a driftless stochastic trend. If it is suspected instead that the series might have a quadratic as well as a linear time trend, the $q_f^{\tau}(k, m)$-statistic is not applicable. Since money growth appears to contain a linear time trend in the current application, the $q_f^{\tau}(k, m)$-statistic is inappropriate either for testing for a single unit root in money or for testing for reduced unit roots in the multivariate systems. We therefore adopt a modification of the Stock–Watson (1986a) $q_f^{\tau}(k, m)$-statistic that can be used to test for k vs. m unit roots when there might be a quadratic term under either the null or the alternative.

This modified statistic, termed $q_f^{\tau^2}(k, m)$ in table 1, is computed exactly as is the $q_f^{\tau}(k, m)$ statistic, except that the data are first detrended using a quadratic as well as a linear time trend. Like $q_f^{\tau}(k, m)$, if the process follows an AR(p), then the limiting distribution of $q_f^{\tau^2}(k, m)$ does not depend on nuisance of the parameters of the process. Selected critical values for the statistic with $m = k-1$, computed as described in Stock and Watson (1986a) using 4000 Monte Carlo replications, are given in table 7.

Appendix B

Computation of non-standard p-values

The marginal significance levels for the test statistics with non-standard distributions were computed by integrating numerically the limiting distributions of these statistics. In general, the causality tests of the type considered here with p restrictions have limits of the form

$$F \geq [F_1 + F_2(W(t);\theta)]/p, \tag{B.1}$$

where F_1 has a χ^2_{p-1} distribution, $W(t)$ is n-dimensional Brownian motion, where n is the dimension of the system, and θ is a vector of parameters. Except in special cases, θ contains parameters from the entire VAR, including the error covariance matrix. In addition, $W(t)$ enters F_2 only through a limited number of functionals such as $\int_0^1 W(t)W(t)'\,\mathrm{d}t \equiv \Gamma$ and $\int_0^1 W(t)\,\mathrm{d}W(t)' \equiv \Psi$.

[Explicit formulas for F_2 are given in Sims, Stock and Watson (1986, theorem 2).]

This representation, combined with theorem 2.2 of Chan and Wei (1988) which implies that F_1 and $W(t)$ (and thus F_1 and F_2) are independent, provides a convenient formulation for Monte Carlo integration of the asymptotic distribution. The specific procedure we adopt is an extension of Stock's (1987b) suggestion for computing the asymptotic distribution of least squares estimators of cointegrating vectors when there is zero drift. Specifically, the non-standard p-values were estimated using the following algorithm: (i) Generate the functionals such as Γ and Ψ using sample path realizations; for example, a draw of Γ is $\Gamma_T = T^{-2}\sum_{t=1}^{T} Y_t Y_t'$, where Y_t is a driftless n-dimensional random walk with $Y_0 = 0$ and an identity innovation covariance matrix. We chose T to be 1000 and used a normal random number generator to construct 4000 pseudo-random realizations of these functionals. (ii) Compute an estimate $\hat{\theta}$ of the parameters θ for the VAR in question. (iii) Drawing sequentially from the previously generated random matrices (Γ_T, etc.), use $\hat{\theta}$ to construct a random realization of F_2; add this to an independently drawn χ^2_{p-1} variate and divide the sum by p to construct a draw of F in (B.1). Step (iii) was repeated 4000 times for each estimated model, and the (non-standard) p-value was computed using the resultant empirical distribution of F.

References

Berk, K.N., 1974, Consistent autoregressive spectral estimates, Annals of Statistics 2, 489–502.

Bernanke, B.S., 1986, Alternative explanations of the money–income correlation, Carnegie–Rochester Conference Series on Public Policy 25, 49–100.

Blanchard, O.J. and M.W. Watson, 1986, Are business cycles all alike? in: R.J. Gordon, ed., The American business cycle (University of Chicago Press, Chicago, IL).

Chan, N.H. and C.Z. Wei, 1988, Limiting distributions of least squares estimates of unstable autoregressive processes, Annals of Statistics 16, 367–401.

Christiano, L.J. and M. Eichenbaum, 1987, Temporal aggregation and structural inference in macroeconomics, Carnegie–Rochester Conference Series on Public Policy 26, 63–130.

Christiano, L.J. and L. Ljungqvist, 1987, Money does Granger-cause output in the bivariate output–money relation, Research Department working paper, Jan. (Federal Reserve Bank of Minneapolis, Minneapolis, MN).

Dickey, D.A. and W.A. Fuller, 1979, Distribution of the estimators for autoregressive time series with a unit root, Journal of the American Statistical Association 74, no. 366, 427–431.

Eichenbaum, M. and K.J. Singleton, 1986, Do equilibrium real business cycle theories explain postwar U.S. business cycles, NBER Macroeconomics Annual 1986, 91–134.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error correction: Representation, estimation and testing, Econometrica 55, 251–276.

Fuller, W.A., 1976, Introduction to statistical time series (Wiley, New York).

Geweke, J., 1984, Inference and causality in economic time series models, Ch. 19 in: Zvi Griliches and M.D. Intriligator, eds., Handbook of econometrics, Vol. 2 (North-Holland, Amsterdam).

Geweke, J., 1986a, The superneutrality of money in the United States: An interpretation of the evidence, Econometrica 54, 1–22.

Geweke, J., 1986b, The secular and cyclical behavior of real GDP in nineteen OECD countries, 1957–1983, Manuscript (Duke University, Durham, NC).

King, R.G. and C.I. Plosser, 1984, Money, credit and prices in a real business cycle, American Economic Review 74, 363–380.

King, S., 1984, Macroeconomic activity and the real rate of interest, Ph.D. thesis (Department of Economics, Northwestern University, Evanston, IL).

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

Nelson, C.R. and C.I. Plosser, 1982, Trends and random walks in macroeconomic time series, Journal of Monetary Economics 10, I 9–162.

Ohanian, L.E., 1986, The spurious effects of unit roots on vector autoregressions: A Monte Carlo study, Manuscript (University of Southern California, Los Angeles, CA).

Perron, P., 1986, Tests of joint hypotheses for time series regression with a unit root, C.R.D.E. discussion paper no. 2086 (Université de Montréal, Montréal).

Perron, P. and P.C.B. Phillips, 1986, Does GNP have a unit root? A reevaluation, Economics Letters, forthcoming.

Runkle, D., 1987, Vector autoregressions and reality, Journal of Business and Economic Statistics 5, 437–442.

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

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

Sims, C.A., 1980b, Comparison of interwar and postwar business cycles: Monetarism reconsidered, American Economic Review 70, 250–257.

Sims, C.A., J.H. Stock and M.W. Watson, 1986, Inference in linear time series models with some unit roots, Manuscript (Stanford University, Stanford, CA).

Stock, J.H., 1987a, Temporal aggregation and structural inference in macroeconomics: A comment, Carnegie–Rochester Conference Series on Public Policy 26, 131–140.

Stock, J.H., 1987b, Asymptotic properties of least squares estimators of cointegrating vectors, Econometrica 55, 1035–1056.

Stock, J.H. and M.W. Watson, 1986a, Testing for common trends, Harvard Institute for Economic Research discussion paper no. 1222, forthcoming in the Journal of the American Statistical Association.

Stock, J.H. and M.W. Watson, 1986b, Does real GNP have a unit root?, Economics Letters 22, 147–151.