

MARKOV SWITCHING CAUSALITY AND THE MONEY–OUTPUT RELATIONSHIP

ZACHARIAS PSARADAKIS,^{a,*} MORTEN O. RAVN^{b,c} AND MARTIN SOLA^{a,d}

^a *School of Economics, Mathematics & Statistics, Birkbeck College, University of London, UK*

^b *Department of Economics, London Business School, UK*

^c *Centre for Economic Policy Research, London, UK*

^d *Department of Economics, Universidad Torcuato Di Tella, Buenos Aires, Argentina*

SUMMARY

The causal link between monetary variables and output is one of the most studied issues in macroeconomics. One puzzle from this literature is that the results of causality tests appear to be sensitive with respect to the sample period that one considers. As a way of overcoming this difficulty, we propose a method for analysing Granger causality which is based on a vector autoregressive model with time-varying parameters. We model parameter time-variation so as to reflect changes in Granger causality, and assume that these changes are stochastic and governed by an unobservable Markov chain. When applied to US data, our methodology allows us to reconcile previous puzzling differences in the outcome of conventional tests for money–output causality. Copyright © 2005 John Wiley & Sons, Ltd.

1. INTRODUCTION

The relationship between money and output has attracted a phenomenal amount of interest over the years in both empirical and theoretical work. An important strand of the literature has examined the ‘causal’ links between money and output, testing for Granger (or Sims) causality from money to output.¹ The interest in this issue is hardly surprising given the vital role typically ascribed to monetary policy over the business cycle. The conclusions drawn from these exercises, however, have changed a number of times over the last three decades as researchers have modified their empirical models in the light of developments in measurement, economic theory and econometric techniques.

Our interest in this issue is partly due to the observation that empirical evidence on money–output causality appears to be ‘unstable’. The particular type of instability that we focus upon is sample dependence (or time dependence) of the results from causality tests.² An example of this type of instability is the contrast between the empirical results of Eichenbaum and Singleton

* Correspondence to: Professor Zacharias Psaradakis, School of Economics, Mathematics and Statistics, Birkbeck College, Malet Street, London WC1E 7HX, UK. E-mail: z.psaradakis@bbk.ac.uk

Contract/grant sponsor: UK Economic and Social Research Council; Contract/grant number: L138251003.

¹ See, among many others, Sims (1972, 1980), Litterman and Weiss (1985), Bernanke (1986), Eichenbaum and Singleton (1986), Christiano and Ljungqvist (1988), Krol and Ohanian (1990), Stock and Watson (1989), Friedman and Kuttner (1992, 1993), Thoma (1994) and Swanson (1998).

² Another source of instability is the instrument dependence of the results. Sims (1972), for example, found that money causes output in a bivariate relationship, but Sims (1980) argued that money is Granger non-causal for output if interest rates are included in the model used to test the causality hypothesis (a finding later challenged by others; e.g., Bernanke, 1986). Other researchers have documented that the results of causality tests depend on the trend properties of the data. Christiano and Ljungqvist (1988) and Krol and Ohanian (1990) showed that the lack of Granger causality from money to output depends on the trend properties of the variables of interest. In an important contribution, Stock and Watson

(1986) and Stock and Watson (1989) on the one hand and those of Friedman and Kuttner (1992, 1993) on the other. Eichenbaum and Singleton (1986) and Stock and Watson (1989) find that the causal role for money is much weaker in a sample that excludes data from the 1980s than in data sets that include the 1980s. In stark contrast, Friedman and Kuttner (1992, p.472) state that: 'Including data from the 1980's sharply weakens the postwar time-series evidence indicating significant relationships between money (however defined) and nominal output or between money and either real output or prices separately. Focusing on data from 1970 onward destroys this evidence altogether'.³ The sensitivity of results from causality tests with respect to the sample period under consideration was further examined by Thoma (1994) and Swanson (1998), who used recursive and rolling window techniques to analyse Granger causality between money and output. This type of evidence indicates that causal relationships may change over time and/or that such links between money and output are very sensitive to the exact sample period under consideration.

In this paper, we propose a method for analysing causality patterns in environments where there may be good reasons to suspect that causality relationships have changed over the sample period of interest. The main motivation for our approach is that, in addition to the implications for economic theory and policy analysis, non-constancy of causality patterns also poses significant econometric problems in the application of standard tests for Granger causality. Although there may be many reasons for time dependence of causality patterns (changes in operating procedures and other aspects of monetary policy, large shocks to the economy, etc.), unless researchers have strong *a priori* indications of when and why the links between the variables of interest change, accounting for such changes may be difficult. Furthermore, it is by now well recognized (see, e.g., Cooley and LeRoy, 1985; Cooley and Dwyer, 1998; Hamilton, 1994, chap. 11) that there is no direct relationship between causality in an economic sense—one variable being responsible for changes in another variable—and causality in an econometric sense—one variable helping to predict another. In other words, information that is useful in terms of understanding changes in causality in an economic sense may not be useful, or sufficient, for guiding one's choice of dates at which Granger causality changes.

To overcome these difficulties, we propose a framework of analysis based on vector autoregressive (VAR) models with time-varying parameters. Crucially, however, parameter time-variation is modelled so as to reflect directly changes in causality between the variables of interest. The starting point of our analysis is the observation that the number and location of points at which causal links change are typically unknown *a priori*. For this reason, we treat changes in causality as random events and allow the data to select the change points. More specifically, we model changes in causality as being governed by a hidden Markov chain with stationary but unknown transition probabilities. Such a formulation is flexible enough to allow for a variety of stochastic changes in causality, ranging from one-time permanent changes to frequent short-lived changes. What is more, even though these changes are unobservable, inferences about them can be made

(1989) argued that the data-generating process of money supply is best described by money growth having a positive trend and, importantly, that detrended money does Granger-cause output. Friedman and Kuttner (1993), however, demonstrated that this result is not robust either to changes in the sample period or to the inclusion of different short-term interest rates in the model.

³ This instability is also manifest in other aspects of the relationship between monetary variables and the real economy. In a recent contribution, Gavin and Kydland (1999) document instability of the correlations between (cyclical measures of) real and nominal variables in a sample of post-war US data. Backus and Kehoe (1992) find similar evidence for a much longer sample covering a cross-section of countries. More generally, Stock and Watson (1996) argue that there is considerable empirical evidence supporting the view that economic time series are best thought of as being generated by random processes with time-varying parameters.

using likelihood-based methods. Given the well-documented success of Markov switching models in describing the behaviour of economic time series subject to changes in regime—see, for example, Kim and Nelson (1999) and the references therein—there are good reasons to expect such a characterization of the data to yield a good representation of the changes in causality that are often observed in practice.

The new methodology is used to re-examine the empirical evidence on the predictive content of money and interest rates for post-war US real output using quarterly data for the period 1959:1 to 2001:2. We find that the causality patterns have changed over the sample for each of these three variables under consideration (namely, M1, M2 and the Federal Funds rate). We identify two main periods during which M1 growth has had predictive power for output growth. The first is a seven-year period that spans the interval from 1976 to mid-1982, while the second is a brief period around the early 1990s recession. M2 growth is found to have Granger-caused output in an interval from 1970 to 1983 and in a short spell at the beginning of our sample. It is noteworthy that our results reveal money to have lost most of its predictive power at the end of the Volcker disinflation period, which explains the difference between the results of Stock and Watson (1989) and Friedman and Kuttner (1992). We also find that there is a marked tendency for the money-stock growth rates to gain predictive power during, or immediately before, recessions. This may indicate that monetary policy is used more actively during recessions, or just before recessions, in order to prevent a deterioration in the economy, while monetary policy might be more accommodating during expansions. The results for the Federal Funds rate are somewhat different. We find that the Federal Funds rate has significant predictive power for output growth in the post-1985 sample and in the pre-1970 sample. In the intermediary period, from 1970 to 1985, we find more frequent changes in the relationship between interest rates and output growth. Again, the post-1985 results appear natural given the importance of the Federal Funds rate for the conduct and stance of US monetary policy.

The plan of the paper is as follows. In the next section, we briefly discuss alternative approaches to modelling changing causality. In Section 3, we propose a multivariate autoregressive model with Markov switching causality and discuss its interpretation and implications. Our empirical results are presented in Section 4. Finally, Section 5 summarizes and concludes.

2. MODELLING CHANGING GRANGER CAUSALITY

As mentioned before, numerous researchers have noted that results from Granger causality tests tend to be sensitive with respect to changes in the sample period. The most common approach to dealing with such instabilities is to test for causality after splitting the sample under consideration into subsamples. This method requires the researcher to specify the dates at which causal relationships are supposed to have changed. In practice, this approach may not be straightforward. Although there may be cases in which researchers have reliable *a priori* information about dates at which the economy has been subject to a structural change (e.g., changes in operating procedures), such information may not be directly related to Granger causality (i.e., changes in the predictive power of one variable for another). Furthermore, such information may often be incomplete and not be precise enough to allow one to choose the break dates accurately. Hence, break points will often have to be estimated in practice rather than be pre-specified. However, in many cases, there is no reason to assume that only one break takes place, especially when the sample covers a long time period. In the absence of prior information about the location of breaks, a main source of

difficulty is that the number of possible locations of break points quickly becomes very large as the sample size increases (in a sample with T observations and b break points, there are $\binom{T}{b}$ possible locations for the breaks). Procedures such as those proposed by Bai and Perron (1998) and Bai (1999) offer a way to test for the presence of multiple (deterministic) changes in the parameters of the model of interest and estimate their number and locations. Such changes may not, however, be directly related to changes in Granger causality (as the relevant coefficients do not necessarily take a zero value before or after a structural break).

An alternative approach, adopted, for example, by Thoma (1994) and Swanson (1998), involves a rolling-window technique. This consists of analysing only a fixed-length window of observations from the sample and then investigating the stability of the results as the window is rolled through the sample. Although this approach offers a sensible way of identifying changes in causality, a potential difficulty is that results may be sensitive with respect to the size of the window used—a very narrow window makes the results sensitive to outliers and sampling errors, while the use of too wide a window makes it unlikely that short-lived changes will be detected—and there are no firm statistical guidelines available for choosing the window size (see Swanson, 1998 for a discussion). Furthermore, although this technique might enable one to check the time-invariance of the results, it does not provide a statistically sound way of identifying the exact dates at which changes have taken place, let alone test for the statistical significance of such changes.

Our starting point in this paper is to substitute the notion of permanent causal relations with a notion of ‘temporary’ Granger causality, that is causality which holds during some periods but not in others. Our strategy then consists of identifying the periods during which a variable Granger-causes another variable. Our methodology is based on a VAR model with time-varying parameters where, given our objectives, parameter time-variation directly reflects changes in causality. By treating changes in causality as random events governed by an exogenous Markov process, inferences about these changes can be made on the basis of the estimated probability that each observation in the sample comes from a particular causality regime.

It should be noted that VAR models with Markov switching have been used before by many researchers, primarily to examine how information about relevant economic variables affects transitions between alternative states of the economy; see, *inter alia*, Sola and Driffill (1994), Krolzig (1997, 2000) and Warne (1999). The models used in these papers are variants of a VAR model where the mean (or intercept), error variance and/or autoregressive coefficients are subject to Markov switching. However, simple VAR models with switching intercepts and/or error variances are not well-suited to studying regime-dependent causality. Furthermore, even models with switching coefficients are not ideal for the problem in hand for at least three reasons. First, the separation of regimes in such models need not be related to changes in causality and hence a regime may include observations from states of nature associated with very different causal patterns. Second, the endogenous variables in a VAR model may be subject to changes that occur at very different points in the sample, and hence an adequate description of these changes may require the use of a model with a large number of regimes. Finally, a general Markov switching VAR model is not necessarily consistent with our notion of temporary causality, i.e., causality which holds during some periods but not in others, since the coefficients that parameterize Granger causality are not restricted to be zero in one or more of the states of nature.⁴

⁴ Our notion of temporary Granger causality is also different from the notion of Granger causality of regimes discussed in Krolzig (2000), where the point of interest is whether Markov regime switching can improve forecast accuracy (for either univariate or multivariate time series).

The approach advocated here is based on a switching VAR model in which the regimes represent alternative causal states of nature. In such a model, there are always four regimes (if the model is bivariate), which are associated with the four possible causal relationships between the variables. Let us now explain our methodology in some detail.

3. A MODEL OF MARKOV SWITCHING GRANGER CAUSALITY

Consider the problem of analysing the Granger causality links between the components of the bivariate time series $\{\mathbf{X}'_t = [X_{1,t} : X_{2,t}]\}$ conditionally on the scalar time series $\{Z_t\}$.⁵ Our analysis is based on the following Markov switching VAR model:

$$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_{10} + \mu_{11}S_{1,t} \\ \mu_{20} + \mu_{21}S_{2,t} \end{bmatrix} + \sum_{k=1}^{h_1} \begin{bmatrix} \phi_{10}^{(k)} + \phi_{11}^{(k)}S_{1,t} & \psi_1^{(k)}S_{1,t} \\ \psi_2^{(k)}S_{2,t} & \phi_{20}^{(k)} + \phi_{21}^{(k)}S_{2,t} \end{bmatrix} \begin{bmatrix} X_{1,t-k} \\ X_{2,t-k} \end{bmatrix} \\ + \sum_{k=1}^{h_2} \begin{bmatrix} \varphi_{10}^{(k)} + \varphi_{11}^{(k)}S_{1,t} \\ \varphi_{20}^{(k)} + \varphi_{21}^{(k)}S_{2,t} \end{bmatrix} Z_{t-k} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \quad t = 1, 2, \dots, T \quad (1)$$

Here, $S_{1,t}$ and $S_{2,t}$ are latent random variables that reflect the ‘state’ (or ‘regime’) of the system at date t and take their values in the set $\{0, 1\}$, and $\{\varepsilon'_t = [\varepsilon_{1,t} : \varepsilon_{2,t}]\}$ is a white-noise process, independent of $\{S_{1,t}\}$ and $\{S_{2,t}\}$, with mean zero and covariance matrix which depends on $S_{1,t}$ and $S_{2,t}$ in a way to be specified below.

The specification in (1) allows for four alternative states of nature, which may be conveniently indexed by using the following state-indicator variable:

$$S_t = \begin{cases} 1 & \text{if } S_{1,t} = 1 \text{ and } S_{2,t} = 1 \\ 2 & \text{if } S_{1,t} = 0 \text{ and } S_{2,t} = 1 \\ 3 & \text{if } S_{1,t} = 1 \text{ and } S_{2,t} = 0 \\ 4 & \text{if } S_{1,t} = 0 \text{ and } S_{2,t} = 0 \end{cases}$$

The covariance matrix of the disturbances in (1) is then specified as

$$\mathbb{E}(\varepsilon_t \varepsilon'_t | S_t = \ell) = [\sigma_{ij,\ell}], \quad i, j = 1, 2; \ell = 1, \dots, 4$$

Before describing the probabilistic properties of the state variables $S_{1,t}$ and $S_{2,t}$, it is useful to make clear how these variables affect the parameters of the VAR. Observe that the model in (1) implies that

$$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \mu_{10} + \mu_{11} \\ \mu_{20} + \mu_{21} \end{bmatrix} + \sum_{k=1}^{h_1} \begin{bmatrix} \phi_{10}^{(k)} + \phi_{11}^{(k)} & \psi_1^{(k)} \\ \psi_2^{(k)} & \phi_{20}^{(k)} + \phi_{21}^{(k)} \end{bmatrix} \begin{bmatrix} X_{1,t-k} \\ X_{2,t-k} \end{bmatrix} \\ + \sum_{k=1}^{h_2} \begin{bmatrix} \varphi_{10}^{(k)} + \varphi_{11}^{(k)} \\ \varphi_{20}^{(k)} + \varphi_{21}^{(k)} \end{bmatrix} Z_{t-k} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \quad \text{if } S_t = 1$$

⁵ Our discussion is structured around a bivariate example in order to make the exposition clearer, but the methodology can be applied to multivariate models of any dimension without any further complications (apart from heavier computational costs). We also introduce a ‘conditioning’ variable in the model because such variables may be important in practical applications involving tests for Granger causality.

$$\begin{aligned}
\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} &= \begin{bmatrix} \mu_{10} \\ \mu_{20} + \mu_{21} \end{bmatrix} + \sum_{k=1}^{h_1} \begin{bmatrix} \phi_{10}^{(k)} & 0 \\ \psi_2^{(k)} & \phi_{20}^{(k)} + \phi_{21}^{(k)} \end{bmatrix} \begin{bmatrix} X_{1,t-k} \\ X_{2,t-k} \end{bmatrix} \\
&\quad + \sum_{k=1}^{h_2} \begin{bmatrix} \phi_{10}^{(k)} \\ \phi_{20}^{(k)} + \phi_{21}^{(k)} \end{bmatrix} Z_{t-k} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \quad \text{if } S_t = 2 \\
\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} &= \begin{bmatrix} \mu_{10} + \mu_{11} \\ \mu_{20} \end{bmatrix} + \sum_{k=1}^{h_1} \begin{bmatrix} \phi_{10}^{(k)} + \phi_{11}^{(k)} & \psi_1^{(k)} \\ 0 & \phi_{20}^{(k)} \end{bmatrix} \begin{bmatrix} X_{1,t-k} \\ X_{2,t-k} \end{bmatrix} \\
&\quad + \sum_{k=1}^{h_2} \begin{bmatrix} \phi_{10}^{(k)} + \phi_{11}^{(k)} \\ \phi_{20}^{(k)} \end{bmatrix} Z_{t-k} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \quad \text{if } S_t = 3 \\
\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} &= \begin{bmatrix} \mu_{10} \\ \mu_{20} \end{bmatrix} + \sum_{k=1}^{h_1} \begin{bmatrix} \phi_{10}^{(k)} & 0 \\ 0 & \phi_{20}^{(k)} \end{bmatrix} \begin{bmatrix} X_{1,t-k} \\ X_{2,t-k} \end{bmatrix} \\
&\quad + \sum_{k=1}^{h_2} \begin{bmatrix} \phi_{10}^{(k)} \\ \phi_{20}^{(k)} \end{bmatrix} Z_{t-k} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}, \quad \text{if } S_t = 4
\end{aligned}$$

It is evident, therefore, that the regime indicators $S_{1,t}$ and $S_{2,t}$ (and hence S_t) determine the causal links in the model. In particular, $S_{1,t}$ determines whether $X_{2,t}$ is Granger-causal for $X_{1,t}$ while $S_{2,t}$ dictates whether $X_{1,t}$ is Granger-causal for $X_{2,t}$. Given that at least one of the parameters $\psi_1^{(1)}, \dots, \psi_1^{(h_1)}$ is non-zero, $X_{2,t}$ Granger-causes $X_{1,t}$ when $S_{1,t} = 1$ ($S_t = 1$ or $S_t = 3$) and is Granger non-causal for $X_{1,t}$ when $S_{1,t} = 0$ ($S_t = 2$ or $S_t = 4$). Likewise, $X_{1,t}$ Granger-causes $X_{2,t}$ when $S_{2,t} = 1$ ($S_t = 1$ or $S_t = 2$), provided that not all of the parameters $\psi_2^{(1)}, \dots, \psi_2^{(h_2)}$ are zero, and is Granger non-causal when $S_{2,t} = 0$ ($S_t = 3$ or $S_t = 4$). Note that this parameterization also encompasses the case where one or both variables are Granger non-causal for the whole sample period ($\psi_1^{(1)} = \dots = \psi_1^{(h_1)} = 0$ and/or $\psi_2^{(1)} = \dots = \psi_2^{(h_2)} = 0$).

The specification of the model is completed by assuming that nature selects the state of the system at date t with a probability which depends only on the state prevailing at date $t - 1$. More specifically, we assume that the random sequences $\{S_{1,t}\}$ and $\{S_{2,t}\}$ are time-homogeneous, first-order Markov chains with transition probabilities

$$p_{ij}^{(\ell)} = \mathbb{P}(S_{\ell,t+1} = j | S_{\ell,t} = i), \quad i, j = 0, 1; \ell = 1, 2$$

Moreover, it is assumed that $\{S_{1,t}\}$ and $\{S_{2,t}\}$ are independent.⁶ Hence, if \mathbf{P} denotes the stochastic matrix whose (i, j) element is the probability $\mathbb{P}(S_{t+1} = i | S_t = j)$, for $i, j = 1, \dots, 4$, the random sequence $\{S_t\}$ is a time-homogeneous, first-order Markov chain with transition probability matrix

$$\mathbf{P} = \begin{bmatrix} p_{11}^{(1)} p_{11}^{(2)} & p_{11}^{(2)} (1 - p_{00}^{(1)}) & p_{11}^{(1)} (1 - p_{00}^{(2)}) & (1 - p_{00}^{(1)}) (1 - p_{00}^{(2)}) \\ p_{11}^{(2)} (1 - p_{11}^{(1)}) & p_{00}^{(1)} p_{11}^{(2)} & (1 - p_{11}^{(1)}) (1 - p_{00}^{(2)}) & p_{00}^{(1)} (1 - p_{00}^{(2)}) \\ p_{11}^{(1)} (1 - p_{11}^{(2)}) & (1 - p_{00}^{(1)}) (1 - p_{11}^{(2)}) & p_{11}^{(1)} p_{00}^{(2)} & p_{00}^{(2)} (1 - p_{00}^{(1)}) \\ (1 - p_{11}^{(1)}) (1 - p_{11}^{(2)}) & p_{00}^{(1)} (1 - p_{11}^{(2)}) & p_{00}^{(2)} (1 - p_{11}^{(1)}) & p_{00}^{(1)} p_{00}^{(2)} \end{bmatrix}$$

⁶ This assumption can be relaxed, but it is used in the sequel since it is not rejected for our data.

Causality analysis based on the Markov switching model in (1) is attractive for a number of reasons. First, the states of nature are defined directly in terms of the causal relationships between the two variables; this is a natural way of classifying regimes when the focus of the analysis is changes in causality. Second, it allows for the possibility of arbitrarily many changes in causality at unknown locations in the sample. What is more, these changes are parameterized in a parsimonious way, driven as they are by a simple, time-homogeneous Markov chain. Finally, our approach allows one to make probabilistic inferences about the dates at which changes in causality occurred during the sample period. More specifically, although the state variables governing the manner in which regime shifts occur are unobservable, the likelihood of each of the possible regimes being operable at each date in the sample period can be inferred on the basis of the estimated conditional probabilities $\mathbb{P}(S_t = \ell | \mathbf{W}_1, \dots, \mathbf{W}_t; \hat{\theta})$, $\ell = 1, \dots, 4$, where $\mathbf{W}_t' = [\mathbf{X}_t' : Z_{t-1}]$ and $\hat{\theta}$ is an estimator of the unknown parameters in (1). Thus, we can evaluate the extent to which changes in causality have actually occurred and identify the locations of such changes in the sample.

4. MONEY–OUTPUT CAUSALITY

In this section, we use the proposed model for Markov switching Granger causality to address the question of whether financial variables have predictive power for output. Our analysis is based on quarterly US data for the period 1959:1 to 2001:2. We measure output by real GDP and examine three financial variables, namely, M1, M2 and the Federal Funds rate (r). We logarithmically transform real GDP (y), M1 (m_1) and M2 (m_2), and use the growth rate of the consumer price index as a measure of inflation (π). Except for the interest rates, all data are seasonally adjusted.

4.1. Trending Properties of the Data

It is well documented that causality inferences are sensitive with respect to the presence of trends in the data, and correct modelling of the long-run characteristics of the data is hence very important (cf. Christiano and Ljungqvist, 1988; Stock and Watson, 1989). With this in mind, we report in Table I the ordinary least-squares (OLS) estimate of the sum of autoregressive coefficients (ρ) in a univariate autoregression for each of the series of interest and its first difference.⁷ We also report 90% equal-tailed confidence intervals for ρ obtained by Hansen's (1999) grid- t bootstrap method, using 999 bootstrap replications at each of 200 grid-points. In addition, we give 90% symmetric confidence intervals for ρ obtained by the subsampling method of Romano and Wolf (2001) applied to the studentized OLS estimator of ρ .^{8,9}

At least one of the confidence intervals for ρ contains unity in the case of y , m_1 , m_2 and π but not in the case of r , Δy , Δm_1 or Δm_2 . This suggests that y , m_1 and m_2 are integrated of order one, while the interest rate and the differenced series are integrated of order zero. We will,

⁷ Following the recommendation of Ng and Perron (1995), the order of each autoregressive model was determined by means of sequential 10%-level t -type tests for the significance of the coefficient on the longest lag, allowing for a maximum order of 12.

⁸ The subsample size was selected by implementing Algorithm 5.1 of Romano and Wolf (2001) with $b_{small} = 8$, $b_{big} = 32$ and $k = 2$ (in their notation). We compute symmetric subsampling intervals because Romano and Wolf (2001) found them to have more accurate coverage than equal-tailed intervals.

⁹ Note that grid-bootstrap confidence intervals have asymptotically correct coverage for all ρ local to unity (i.e., $\rho = 1 + c/T$ for some constant c). Subsampling confidence intervals are asymptotically correct for any ρ in the interval $(-1, 1]$.

Table I. Trending properties of the data

Series	q	$\hat{\rho}$	SE ($\hat{\rho}$)	90% Confidence interval for ρ		t -Ratio	
				Grid bootstrap	Subsampling	$\hat{\delta}$	$\hat{\mu}$
y	2	0.948	0.018	(0.934, 0.993)	(0.888, 1.008)	2.778	2.904
m_1	5	0.988	0.007	(0.985, 1.007)	(0.966, 1.009)	1.629	1.847
m_2	5	0.995	0.004	(0.993, 1.006)	(0.995, 1.011)	1.155	1.574
r	5	0.901	0.034	(0.875, 0.984)	(0.806, 0.996)	0.014	2.291
Δy	1	0.376	0.093	(0.258, 0.562)	(0.206, 0.546)	-0.759	3.762
Δm_1	4	0.793	0.069	(0.739, 0.959)	(0.642, 0.944)	-0.903	2.284
Δm_2	4	0.766	0.071	(0.701, 0.930)	(0.597, 0.935)	-1.426	2.961
π	4	0.923	0.036	(0.905, 1.017)	(0.831, 1.014)	-1.025	1.950

Note: For a time series $\{Y_t\}$, results are based on the regression equation $Y_t = \mu + \delta t + \rho Y_{t-1} + \sum_{k=1}^q \gamma_k \Delta Y_{t-k} + \eta_t$, where $\Delta Y_t = Y_t - Y_{t-1}$. The notation $\hat{\cdot}$ indicates OLS estimates.

therefore, first-difference y , m_1 and m_2 in subsequent analysis. We will not difference π since the unit autoregressive root suggested by the results in Table I is most likely an artefact due to a substantial shift in the level of inflation in the early 1970s.¹⁰

Granger causality results are also sensitive with respect to the specification of deterministic trends, as Stock and Watson (1989), for example, demonstrated. We investigate this issue in the last two columns of Table I. Unlike Stock and Watson (1989) and Swanson (1998), who report evidence of significant deterministic trends in some of their measures of money growth, we find a significant drift in Δy , Δm_1 , Δm_2 and π but no statistically significant linear trend (at significance levels lower than 16%) over our sample period.¹¹ Hence, no specifications with deterministic trends will be considered in subsequent analysis.

4.2. Results from Linear VAR Models

As a starting point for the analysis, we consider conventional full-sample tests for Granger non-causality from m_1 , m_2 or r to y , based on linear VAR models of the form

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{k=1}^{n_1} \boldsymbol{\Phi}_k \mathbf{X}_{t-k} + \sum_{k=1}^{n_2} \boldsymbol{\beta}_k \pi_{t-k} + \boldsymbol{\xi}_t, \quad t = 1, 2, \dots, T \quad (2)$$

where $\mathbf{X}'_t = [\Delta y_t : \Delta m_{1,t}]$, $\mathbf{X}'_t = [\Delta y_t : \Delta m_{2,t}]$ or $\mathbf{X}'_t = [\Delta y_t : r_t]$, and $\boldsymbol{\xi}_t$ is a vector of disturbances.¹² Following the practice of some authors (e.g., Friedman and Kuttner, 1992), lagged inflation is included as a conditioning variable when testing for Granger causality between the

¹⁰ The recursive and rolling augmented Dickey–Fuller tests discussed in Banerjee *et al.* (1992) do indeed reject the hypothesis that inflation is integrated of order one.

¹¹ The model selection procedure of Phillips and Ploberger (1994) based on their posterior information criterion (PIC) does not support the presence of deterministic trends in these time series either.

¹² It is worth mentioning that y does not cointegrate with either m_1 or m_2 over our sample period.

financial variables and output. We set $n_1 = n_2 = 2$ for $\mathbf{X}'_t = [\Delta y_t : \Delta m_{1,t}]$ or $\mathbf{X}'_t = [\Delta y_t : \Delta m_{2,t}]$ and $n_1 = n_2 = 3$ for $\mathbf{X}'_t = [\Delta y_t : r_t]$.¹³

The p -value for a conventional non-causality F -test (from each of the nominal variables to real output) is 0.1484, 0.0002 and 0 for Δm_1 , Δm_2 and r , respectively. Thus, in the full sample there is evidence in favour of Δm_2 and the Federal Funds rate having significant predictive content for output growth, while Δm_1 does not appear to have any such predictive power. This difference between the predictive content of M1 and M2 growth is in line with previous findings in the literature (see, e.g., Friedman and Kuttner, 1992).

To examine the temporal stability of causal relationships, we now test the constancy of the parameters of the VAR models on which the causality tests are based using the procedures discussed in Hansen (1990), Andrews (1993) and Andrews and Ploberger (1994). These tests are based on functionals of the sequence of Lagrange multiplier statistics which test the null hypothesis of parameter stability against the alternative of a one-time change at each possible break point in the sample. The test statistics are defined as

$$\text{MLM} = \int_{\omega_1}^{\omega_2} \text{LM}(\omega) d\omega, \quad \text{ELM} = \ln \left(\int_{\omega_1}^{\omega_2} \exp\left\{\frac{1}{2}\text{LM}(\omega)\right\} d\omega \right), \quad \text{SLM} = \sup_{\omega_1 \leq \omega \leq \omega_2} \text{LM}(\omega)$$

where, for $0 < \omega < 1$, $\text{LM}(\omega)$ stands for the Lagrange multiplier statistic for testing the null hypothesis of constant parameters against the alternative of a change at date $t = [\omega T]$ (square brackets designating the integer part). Following Andrews (1993), the tests are implemented with $\omega_1 = 1 - \omega_2 = 0.15$. We also consider Nyblom's (1989) test for randomly time-varying parameters (denoted NYB), for which the alternative hypothesis is that the coefficients follow a random walk.

These tests are used to test the constancy of the coefficients in the VAR models defined in (2). In Table II, we report the outcome of the tests for the output equation and the associated p -values. These p -values are obtained from a bootstrap approximation to the null distribution of the test statistics, which was constructed by means of Monte Carlo simulation using 999 artificial samples generated from a VAR model with constant parameters. The sequential Lagrange multiplier tests suggest that there is significant evidence of parameter non-constancy in the output equation of the

Table II. Stability tests for output equation

Test	Δm_1	Δm_2	r
MLM	13.22 (0.038)	11.81 (0.098)	13.26 (0.385)
ELM	10.28 (0.004)	7.129 (0.104)	8.657 (0.268)
SLM	27.21 (0.004)	19.20 (0.107)	23.41 (0.193)
NYB	1.428 (0.162)	1.227 (0.322)	1.119 (0.879)

Note: Figures in parentheses are bootstrap p -values.

¹³ We note that, using the familiar Akaike information criterion (AIC) to select the order of the VAR models (with a maximum allowable order of 8), we obtain $n_1 = n_2 = 3$ for the model with the Federal Funds rate and $n_1 = n_2 = 1$ for the other two models. However, there is evidence of neglected error autocorrelation in first-order models and the p -values for Wald tests for the significance of the coefficients on second-order lags are less than 0.1. For these reasons we have decided to set $n_1 = n_2 = 2$ in the output–money models.

output–M1 and output–M2 models, although this evidence is not very strong in the latter model. No significant evidence of non-constancy is found in the model with the Federal Funds rate. It should be borne in mind, however, that these tests for parameter instability are not designed to be optimal against Markov changes in parameters and may not be particularly powerful in the presence of Markov regime switching of the type discussed in Section 2 (cf. Carrasco, 2002).

To investigate further the potential changes in causality over the sample period, we consider next F -tests for Granger non-causality computed from rolling subsamples of a fixed size (cf. Thoma, 1994; Swanson, 1998). More specifically, the p -values of F -statistics testing the lack of Granger causality from Δm_1 , Δm_2 or r to Δy are computed from the VAR models defined in (2) fitted to rolling windows of 24 observations (six years of data). To avoid the potential inaccuracies of conventional asymptotic inference procedures in small samples (only 24 observations in our case), the p -values were obtained from a bootstrap approximation to the sampling distribution of the test statistics. The bootstrap approximation was constructed, as before, by means of Monte Carlo simulation, using 999 artificial samples generated from a VAR model which satisfies the null hypothesis under test and the disturbances of which are obtained from the least-squares VAR residuals by equiprobable sampling with replacement.

The bootstrap p -values of the rolling test statistics are shown in Figure 1 (the horizontal axes show the final observation in each of the six-year rolling windows). In line with the full-sample results, the plots indicate that, at least during certain parts of the sample, the monetary variables have predictive power for output. However, the plots also suggest that there have been important changes in causal links over the sample period. We find that there is only a brief period during which M1 appears to have had predictive content for output. This period consists of a four-year period from 1983 to 1987. During the rest of the sample, the results do not indicate a significant causal role for M1. The results for M2 are somewhat similar, but for this variable the evidence suggests that M2 had predictive power for output mainly in the period 1973–1978. For both money-stock measures, however, we also observe that the p -values change substantially over the sample, although the causality from money to output cannot be rejected at the 10% level during most of the sample. Finally, the rolling-sample results indicate that the Federal Funds rate was Granger-causal during a long period from 1970 to around 1987.

4.3. Results from Markov Switching Models

Let us now consider the results from estimating Markov switching VAR models of the type specified in (1) with $\mathbf{X}'_t = [\Delta y_t : \Delta m_{1,t}]$, $\mathbf{X}'_t = [\Delta y_t : \Delta m_{2,t}]$ or $\mathbf{X}'_t = [\Delta y_t : r_t]$ and $Z_t = \pi_t$. We set $h_1 = h_2 = 2$ in all three cases, a choice of lag length which is supported by the data since the resulting models have standardized residuals which exhibit no significant autocorrelation (in either their levels or their squares) according to the portmanteau test of Ljung and Box (1978).¹⁴

The parameters of Markov switching models are estimated by maximum likelihood (ML), assuming that the conditional distribution of \mathbf{X}_t given $\{\mathbf{W}_{t-1}, \dots, \mathbf{W}_1, S_t, S_{t-1}, \dots, S_0\}$ is normal. The likelihood function is evaluated by means of an iterative filtering algorithm similar to the one discussed in Hamilton (1994, chap. 22), and the ML estimates are found by a quasi-Newton optimization algorithm which uses the Broyden–Fletcher–Goldfarb–Shanno secant update to the

¹⁴ We note that the choice $h_1 = h_2 = 2$ is also supported by the AIC. Results on the properties of order selection procedures based on complexity-penalized likelihood criteria for Markov switching autoregressive models can be found in Kapetanios (2001).

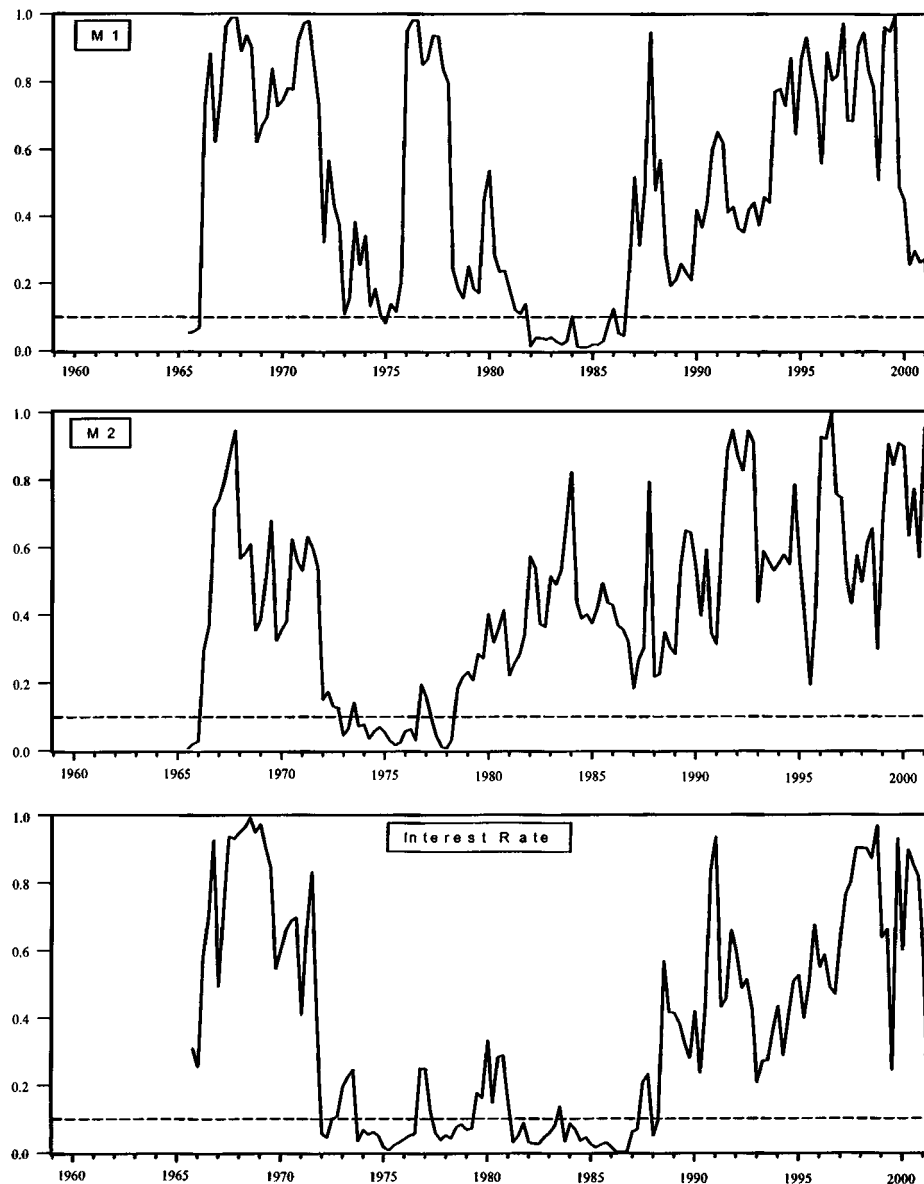


Figure 1. Bootstrap p -values for rolling Granger non-causality F -tests

hessian. The estimates of all 40 parameters of each switching VAR model and the associated asymptotic standard errors—computed from the inverse of the empirical Hessian—are reported in Tables III–V. (The output equation is always the first equation of the VAR.)

In all three bivariate models there is significant evidence of regime shifts for some of the parameters. Importantly, we find that $\psi_1^{(1)}$ and/or $\psi_1^{(2)}$, the parameters which determine whether

Table III. Estimates of parameters of the model for output and money (M1)

	Estimate	Std. error		Estimate	Std. error
$\phi_{10}^{(1)}$	0.027	0.095	$\phi_{20}^{(1)}$	-0.049	0.068
$\phi_{11}^{(1)}$	0.052	0.105	$\phi_{21}^{(1)}$	0.273	0.109
$\phi_{10}^{(2)}$	0.088	0.089	$\phi_{20}^{(2)}$	-0.084	0.084
$\phi_{11}^{(2)}$	0.097	0.117	$\phi_{21}^{(2)}$	-0.112	0.105
$\psi_1^{(1)}$	0.418	0.140	$\psi_2^{(1)}$	0.737	0.086
$\psi_1^{(2)}$	0.331	0.139	$\psi_2^{(2)}$	0.008	0.092
$\varphi_{10}^{(1)}$	0.044	0.327	$\varphi_{20}^{(1)}$	0.636	0.239
$\varphi_{11}^{(1)}$	0.013	0.720	$\varphi_{21}^{(1)}$	-0.257	0.202
$\varphi_{10}^{(2)}$	-0.892	0.357	$\varphi_{20}^{(2)}$	0.174	0.209
$\varphi_{11}^{(2)}$	0.177	0.524	$\varphi_{21}^{(2)}$	0.285	0.223
μ_{10}	0.650	0.189	μ_{20}	0.325	0.133
μ_{11}	0.461	0.366	μ_{21}	0.256	0.135
$\sigma_{11,1}$	0.979	0.373	$\sigma_{22,1}$	1.269	0.856
$\sigma_{11,2}$	0.608	0.125	$\sigma_{22,2}$	0.362	0.072
$\sigma_{11,3}$	0.964	0.328	$\sigma_{22,3}$	0.240	0.037
$\sigma_{11,4}$	0.191	0.036	$\sigma_{22,4}$	0.402	0.262
$\sigma_{12,1}$	1.653	0.402	$\sigma_{12,3}$	0.079	0.091
$\sigma_{12,2}$	0.237	0.077	$\sigma_{12,4}$	0.876	0.081
$p_{00}^{(1)}$	0.975	0.019	$p_{00}^{(2)}$	0.931	0.043
$p_{11}^{(1)}$	0.977	0.018	$p_{11}^{(2)}$	0.979	0.014
log-likelihood	-48.646				

there are changes in causality, are significantly different from zero.¹⁵ We also notice that the large estimated transition probabilities imply that the two regimes are very persistent. Further information about changes in causality throughout the sample period is given in Figures 2–4, where we plot the sum of $\mathbb{P}(S_t = 1 | \mathbf{W}_1, \dots, \mathbf{W}_t; \hat{\theta})$ and $\mathbb{P}(S_t = 3 | \mathbf{W}_1, \dots, \mathbf{W}_t; \hat{\theta})$, which is the estimated probability that money or the interest rate is Granger-caused for output at each point in the sample.¹⁶ The shaded areas correspond to recessions according to the National Bureau of Economic Research (NBER) business-cycle peaks and troughs (see <http://www.nber.org/cycles.html>).

In the case of M1, the results in Table III show that both $\psi_1^{(1)}$ and $\psi_1^{(2)}$ are significantly different from zero so M1 growth Granger-causes real GDP growth when $S_t = 1$ or $S_t = 3$. From Figure 2, we see that the strongest evidence in favour of money–output causality is obtained for the period after the first oil-price crisis, namely from 1977 up to the end of the Volcker disinflation. We also find evidence of money–output causality in a short period around the early 1990s recession, and less significantly around the early 1960s recession.

These results are interesting for many reasons. Firstly, while full-sample evidence does not support the hypothesis that M1 is Granger-causal for output, the results based on the switching model clearly show that M1 growth did have predictive power for output during various subperiods

¹⁵ The null hypothesis that output is not Granger-caused by the other endogenous variable in the system is rejected at level α if the largest (in absolute value) of the t -ratios associated with $\hat{\psi}_1^{(1)}$ and $\hat{\psi}_1^{(2)}$ exceeds the $(1 - \alpha/4)$ th quantile of the standard normal distribution. The test for Granger non-causality induced by the separate tests based on the individual t -ratios has an asymptotic significance level no larger than α (cf. Savin, 1980).

¹⁶ These plots show filtered probabilities, that is probabilities estimated using sample information up to the current observation date. Almost identical results are obtained using smoothed probabilities based on information available through the end of the sample.

Table IV. Estimates of parameters of the model for output and money (M2)

	Estimate	Std. error		Estimate	Std. error
$\phi_{10}^{(1)}$	0.047	0.101	$\phi_{20}^{(1)}$	0.562	0.131
$\phi_{11}^{(1)}$	0.314	0.110	$\phi_{21}^{(1)}$	0.834	0.091
$\phi_{10}^{(2)}$	-0.019	0.093	$\phi_{20}^{(2)}$	0.076	1.137
$\phi_{11}^{(2)}$	0.244	0.114	$\phi_{21}^{(2)}$	-0.167	0.087
$\psi_1^{(1)}$	0.146	0.191	$\psi_2^{(1)}$	-0.027	0.074
$\psi_1^{(2)}$	0.390	0.162	$\psi_2^{(2)}$	-0.056	0.071
$\varphi_{10}^{(1)}$	-0.138	0.316	$\varphi_{20}^{(1)}$	-0.294	0.222
$\varphi_{11}^{(1)}$	0.005	0.244	$\varphi_{21}^{(1)}$	-0.382	0.186
$\varphi_{10}^{(2)}$	-0.394	0.314	$\varphi_{20}^{(2)}$	0.373	0.225
$\varphi_{11}^{(2)}$	-0.205	0.272	$\varphi_{21}^{(2)}$	0.089	0.200
μ_{10}	0.479	0.182	μ_{20}	0.625	0.142
μ_{11}	0.316	0.359	μ_{21}	0.749	0.155
$\sigma_{11,1}$	1.124	0.223	$\sigma_{22,1}$	0.610	0.090
$\sigma_{11,2}$	0.423	0.116	$\sigma_{22,2}$	0.207	0.038
$\sigma_{11,3}$	1.807	1.875	$\sigma_{22,3}$	2.620	4.765
$\sigma_{11,4}$	0.234	0.039	$\sigma_{22,4}$	0.238	0.063
$\sigma_{12,1}$	0.331	0.104	$\sigma_{12,3}$	3.869	2.291
$\sigma_{12,2}$	0.103	0.053	$\sigma_{12,4}$	0.248	0.039
$p_{00}^{(1)}$	0.993	0.008	$p_{00}^{(2)}$	0.973	0.021
$p_{11}^{(1)}$	0.992	0.010	$p_{11}^{(2)}$	0.983	0.012
log-likelihood	0.816				

in the sample. Secondly, we find, as expected, that the predictive power of M1 is highest around the Volcker period. Furthermore, the results indicate that monetary policies directed towards instruments that affect high-powered money may have been used actively during the two recessions in the early 1980s and the recession in the early 1990s, and that these policies have had some effect on the predictive power of money for output growth. For the early 1980s episodes, this result seems natural given the money growth policies adopted under Paul Volcker's disinflation and the subsequent recessions. Finally, the results also make it clear why Friedman and Kuttner (1992) found that the evidence on money–output causality is much weaker if data for the 1980s are considered, while Stock and Watson (1989) documented the opposite result. We find that the 1980s are divided into two distinct periods, with evidence in favour of causality from money to output found in the first half and evidence of non-causality found in the second. Thus, when using conventional full-sample tests of Granger causality, it is crucial to the results whether one includes in the sample data from only the first half of the 1980s, as in Stock and Watson (1989), or data from the entire 1980s, as in Friedman and Kuttner (1992).

The results for M2, shown in Table IV, are in many ways similar to those for M1. We find that $\psi_1^{(2)}$ is significantly different from zero so that M2 growth causes real GDP growth in the states represented by $S_t = 1$ and $S_t = 3$. The plot of the estimated probability that M2 growth is Granger-causal for output, shown in Figure 3, is similar to that for M1 growth with the following exceptions. First, the period where M2 growth causes output growth is longer than the period identified for M1 growth. For M2, money causes output during most of the period 1970–1983, apart from a very short period that coincides with the first oil-price crisis. Second, we no longer find causality around the early 1990s recession. These differences may occur for a number of

Table V. Estimates of parameters of the model for output and interest rate

	Estimate	Std. error		Estimate	Std. error
$\phi_{10}^{(1)}$	0.218	0.088	$\phi_{20}^{(1)}$	0.179	0.043
$\phi_{11}^{(1)}$	0.053	0.094	$\phi_{21}^{(1)}$	1.189	0.067
$\phi_{10}^{(2)}$	0.117	0.096	$\phi_{20}^{(2)}$	0.056	0.044
$\phi_{11}^{(2)}$	0.180	0.072	$\phi_{21}^{(2)}$	-0.259	0.059
$\psi_1^{(1)}$	0.121	0.055	$\psi_2^{(1)}$	0.699	0.175
$\psi_1^{(2)}$	-0.238	0.061	$\psi_2^{(2)}$	0.067	0.195
$\varphi_{10}^{(1)}$	-0.173	0.380	$\varphi_{20}^{(1)}$	0.270	0.135
$\varphi_{11}^{(1)}$	0.248	0.230	$\varphi_{21}^{(1)}$	-0.781	1.350
$\varphi_{10}^{(2)}$	0.749	0.362	$\varphi_{20}^{(2)}$	0.174	0.209
$\varphi_{11}^{(2)}$	-0.594	0.240	$\varphi_{21}^{(2)}$	1.903	1.380
μ_{10}	0.480	0.299	μ_{20}	-0.121	0.154
μ_{11}	0.887	0.145	μ_{21}	-0.137	0.108
$\sigma_{11,1}$	0.735	0.137	$\sigma_{22,1}$	0.204	0.017
$\sigma_{11,2}$	0.351	0.225	$\sigma_{22,2}$	0.972	2.555
$\sigma_{11,3}$	0.172	0.048	$\sigma_{22,3}$	0.291	0.166
$\sigma_{11,4}$	1.704	0.734	$\sigma_{22,4}$	3.444	5.296
$\sigma_{12,1}$	0.074	0.039	$\sigma_{12,3}$	0.546	0.061
$\sigma_{12,2}$	4.491	0.420	$\sigma_{12,4}$	11.540	1.713
$p_{00}^{(1)}$	0.834	0.061	$p_{00}^{(2)}$	0.973	0.016
$p_{11}^{(1)}$	0.810	0.070	$p_{11}^{(2)}$	0.871	0.070
log-likelihood	-47.106				

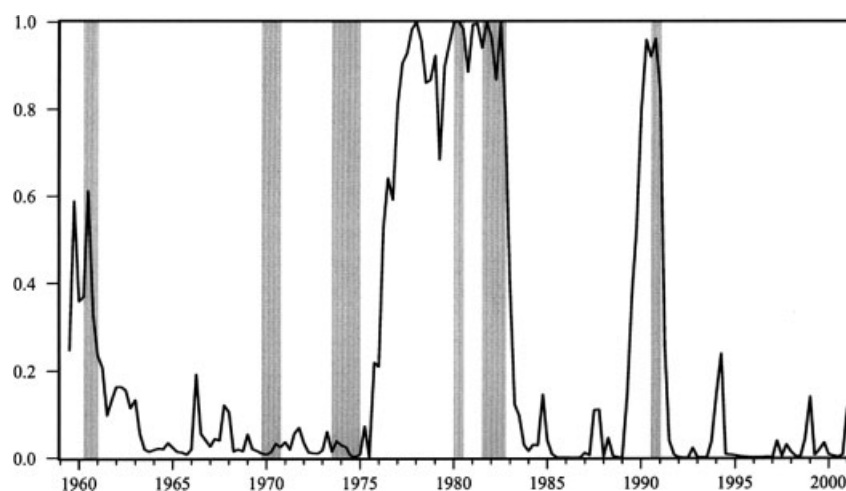


Figure 2. Inferred probability of M1 Granger-causing output

reasons but, given the difference in the content of M1 and M2, it is not surprising that these money stock measures may behave differently. We note again that the full-sample results are somewhat misleading in so far as they would lead one to conclude that M2 growth has predictive power over output when we find that the predictive content of M2 has in fact vanished since the end of the Volcker disinflation.

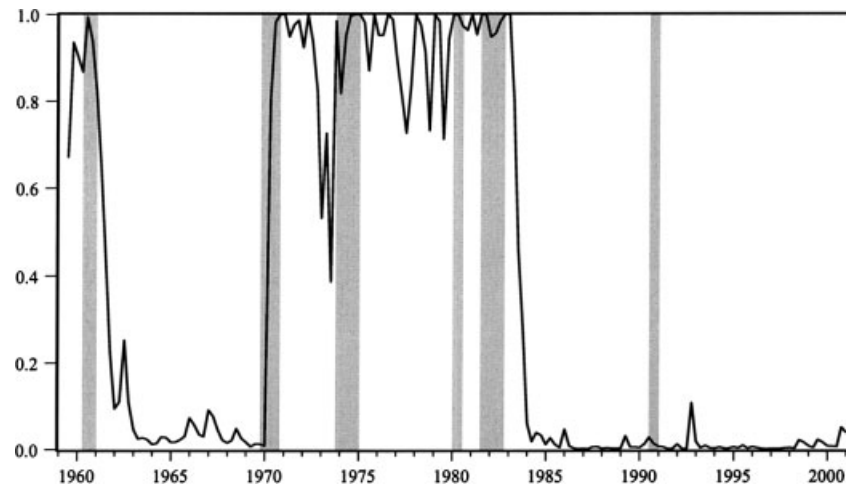


Figure 3. Inferred probability of M2 Granger-causing output

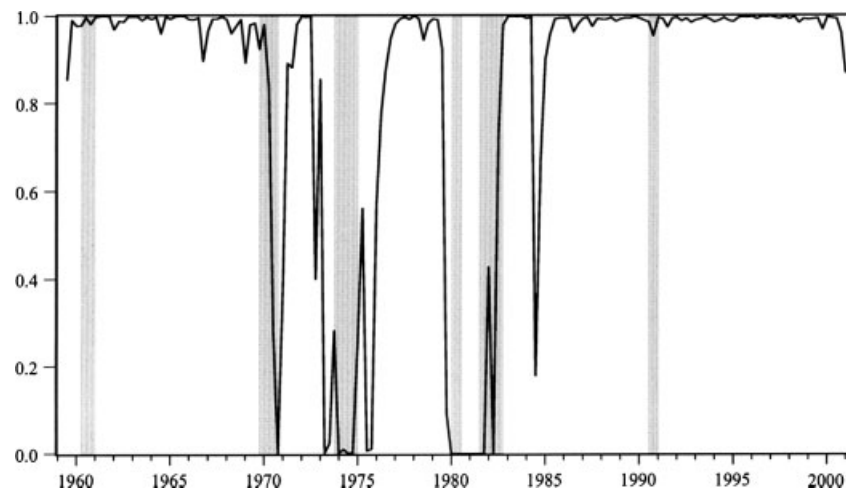


Figure 4. Inferred probability of the interest rate Granger-causing output

It is interesting to note that our results suggest that money-stock growth is more likely to have predictive power for output growth during recessions than during expansions. In particular, M1 growth has predictive power for output growth during four of the six recessions in the sample and very little predictive power during any expansionary periods. M2 growth has predictive power for output growth during all recessions apart from the early 1990s recession, while during expansions M2 growth has predictive power only during parts of the 1970s. Thus, there seems to be asymmetry in the causal links between money and output related to the state of the business cycle. A possible explanation for this result is that monetary policy might assume a more active role during—or just before—recessions in order to prevent further deteriorations in the economy, while more accommodating policies are applied during expansions.

The results for the Federal Funds rate are shown in Table V and the estimated probabilities that the interest rate is Granger-causal for output are plotted in Figure 4. Evidently, there is a clear division of the sample period into periods of causality from interest rates to output. The interest rate Granger-causes output in the period 1960–1969 and then again from 1983 onwards. In the intermediate period, we find that the probability that the interest rate causes output drops in each of the recessions in the late 1960s, the oil-price crises in the 1970s and the early 1980s recession (when the Fed targeted money growth and the Federal Funds rate soared). Furthermore, the interest rate appears to regain predictive power for output in each of the intermediate periods between these recessions. It is worth noting that the Federal Funds rate maintains its predictive power in the 1990s when the money stocks lose their predictive ability. This may reflect the fact that the Federal Funds rate is endogenous and hence has predictive power for output growth—to the extent that the demand and supply of Federal Funds have some predictive power for output (over and above the predictive power of past output growth itself). Notwithstanding the lack of a direct link between predictive ability and causation in an economic sense, this evidence may offer an explanation as to why the Federal Funds rate seems to work better than money stocks as a basis for evaluating the stance and effects of US monetary policy.¹⁷

4.4. Some Simulation Results

One issue that we have not addressed so far is the effectiveness of our method as a means of detecting changes in causality. In order to get some insight into the ‘precision’ of the method in identifying the different regimes in cases where there are periods of causality and non-causality, we carry out a few Monte Carlo experiments which are based on the empirical results reported in Tables III–V.

To be more precise, the basis of our calculations is 500 independent samples of size $T = 170$ from a bivariate Markov switching model like (1) with Gaussian errors and $h_1 = h_2 = 2$. To ensure the empirical relevance of the simulations, the coefficients and error covariance matrix of the model are chosen to be the ML estimates reported in Tables III–V (referred to in the sequel as models 2, 3 and 4, respectively). In all cases, observed historical values of inflation are used as $\{Z_t\}$. The transition probabilities of $\{S_{1,t}\}$ and $\{S_{2,t}\}$ are chosen as: (i) $p_{00}^{(1)} = p_{11}^{(1)} = 0.9$, $p_{00}^{(2)} = p_{11}^{(2)} = 0.96$; or (ii) $p_{00}^{(1)} = p_{11}^{(1)} = 0.98$, $p_{00}^{(2)} = p_{11}^{(2)} = 0.96$. The difference between the two sets of transition probabilities is that the expected duration of the regime in which $X_{2,t}$ is Granger-causal for $X_{1,t}$ is much higher in case (ii) (50 time periods, compared with an expected duration of 10 periods for case (i)).

For each artificial sample, ML estimates of the parameters of model (1) (with $h_1 = h_2 = 2$) are obtained and the probabilities $\mathbb{P}(S_t = \ell | \mathbf{W}_1, \dots, \mathbf{W}_t; \hat{\theta})$, $\ell = 1, \dots, 4$, are computed using these estimates. To assess how accurately the four regimes can be classified by using the filtered probabilities, we then compute the value of the criterion

$$C_\ell = (1/T) \sum_{t=1}^T \{\mathbb{P}(S_t = \ell | \mathbf{W}_1, \dots, \mathbf{W}_t; \hat{\theta}) - \mathbb{I}(S_t = \ell)\}^2, \quad \ell = 1, \dots, 4$$

¹⁷ It is perhaps worth mentioning that we repeated the analysis excluding inflation from the switching VAR equations. The results changed by very little and the filtered probabilities identified virtually the same dates as dates at which changes in causality took place.

Table VI. Monte Carlo mean of C_ℓ

	$p_{00}^{(1)}, p_{11}^{(1)}$	$p_{00}^{(2)}, p_{11}^{(2)}$	C_1	C_2	C_3	C_4
Model 2	0.90	0.96	0.14	0.07	0.07	0.11
	0.98	0.96	0.08	0.03	0.05	0.10
Model 3	0.90	0.96	0.11	0.04	0.07	0.08
	0.98	0.96	0.07	0.06	0.07	0.07
Model 4	0.90	0.96	0.12	0.11	0.08	0.10
	0.98	0.96	0.07	0.05	0.07	0.09

where $\mathbb{I}(A)$ denotes the indicator of the event A . Note that $0 \leq C_\ell \leq 1$ and low values of C_ℓ imply that the method works well while high values imply inaccurate classification of regimes.

Table VI reports the average value of C_ℓ across the 500 Monte Carlo replications. For all design points, the average value of the regime-classification indicator is fairly small, suggesting that our method is quite accurate in identifying the points in the sample at which changes in causality have occurred. Combined with the other attractive features of our approach discussed earlier, these results reinforce the case for using Markov switching models of the form (1) to analyse causality patterns which are suspected to be unstable over the sample period.

5. SUMMARY

In this paper we have proposed using a VAR model with time-varying parameters to carry out Granger causality analysis in situations in which causality is non-permanent. Parameter changes are modelled so as to be directly related to changes in causality. Furthermore, changes in parameters (and hence causality regimes) are viewed as random events that are governed by an exogenous, unobservable Markov chain. Such a specification offers a relatively parsimonious way of allowing for multiple stochastic changes in the causal relationships between variables of interest, and permits probabilistic inferences about the causality regime that is operable at each date in the sample period to be made.

The proposed methodology has been used to investigate the causal relationships between US money, interest rates and output. Our empirical findings suggest that there have been significant changes in the predictive content of M1, M2 and the Federal Funds rate for real GDP growth over the period 1959–2001. We believe that such evidence on the time-varying nature of the predictive power of financial variables is not surprising given the important changes that have occurred in the conduct of monetary policy and operating procedures as well as the developments that have taken place in the financial sector.

ACKNOWLEDGEMENTS

The authors are grateful to John Driffill, James Hamilton, Andrew Scott, Ron Smith and two anonymous referees for helpful comments on an earlier draft. Martin Sola gratefully acknowledges the financial support of the UK Economic and Social Research Council through grant L138251003.

REFERENCES

- Andrews DWK. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* **61**: 821–856.
- Andrews DWK, Ploberger W. 1994. Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica* **62**: 1383–1414.
- Backus DK, Kehoe PJ. 1992. International evidence on the historical properties of business cycles. *American Economic Review* **82**: 864–888.
- Bai J. 1999. Likelihood ratio tests for multiple structural changes. *Journal of Econometrics* **91**: 299–323.
- Bai J, Perron P. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* **66**: 47–78.
- Banerjee A, Lumsdaine RL, Stock JH. 1992. Recursive and sequential tests of the unit root and trend-break hypotheses: theory and international evidence. *Journal of Business and Economic Statistics* **10**: 271–287.
- Bernanke BS. 1986. Alternative explanations of the money–income correlation. *Carnegie-Rochester Conference Series on Public Policy* **25**: 49–99.
- Carrasco M. 2002. Misspecified structural change, threshold, and Markov-switching models. *Journal of Econometrics* **109**: 239–273.
- Christiano LJ, Ljungqvist L. 1988. Money does Granger-cause output in the bivariate money–output relation. *Journal of Monetary Economics* **22**: 217–235.
- Cooley TF, Dwyer M. 1998. Business cycle analysis without much theory: a look at structural VARs. *Journal of Econometrics* **83**: 57–88.
- Cooley TF, LeRoy SF. 1985. Atheoretical macroeconomics: a critique. *Journal of Monetary Economics* **16**: 283–308.
- Eichenbaum M, Singleton KJ. 1986. Do equilibrium real business cycle theories explain postwar U.S. business cycles? In *NBER Macroeconomics Annual 1986*, Fischer S (ed.). MIT Press: Cambridge, MA; 91–135.
- Friedman BM, Kuttner KN. 1992. Money, income, prices, and interest rates. *American Economic Review* **82**: 472–492.
- Friedman BM, Kuttner KN. 1993. Another look at the evidence on money–income causality. *Journal of Econometrics* **57**: 189–203.
- Gavin WT, Kydland FE. 1999. Endogenous money supply and the business cycle. *Review of Economic Dynamics* **2**: 347–369.
- Hamilton JD. 1994. *Time Series Analysis*. Princeton University Press: Princeton, NJ.
- Hansen BE. 1990. Lagrange multiplier tests of parameter instability in non-linear models. Unpublished manuscript, Department of Economics, University of Rochester.
- Hansen BE. 1999. The grid bootstrap and the autoregressive model. *Review of Economics and Statistics* **81**: 594–607.
- Kapetanios G. 2001. Model selection in threshold models. *Journal of Time Series Analysis* **22**: 733–754.
- Kim C-J, Nelson CR. 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press: Cambridge, MA.
- Krol R, Ohanian LE. 1990. The impact of stochastic and deterministic trends on money–output causality: a multi-country investigation. *Journal of Econometrics* **45**: 291–308.
- Krolzig H-M. 1997. *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis. Lecture Notes in Economics and Mathematical Systems* No. 454. Springer-Verlag: Berlin.
- Krolzig H-M. 2000. Predicting Markov-switching vector autoregressive processes. Unpublished manuscript, Nuffield College, University of Oxford.
- Litterman RB, Weiss L. 1985. Money, real-interest rates, and output: a reinterpretation of postwar U.S. data. *Econometrica* **53**: 128–156.
- Ljung GM, Box GEP. 1978. On a measure of lack of fit in time series models. *Biometrika* **65**: 297–303.
- Ng S, Perron P. 1995. Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association* **90**: 268–281.
- Nyblom J. 1989. Testing for the constancy of parameters over time. *Journal of the American Statistical Association* **84**: 223–230.
- Phillips PCB, Ploberger W. 1994. Posterior odds testing for a unit root with data-based model selection. *Econometric Theory* **10**: 774–808.

- Romano JP, Wolf M. 2001. Subsampling intervals in autoregressive models with linear time trend. *Econometrica* **69**: 1283–1314.
- Savin NE. 1980. The Bonferroni and the Scheffé multiple comparison procedures. *Review of Economic Studies* **47**: 255–273.
- Sims CA. 1972. Money, income and causality. *American Economic Review* **62**: 540–552.
- Sims CA. 1980. Macroeconomics and reality. *Econometrica* **48**: 1–48.
- Sola M, Driffill J. 1994. Testing the term structure of interest rates using a stationary vector autoregression with regime switching. *Journal of Economic Dynamics and Control* **18**: 601–628.
- Stock JH, Watson MW. 1989. Interpreting the evidence on money–income causality. *Journal of Econometrics* **40**: 161–181.
- Stock JH, Watson MW. 1996. Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* **14**: 11–30.
- Swanson NR. 1998. Money and output through a rolling window. *Journal of Monetary Economics* **41**: 455–473.
- Thoma MA. 1994. Subsample instability and asymmetries in money–income causality. *Journal of Econometrics* **64**: 279–306.
- Warne A. 1999. Causality and regime inference in a Markov switching VAR. Unpublished manuscript, Research Department, Sveriges Riksbank, Stockholm.