ELSEVIER

Journal of Econometrics 64 (1994) 279–306

JOURNAL OF
Econometrics

Subsample instability and asymmetries in money–income causality

Mark A. Thoma

Department of Economics, University of Oregon, Eugene, OR 97403-1285, USA

(Received July 1991; final version received August 1993)

Abstract

One puzzle in the money–income causality literature is the variability in results across sample periods, particularly with regard to the 1980's. This paper shows that the p-values of money–income causality tests across sample periods are highly correlated with the level of real activity. When real activity declines, money–income causality becomes stronger, and when real activity increases, the relationship becomes weaker. The investigation proceeds in three stages. First, rolling regressions are used to document the subsample instability and the correlation with real activity as well as the variability of the results across model specifications, trend specifications, and over the business cycle. The second stage uses bootstrap simulations to show that the variability across sample periods and the correlation with real activity is significant, and the third stage builds a combination dummy variable and varying coefficients VAR model that captures the asymmetries in the relationship between money and income.

Key words: Causality; Money; Asymmetries; Subsample instability
JEL classification: E51

1. Introduction

This paper documents a striking empirical regularity in the outcomes of money–income causality tests. The level of real activity is highly correlated with the p-values of money–income causality tests. When real activity experiences a cyclical decline, the relationship between money and income becomes

Thanks to Jo Anna Gray and two anonymous referees for helpful suggestions.

SSDI 0304407693015 86 B

stronger, and when an upswing in real activity occurs, the relationship between money and income becomes weaker. This striking regularity offers a fresh perspective on the problem of subsample instability in money–income causality studies and provides yet another challenge to existing explanations of the business cycle.

Subsample instability in money–income causality is particularly evident in papers using data sets that include the 1980's. Eichenbaum and Singleton (1986) note that the evidence supporting money–income causality is much weaker when the 1980's are excluded from the data than when the 1980's are included. Stock and Watson (1989) claim to resolve this instability through careful treatment of the unit root and trend properties of the data. Unlike Eichenbaum and Singleton, who use a detrended log levels specification, Stock and Watson use a detrended log differences specification. Using this specification they find that the role of money in predicting output is still reduced when the 1980's are excluded from the sample. However, they argue that the difference in results is due to increased variability of the data in the 1980's which leads to more powerful test statistics, and the simple fact that the longer sample contains more observations leading to more precise estimates.

A more recent paper by Friedman and Kuttner (1993) claims that one of the reasons the Stock and Watson results are important is that, contrary to the view that the money–income relationship has broken down in the 1980's, the Stock and Watson results show a significant relationship in data extending through 1985. Friedman and Kuttner then show that the Stock and Watson results are not robust to extending the sample through 1988. Extending the sample period through 1988 results in an insignificant relationship between money and income leading Friedman and Kuttner to conclude that the money–income relationship does appear to have weakened in the 1980's.

This paper provides a detailed analysis of the stability of the money–income causality relationship over the business cycle, across sample periods, over different model specifications, highlights episodes within the overall sample for which money–income causality receives the most support, and builds a model capturing asymmetries and state dependency in the relationship between money and income. The analysis is conducted in three stages. First, rolling regressions are used to show how causality changes as the sample period is extended by adding successive observations to the data set. The results show that the p-values for the causality tests are highly variable across sample periods, and that there are two episodes since 1959 over which the money–income causality hypothesis receives the most support, 1969 through mid-1973 and 1978 through 1982. It is also shown that when output is decreasing, the relationship between money and income becomes stronger, and when output is increasing, the relationship becomes weaker. When the variables are measured in differences, the pattern of results across sample periods is shown to be robust to the exclusion of the interest rate from the model used to assess causality, and to the inclusion of linear trends. However, when the model is specified in levels, the results are not as robust.

The second stage of the investigation examines the significance of the patterns in the causality statistics generated with the rolling regression technique. Because no distribution theory exists for testing the significance of the patterns in the causality statistics from rolling regressions, bootstrap simulations are employed. The results of the bootstrap simulations show that the amount of variability in the outcomes of the causality tests and the correlation between the p-values and real activity are both significant.

The first two stages of the investigation imply that there are asymmetries in the relationship between money and output. The third stage of the investigation builds an empirical model that captures these asymmetries. The model generalizes a VAR model to allow positive and negative changes in money growth to have different effects on output, and allows the effects to vary over the business cycle. The results show a strong relationship between money and income when money growth and output are declining, and a weaker relationship when money growth and output are increasing, a result that agrees with the conclusions from stages 1 and 2. The results also provide strong evidence that the effects of negative changes in money growth on output are variable over the business cycle with the effects strongest when output is high and weakest when output is low.

The results of the third stage also show that when a traditional VAR model with differenced data is used, the relationship between money and income is insignificant. Thus, separating money into positive and negative components and allowing the effects to vary over the business cycle overturns the insignificance of the money–income relationship obtained with the traditional approach, a result that may explain why many authors (e.g., Eichenbaum and Singleton, 1986; Friedman and Kuttner, 1993) who use the traditional approach find weak links between money and income.

Finally, the paper concludes with a comparison of the forecasting abilities of a traditional VAR model with a VAR model that allows positive and negative changes in money to have asymmetric effects and allows the effects to vary over the business cycle. The results indicate that the model with asymmetries and varying effects of money over the business cycle generally has superior forecasting abilities.

2. Robustness of causality tests

This section examines the robustness of the outcome of causality tests to changes in the sample period, the number of variables included in the model and the removal of stochastic and deterministic trends.

2.1. Methodology

Following Sims (1980), Litterman and Weiss (1985), Stock and Watson (1989), Friedman and Kuttner (1993), Krol and Ohanian (1990), and many others, money–income causality is assessed using a four variable VAR model. A trivariate model is also examined. The basic model is

$$X_{1,t} = f(t) + \sum_{i=1}^{12} a_{1i} X_{1,t-i} + \sum_{i=1}^{K} a_{2i} X_{2,t-i} + \sum_{i=1}^{12} a_{3i} X_{3,t-i} + \sum_{i=1}^{12} a_{4i} X_{4,t-i} + e_t, \tag{1}$$

where $f(t)$ is the trend specification, $X_{1,t}$ is the log of industrial production ($X_{1,t} \equiv Y_t$) or the growth rate of industrial production ($X_{1,t} \equiv \Delta Y_t$), $X_{2,t}$ is the log of M1 ($X_{2,t} \equiv M_t$) or the growth rate of M1 ($X_{2,t} \equiv \Delta M_t$), $X_{3,t}$ is the log of the price level ($X_{3,t} \equiv P_t$) or the inflation rate ($X_{3,t} \equiv \Delta P_t$), and $X_{4,t}$ is the interest rate ($X_{4,t} \equiv R_t$) or the change in the interest rate ($X_{4,t} \equiv \Delta R_t$). Following Stock and Watson, the number of lags on the measure of money, K, is either 6 or 12, while the number of lags on the other variables is 12.

There are 16 separate specifications of model (1) considered in this paper. A summary of the specifications follows:

Differences

$$\Delta Y_t = f(t) + g(\Delta Y_t, \Delta M_t, \Delta P_t, \Delta R_t) + e_t, \tag{2}$$

$$\Delta Y_t = f(t) + g(\Delta Y_t, \Delta M_t, \Delta P_t) + e_t; \tag{3}$$

Levels

$$Y_t = f(t) + g(Y_t, M_t, P_t, R_t) + e_t, \tag{4}$$

$$Y_t = f(t) + g(Y_t, M_t, P_t) + e_t. \tag{5}$$

There are 4 models listed. Two specifications of the number of lags of the money measure, 6 or 12, are examined for each model bringing the total number of models to 8. Additionally, 2 specifications of the trend term, $f(t)$, are examined for each model bringing the total to 16. The trend specifications are $f(t) = f_0$ or $f(t) = f_0 + f_1 t$.

One of the goals of the paper is to examine the stability of money–income causality statistics across sample periods. To accomplish this, each of the 16 specifications of the model described above is estimated using a rolling regression technique. First, a Granger causality test is performed using monthly data from 1960:2 to 1966:1.[1] Then, one month is added to the data set so that the data covers 1960:2 to 1966:2 and the test is repeated. The process of adding one month to the data and repeating the Granger causality test is continued until the entire data set, 1960:2 to 1989:12 is used to perform the test. The result is 300 F-statistics for causality,[2] each with a different number of degrees of freedom. Finally, the p-values of the F-statistics, which are comparable across sample periods and degrees of freedom, are plotted with the date that the sample ends on the horizontal axis and the p-value for the F-test for the hypothesis of no Granger causality from money to income on the vertical axis.

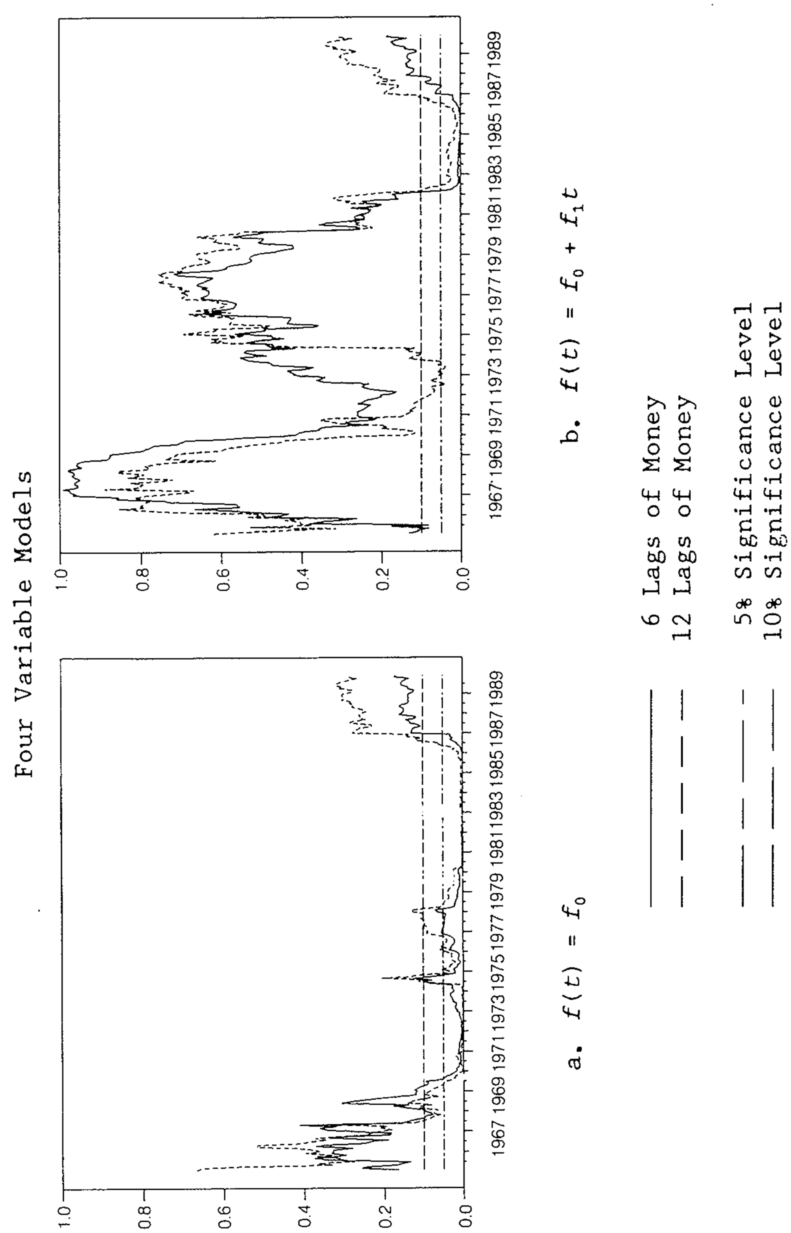
2.2. Results

Consider first the results for the three- and four-variable differenced models [specifications (2) and (3) above]. In Fig. 1 four plots are shown for these specifications, with the top half of the figure showing the plots for the four-variable differenced model, and the bottom half showing the plots for the three-variable differenced model. Figs. 1a and 1c include a constant in the VAR model, and Figs. 1b and 1d a constant and a linear time trend (1b is the Stock and Watson specification). There are two sets of results shown in each panel, one for the model with 6 lags of money and one with 12 lags of money.

The results for the four-variable differences specification with a time trend in Fig. 1b show two subsamples when support for money–income causality is the strongest. For samples ending in 1973 through 1974 the results are significant at the 10% level for the six-lag model, but not for the twelve-lag model, and for samples ending from 1982 to 1987 the results are significant at the 5% level for both the 6 and 12 lag specifications. The figures also reveal which time periods support the money–income hypothesis and which time periods are contrary to the hypothesis. If the p-value increases (i.e., moves toward 1.0) when an additional observation is added to the data set, then this observation is not supportive of the hypothesis. If the p-value decreases (i.e., moves toward 0.0), then the observation supports the hypothesis. There are two time periods in which the p-values show a general decrease. Beginning with the sample ending in 1969, the

[1] All data are from the *Citibase* data bank. The data are monthly and cover the sample period 1959:1 to 1989:12. Allowing for lags and differencing causes the data to begin at 1960:2. Money is measured as M1, the interest rate is the rate on three-month T-bills, prices are measured as the CPI, and output is measured as industrial production.

[2] As in Stock and Watson (1989) and Friedman and Kuttner (1993), F-statistics rather than variance decompositions and impulse response functions are used to assess causality. This is so that the results here can be compared directly to the results of Stock and Watson, Friedman and Kuttner, and others. The limitations inherent in the use of F-statistics to assess causality in multivariate VAR models (i.e., with more than two variables) are acknowledged.


Four Variable Models
1.0
0.8
0.6
0.4
0.2
0.0
1967 1969 1971 1973 1975 1977 1979 1981 1983 1985 1987 1989
a. f(t) = f0
b. f(t) = f0 + f1t
6 Lags of Money
12 Lags of Money
5% Significance Level
10% Significance Level

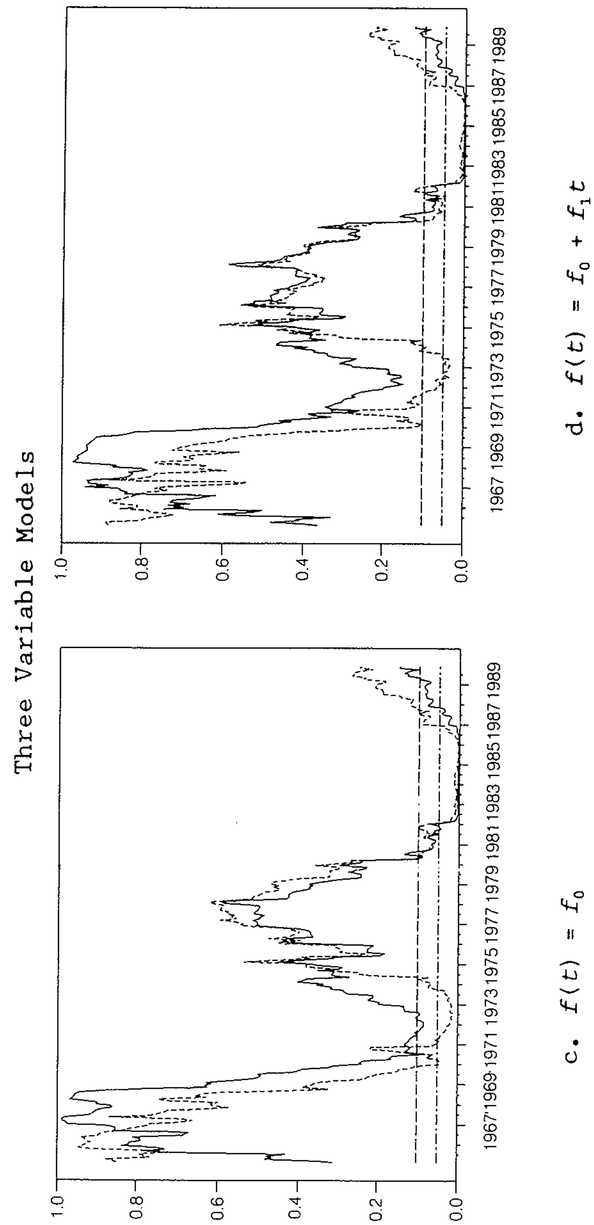
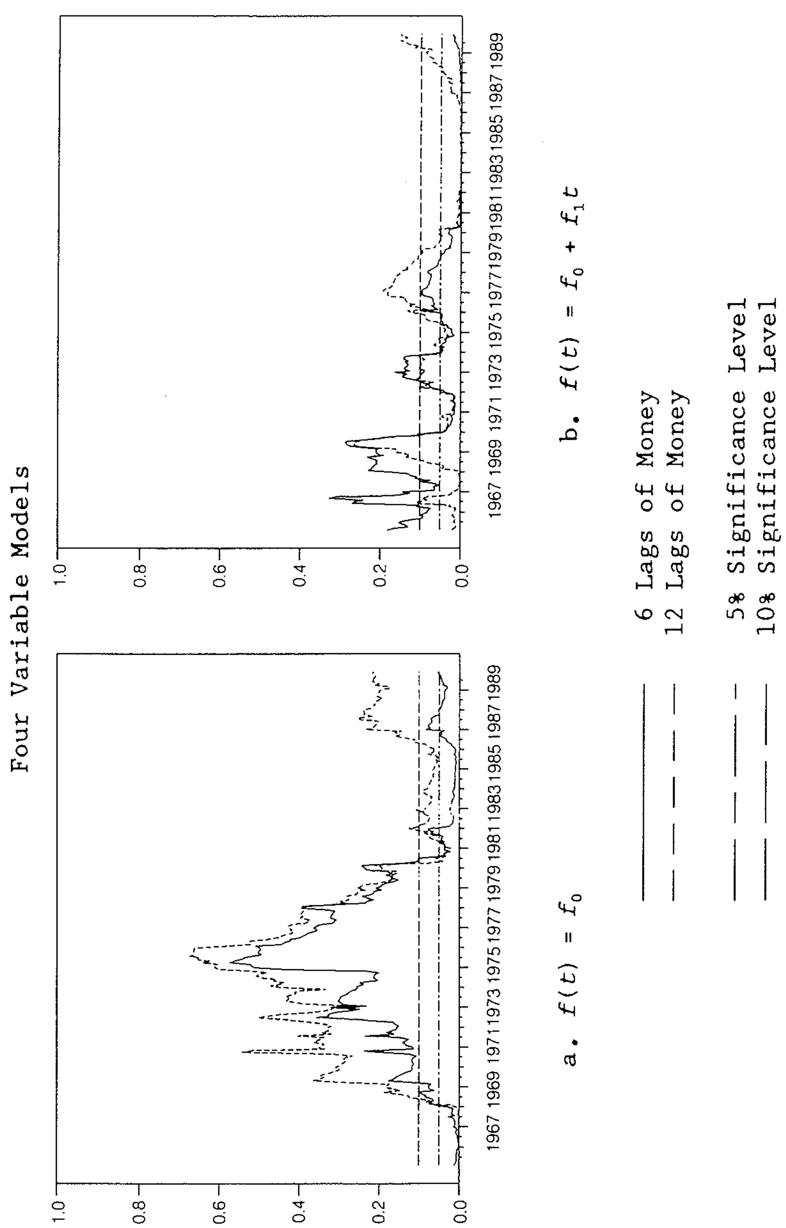


Fig. 1. Three- and four-variable models – Differences.


Four Variable Models
1.0
0.8
0.6
0.4
0.2
0.0
1967 1969 1971 1973 1975 1977 1979 1981 1983 1985 1987 1989
a. f(t) = f0
b. f(t) = f0 + f1t
6 Lags of Money
12 Lags of Money
5% Significance Level
10% Significance Level

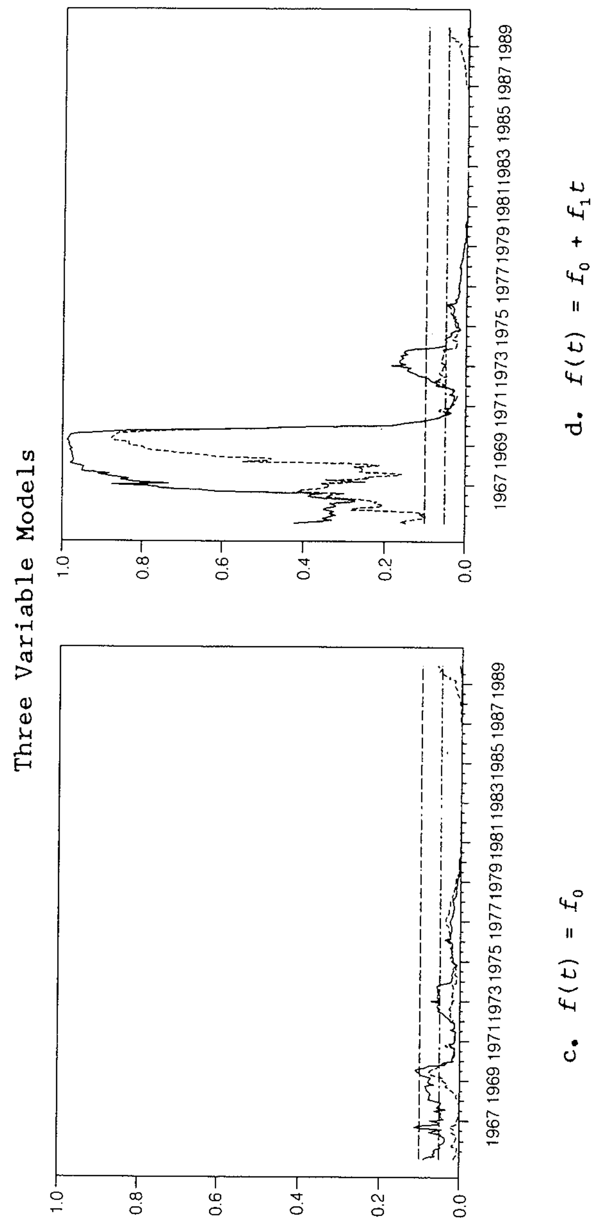


Fig. 2. Three- and four-variable models – Levels.

p-values decrease reaching a minimum for samples ending at around mid-1973. Thus, the time period 1969 to mid-1973 appears to be a period when money is causing income. Similarly, the p-values show a consistent decrease for samples ending from 1978 to 1982 and level off for samples ending from 1983 to 1987. Thus, the two time periods 1969 to mid-1973 and 1978 to 1982 appear to be the time periods when the relationship between money and income is the strongest. In the model shown in Fig. 1a, where only a constant term is included, the same pattern is evident, but it is highly damped. The results generally attain a higher level of significance than for the differences specification with a time trend. The relative stability of the results and the higher levels of significance in this case are noteworthy because the log-differenced model without a trend is one of the most common specifications used in the literature.

We now turn to the robustness of the results to the inclusion of interest rates, and to specifying the model in levels. The bottom half of Fig. 1 shows the differenced version of the model that excludes the interest rate [model (3) above]. The results are nearly identical to the four-variable differenced model except that, unlike the four-variable model, the three-variable model does not behave differently when only a constant is included in the trend specification. Overall, Fig. 1 documents a striking consistency in the movements of the p-values as the sample period is varied.

Fig. 2 shows the results for the models estimated in levels rather than differences.[3] These results do not exhibit the same degree of consistency as the differenced models. In the four- and three-variable models shown in Fig. 2 [models (4) and (5)], there is substantial variation across the four plots. The one consistent feature of the results for the levels models is that samples ending in 1982 through 1987 always exhibit significant money–income causality (the same is true for the differenced models in Fig. 1).[4] The most notable feature of the results for the levels specifications is the generally higher level of support for the money–income causality hypothesis than with the corresponding differences specifications.[5]

Many of the puzzles in the money–income causality literature are evident in Figs. 1 and 2. The Sims (1980) result that adding an interest rate to a VAR model

[3] As Stock and Watson note, if the variables contain unit roots, then causality tests computed from models specified in levels will typically have nonstandard F-distributions. However, the goal here is to examine the robustness of the causality tests to variations in sample period and model specification. Thus, the significance levels reported here are not corrected to account for the nonstandard distributions.

[4] A two-variable model including only money and output was also examined using data measured both in differences and levels. The results were similar to those obtained for the three-variable models. Also, quadratic and cubic trend specifications were examined for the two-, three-, and four-variable models. The results differed little from the linear trend specification.

[5] This statement should be interpreted in light of the qualification noted in Footnote 3.

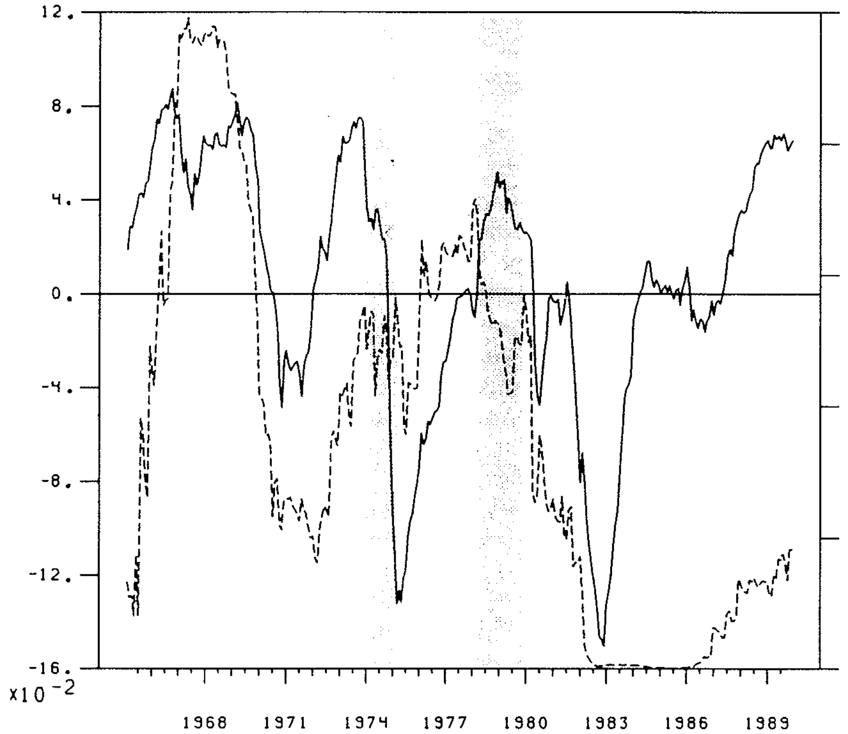


Fig. 3. Plot of p-values and business cycles.

specified in log levels overturns the result that money causes income can be seen by comparing Figs. 2a and 2c for the sample ending in 1978. The Eichenbaum and Singleton result that a log levels specification is more favorable to the money–income causality hypothesis than a differences specification is noted above and can be seen by comparing Figs. 1 and 2. The Eichenbaum and Singleton result that adding the early 1980's to the sample results in more significant money–income causality is evident in the figures which show that the p-values typically decrease as the early 1980's are included in the data. The significance of the results across Figs. 1 and 2 for samples ending in 1985 is consistent with the Stock and Watson results. The Friedman and Kuttner result that extending the sample to 1988 overturns the Stock and Watson result that money causes income can be seen in Fig. 1b, the Stock and Watson specification, as well as generally across Figs. 1 and 2.

The results displayed in Fig. 1 for the differenced models are generally similar. They do, however, indicate substantial variability in the p-values as the sample period changes. What is responsible for this variability? In Fig. 3 the p-values from the Stock and Watson specification (the six-lag model of Fig. 1b) are plotted along with deviations of the log of industrial production from a quadratic trend.[6] The correlation in the two series is striking. When output rises, the p-value for the causality statistic tends to rise (i.e., increase toward 1.0). Conversely, as output falls, the p-value for the causality statistic tends to fall, attaining its lowest level at the trough of cycles. Thus, Fig. 3 shows that time periods when output is falling are favorable to the money–income causality hypothesis, while time periods when output is rising are unfavorable. Generally, money causes income only when income is falling, not when income is rising.

There are two periods over which it is evident that cycles in industrial production are not positively correlated with the p-values. These time periods, which are shaded in Fig. 3, are 1973:11 to 1975:1 and 1978:4 to 1979:6, both of which are associated with large increases in the price of oil. At first glance, the time period 1983–1986 also appears to be anomalous. After a sharp decrease in the p-values during the recession of the early 1980's, the p-values are fairly constant from 1983 to 1986, even though output rises dramatically over the first half of this period. However, closer inspection reveals that the p-values do rise during this period, but the movement is extremely small.[7] Thus, except for two time periods when supply shocks were important, the two series display a marked positive correlation.

In sum, the results reported in this section document a high degree of variability in the significance of the relationship between money and income as

[6] Stock and Watson's conclusion that money causes income and that most of the puzzles in the money–income literature are resolved has proven to be controversial. See, for example, Friedman and Kuttner (1993) and Krol and Ohanian (1990). However, Stock and Watson's techniques used to specify the VAR model are widely, though not universally, accepted (see Sims, 1988). Thus, it seems reasonable to use the Stock and Watson specification as representative of the results for the models specified in differences. The measure of business cycles used is deviations of the log of industrial production from a quadratic trend. This is because the Stock and Watson specification includes a linear trend in log differences. This implies a quadratic trend in log levels.

[7] One explanation for small changes in significance levels is a form of persistence caused by the extremely high levels of significance attained in the early 1980's. That is, the relationship between money and income is so strong that a large number of additional observations that disagree with the hypothesis must be added to the data set before the relationship is noticeably weakened.

the sample period varies. The variation is systematic and consistent with the results of a number of earlier studies. Furthermore, a striking positive correlation between the p-values for the money–income causality relationship and the business cycle is uncovered. In the next section, the problem of assessing the significance of the amount of variation shown in Figs. 1 and 2 and the correlation shown in Fig. 3 is addressed. Then, in Section 4, a model that explicitly allows for the nonlinearities suggested in Fig. 3 is estimated and more direct tests for the nonlinearities are performed.

3. Bootstrap simulation

This section examines the significance of the variability in p-values and the correlation of the p-values with the level of real activity displayed in Fig. 3. The significance of the results in Fig. 3 is difficult to assess because no distribution theory exists that allows statistical tests to be performed. This is overcome through the use of bootstrap simulations to test two hypotheses, that the variability in the p-values is not statistically larger than the mean or expected variability, and that the correlation of the p-values with the level of real activity displayed in Fig. 3 is not statistically different from zero.

3.1. Methodology

The technique is to first estimate the constant coefficient VAR used in Fig. 3 over the entire sample period and save the 4×1 vector of estimated residuals $\hat{e}_t$, $t = 1, 2, \ldots, T$.[8] That is, estimate

$$X_t = \hat{F}(t) + \hat{A}(L)\, X_{t-1} + \hat{e}_t, \tag{6}$$

where $X_t = [\Delta Y_t, \Delta M_t, \Delta P_t, \Delta R_t]^T$, $A(L)$ is a matrix of polynomials in the lag operator, $F(t)$ is a 4×1 vector of linear trends, and $\hat{e}_t$ is a vector of estimated residuals. Next, randomly draw with replacement T vectors of residuals from the $\hat{e}_t$, $t = 1, 2, \ldots, T$, and form the new set of residuals $\hat{e}_t^*$. Then form the artificial observations

$$X_t^* = \hat{F}(t) + \hat{A}(L)\, X_{t-1}^* + \hat{e}_t^*, \tag{7}$$

where the actual values of X_t are used to prime the model and obtain the first twelve artifical observations. Finally, use the rolling regression technique described above on the bootstrapped data to generate a series of p-values and

[8] See Christiano and Ljungqvist (1988) for a justification of the bootstrap simulation procedures used here.

Table 1
Frequency that simulated statistics exceed empirical statistics

Statistic	Mean	Standard error	Empirical value	Frequency
σ_c^2	0.0381	0.0247	0.0871	4.72%
d_c	0.7214	0.1701	0.9912	0.28%
ρ_{yc}	− 0.0024	0.3291	0.5513	4.26%

σ_c^2 = the variance, d_c = the difference between the minimum and maximum values, and ρ_{yc} = the correlation between the significance levels and the deviations of industrial production from a quadratic trend. The mean and standard error are for the simulated statistics. The empirical values are the values from the results shown in Fig. 3. The frequency is the percentage of the simulated metrics that exceed the empirical metric. The statistics are based upon 5000 simulated values.

deviations of industrial production from a quadratic trend. This process is repeated 5000 times so that 5000 sets of p-values and deviations of industrial production from trend based on the artificial data are obtained.

To evaluate hypotheses, a metric over the 5000 sets of p-values and deviations of industrial production from trend is required. One provocative feature of Fig. 3 is the degree of variation in the p-values for the causality tests. Two metrics measuring the degree of dispersion are constructed to examine the question of whether the degree of variation in the p-values is larger than expected, the variance of the p-values (σ_c^2) and the difference between the minimum and maximum values (d_c) .

The second notable feature of Fig. 3 is the apparent high degree of association between the p-values and the deviations of industrial production from a quadratic trend. The metric used to examine the question of whether the degree of association is significant is the correlation between the two series (ρ_{yc}) .

The metrics are evaluated as follows. Calculate each of the metrics for each of the 5000 sets of artificial data. This gives 5000 observations of σ_c^2, d_c, and ρ_{yc}. These 5000 observations of each metric are then used to construct frequency distributions which are used to evaluate the significance of the metrics calculated using the actual data. There are two hypotheses of interest. First, that the degree of dispersion, measured as σ_c^2 or d_c, is not larger than the expected degree of dispersion (i.e., H_0: $\sigma_c^2 = \bar{\sigma}_c^2$ vs. H_1: $\sigma_c^2 > \bar{\sigma}_c^2$ or H_0: $d_c = \bar{d}_c$ vs. H_1: $d_c > \bar{d}_c$, where a bar indicates the expected value of the statistic). Second, that the correlation, measured as ρ_{yc}, is not different from zero (i.e., H_0: $\rho_{yc} = 0$ vs. H_1: $\rho_{yc} > 0$). These hypotheses are rejected at the 5% level if the statistic calculated using actual data lies within the right-hand tail of the frequency distribution of the statistic calculated from the bootstrapped data containing 5% of the area.

3.2. Results

Table 1 shows the results of the simulations and hypothesis tests. The table shows the values of the empirical metrics σ_c^2, d_c, and ρ_{yc}, the mean and variance of the simulated metrics, and the percentage of the simulated metrics greater than the empirical metrics. Figs. 4a–4c show the corresponding frequency distributions. The result that only 4.7% of the simulated σ_c^2's and 0.3% of the simulated d_c's exceed the empirical values indicates that the variation in p-values is significantly larger than the mean variation at the 5% level. For the correlation between the p-values and the deviations of industrial production from a quadratic trend, 4.3% of the simulated ρ_{yc}'s exceed the empirical value. Therefore, the correlation is significantly greater than zero at the 5% level.[9] Thus, the bootstrap simulations establish the significance of the variability in the outcomes of money–income causality tests across sample periods and the association between money–income causality statistics and real activity.

4. Asymmetries in the effects of money on output

Fig. 3 in conjunction with the bootstrap simulation results of the preceding section suggest that asymmetries are present in the relationship between money and industrial production. In this section, an explicit model incorporating the asymmetries is developed, tests for asymmetries are conducted, and forecasting exercises are performed. The technique used is to separate money growth into two groups, one containing positive changes in money growth and the other containing negative changes in money growth. The positive and negative changes in money growth are allowed to have differing effects on industrial production. Additionally, the coefficients measuring the impact of positive and negative changes in money growth are allowed to vary over the business cycle. The results of hypothesis tests and impulse response functions show that there is strong evidence for the presence of asymmetries and state dependency, a result that is supported by the forecasting exercises.

4.1. The empirical model

Two modifications to the basic VAR model are introduced in order to investigate the possible existence of asymmetries. The first allows real activity to respond asymmetrically to positive and negative changes in money growth. The

[9] The t-statistic from a regression of deviations of industrial production from trend on the significance levels gives exactly the same results as the correlation. The t-statistic from this regression is 11.16.

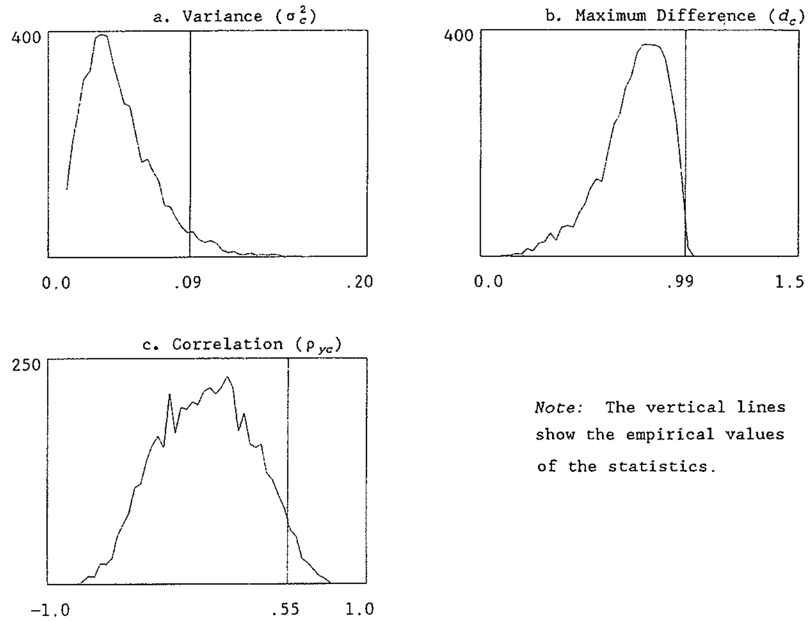


Fig. 4. Frequency distributions.

second allows the response of real activity to changes in money growth to vary systematically over the business cycle, where the business cycle is defined as deviations of output from a quadratic trend.

The modifications are carried out in two steps. First, the money growth series, ΔM_t, is divided into subsamples separating positive and negative changes in the money growth rate. To accomplish this, partition ΔM_t according to

$$\Delta M_t^+ = \begin{cases} \Delta M_t & \text{if } \Delta M_t - \Delta M_{t-1} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

and (8)

$$\Delta M_t^- = \begin{cases} \Delta M_t & \text{if } \Delta M_t - \Delta M_{t-1} \leqslant 0, \\ 0 & \text{otherwise.} \end{cases}$$

Now write the output equation in the VAR model (2) as

$$\Delta Y_t = f(t) + A_{YY}(L)\Delta Y_t + A_{YM}^+(L)\Delta M_t^+ + A_{YM}^-(L)\Delta M_t^- + A_{YP}(L)\Delta P_t + A_{YR}(L)\Delta R_t + e_t^Y, \tag{9}$$

where $f(t)$ is a linear trend. Allowing the coefficients on ΔM_t^+ and ΔM_t^- to differ allows for asymmetries in the response of output to positive and negative changes in money growth. Tests are conducted on this equation to determine whether the effects of positive changes in money growth differ from the effects of negative changes.

The second step in modifying the basic model is to allow the coefficients of the output equation to vary over the business cycle. To incorporate this effect, the coefficients on the money terms in the VAR model $A_{ym}^+(L)$ and $A_{ym}^-(L)$ are allowed to depend linearly on the state of the economy, S_t, defined as the deviation of the growth rate of the economy from a linear trend. This measure is consistent with the cyclical output measure shown in Fig. 3.[10] Note that

$$A_{YM}^+(L) = \sum_{k=1}^{K} a_{YMK}^+ L^k, \qquad A_{YM}^-(L) = \sum_{k=1}^{K} a_{YMK}^- L^k. \tag{10}$$

If we let

$$a_{YMk}^+ = b_{YMk}^+ + c_{YMk}^+ S_{t-k} \quad \text{and} \quad a_{YMk}^- = b_{YMk}^- + c_{YMk}^- S_{t-k}, \tag{11}$$

then

$$A_{YM}^+(L) = \sum_{k=1}^{k} (b_{YMk}^+ + c_{YMk}^+ S_{t-k})\, L^k = B_{YM}^+(L) + C_{YM}^+(L)\, S_t$$

and (12)

$$A_{YM}^-(L) = \sum_{k=1}^{k} (b_{YMk}^- + c_{YMk}^- S_{t-k})\, L^k = B_{YM}^-(L) + C_{YM}^-(L)\, S_t.$$

Therefore, the output equation in the VAR model becomes

$$\begin{aligned}\Delta Y_t = {} & f(t) + A_{YY}(L)\,\Delta Y_t + [B_{YM}^+(L) + C_{YM}^+(L)\, S_t]\Delta M_t^+ \\ & + [B_{YM}^-(L) + C_{YM}^-(L)\, S_t]\Delta M_t^- + A_{YP}(L)\,\Delta P_t \\ & + A_{YR}(L)\,\Delta R_t + e_t^Y.\end{aligned} \tag{13}$$

Rearranging terms gives

$$\Delta Y_t = f(t) + A_{YY}(L)\,\Delta Y_t + B_{YM}^+(L)\,\Delta M_t^+ + B_{YM}^-(L)\,\Delta M_t^-$$

[10] The measure of the cyclical position of the economy, S_t, is constructed from the residual of a regression of the growth rate of industrial production on a constant and time. Detrending the growth of industrial production prior to its use as regressor makes the procedure used here a two-step procedure. As such, it is subject to the problems associated with the use of generated regressors. This is discussed further in Footnote 13 below.

$$+ C_{YM}^{+}(L)\, S_t \Delta M_t^{+} + C_{YM}^{-}(L)\, S_t \Delta M_t^{-} + A_{YP}(L)\, \Delta P_t + A_{YR}(L)\, \Delta R_t + e_t^{Y}. \tag{14}$$

This suggests the VAR model

$$W_t = F(t) + D(L)\, W_t + e_t, \tag{15}$$

where $W_t = [\Delta Y_t, \Delta M_t^{+}, \Delta M_t^{-}, X_t^{+}, X_t^{-}, \Delta P_t, \Delta R_t]^T$, $X_t^{+} \equiv S_t \Delta M_t^{+}$, $X_t^{-} \equiv S_t \Delta M_t^{-}$, $D(L)$ is a 7×7 matrix of polynomials in the lag operator, and $F(t)$ is a vector of trend terms.

The hypotheses to be tested concern asymmetries in the effects of positive and negative changes in money growth on output as well as differential effects over the business cycle. Hypothesis tests are conducted as follows. First a VAR model including the variables Δy, ΔM^{+}, ΔM^{-}, ΔP, and ΔR is estimated. In this model, the effects of positive and negative changes in money growth are allowed to differ, but the effects are constant over the business cycle. The output growth equation,

$$\Delta Y_t = f(t) + D_{YY}(L)\, \Delta Y_t + D_{YM}^{+}(L)\, \Delta M_t^{+} + D_{YM}^{-}(L) \Delta M_t^{-} + D_{YP}(L)\, \Delta P_t + D_{YR}(L)\, \Delta R_t + e_t^{Y}, \tag{16}$$

is then subjected to the following coefficient restrictions

(i) H_0: $D_{YM}^{+}(L) = D_{YM}^{-}(L)$,

(ii) H_0: $D_{YM}^{+}(L) = 0$,

(iii) H_0: $D_{YM}^{-}(L) = 0$,

(iv) H_0: $D_{YM}^{+}(L) = D_{YM}^{-}(L) = 0$.

The first test examines whether the model allowing asymmetries in the response of output growth to positive and negative money growth changes is significantly different from the baseline model where asymmetries are not present. The second and third tests examine whether the positive and negative monetary growth terms are significant in explaining movements in output growth. The fourth test is a joint test of the significance of the positive and negative money growth terms for comparison with the significance of money growth in the baseline model.[11]

[11] The term baseline model refers to the constant coefficient, no asymmetries model used to generate the p-values shown in Fig. 3.

Second, a VAR model including the variables ΔY, ΔM^+, ΔM^-, X^+, X^-, ΔP, and ΔR is estimated. The output growth equation,

$$\Delta Y_t = f(t) + D_{YY}(L)\,\Delta Y_t + D_{YM}^+(L)\,\Delta M_t^+ + D_{YM}^-(L)\,\Delta M_t^- + D_{YX}^+(L)X_t^+ + D_{YX}^-(L)\,X_t^- + D_{YP}(L)\,\Delta P_t + D_{YR}(L)\,\Delta R_t + e_t^Y, \tag{17}$$

is then subjected to the coefficient restrictions

(i) H_0: $D_{YM}^+(L) = D_{YM}^-(L)$ and $D_{YX}^+(L) = D_{YX}^-(L) = 0$,

(ii) H_0: $D_{YM}^+(L) = D_{YX}^+(L) = 0$,

(iii) H_0: $D_{YM}^-(L) = D_{YX}^-(L) = 0$,

(iv) H_0: $D_{YM}^+(L) = D_{YX}^+(L) = D_{YM}^-(L) = D_{YX}^-(L) = 0$.

These tests parallel those described in the previous paragraph. The first test examines whether the model allowing asymmetries in the effects of positive and negative changes in money growth as well as asymmetries in the effects of changes in money growth over the business cycle is significantly different from a baseline model where asymmetries are not present. The second and third tests examine whether the positive and negative money growth terms are significant in explaining movements in output growth. The fourth test is a joint test of the significance of the positive and negative money growth terms for comparison with the significance of money growth in the baseline model.

Examination of the hypothesis that positive and negative changes in money growth have differing effects on real activity over the business cycle is accomplished through the use of impulse response functions. Eq. (13) is useful for explaining the procedure used to obtain the impulse response functions. In this equation, the coefficients on ΔM_t^+ and ΔM_t^- are $B_{YM}^+(L) + C_{YM}^+(L)S_t$ and $B_{YM}^-(L) + C_{YM}^-(L)S_t$. Because these coefficients depend upon the state of the economy, S_{t-j}, $j = 1, 2, \ldots, K$, the impulse response functions also depend upon the state of the economy. By choosing three different values for the S_{t_0-j}, $j = 1, 2, \ldots, K$, where t_0 is the time period in which the shock occurs, impulse response functions are obtained for three initial states of the economy. The initial values chosen correspond to real activity at its potential (S_{t_0-j} = two standard deviations above the mean, $j = 1, 2, \ldots, K$), at its average level (S_{t_0-j} = the mean of the S_t, $j = 1, 2, \ldots, K$), and in a recession (S_{t_0-j} = two standard deviations below the mean, $j = 1, 2, \ldots, K$).

The impulse response functions show the direction of the effect of positive and negative changes in money growth on real activity, and also show how the effects of positive and negative changes in money growth vary over the business cycle.

4.2. Results for the output equation allowing asymmetries

The VAR models allowing asymmetries described in the previous section are estimated and the hypotheses are tested using the same data, variables, trend specification, and the longest sample period used to obtain the results shown in Fig. 3.

4.2.1. Does money growth have asymmetric effects on output?

Table 2 presents the hypothesis tests. In model A, the effects of positive and negative changes in money growth are allowed to differ, but the effects do not vary over the business cycle. In model B the effects of positive and negative money growth changes are allowed to differ as in model A, and the effects are allowed to vary over the business cycle.

Five F-statistics corresponding to the hypotheses described in Section 4.1 are presented for each model. The top panel of Table 2 shows the results for model A. Note first the p-value of 0.182 for money in the baseline model in which no asymmetries are present; the data fail to reject the hypothesis that money does not cause output. Next is the p-value of the test of the joint hypothesis that the positive and negative money growth terms are jointly zero. With a p-value of 0.142, the data also fail to reject this hypothesis. Thus, allowing the effects of the positive and negative money terms to differ does not overturn the insignificance of money at the 5% level in the baseline model. As the next two statistics reported in the panel show, individual tests of the significance of the positive and negative money terms do not overturn this result either. The p-values for these tests are 0.327 and 0.120. Finally, the last entry in the panel indicates that the model in which the effects of positive and negative changes in money growth are allowed to differ is not significantly different from the baseline model, which allows no asymmetries. The p-value for the test of the restrictions necessary to obtain the baseline model is 0.213.

The results for model B, which allows the effects of positive and negative changes in money growth to vary over the business cycle are more interesting. At a 5% level of significance, the data fail to reject only the hypothesis that positive changes in money growth do not affect output. The p-value for the test of the hypothesis that the effects of positive and negative changes in money growth are jointly zero is 0.015, much higher than for the baseline model. Thus, allowing for cycle dependent asymmetries overturns the result that changes in money growth do not affect the growth rate of output. The p-values for tests of the next two hypotheses, that positive and negative changes in money growth individually do not affect output growth, are 0.074 and 0.025. This indicates that it is negative changes in money growth that are responsible for the movements in output. Finally, the last hypothesis examined in the table, that the restrictions necessary to

Table 2

Tests for asymmetries in the output equation

(A) $\Delta Y_t = f(t) + g(\Delta Y, \Delta M^+, \Delta M^-, \Delta P, \Delta R)$

Money in baseline ($\Delta M = 0$)	Pos. and neg. shocks zero ($\Delta M^+ = \Delta M^- = 0$)	Pos. shocks zero ($\Delta M^+ = 0$)	Neg. shocks zero ($\Delta M^- = 0$)	Relative to baseline ($\Delta M^+ = \Delta M^-$)
$F_{6,315} = 1.487$ (0.182)	$F_{12,309} = 1.451$ (0.142)	$F_{6,309} = 1.162$ (0.327)	$F_{6,309} = 1.702$ (0.120)	$F_{6,309} = 1.402$ (0.213)

(B) $\Delta Y_t = f(t) + g(\Delta Y, \Delta M^+, \Delta M^-, X^+, X^-, \Delta P, \Delta R)$

Money in baseline ($\Delta M = 0$)	Pos. and neg. shocks zero ($\Delta M^+ = \Delta M^- = X^+ = X^- = 0$)	Pos. shocks zero ($\Delta M^+ = X^+ = 0$)	Neg. shocks zero ($\Delta M^- = X^- = 0$)	Relative to baseline ($\Delta M^+ = \Delta M^-, X^+ = X^- = 0$)
$F_{6,315} = 1.487$ (0.182)	$F_{24,297} = 1.781$ (0.015)	$F_{12,297} = 1.664$ (0.074)	$F_{12,297} = 1.992$ (0.025)	$F_{18,297} = 1.855$ (0.019)

The baseline model is $\Delta Y_t = f(t) + g(\Delta Y, \Delta M, \Delta P, \Delta R)$. The p-values are in parentheses.

obtain the baseline model are satisfied, is rejected at the 5% level. The p-value for this test is 0.019.[12]

The results shown in Table 2 for model B verify the presence in the data of the asymmetries suggested in Fig. 3. Fig. 3 shows that the effects of money strengthens as output falls and weakens as output rises. Table 2 shows that negative changes in money growth, which are shown in the next section to cause output to fall, significantly affect output growth, while positive shocks do not. The results also suggest that simply separating money growth according to whether it is rising or falling by itself does not overturn the insignificance of money in the baseline model. The effects of money must also be allowed to vary over the business cycle.[13]

4.2.2. *Do the effects of money vary over the business cycle?*

This section presents impulse response functions that allow a more complete assessment of the effects of the positive and negative changes in money growth on output. Figs. 5a–5f show the effects of both positive and negative changes in money growth on output over a 48-month horizon. Three separate impulse response functions along with two standard deviation confidence bands are shown for each case. The three functions represent three initial states of the economy, industrial production at its potential level, in a typical recession, and at its average level.

The results of Table 2 are evident in Fig. 5. Positive changes in money growth, shown in Figs. 5a–5c, have very little impact on the level of output, and the initial state of the economy has little impact on the impulse responses. In no case can the hypothesis of a zero response of output be rejected based upon the confidence bands. However, negative changes in money growth, shown in

[12] Because F-statistics are quite sensitive to the presence of serial correlation in the residuals, the models used to construct Table 2 are subjected to Ljung–Box Q tests. The presence of first-, sixth-, and twelfth-order autocorrelation in the residuals is not detected in any case. The p-values for these tests are 0.97, 0.95, and 0.96 for model A, 0.94, 0.99, and 0.96 for model B, and 0.91, 0.91, and 0.97 for the baseline model. Similar tests applied to the other models used in the paper (e.g., the models used to construct Figs. 1 and 2) yield similar results.

[13] As noted above, the procedure used here is a two-step procedure where the cyclical measure for the economy, S_t, is generated in the first step for use as a regressor in the output equation in the second step. With the two-step procedure, the error in estimating the coefficients in the first step is not accounted for in the second step. Since there is error that is unaccounted for, the estimated standard errors in the second step are too small causing the null hypothesis to be rejected too often. Because the interesting results reported in Table 2 involve rejections of the null hypothesis, the use of a generated regressor could constitute a problem. To check for this, a nonlinear least squares estimator that jointly estimates the cyclical position of the economy, S_t, and the coefficients of the output equation and thus avoids the generated regressors problem was used to estimate model B of Table 2 and the hypotheses tests were repeated using likelihood ratio tests. The conclusions were identical to those obtained with the two-step approach.

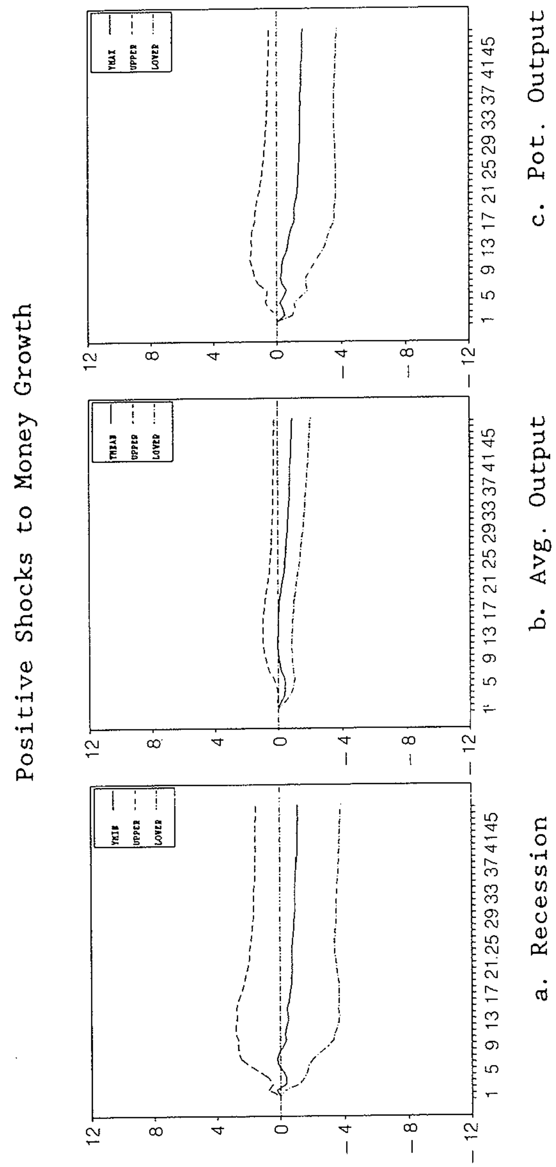
Figs. 5d–5f, have a much larger impact on the level of output, and the initial state of the economy affects the impulse response. When the economy is in a recession or at average output as in Figs. 5d and 5e, a negative change in money growth does not have an effect on output that is significantly different from zero. When industrial production is at its potential level as in Fig. 5f, a negative change in money growth brings about a decline in output for 12 months followed by a return to its original level at 48 months, and the decline in output is statistically different from zero from 7 to 17 months after the shock.

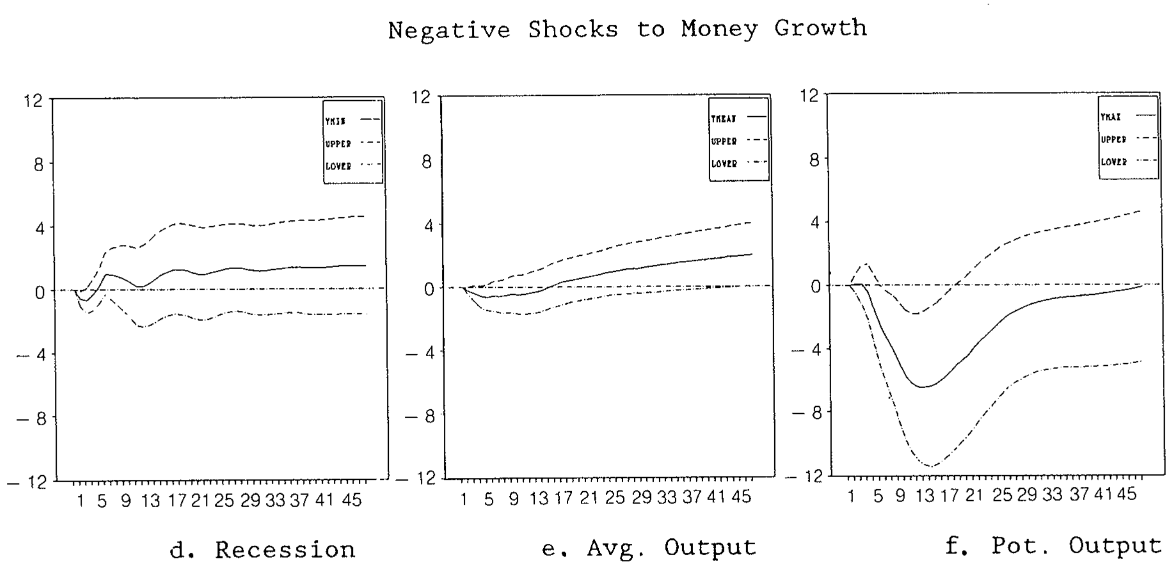
4.3. Relative forecasting ability

The model that allows the effects of positive and negative changes in money growth to differ, and also allows the effects to vary with the state of the economy, appears to fit the data better than a traditional VAR model. However, finding a more flexible model that fits the data better in sample does not guarantee that the model will produce better out of sample forecasts. This section compares the out of sample predictive power of the three models discussed earlier in this section, a traditional VAR model as shown in (2), the model shown in (9) where positive and negative changes in money growth can have different effects, and the model in (15) where positive and negative changes in money growth can have different effects and the effects can be state-dependent. Comparing the predictive power of these models allows assessment of whether the more flexible models are associated with a decline in predictive power.

The forecasting exercise is conducted by first estimating each model through 1984:12, then producing forecasts from each model from 1 to 24 months ahead. Then, the models are estimated through 1985:1, and forecasts from 1 to 24 months ahead are generated. This process of adding one month to the sample and producing forecasts from 1 to 24 months ahead is continued until the models are estimated through 1987:12, and forecasts for 1988:1 through 1989:12 (the end of the sample) are produced. This yields 36 observations for each forecast step for each of the three models. Finally, root mean square forecast errors are calculated for each of the 24 forecasting steps for each model. In this procedure, the beginning of the forecast period is 1985:1. The exercise is repeated for a forecast period beginning at 1982:1.

Table 3 presents the root mean square forecast errors for three models, model (a) is a baseline standard VAR model, model (b) allows the effects of positive and negative changes in money growth to differ, and model (c) allows the effects of positive and negative changes in money growth to differ and allows for state dependency in the effects of positive and negative changes in money growth. Two out of sample forecast periods are examined for each model, the first beginning in 1985:1, and the second in 1982:1. Forecasts for 1, 4, 8, 12, 16, 20 and 24 steps ahead are presented for each case. Table 3 shows that the root mean square forecast errors are generally lowest for model (c), but not uniformly.


Positive Shocks to Money Growth
12
8
4
0
−4
−8
−12
1 5 9 13 17 21 25 29 33 37 41 45
a. Recession
b. Avg. Output
c. Pot. Output
UPPER
LOWER



Note: The responses shown are in levels (produced by integrating the response of the growth rate)

Fig. 5. Impulse response functions for output.

Table 3
Root mean square forecast errors

Beginning of out-of-sample forecast period						
	1985:1			1982:1		
	Model			Model		
Step	(a)	(b)	(c)	(a)	(b)	(c)
1	0.0878	0.0876	0.0821	0.1013	0.1014	0.0995
4	0.0946	0.1013	0.0991	0.1045	0.1071	0.1020
8	0.0861	0.0875	0.0834	0.1000	0.0987	0.0969
12	0.0759	0.0748	0.0717	0.0782	0.0764	0.0736
16	0.0534	0.0532	0.0529	0.0725	0.0711	0.0675
20	0.0484	0.0481	0.0486	0.0665	0.0662	0.0672
24	0.0460	0.0452	0.0441	0.0605	0.0603	0.0603

(a) indicates the baseline model, (b) the model with positive and negative changes in money growth, and (c) the model with both positive and negative changes in money growth and state dependency in the effects of money. The entries in the table are multiplied by a factor of 10.

Table 4
Percentage change in the root mean square forecast errors

	% change for models (a) and (b)		% change for models (a) and (c)	
Step	1985:1	1982:1	1985:1	1982:1
1	0.30	− 0.12	6.53	1.68
4	− 7.02	− 2.50	− 4.73	2.43
8	− 1.60	1.39	3.10	3.18
12	1.51	2.41	5.51	5.98
16	0.45	1.93	0.92	6.92
20	0.69	0.38	− 0.24	− 1.12
24	1.82	0.24	4.10	0.29

The definitions of (a), (b), and (c) are the same as in Table 3. The entries on the left side of the table are the percentage change between columns (b) and (a) of Table 3, and the entries in the right side are the percentage between columns (c) and (a). The column headings indicate the beginning of the forecast period.

This can be seen more clearly in Table 4 where the percentage change in the root mean square errors relative to baseline are presented. The left-hand side of the table shows the percentage change in the root mean square error from using model (b) rather than model (a), and the right-handside the percentage change from using model (c) rather than model (a). A positive entry indicates that model (b) or (c) has a lower mean square error than model (a), and a negative entry indicates that model (a), the traditional VAR model, has the lower root mean square error.

On the left-hand side of the table, where the predictive power of model (b) is compared to the predictive power of model (a), the results are mixed. Model (b) forecasts better at longer horizons of 12–24 months, but not at shorter horizons of 1–8 months. In all cases where model (b) forecasts better, the percentage changes are modest. Overall, it is difficult to make a case that model (b) forecasts better than model (a).

On the right-hand side of the table, where model (c) is compared to model (a), the results are stronger. Across both forecasting periods, and across the forecasting horizons model (c), which allows the effects of positive and negative changes in money to differ and allows for state dependencies, is generally superior.

These forecasting results support earlier findings in this paper (e.g., see Table 2) that separating money into positive and negative components does not by itself produce a model that explains the data better than a traditional VAR. It is also necessary to allow for state dependencies in the effects of money.

5. Conclusions

This paper provides a detailed examination of the stability of money–income causality tests across sample periods, over the business cycle, and across the specification of the VAR model used in the tests. When the model is specified in differences, the causality tests exhibit a substantial amount of variation across sample periods. The results are generally robust to variations in the specification of the trend, the number of lags, and to the exclusion of interest rates. The exception is the popular four-variable log difference model without a trend where the results are relatively stable. When the model is specified in levels, the results are generally more significant than with the differences specification, but the p-values of the causality tests vary substantially across sample periods, the specification of the trend, and to the exclusion of interest rates. The one robust conclusion that is exhibited across all specifications is the significance of the relationship between money and income for samples ending from 1982 to 1987.

Assessment of the importance of these results is conducted through bootstrap simulations, the construction of an empirical model where asymmetries exist in the relationship between money and real activity, and forecasting exercises. The results support the conclusions that there is a high degree of variability in the outcome of causality tests across sample periods, and that the variability is explained by movements in real activity. The results also indicate that the effects of money vary over the business cycle.

Many of the puzzles present in the money–income causality literature are due to the subsample instability examined in this paper. For example, the Eichenbaum and Singleton result that adding the early 1980's to the data makes the money–income relationship stronger and the Friedman and Kuttner result that adding the late 1980's weakens the relationship are both a result of the lack of robustness in the causality statistics to variations in the sample period. Thus, providing an explanation as to why the test statistics vary as the sample period is extended is of paramount importance in the resolution of these puzzles. Whether the most recent vintages of the models can explain the empirical regularity illustrated in Fig. 3 and in Table 2 is not entirely obvious, but surely the results presented here pose an interesting challenge to these models.

References

Christiano, Lawrence J. and Lars Ljungqvist, 1988, Money does Granger-cause output in the bivariate money–output relation, Journal of Monetary Economics 22, 217–235.

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

Friedman, Benjamin M. and Kenneth N. Kuttner, 1993, Another look at the evidence on money–income causality, Journal of Econometrics 57, 189–203.

Krol, Robert and Lee E. Ohanian, 1990, The impact of stochastic and deterministic trends on money–output causality: A multi-country investigation, Journal of Econometrics 45, 291–308.

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

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

Sims, Christopher A., 1988, Bayesian skepticism on unit root econometrics, Journal of Economic Dynamics and Control 12, 463–474.

Stock, James H. and Mark W. Watson, 1989, Interpreting the evidence on money–income causality, Journal of Econometrics 40, 161–181.