DANIEL L. THORNTON
DALLAS S. BATTEN

Lag-Length Selection and Tests of Granger Causality between Money and Income

1. INTRODUCTION

ONE OF THE MOST WIDELY ACCEPTED RELATIONSHIPS IN macroeconomics is the positive relationship between the rate of growth of the money supply and the rate of growth of income. A diversity of opinion arises, however, with the attempt to identify a causal relationship between these variables. While Friedman and Schwartz (1963) initiated the empirical investigation of money/income causality, Sims (1972) made the most significant contribution to this literature with a practical application of the criterion developed by Granger (1969) for identifying a causal relationship. Since the appearance of Sims's approach, and an alternative one advanced by Sargent (1976), many researchers have applied these techniques to the investigation of the causal relationship between money and income.[1] With two exceptions, Hsiao (1981) and McMillin and Fackler (1984), these researchers have arbitrarily chosen the lengths of their distributed lags.

The purpose of this paper is to test for Granger causality between money and income. This paper differs from previous work in that considerably more attention is focused on the effect of the lag specification on the test results. In particular, we investigate the extent to which rules of thumb, arbitrary lag specifications, and

The authors would like to thank Rik Hafer, Stan Johnson, Doug Pearce, and two anonymous referees for helpful comments and John Schulte for his research assistance. The views expressed here are the authors' and do not necessarily reflect those of the Federal Reserve Bank of St. Louis or the Board of Governors of the Federal Reserve System.

[1]See, for example, Elliott (1975), Ciccolo (1978), Mehra (1978), Feige and Pearce (1979), Mehra and Spencer (1979), Pautler and Rivard (1979), Sims (1980a, 1980b), Carlson and Hein (1980), Hafer (1981), Hsiao (1981), Davidson and Hafer (1983), and McMillin and Fackler (1984).

DANIEL L. THORNTON and DALLAS S. BATTEN *are senior economist and research officer, respectively, Federal Reserve Bank of St. Louis.*

Journal of Money, Credit, and Banking, Vol. 17, No. 2 (May 1985)

statistically determined lag structures can be relied upon in testing Granger causality between money and income. We find that ad hoc approaches, such as considering a few arbitrary lag structures or employing some "rule of thumb," can produce misleading results. Moreover, these approaches ignore the prominent role that model specification should play in causality testing. We also show that the problem of lag specification is not circumvented by using statistical criteria for determining the lag structure. Not surprisingly, models selected by different statistical criteria yield contradictory conclusions concerning the Granger-causal ordering between money and income.

The remainder of the paper is composed of three sections. Section 2 contains a brief presentation of Granger-causality testing with special emphasis given to the specification of the lag structure. The results of Granger-type causality tests for the money/income relationship are presented in section 3. These tests are conducted using lag structures determined by three model-specification criteria, as well as over an extensive portion of the lag space. Some conclusions are presented in section 4.

2. THE MODEL

A standard bivariate dynamic structural model on which the Granger test is based can be expressed as

$$\begin{aligned} y_t + \alpha x_t &= L(\beta)^H y_{t-1} + L(\delta)^K x_{t-1} + \varepsilon_{1t} \\ x_t + \delta y_t &= L(\lambda)^I x_{t-1} + L(\mu)^N y_{t-1} + \varepsilon_{2t}\,. \end{aligned} \tag{1}$$

In the maintained structure, x and y are jointly determined endogenous variables, and ε_1 and ε_2 are assumed to be $iid(0, \sigma_i^2)$, $i = 1, 2$. For the sake of simplicity let $E(\varepsilon_{1t}\varepsilon_{2t'}) = 0$ for all t and t', and let $L(\cdot)^J$ denote the polynomial lag operator of order J, for example, $L(\beta)^J y_{t-1} = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_J y_{t-J}$. The reduced form of this model is

$$\begin{aligned} y_t &= L(\pi_{11})^G y_{t-1} + L(\pi_{12})^P x_{t-1} + u_{1t} \\ x_t &= L(\pi_{21})^G y_{t-1} + L(\pi_{22})^P x_{t-1} + u_{2t}\,, \end{aligned} \tag{2}$$

where G and P are the larger of H or N and K or I, respectively, and the πs are nonlinear functions of the structural parameters in (1). The equations in (2) are the standard equations for the Granger variant of the causality test.

To examine Granger causality between x and y, the following hypotheses are tested: $L(\pi_{12})^P = 0$ and $L(\pi_{21})^G = 0$. If neither can be rejected, then x and y are independent series. If both are rejected, then there is "feedback" between x and y. If the former hypothesis is rejected but the latter is not, there is unidirectional causality running from x to y, whereas if the latter is rejected and the former is not, the reverse is true.

Tests of the above hypotheses are critically dependent on the unknown parameters P and G. It is always possible to fail to reject either of the above hypotheses by selecting sufficiently large values for these parameters. Moreover, it is reasonable to suspect that the same might be true by choosing sufficiently small values for these parameters as well. Since these parameters are generally unknown, there are several approaches that could be taken in testing for Granger causality. An agnostic approach would be to accept the hypothesis of unidirectional causality only if either of the hypotheses, $L(\pi_{12})^P = 0$ or $L(\pi_{21})^G = 0$, can be rejected for all values of P and G over some reasonable range of values for these parameters; independence would hold if neither can be rejected.[2] While such a search can be conducted in a reasonably efficient manner for the bivariate case, some may consider it too burdensome.[3] Furthermore, there may exist some lag specifications for which the null hypothesis can be rejected, at say the 5 percent level, by chance alone.

Two alternatives to a complete search of the lag space are (a) to identify the "appropriate" lag structure for the equations in (2) based on some statistical criterion or (b) to specify a few alternative lag structures (the most common are 4-4 and 8-8). The former approach is followed by Hsiao(1981). There are a number of statistical criteria, however, on which the selection of these parameters can be based.[4] Consequently, different causality test results could be obtained from the same data due to the use of different statistical criteria for choosing values of P and G. Thus, individuals could arrive at different, but equally legitimate, conclusions concerning the Granger-causal relationship between time series due solely to differences in their lag-length selection criteria.[5]

3. LAG-LENGTH SELECTION

To identify the appropriate lag structure one must first choose among a wide variety of model selection criteria. The typical model selection criteria trade off the bias associated with a parsimonious parameterization against the inefficiency associated with overparameterization. Because various criteria give different weights to the bias/efficiency trade-off, they can select quite different lag structures. In general, if either P or G (or both) is too large, the estimates will be unbiased but inefficient.

[2]In the classical framework of hypothesis testing, the null hypothesis is constructed as a proposition which the researcher wants to reject. In the case of causality testing, however, the null hypothesis cannot always be specified in this manner. Thus, an agnostic, who does not care about the order of the model within which the hypothesis is rejected, could be characterized as possessing priors toward finding evidence of causality.

[3]The F-statistics for a bivariate model can be calculated in an efficient manner using orthogonal regressions. See Batten and Thornton (1983a).

[4]For a discussion of some of these and references to others, see Judge, et al. (1980, chap. 11).

[5]A question arises naturally: why should different criteria be used for determining the lag length and for the tests of Granger causality? Certainly the same classical hypothesis-testing framework that normally is employed in tests of Granger causality could be employed in lag-length selection. Likewise, if one preferred an alternative criterion, say mean square error, the same criterion could be applied to tests of Granger causality. It may be, however, that Granger causality is only one of several questions to be addressed. Furthermore, some of the commonly used criteria for identifying the order of a model do not have a hypothesis-testing counterpart. Whatever the reason, the case where the objective functions differ is considered here.

If either P or G (or both) is too small, the estimates will be biased but will have a smaller variance. If one is too large and the other too small, the estimates will be biased and may be inefficient.

Three different criteria for specifying the lag length are used here. The first is Akaike's (1970) final prediction error (FPE) suggested by Hsiao (1981). It gives relatively more importance to unbiasedness over efficiency, but is asymptotically inefficient in that, on average, it selects lags that are too long in large samples. The second is the Bayesian estimation criterion (BEC) suggested by Geweke and Meese (1981). On average it selects the correct lag asymptotically and, hence, is asymptotically efficient. The experience of, for example, Batten and Thornton (1983b) and Geweke and Meese (1981) indicates that it tends to select lags that are too short in finite samples. These criteria are chosen because they are commonly suggested and because they provide interesting extremes in the bias/efficiency trade-off.[6] A third technique, suggested by Pagano and Hartley (P-H) (1981), is also used. This method has performed well in the past and provides some computational efficiency useful in searching the lag space. In this application, the P-H technique is similar to a standard F-test.[7]

The above criteria are applied to bivariate Granger-type equations involving nominal income (Y) and the monetary aggregates, *M1*, *M2*, *M2* net of *M1* (*NM2*), and the adjusted monetary base (*MB*). All variables are expressed in compounded annual growth rates. The equations are estimated over the period from 1962.II–1982.III.[8] The autocorrelation functions suggest that all of the univariate time series, except *MB*, are covariance stationary; nevertheless, a time trend is included in all regressions.

The lag-length selection results are reported in Table 1. They show that the BEC selects very short lag structures. As expected, the FPE criterion selects longer lags than those selected by the BEC, but, in general, somewhat shorter than those selected by the P-H technique.

A natural question is: How well do these criteria perform in model specification? Rather than considering all possible combinations, the FPE-selected models are compared with those selected by the other criteria and with the commonly used arbitrary lags of 4-4 and 8-8. The results of this comparison are reported in Table 2. The models that are not nested are identified by NN. If the FPE structure is of a lower order than the alternative, it is denoted by (L); if it is of a higher order, it is denoted by (H). If the FPE criterion always picks the model that is preferred by this classical hypothesis testing norm at the 5 percent significance level, all the Fs designated

[6]If there are only two alternative lag specifications to be considered, p and $p + q$, the FPE criterion would pick p if $F < 2T/(T + p + 1)$, while the BEC criterion would select p if $F < 1 + ((T - p - 1) \ln T/(T - p - q - 1))$. Thus, the FPE is the more powerful test procedure of the two. Clearly, if one wishes to give more weight to bias in the bias/efficiency trade-off, the FPE criterion would be preferred to the BEC. Unfortunately, the relative power of the FPE and P-H procedures cannot be determined.

[7]See the appendix to Batten and Thornton (1983a) for details.

[8]It appears that there may have been a change in the usual income-money relationship in 1982.IV and 1983.I; see Batten and Thornton (1983c). The sample ended with 1982.III so that the results would not be affected by this development. The beginning of the sample was determined by the availability of the *M2* series. A maximum of twelve lags was considered for each variable in determining these lag specifications.

TABLE 1

LAG LENGTHS SELECTED BY VARIOUS CRITERIA

Dependent Variable/ Independent Variable	FPE	BEC	P-H
Y/M1	0/2	0/1	7/10
M1/Y	8/7	0/0	6/6
Y/M2	0/1	0/1	11/6
M2/Y	1/12	1/0	1/12
Y/NM2	0/1	0/1	11/1
NM2/Y	1/12	1/0	1/12
Y/MB	0/3	0/1	0/3
MB/Y	1/0	1/0	1/11

TABLE 2

F-TESTS OF THE FPE LAG SPECIFICATIONS

Dependent Variable/ Independent Variable	4-4	8-8	BEC	P-H
Y/M1	0.09 (*L*)	0.71 (*L*)	5.72*(*H*)	0.97 (*L*)
M1/Y	3.46*(*H*)	0.33 (*L*)	2.33*(*H*)	2.38 (*L*)
Y/M2	0.37 (*L*)	0.75 (*L*)	—	1.80 (*L*)
M2/Y	NN	NN	3.04*(*H*)	—
Y/NM2	0.35 (*L*)	0.60 (*L*)	—	1.66 (*L*)
NM2/Y	NN	NN	3.06*(*H*)	—
Y/MB	0.22 (*L*)	0.70 (*L*)	2.82 (*H*)	—
MB/Y	0.93 (*L*)	0.84 (*L*)	—	1.40 (*L*)

Notes: NN = not tested. A dash indicates that the lag is the same as that chosen by the FPE.
*Significance at the 5 percent level.

H would be significant, while those designated *L* would not. Since this is true with only one exception, the FPE criterion appears to perform well, at least for these data.[9]

4. TESTS OF GRANGER CAUSALITY

Since we are interested in how the causality test results may vary across various specifications, standard F-tests for Granger causality were performed on all of the lag-length specifications of Table 1 and for the arbitrarily chosen lags of 4-4 and 8-8; the results are presented in Table 3.[10] The results of the F-tests vary considerably across specifications. For example, in the test of nominal income on *M1*, the hypothesis that the four lag coefficients on *Y* are simultaneously zero cannot be rejected for the 4-4 specification, while the test of the eight lags on *Y* are zero is rejected for the 8-8 specification. Moreover, the BEC criterion selects no lag for *Y*, suggesting that nominal income does not Granger-cause *M1* by this criterion; however the F-tests are significant, indicating Granger causality from *Y* to *M1*, for both FPE- and PH-determined lag structures.

[9]Similar results are found in Batten and Thornton (1983b) and Hsiao (1981).
[10]There are a number of other ways this test can be performed; however, results by Geweke, Meese, and Dent (1983) and Guilkey and Salemi (1982) indicate that the Granger variant of the test is preferred.

TABLE 3

Granger Causality F-statistics for Various Lag Specifications

Dependent Variable/ Independent Variable	4-4	8-8	Model Chosen By BEC	FPE	P-H
Y/M1	4.10*	2.91*	14.46*	10.52*	3.06*
M1/Y	1.43	2.32*	NL	2.63*	2.88*
Y/M2	2.54*	2.03	11.68*	11.68*	4.72*
M2/Y	1.50	2.31*	NL	3.04*	3.04*
Y/NM2	1.46	1.24	7.04*	7.04*	13.34*
NM2/Y	1.84	1.85	NL	3.06*	3.06*
Y/MB	4.87*	3.06*	16.95*	7.79*	7.79*
MB/Y	0.95	0.81	NL	NL	1.40

NL = no lag chosen for the independent variable.
*Significant at the 5 percent level.

There are several interesting general observations that emerge from the results in Table 3.[11] First, there are three instances when the commonly used lags of 4-4 and 8-8 produce contradictory results. Second, there are two instances in which the null hypothesis is not rejected for the arbitrarily chosen lag structures but is rejected for both the FPE- and PH-determined lag structures. Third, the BEC indicates unidirectional causality running from each monetary aggregate to nominal GNP. There is only one monetary aggregate for which this result is borne out by the other lag specifications—the case of the adjusted monetary base.

In order to investigate further the sensitivity of these tests to lag-length specification (and to consider the results that would be obtained if the agnostic approach were employed), the significance levels corresponding to calculated F-statistics of Granger tests over a large segment of the lag space are presented in Tables 4–7. The significance levels are given for tests that all the coefficients on the hypothesized exogenous variables are zero for all possible lags of the dependent variable up to order twelve. The outcomes that yield a rejection of the null hypothesis at the 5 percent level are indicated with an asterisk. (The reader may wish to note how these regions would have changed had a 10 percent significance level been chosen.) The significance levels are presented rather than the corresponding F-statistics because the latter are not degree-of-freedom invariant.

These results show that the outcome of causality tests between money and income are unquestionably sensitive to the lag specification. Also, they suggest that researchers should not rely on arbitrarily chosen lag lengths, so common in applications of these tests. For example, in tests of *M2* on *Y* the lag structure of 4-4 suggests that income does not Granger cause *M2*, while the 8-8 lag structure yields the opposite conclusion. Had slightly longer lags been used, say 5-5 and 9-9, these conclusions would have been reversed. In other cases the null hypothesis cannot be

[11]The reader familiar with Hsiao's work may find it odd that we performed F-tests on the FPE-selected lags. Hsiao has indicated that conditions for Granger causality are satisfied if $\mathrm{FPE}(X, Y) < \mathrm{FPE}(X, \mathrm{O})$. The FPE criteria, however, can be interpreted as simply applying higher than conventional significance levels. Thus, it does not, as Hsiao implies, eliminate the need for statistical testing if one wishes to maintain conventional significance levels; see note 15 below.

rejected for one or both of the arbitrary lag specifications, even though it can be rejected over a considerable portion of the lag space.[12]

The evidence also suggests that the intuitively appealing rule of thumb — that the lag on the dependent variable be relatively long to account for possible autocorrelation, while the lag on the hypothesized independent variable be relatively short to conserve degrees of freedom — does not perform well either. The rationale for this approach is that if the proposed independent variable exerts a significant effect, it should be captured by a few lags. If one is content to reject the null hypothesis if it can be rejected over any part of the lag space, this rule of thumb would have yielded unsatisfactory conclusions for the tests of *M1, M2,* and *NM2* on *Y*.[13]

The evidence based on the specifications chosen by either the FPE or P-H and on the results in Tables 4–7 suggests that there is bidirectional Granger causality between *M1, M2,* and *NM2* and nominal GNP over this period.[14] These results are in contrast to those recently reported by Hsiao (1981) and McMillin and Fackler (1984), which show bidirectional causality between *M1* and income and unidirectional causality running from *M2* to income.[15]

The only case of unidirectional causality is for the monetary base. There is no portion of the lag space where GNP exerted a significant effect on the monetary base, while the reverse is true for nearly every lag structure considered. As noted above, however, *MB* does not appear to be covariance stationary. Consequently, the tests were repeated using first differences of the growth rates of these variables.[16] The results appear in Table 8. These results indicate that the relevant hypotheses can be rejected over some segment of the lag space. The agnostic may conclude that the tests indicate bidirectional causality between all the monetary aggregates and income. The Box-Ljung Q-test on the specifications of ΔMB on ΔY with no own lags, however, indicates serially correlated residuals. Consequently, conducting hypothesis tests using these residuals is inappropriate.[17]

The remaining two specifications do not exhibit this problem, but OLS estimates of these equations produce results that are difficult to reconcile with the F-tests.

[12]Some other common arbitrary specifications are 4-8 and 8-4 and some combination of 4-4, 6-6, or 8-8. The reader can see that these choices do not fare any better.

[13]It could be argued that tests where the lag on either variable is short lack power because the short lags introduce serial correlation into the residuals. This observation does not militate against our result, but against the usefulness of such a rule. Nevertheless, Box-Ljung Q-tests of a sampling of these shorter-lag specifications produced no evidence of misspecification.

[14]This result gives further evidence to a view expressed by Gordon (1984) and others that the standard, short-run, money-demand equation may be a confluence of money-demand and supply-side forces.

[15]Actually, Hsiao finds bidirectional causality between *M1* and *Y* only because the use of the FPE criterion effectively increases the significance levels above those commonly used. The coefficient on the one lag of *Y* on *M1* had a t-ratio of only 1.618; Hsiao (1981, p. 96).

Hsiao also notes that the series may not be stationary if trended with time. Spot checks of regressions revealed some significant coefficients on the time trend. In order to guard against the possibility that the results in Tables 4–6 are spurious, we performed the F-tests of these tables using first differences of *M1, M2, NM2,* and *Y*. The results for *M1* and *M2* were qualitatively the same as those in Tables 4 and 5, though the shaded areas changed significantly. The results for net *M2* suggested unidirectional causality running from *Y* to *NM2*.

[16]The first difference of the growth rate of the monetary base does appear to be covariance stationary; however, ΔY does not.

[17]Since the distributed lags of the Granger equations are theoretically of infinite order, the specification with no lags of the left-hand-side variable may not be a valid specification for causality testing.

TABLE 4

SIGNIFICANCE LEVELS FOR GRANGER CAUSALITY TESTS OF *M1* AND *Y*

M1 on Y

Lags of *M1*	Lags of *Y*: 1	2	3	4	5	6	7	8	9	10	11	12
0	0.967	0.977	0.526	0.455	0.546	0.177	0.203	0.213	0.198	0.149	0.045*	0.066
1	0.882	0.959	0.515	0.404	0.465	0.156	0.150	0.146	0.126	0.111	0.022*	0.035*
2	0.936	0.954	0.499	0.395	0.444	0.164	0.158	0.149	0.120	0.112	0.024*	0.038*
3	0.947	0.946	0.473	0.370	0.418	0.156	0.159	0.152	0.112	0.116	0.027*	0.042*
4	0.813	0.700	0.616	0.233	0.257	0.105	0.106	0.124	0.099	0.128	0.048*	0.073
5	0.736	0.532	0.283	0.223	0.093	0.044*	0.047*	0.053	0.055	0.070	0.043*	0.066
6	0.929	0.588	0.377	0.342	0.114	0.015*	0.016*	0.018*	0.021*	0.032*	0.021*	0.034*
7	0.828	0.810	0.454	0.490	0.280	0.059	0.025*	0.025*	0.025*	0.040*	0.029*	0.046*
8	0.812	0.850	0.343	0.414	0.188	0.063	0.019*	0.030*	0.033*	0.050	0.034*	0.054
9	0.812	0.852	0.348	0.410	0.194	0.067	0.020*	0.032*	0.027*	0.041*	0.031*	0.049*
10	0.767	0.825	0.454	0.500	0.268	0.103	0.031*	0.049*	0.041*	0.059	0.046*	0.072
11	0.802	0.844	0.468	0.506	0.280	0.107	0.035*	0.054	0.045*	0.064	0.048*	0.074
12	0.803	0.921	0.508	0.546	0.238	0.062	0.026*	0.038*	0.027*	0.034*	0.028*	0.040*

Y on M1

Lags of *Y*	Lags of *M1*: 1	2	3	4	5	6	7	8	9	10	11	12
0	0.000*	0.000*	0.000*	0.001*	0.001*	0.003*	0.005*	0.009*	0.007*	0.005*	0.009*	0.013*
1	0.001*	0.000*	0.001*	0.003*	0.004*	0.009*	0.013*	0.023*	0.018*	0.013*	0.021*	0.029*
2	0.001*	0.000*	0.001*	0.003*	0.005*	0.010*	0.014*	0.025*	0.019*	0.013*	0.022*	0.029*
3	0.001*	0.000*	0.001*	0.004*	0.006*	0.012*	0.017*	0.029*	0.022*	0.014*	0.023*	0.030*
4	0.001*	0.001*	0.002*	0.005*	0.008*	0.016*	0.021*	0.035*	0.028*	0.018*	0.029*	0.037*
5	0.001*	0.000*	0.001*	0.004*	0.008*	0.016*	0.024*	0.039*	0.032*	0.018*	0.031*	0.035*
6	0.001*	0.000*	0.002*	0.004*	0.008*	0.016*	0.025*	0.041*	0.035*	0.019*	0.032*	0.034*
7	0.000*	0.000*	0.001*	0.002*	0.004*	0.007*	0.005*	0.008*	0.009*	0.003*	0.005*	0.008*
8	0.000*	0.000*	0.001*	0.002*	0.004*	0.007*	0.004*	0.008*	0.007*	0.003*	0.005*	0.006*
9	0.000*	0.000*	0.001*	0.002*	0.004*	0.008*	0.004*	0.008*	0.008*	0.004*	0.006*	0.008*
10	0.000*	0.000*	0.001*	0.002*	0.004*	0.009*	0.005*	0.010*	0.010*	0.002*	0.004*	0.003*
11	0.000*	0.001*	0.003*	0.006*	0.013*	0.025*	0.013*	0.024*	0.025*	0.004*	0.007*	0.006*
12	0.004*	0.006*	0.015*	0.033*	0.065	0.109	0.059	0.095	0.099	0.014*	0.023*	0.026*

*Indicates significance level less than 0.05

TABLE 5

SIGNIFICANCE LEVELS FOR GRANGER CAUSALITY TESTS OF $M2$ AND Y

M2 on Y

	Lags of Y											
Lags of $M2$	1	2	3	4	5	6	7	8	9	10	11	12
0	0.803	0.777	0.604	0.565	0.065	0.005*	0.005*	0.001*	0.002*	0.001*	0.000*	0.000*
1	0.036*	0.103	0.130	0.201	0.021*	0.015*	0.027*	0.012*	0.020*	0.021*	0.009*	0.002*
2	0.046*	0.130	0.158	0.239	0.026*	0.019*	0.033*	0.013*	0.023*	0.020*	0.008*	0.002*
3	0.044*	0.120	0.117	0.187	0.020*	0.014*	0.026*	0.012*	0.020*	0.017*	0.009*	0.002*
4	0.041*	0.116	0.131	0.213	0.024*	0.017*	0.031*	0.015*	0.024*	0.020*	0.011*	0.003*
5	0.042*	0.119	0.130	0.204	0.009*	0.007*	0.013*	0.006*	0.011*	0.010*	0.008*	0.002*
6	0.045*	0.130	0.120	0.204	0.018*	0.014*	0.027*	0.013*	0.021*	0.020*	0.016*	0.004*
7	0.085	0.216	0.186	0.273	0.037*	0.037*	0.055	0.022*	0.036*	0.033*	0.026*	0.008*
8	0.108	0.267	0.223	0.317	0.036*	0.042*	0.056	0.031*	0.048*	0.044*	0.035*	0.011*
9	0.148	0.350	0.336	0.452	0.056	0.063	0.078	0.050	0.071	0.065	0.054	0.018*
10	0.149	0.352	0.343	0.458	0.055	0.064	0.080	0.049*	0.073	0.050	0.040*	0.012*
11	0.136	0.332	0.344	0.455	0.061	0.073	0.093	0.059	0.085	0.060	0.042*	0.015*
12	0.175	0.399	0.408	0.525	0.074	0.089	0.109	0.064	0.090	0.070	0.051	0.013*

Y on M2

	Lags of $M2$											
Lags of Y	1	2	3	4	5	6	7	8	9	10	11	12
0	0.001*	0.003*	0.007*	0.016*	0.032*	0.021*	0.022*	0.032*	0.025*	0.032*	0.016*	0.021*
1	0.003*	0.009*	0.019*	0.038*	0.071	0.046*	0.052	0.073	0.058	0.070	0.035*	0.046*
2	0.003*	0.009*	0.020*	0.039*	0.072	0.049*	0.055	0.078	0.063	0.077	0.038*	0.050
3	0.003*	0.011*	0.023*	0.044*	0.081	0.055	0.063	0.088	0.068	0.081	0.039*	0.050
4	0.004*	0.013*	0.027*	0.047*	0.081	0.047*	0.060	0.085	0.061	0.077	0.038*	0.049*
5	0.005*	0.017*	0.034*	0.058	0.088	0.035*	0.039*	0.061	0.044*	0.061	0.026*	0.034*
6	0.004*	0.013*	0.026*	0.048*	0.075	0.038*	0.040*	0.061	0.044*	0.061	0.024*	0.031*
7	0.001*	0.005*	0.010*	0.020*	0.034*	0.017*	0.028*	0.047*	0.039*	0.052	0.018*	0.025*
8	0.002*	0.009*	0.016*	0.030*	0.051	0.025*	0.041*	0.057	0.032*	0.046*	0.013*	0.020*
9	0.001*	0.005*	0.011*	0.020*	0.037*	0.020*	0.034*	0.046*	0.036*	0.051	0.016*	0.024*
10	0.001*	0.002*	0.005*	0.007*	0.012*	0.009*	0.016*	0.023*	0.020*	0.016*	0.007*	0.012*
11	0.000*	0.000*	0.001*	0.001*	0.001*	0.000*	0.001*	0.002*	0.002*	0.001*	0.002*	0.003*
12	0.001*	0.003*	0.008*	0.014*	0.011*	0.005*	0.010*	0.014*	0.013*	0.008*	0.011*	0.019*

*Indicates significance level less than 0.05.

TABLE 6
SIGNIFICANCE LEVELS FOR GRANGER CAUSALITY TESTS OF *NM2* AND *Y*

NM2 on Y

	Lags of *Y*											
Lags of *NM2*	1	2	3	4	5	6	7	8	9	10	11	12
0	0.702	0.666	0.656	0.421	0.035*	0.005*	0.002*	0.000*	0.000*	0.000*	0.000*	0.000*
1	0.016*	0.047*	0.090	0.067	0.009*	0.014*	0.026*	0.018*	0.031*	0.020*	0.021*	0.002*
2	0.033*	0.098	0.168	0.116	0.019*	0.030*	0.052	0.028*	0.047*	0.020*	0.021*	0.003*
3	0.030*	0.085	0.124	0.088	0.015*	0.023*	0.041*	0.027*	0.045*	0.020*	0.024*	0.004*
4	0.029*	0.086	0.145	0.130	0.024*	0.035*	0.060	0.040*	0.065	0.028*	0.034*	0.006*
5	0.031*	0.091	0.156	0.139	0.017*	0.025*	0.044*	0.028*	0.047*	0.023*	0.030*	0.005*
6	0.040*	0.114	0.178	0.176	0.034*	0.052	0.085	0.054	0.087	0.044*	0.056	0.009*
7	0.060	0.163	0.240	0.199	0.048*	0.079	0.129	0.079	0.122	0.066	0.085	0.019*
8	0.066	0.177	0.258	0.215	0.051	0.086	0.138	0.085	0.130	0.071	0.091	0.021*
9	0.081	0.216	0.335	0.302	0.074	0.120	0.185	0.131	0.190	0.110	0.136	0.034*
10	0.084	0.222	0.343	0.313	0.077	0.124	0.191	0.135	0.197	0.087	0.108	0.024*
11	0.081	0.216	0.337	0.306	0.081	0.131	0.201	0.145	0.209	0.095	0.109	0.028*
12	0.082	0.219	0.342	0.313	0.085	0.138	0.210	0.154	0.221	0.102	0.117	0.028*

Y on NM2

	Lags on *NM2*											
Lags of *Y*	1	2	3	4	5	6	7	8	9	10	11	12
0	0.010*	0.030*	0.069	0.123	0.165	0.138	0.144	0.207	0.240	0.284	0.191	0.216
1	0.019*	0.062	0.125	0.207	0.268	0.234	0.253	0.344	0.390	0.437	0.290	0.334
2	0.019*	0.063	0.124	0.208	0.267	0.237	0.258	0.351	0.402	0.449	0.304	0.347
3	0.021*	0.070	0.139	0.225	0.282	0.255	0.277	0.372	0.415	0.462	0.314	0.358
4	0.024*	0.079	0.154	0.225	0.249	0.209	0.244	0.334	0.354	0.413	0.284	0.328
5	0.032*	0.099	0.186	0.264	0.246	0.157	0.175	0.253	0.273	0.344	0.213	0.253
6	0.025*	0.081	0.152	0.228	0.219	0.164	0.182	0.261	0.282	0.353	0.213	0.251
7	0.014*	0.048*	0.091	0.152	0.154	0.118	0.161	0.235	0.262	0.328	0.185	0.231
8	0.027*	0.085	0.141	0.214	0.204	0.156	0.210	0.294	0.278	0.350	0.188	0.242
9	0.015*	0.053	0.104	0.155	0.167	0.140	0.191	0.269	0.293	0.359	0.201	0.258
10	0.008*	0.025*	0.054	0.075	0.074	0.076	0.114	0.170	0.203	0.197	0.145	0.194
11	0.001*	0.002*	0.006*	0.010*	0.005*	0.004*	0.007*	0.014*	0.021*	0.017*	0.027*	0.043*
12	0.004*	0.017*	0.044*	0.073	0.037*	0.029*	0.051	0.083	0.107	0.084	0.112	0.159

*Indicates significance level less than 0.05.

TABLE 7

SIGNIFICANCE LEVELS FOR GRANGER CAUSALITY TESTS OF *MB* AND *Y*

MB on Y

	Lags of *Y*											
Lags of *MB*	1	2	3	4	5	6	7	8	9	10	11	12
0	0.944	0.930	0.581	0.721	0.812	0.788	0.867	0.758	0.774	0.539	0.483	0.500
1	0.569	0.718	0.437	0.441	0.583	0.511	0.599	0.513	0.611	0.428	0.193	0.260
2	0.468	0.678	0.431	0.447	0.586	0.498	0.592	0.488	0.589	0.439	0.205	0.273
3	0.379	0.678	0.314	0.344	0.483	0.449	0.558	0.484	0.576	0.461	0.164	0.224
4	0.461	0.756	0.473	0.438	0.583	0.572	0.678	0.547	0.633	0.558	0.237	0.311
5	0.523	0.815	0.521	0.567	0.709	0.668	0.776	0.638	0.731	0.642	0.339	0.426
6	0.486	0.783	0.482	0.539	0.674	0.578	0.696	0.610	0.707	0.646	0.361	0.450
7	0.473	0.770	0.439	0.496	0.636	0.578	0.693	0.596	0.690	0.645	0.381	0.470
8	0.466	0.765	0.446	0.508	0.648	0.589	0.703	0.597	0.691	0.626	0.363	0.452
9	0.438	0.733	0.465	0.534	0.679	0.606	0.721	0.635	0.732	0.670	0.356	0.445
10	0.499	0.794	0.636	0.621	0.750	0.758	0.847	0.801	0.871	0.861	0.507	0.595
11	0.503	0.795	0.636	0.590	0.728	0.732	0.828	0.803	0.873	0.866	0.452	0.535
12	0.472	0.768	0.607	0.562	0.692	0.666	0.774	0.784	0.858	0.834	0.430	0.523

Y on MB

	Lags of *MB*											
Lags of *Y*	1	2	3	4	5	6	7	8	9	10	11	12
0	0.000*	0.000*	0.000*	0.000*	0.001*	0.002*	0.003*	0.006*	0.012*	0.021*	0.031*	0.045*
1	0.000*	0.001*	0.000*	0.001*	0.003*	0.006*	0.009*	0.016*	0.028*	0.046*	0.065	0.090
2	0.000*	0.001*	0.000*	0.001*	0.003*	0.006*	0.010*	0.017*	0.029*	0.048*	0.067	0.093
3	0.000*	0.001*	0.000*	0.001*	0.002*	0.005*	0.008*	0.014*	0.024*	0.040*	0.057	0.080
4	0.000*	0.001*	0.001*	0.002*	0.003*	0.006*	0.010*	0.017*	0.030*	0.049*	0.068	0.094
5	0.000*	0.001*	0.000*	0.001*	0.002*	0.004*	0.006*	0.011*	0.018*	0.031*	0.041*	0.052
6	0.000*	0.001*	0.000*	0.001*	0.002*	0.004*	0.007*	0.012*	0.020*	0.034*	0.044*	0.057
7	0.000*	0.001*	0.000*	0.001*	0.002*	0.005*	0.008*	0.014*	0.022*	0.037*	0.049*	0.062
8	0.000*	0.001*	0.000*	0.001*	0.001*	0.002*	0.004*	0.006*	0.010*	0.017*	0.027*	0.029*
9	0.000*	0.001*	0.000*	0.001*	0.002*	0.003*	0.005*	0.007*	0.012*	0.021*	0.032*	0.033*
10	0.000*	0.002*	0.001*	0.001*	0.002*	0.004*	0.008*	0.010*	0.017*	0.029*	0.044*	0.045*
11	0.001*	0.004*	0.002*	0.006*	0.007*	0.012*	0.019*	0.023*	0.034*	0.056	0.072	0.068
12	0.004*	0.012*	0.006*	0.013*	0.014*	0.024*	0.029*	0.029*	0.038*	0.056	0.078	0.098

*Indicates significance level less than 0.05.

TABLE 8

SIGNIFICANCE LEVELS FOR GRANGER CAUSALITY TESTS OF ΔMB AND ΔY

ΔMB *on* ΔY

	Lags of ΔY											
Lags of ΔMB	1	2	3	4	5	6	7	8	9	10	11	12
0	0.140	0.288	0.076	0.144	0.168	0.179	0.190	0.250	0.250	0.048*	0.047*	0.070
1	0.105	0.271	0.074	0.107	0.122	0.139	0.084	0.131	0.147	0.052	0.076	0.110
2	0.354	0.638	0.073	0.099	0.166	0.192	0.114	0.165	0.169	0.040*	0.058	0.084
3	0.359	0.628	0.081	0.108	0.179	0.201	0.124	0.177	0.163	0.041*	0.060	0.086
4	0.563	0.765	0.322	0.448	0.591	0.652	0.366	0.468	0.350	0.186	0.247	0.298
5	0.488	0.669	0.308	0.461	0.573	0.629	0.409	0.501	0.397	0.246	0.305	0.360
6	0.518	0.699	0.339	0.498	0.605	0.667	0.442	0.536	0.425	0.264	0.322	0.379
7	0.530	0.673	0.316	0.474	0.569	0.649	0.476	0.570	0.464	0.293	0.351	0.403
8	0.450	0.630	0.354	0.509	0.578	0.676	0.534	0.644	0.526	0.291	0.337	0.376
9	0.571	0.784	0.413	0.584	0.696	0.807	0.731	0.821	0.794	0.487	0.570	0.623
10	0.559	0.760	0.369	0.537	0.630	0.746	0.720	0.812	0.797	0.401	0.490	0.514
11	0.541	0.753	0.349	0.514	0.604	0.717	0.715	0.809	0.771	0.396	0.469	0.502
12	0.544	0.735	0.354	0.518	0.622	0.731	0.745	0.832	0.810	0.387	0.465	0.456

ΔY *on* ΔMB

	Lags of ΔMB											
Lags of ΔY	1	2	3	4	5	6	7	8	9	10	11	12
0	0.018*	0.060	0.080	0.121	0.185	0.266	0.321	0.392	0.481	0.559	0.653	0.635
1	0.011*	0.028*	0.034*	0.069	0.121	0.153	0.199	0.279	0.372	0.452	0.540	0.507
2	0.028*	0.052	0.028*	0.060	0.101	0.136	0.200	0.281	0.369	0.434	0.520	0.516
3	0.020*	0.048*	0.016*	0.033*	0.057	0.065	0.100	0.145	0.206	0.270	0.333	0.343
4	0.041*	0.088	0.030*	0.058	0.089	0.094	0.141	0.202	0.280	0.356	0.431	0.447
5	0.038*	0.111	0.039*	0.075	0.090	0.069	0.111	0.153	0.209	0.253	0.323	0.297
6	0.030*	0.090	0.046*	0.078	0.094	0.039*	0.067	0.087	0.134	0.182	0.229	0.197
7	0.026*	0.080	0.044*	0.075	0.094	0.040*	0.069	0.091	0.139	0.191	0.237	0.198
8	0.038*	0.118	0.065	0.099	0.133	0.051	0.086	0.084	0.129	0.169	0.191	0.128
9	0.025*	0.082	0.099	0.156	0.186	0.091	0.146	0.139	0.196	0.239	0.229	0.187
10	0.061	0.175	0.194	0.320	0.392	0.159	0.237	0.162	0.223	0.202	0.173	0.187
11	0.157	0.366	0.411	0.582	0.591	0.192	0.269	0.151	0.219	0.216	0.237	0.239
12	0.181	0.403	0.445	0.617	0.621	0.212	0.288	0.159	0.228	0.212	0.229	0.251

*Indicates significance level less than 0.05.

TABLE 9

ORDINARY LEAST SQUARES ESTIMATES OF ΔMB ON ΔY

Lag	ΔMB	ΔY
1	−0.453* (3.98)	−0.068 (1.43)
2	−0.338* (2.92)	−0.068 (1.16)
3		−0.167* (2.59)
4		−0.080 (1.09)
5		−0.039 (0.54)
6		−0.073 (0.99)
7		−0.053 (0.74)
8		0.013 (0.19)
9		0.034 (0.58)
10		0.125* (2.49)
Constant	0.272 (0.68)	
Time	−0.005 (0.56)	
$\bar{R}^2 = 0.27$	D-W = 1.98	

NOTE. Absolute value of the t-statistic is in parentheses.
*Significant at the 5 percent level.

Ordinary least squares estimates of one of these specifications are reported in Table 9.[18] Only two of the ten lags on ΔY have t-ratios greater than two. Indeed, in most cases they are substantially below two. This is difficult to reconcile with the F-test results unless ΔY exhibits significant autocorrelation. Tests indicate significant negative autocorrelation; thus, it appears that this series is overdifferenced. (This probably accounts for the marked change in the tests with income as the dependent variable between Tables 7 and 8.) Furthermore, the results indicate a much longer lag from income to the base than most devotees of a Federal Reserve reaction function would find acceptable. Consequently, the evidence of Granger causality running from income to the adjusted monetary base is extremely weak at best.

5. SUMMARY AND CONCLUSIONS

This paper has employed Granger tests to investigate the causal relationship between money and income based on arbitrarily chosen lags, lags chosen by three statistical criteria, and an extensive search of the lag space. The results suggest bidirectional causality between *M1*, *M2*, and the non-*M1* components of *M2* and nominal income. Although we found some evidence of bidirectional causality between the adjusted monetary base and income, a reasonable interpretation of the data would suggest unidirectional causality running from base to income.

The results also suggest that the arbitrary lag-length specifications fail to ensure that the results are not based on the capricious choice of the lag structure. Furthermore, a frequently suggested rule of thumb does not appear to fare any better. Indeed, if one wishes to be agnostic about the order of the model, it appears that the

[18]Results for the other equation are not appreciably different, so only one is presented here.

safest approach is to perform an extensive search of the lag space. Even when this approach is taken, however, care must be taken in interpreting the results.

A comparison of several commonly used criteria for lag-length selection suggests that based on a standard, classical, hypothesis-testing norm Akaike's FPE criterion performed well in selecting the model relative to the others. As a result, it did a reasonably good job of finding an order of the model which gave evidence of Granger causality, when such an order existed. Unfortunately, we do not know the extent to which these results are data specific. Furthermore, the FPE criterion may not conform to all researchers' prior beliefs about the appropriate trade-off between bias and efficiency. Clearly, more work needs to be done on this problem. As a generalization, however, there appears to be no substitute for selecting a model specification criterion ex ante or for an extensive search of the lag space if one is to ensure that the causality test results are not critically dependent on the judicious (or perhaps fortuitous) choice of the lag structure.

LITERATURE CITED

Akaike, H. "Statistical Predictor Identification." *Annals of the Institute of Statistical Mathematics* 22 (1970), 203–17.

Batten, Dallas S., and Daniel L. Thornton. "Polynomial Distributed Lags and the Estimation of the St. Louis Equation." Federal Reserve Bank of St. Louis *Review* 65 (April 1983), 13–25 (a).

______. "Lag-Length Selection Criteria: Empirical Results from the St. Louis Equation." Federal Reserve Bank of St. Louis Research Paper 83-008 (1983) (b).

______. "M1 or M2: Which Is the Better Monetary Target?" Federal Reserve Bank of St. Louis *Review* 65 (June/July 1983), 36–42 (c).

Carlson, Keith M., and Scott E. Hein. "Monetary Aggregates as Monetary Indicators." Federal Reserve Bank of St. Louis *Review* 62 (November 1980), 12–21.

Ciccolo, John H., Jr. "Money, Equity Values, and Income: Tests for Exogeneity." *Journal of Money, Credit, and Banking* 10 (February 1978), 46–64.

Davidson, Lawrence S., and R. W. Hafer. "Some Evidence on Selecting an Intermediate Target for Monetary Policy." *Southern Economic Journal* 50 (October 1983), 406–21.

Elliott, J. Walter. "The Influence of Monetary and Fiscal Actions on Total Spending: The St. Louis Total Spending Equation Revisited." *Journal of Money, Credit, and Banking* 7 (May 1975), 181–92.

Feige, Edgar L., and Douglas K. Pearce. "The Casual Causal Relationship between Money and Income: Some Caveats for Time Series Analysis." *Review of Economics and Statistics* 61 (November 1979), 521–33.

Friedman, Milton, and Anna J. Schwartz. *A Monetary History of the United States, 1867–1960*. Princeton, N.J.: Princeton University Press, 1963.

Geweke, John, and Richard Meese. "Estimating Regression Models of Finite but Unknown Order." *International Economic Review* 22 (February 1981), 55–70.

Geweke, John, Richard Meese, and Warren Dent. "Comparing Alternative Tests of Causality in Temporal Systems: Analytic Results and Experimental Evidence." *Journal of Econometrics* 21 (February 1983), 161–94.

Gordon, Robert J. "The Short-Run Demand for Money: A Reconsideration." *Journal of Money, Credit, and Banking* 16 (November 1984), 403–34.

Granger, Clive W. J. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (July 1969), 424–38.

Guilkey, David K., and Michael K. Salemi. "Small Sample Properties of Three Tests for Granger-Causal Ordering in a Bivariate Stochastic System." *Review of Economics and Statistics* 64 (November 1982), 668–80.

Hafer, R. W. "Selecting a Monetary Indicator: A Test of the New Monetary Aggregates." Federal Reserve Bank of St. Louis *Review* 63 (February 1981), 12–18.

Hsiao, Cheng. "Autoregressive Modelling and Money-Income Causality Detection." *Journal of Monetary Economics* 7 (January 1981), 85–106.

Judge, George G., William E. Griffiths, R. Carter Hill, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*. New York: Wiley, 1980.

McMillin, W. Douglas, and James S. Fackler. "Monetary vs. Credit Aggregates: An Evaluation of Monetary Policy Targets." *Southern Economic Journal* 50 (January 1984), 711–23.

Mehra, Yash P. "An Empirical Note on Some Monetarist Propositions." *Southern Economic Journal* 45 (July 1978), 154–67.

Mehra, Yash P., and David E. Spencer. "The St. Louis Equation and Reverse Causation: The Evidence Reexamined." *Southern Economic Journal* 45 (April 1979), 1104–20.

Pagano, Marcello, and Michael J. Hartley. "On Fitting Distributed Lag Models Subject to Polynomial Restrictions." *Journal of Econometrics* 16 (June 1981), 171–98.

Pautler, Paul A., and Richard J. Rivard. "Choosing a Monetary Aggregate: Causal Relationship as a Criterion." *Review of Business and Economic Research* 15 (Fall 1979), 1–18.

Sargent, Thomas J. "A Classical Macroeconometric Model of the United States." *Journal of Political Economy* 84 (April 1976), 207–37.

Sims, Christopher A. "Money, Income, and Causality." *American Economic Review* 62 (September 1972), 540–52.

______. "Macroeconomics and Reality." *Econometrica* 48 (January 1980), 1–48 (a).

______. "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered." *American Economic Review* 70 (May 1980), 250–57 (b).