

Money-output Causality Revisited – A Bayesian Logistic Smooth Transition VECM Perspective*

DEBORAH GEFANG

*Department of Economics, Lancaster University Business School, Lancaster University,
Lancaster LA1 4YX, UK (e-mail: d.gefang@lancaster.ac.uk)*

Abstract

This article proposes a Bayesian approach to examining money-output causality within the context of a logistic smooth transition vector error correction model. Our empirical results provide substantial evidence that the postwar US money-output relationship is nonlinear, with regime changes mainly governed by the output growth and price levels. Furthermore, we obtain strong support for nonlinear Granger causality from money to output, although there is also some evidence for models indicating that money is not Granger causal or long-run causal to output.

I. Introduction

The effect of money on output appears to be of great importance both to economists and policymakers. The issue has been intensively investigated in the literature (see Sims, 1972, 1980; Stock and Watson, 1989; King and Watson, 1997). Yet, from the late 1980s through the early 2000s, with the prevalence of the Taylor rule (Taylor, 1993), the role of money (monetary base or monetary aggregates) has been de-emphasized or overlooked in many studies (see Taylor, 1999; Clarida, Galí and Gertler, 2000). Recent years have witnessed a renewed interest in the effect of money. Meltzer (2001), Nelson (2002, 2003), Duca and VanHoose (2004), among others, raise the issue that money constitutes a crucial channel for the transmission mechanism of monetary policy, and the role of money cannot be simply replaced by other policy instruments. In particular, money re-emerges as an important variable in a number of recent empirical studies (see Wang and Wen, 2005; Sims and Zha, 2006; Hill, 2007). Moreover, the policy responses to the current economic crises indicate that money remains to be a vital policy instrument, even though the full consequences of the unprecedented money injections are still uncertain.

This article contributes to the discussion on whether money matters by revisiting an old topic: the causal effects from money to output in postwar US data. However, the present

*The author is indebted to Gary Koop, Anders Rahbek and two anonymous referees for a number of constructive suggestions. The author would also like to thank Christopher Bowdler, Stephen Hall, Roberto Leon-Gonzalez, Roderick McCrorie, Emi Mise, Rodney Strachan and Jim Taylor for helpful comments. Any remaining errors are the author's responsibility. Financial support from the Department of Economics, University of Leicester is gratefully acknowledged.

JEL Classification numbers: C11, C32, E42, E52.

article departs from the literature in two main aspects. First, to capture the possible regime changes in US monetary policy, we adopt a logistic smooth transition vector error correction model (STVECM) incorporating cointegration of an unknown form. Second, we develop a simple Bayesian approach to investigating the causal effects in the STVECM context.

Single-equation smooth transition error correction models have been widely used in the literature to capture the possible nonlinear money-output relationship (see Lütkepohl, Teräsvirta and Wolters, 1999; Teräsvirta and Eliasson, 2001; Escribano, 2004; Haug and Tam, 2007). However, considering the interaction between endogenously determined money, interest rates and the ultimate policy targets of output and inflation, we believe STVECM can be more effective in capturing both the long- and short-run dynamics in the linkage among all the variables. Perhaps the reason why researchers have not followed this route is because of the lack of a fully developed statistics tool that can directly test the cointegration (or no cointegration) null in a nonlinear VECM against its both linear and nonlinear alternatives (see Seo, 2004, 2006; Kapetanios, Shin and Snell, 2006, for details). In the literature, only Rothman, van Dijk and Franses (2001) apply a multivariate STVECM framework, which is closest to us to study the money-output relationship.¹ Yet Rothman *et al.* (2001) pre-impose a theory-based long-run cointegrating relationship in their estimation. While recognizing that the actual money-output interrelationship can be rather complex, unlike Rothman, *et al.* (2001), we let both the cointegration rank and cointegrating vectors be determined by the data. In particular, we allow for different cointegrating ranks across the upper and lower regimes.² To our knowledge, we are among the first in the literature to adopt this more flexible modelling framework in the present context.

Our estimation technique is Bayesian. Specifically, we extend the Bayesian cointegration space approach introduced in Strachan and Inder (2004) and the collapsed Gibbs sampler developed in Koop, Leon-Gonzalez and Strachan (2005) into the nonlinear framework. Our method jointly captures the equilibrium and presence of nonlinearity in the STVECM in a single step. Compared with available classical estimation techniques which often require multiple steps and Taylor expansions, our Bayesian approach is less susceptible to the sequential testing and inaccurate approximation problems. Furthermore, the commonly used maximum likelihood estimation in classical works is subject to the multi-mode problem caused by the nuisance parameters in the transition function of the STVECM. Yet, jagged likelihood functions do not create any particular problems in our Gibbs sampling scheme.

Adopting a Bayesian perspective, we treat models as random variables. Considering that the large model we employed is subject to the criticism of being too parameter rich, we use Bayes factors for model comparison to reward more parsimonious models.³ Alternative models are specified by placing zero restrictions on certain parameters of the unrestricted STVECM. Our approach to examining whether money in the long-run causes output is in spirit closest to that in Hall and Wickens (1993), Hall and Milne (1994) and Granger and

¹Rothman *et al.* (2001) test Granger causality from money to output in a classical context involving rolling window forecasting.

²I thank an anonymous referee for pointing out this possibility.

³Bayes factors include an automatic penalty for more complex models (see Koop and Potter, 1999a,b for details).

Lin (1995). We specify that money is not long-run causal to output if departures from the long-run cointegration equilibrium involving money do not enter the output equation. With regards to Granger causality from money to output, following Rothman *et al.* (2001), we consider the following two cases. First, the disequilibria from the long-run cointegration relationship involving the money variables and the first differences of money may enter the output equation as regressors. Second, money or the growth of money may be identified as the transition indicator which triggers switching between regimes. As a complement to Rothman *et al.* (2001), we also define money as Granger causal to output by identifying money velocity as the transition variable in the STVECM model.

Consistent with the in-sample testing results in Rothman *et al.* (2001), we find compelling evidence that money is nonlinearly Granger causal to output. More specifically, we find that STVECM models with lagged money variables entering the output equation as regressors receive an overwhelming posterior support. However, there is also some support for the models indicating that money is neither Granger causal nor long-run causal to output. Experimenting with 18 candidate transition indicators, we find that regime changes are mainly governed by the lagged monthly output growth and price level. In contrast, we find little evidence that either money or money velocity may be identified as the transition variable.

The model comparison results show that there is a great degree of model uncertainty about the the postwar US money-output relationship. To account for this uncertainty, we use Bayesian model averaging (BMA) to calculate the general impulse response functions. The impulse response analysis shows that the dynamics of output growth given an unexpected money shock are rather complex: both the sign and magnitude of the initial shock matter a lot. This result is not surprising given that the nonlinear models together receive a majority of the posterior mass.

The outline of this article is as follows. Section II describes the model and the Bayesian method. Section III reports the empirical results. Section IV concludes. The technical details and evidence on prior sensitivity analysis are provided in the Appendices.

II. STVECM model and Bayesian inference

Aside from the STVECM, various types of nonlinear VECM have been proposed in the literature, for instance, the threshold autoregressive types of models (see Hansen and Seo, 2002; Aslanidis and Kouretas, 2005), the Markov switching types of models (see Francis and Owyang, 2005; Sugita, 2008), the autoregressive conditional root models (see Bec and Rahbek, 2004; Bec, Rahbek and Shephard, 2008) and the mixture autoregressive models (see Saikkonen, 2005; Fong, Li and Wong, 2007). Considering that regime changes in macroeconomic data can be either gradual or abrupt, in this article, we adopt an STVECM since it allows for both the smooth and discrete adjustment mechanisms, while most of the other frameworks tend to suggest discontinuous regime changes.

The STVECM we adopted is largely in line with Rothman *et al.* (2001). More specifically, we investigate the money-output linkage within a nonlinear interdependent system among output (y_t), money (m_t), prices (p_t) and interest rates (i_t). Letting $x_t = (y_t, m_t, p_t, i_t)$, the STVECM of the $1 \times n$ (here $n = 4$) vector time series process x_t , $t = 1, \dots, T$, conditioning on the p observations $t = -p + 1, \dots, 0$, can be specified as:

$$\Delta x_t = x_{t-1}\beta\alpha + \sum_{h=1}^p \Delta x_{t-h}\Gamma_h + d_t\xi + F(z_t) \left(x_{t-1}\beta^z\alpha^z + \sum_{h=1}^p \Delta x_{t-h}\Gamma_h^z + d_t\xi^z \right) + \varepsilon_t, \quad (1)$$

where ε_t is a Gaussian white noise process with $E(\varepsilon_t) = 0$, $E(\varepsilon'_s\varepsilon_t) = \Sigma$ for $s = t$, and $E(\varepsilon'_s\varepsilon_t) = 0$ for $s \neq t$. Note that $\Delta x_t = x_t - x_{t-1}$ and $d_t = (1, t)$.⁴ The dimensions of Γ_h and Γ_h^z are $n \times n$, and the dimensions of ξ and ξ^z are $2 \times n$. Following the suggestions of an anonymous referee, we specify the dimensions of β and α' are $n \times r_1$, and the dimensions of β^z and α'^z are $n \times r_2$, where r_1 is not necessarily equal to r_2 . Note that this specification allows for a more flexible long-run cointegration relationship.

In model (1), the dynamics of the regime changes are assumed to be captured by the first-order logistic smooth transition function introduced in Granger and Teräsvirta (1993) and Teräsvirta (1994):

$$F(z_t) = \{1 + \exp[-\gamma(z_t - c)/\sigma]\}^{-1}, \quad (2)$$

where z_t is the transition variable determining the regimes. Note that z_t can be any exogenous or lagged endogenous variables of interest. Using the Bayesian approach, we are able to search over large numbers of candidate choices for z_t (or average over them).

The transition function $F(z_t)$ is bounded by 0 and 1. As convention, we define $F(z_t) = 0$ and $F(z_t) = 1$ corresponding to the lower and upper regimes, respectively. In function (2), the smoothing parameter γ (which is non-negative) determines the speed of the smooth transition. Observe that when $\gamma \rightarrow \infty$, the transition function becomes a Dirac function, model (1) becomes a two-regime threshold VECM along the lines of Tong (1983). When $\gamma = 0$, the logistic function becomes a constant (equal to 0.5), and the nonlinear model (1) collapses into a linear VECM. The transition parameter c is the threshold around which the dynamics of the model change. The value for the parameter σ is chosen by the researcher and could reasonably be set to 1. In this study, we set σ equal to the standard deviation of the process z_t . This effectively normalizes γ such that we can give γ an interpretation in terms of the inverse of the number of standard deviations of z_t .

The transition from one extreme regime to the other is smooth for reasonable values of γ . At the lower regime, the dynamics of model (1) is determined by

$$\Delta x_t = x_{t-1}\beta\alpha + \sum_{h=1}^p \Delta x_{t-h}\Gamma_h + d_t\xi + \varepsilon_t. \quad (3)$$

While at the upper regime, model (1) turns into

$$\Delta x_t = x_{t-1}(\beta\alpha + \beta^z\alpha^z) + \sum_{h=1}^p \Delta x_{t-h}(\Gamma_h + \Gamma_h^z) + d_t(\xi + \xi^z) + \varepsilon_t. \quad (4)$$

Similar to Rothman *et al.* (2001), we specify that if money does not Granger-cause output, deviations from the cointegration relationships associated with money and lagged money growth do not enter the equation for output, and money and the velocity of money cannot be identified as the transition variables triggering regime changes. In terms of long-

⁴Different forms of the deterministic terms described in Johansen (1995, ch. 5) can be achieved by putting restrictions in model (1). We do not investigate those exact forms in this article since their implications are not crucial in studying our main interest of concern.

run causality, following Hall and Wickens (1993), Hall and Milne (1994) and Granger and Lin (1995), we specify that if money is not long-run causal (or weakly causal) for output, departures from the equilibrium of the long-run cointegration relationships associated with money do not affect output.

It is seen that model (1) encompasses a set of models distinguished by the existence of the long-run cointegration relationships, the ranks of the cointegrating vectors, the order of the autoregressive process, the existence of the nonlinear effects, the choice of the transition variables, and whether Granger non-causality or long-run non-causality from money to output is imposed.

The asymptotic theories on testing and estimating the smooth transition types of VECM models have attracted a lot of attention in the classical context, latest development see Kristensen and Rahbek (2008), Saikkonen (2008), among others. Our econometric method is Bayesian. Compared with the available classical approaches which often require multiple tests and Taylor expansions, our Bayesian method is more straightforward and less subject to the sequential testing problems. The detailed Bayesian techniques are provided in appendix A. Here we only sketch four of the major points. First, following Strachan (2003), Strachan and Inder (2004) and Strachan and van Dijk (2004, 2006), we elicit uniform priors on the cointegration space to circumvent the local and global non-identification problems associated with the long-run multipliers $\beta\alpha$ and $\beta^z\alpha^z$. Second, we incorporate the collapsed Gibbs sampler developed in Koop *et al.* (2005) into our posterior simulation algorithm to improve the computing efficiency. Third, following Bauwens, Lubrano and Richard (1999) and Lubrano (1999a,b), we elicit our priors for the nuisance parameter γ in the transition function as a Gamma distribution to avoid the non-identification problem caused by $\gamma = 0$. Finally, to reward more parsimonious models, we use Bayes factors derived from the Savage–Dickey density ratios (SDDRs) to compute posterior model probabilities. Taking a Bayesian approach we have a number of options for obtaining inference. If a single model has dominant support, we can model the data generating process via this most preferred model. However, if there is considerable model uncertainty then it would make sense to use BMA and weigh features of interest across different models using posterior model probabilities (as suggested by Leamer, 1978).

III. Empirical results

We use the monthly US data spanning from 1959:1 to 2006:12. The data are obtained from the database of the Federal Reserve Bank of St. Louis. Various measures of output, money, prices and interest rates are used in the literature. In this article, we adopt the seasonally adjusted industrial production index (y_t), the seasonally adjusted M2 money stock (m_t), the producer price index for all commodities (p_t) and the secondary market rate on 3-month Treasury bills (i_t) for the measures of output, money, prices and interest rates, respectively. All variables are in logarithms except for interest rates which are in per cent.

In empirical work, we allow the cointegration rank of the unrestricted model (1) to vary from 1 to 3.⁵ Remember that our econometric model indicates that the cointegration

⁵We do not consider unrestricted models with rank 0 since they can be derived by imposing zero restrictions on the long-run adjustment parameters of the unrestricted models with rank 1, 2 or 3. In addition, we rule out the possibility that the cointegration rank is equal to 4 for that can only happen when the time series i_t , m_t , p_t and r_t are stationary.

ranks in the upper and lower regimes are not necessarily the same. Thus, with regard to the cointegration relationship, we have nine cases to investigate depending on the combinations of r_1 and r_2 in the lower and upper regimes.

- ($r = 1$), where $r_1 = 1, r_2 = 1$
- ($r = 2$), where $r_1 = 2, r_2 = 2$
- ($r = 3$), where $r_1 = 3, r_2 = 3$
- ($r = 4$), where $r_1 = 1, r_2 = 2$
- ($r = 5$), where $r_1 = 1, r_2 = 3$
- ($r = 6$), where $r_1 = 2, r_2 = 1$
- ($r = 7$), where $r_1 = 2, r_2 = 2$
- ($r = 8$), where $r_1 = 3, r_2 = 1$
- ($r = 9$), where $r_1 = 3, r_2 = 2$.

As explained before, the STVECM allows for any exogenous or lagged endogenous variables of interest to be the transition variables to trigger the regime changes. In this article, we consider the following 18 possible transition variables.⁶

- ($f = 1$) $z_t = y_{t-1}$, output level
- ($f = 2$) $z_t = m_{t-1}$, money
- ($f = 3$) $z_t = p_{t-1}$, price level
- ($f = 4$) $z_t = i_{t-1}$, interest rate level
- ($f = 5$) $z_t = \Delta^1 y_{t-1}$, monthly output growth
- ($f = 6$) $z_t = \Delta^1 m_{t-1}$, monthly money growth
- ($f = 7$) $z_t = \Delta^1 p_{t-1}$, monthly inflation rate
- ($f = 8$) $z_t = \Delta^1 i_{t-1}$, monthly interest rate growth
- ($f = 9$) $z_t = \Delta^4 y_{t-1}$, quarterly output growth
- ($f = 10$) $z_t = \Delta^4 m_{t-1}$, quarterly money growth
- ($f = 11$) $z_t = \Delta^4 p_{t-1}$, quarterly inflation
- ($f = 12$) $z_t = \Delta^4 i_{t-1}$, quarterly interest rate growth
- ($f = 13$) $z_t = \Delta^{12} y_{t-1}$, yearly output growth
- ($f = 14$) $z_t = \Delta^{12} m_{t-1}$, yearly money growth
- ($f = 15$) $z_t = \Delta^{12} p_{t-1}$, yearly inflation
- ($f = 16$) $z_t = \Delta^{12} i_{t-1}$, yearly interest rate growth
- ($f = 17$) $z_t = y_{t-1} - m_{t-1} - p_{t-1}$, the velocity of money
- ($f = 18$) $z_t = t$, the natural time ordering.

Note that the candidate variables from the 1st to the 17th are often used to measure the state of the economy in the literature (see Weise, 1999; Rothman *et al.*, 2001). The reason why we consider time ordering as a possible transition indicator is that we expect to capture the possible periodical changes in US monetary policy (see Clarida *et al.*, 2000; Dolado, Pedrero and Ruge-Murcia, 2004).

Using standard Johansen techniques, both the Akaike's information criterion (AIC) and Schwarz–Bayesian information criterion (SBIC) choose four lags for a linear VECM. Here, we set the maximum lag length of the STVECM to be 6. We neither want to end up with a parameter-rich model with too many lags nor do we want to compromise with

⁶We thank an anonymous referee for pointing out this necessity.

fewer than enough lags. For simplification, we assume the same autoregressive lag lengths across regimes.

Altogether, we run 162 Gibbs sampling schemes for every unrestricted STVECM. Each Gibbs sampler is run for 12,000 passes with the first 2,000 discarded. The convergence of the sequence draws is checked by the convergence diagnostic measure introduced by Geweke (1992). We use the MATLAB program from LeSage's Econometrics Toolbox (LeSage, 1999) for the diagnostics.

There are a large number of models and many parameters and it is hard to interpret each of them. Instead of providing detailed discussions of each model, we focus on the model comparison results and the general impulse response functions in our analysis.

Model comparison results

In this section, we report the model comparison results associated with a set of 3,138 possible models nested in the original 162 unrestricted models. Bayesian posterior probabilities are calculated from the Bayes factors to reward parsimony. Precise details on how to calculate the Bayes factors are described in appendix A. Assuming all the 3,138 models are mutually independent and exhaustive, we allocate the same prior weight to each of them.

Comparing across the posterior support received by each of the 3,138 models, we find that there is a high degree of model uncertainty. The posterior probabilities of the top 10 models are presented in Table 1. Note that the posterior probability associated with the most preferred model is only 9.79%. Even though the top 10 models together account for 48.35% of the total posterior mass, we also find another 104 models receive non-negligible support from the data. As a result, we argue that it can be misleading to draw empirical conclusions based on any single model.

Overall, we find overwhelming support for models indicating money is nonlinearly Granger causal for output. Table 2 summarizes the posterior model probabilities for different model types. Observe that the models without restricting non-Granger causal or non-long-run causal from money to output receive a total of 91.28% of the posterior mass. Among them, the posterior probabilities received by the nonlinear VECM models, the linear VECM models and the nonlinear VAR models are 87.44%, 3.83% and 0.01%, respectively.⁷ On the contrary, the models indicating money is not Granger causal for output or not long-run causal to output only account for 8.72% of the total posterior mass. Among them, the linear VAR, nonlinear VAR, and nonlinear VECM models implying non-Granger causality from money to output each receives 0.05%, 2.64% and 4.99% posterior model probabilities, and models indicating money is only not long-run causal to output receive 1.03% posterior support. Our finding that money nonlinearly Granger causes output is largely in accord with the in-sample evidence in Rothman *et al.* (2001).

Given the substantial support (96.12% altogether) for nonlinear models, it is interesting to examine which transition variable plays a more important role in triggering regime changes. We present the posterior support for the 18 possible transition indicators we con-

⁷Note that here money nonlinearly Granger causes output mainly through entering the output equation as regressors in the STVECM.

TABLE 1
 Posterior model probabilities for the
 top 10 models

No.	Model	$Pr(m_{No.} data)(\%)$
1	STVECM _{$r=2, f=14, p=1$}	9.79
2	STVECM _{$r=5, f=14, p=1$}	7.67
3	STVECM _{$r=4, f=3, p=1$}	5.86
4	STVECM _{$r=4, f=2, p=1$}	4.91
5	VECM _{$r=1, p=2$}	3.82
6	STVECM _{$r=5, f=17, p=2$}	3.73
7	STVECM _{$r=5, f=13, p=1$}	3.51
8	STVECM _{$r=2, f=10, p=1$}	3.29
9	STVECM _{$r=2, f=3, p=1$}	2.98
10	STVECM _{$r=4, f=15, p=1$}	2.78

Notes: VECM, vector error correction model; STVECM, smooth transition vector error correction model.

TABLE 2
 Posterior probabilities for different types of models

M_i	Description	Number of models	$Pr(M_i data)(\%)$
1	Unrestricted VAR, money Granger causes output	6	0.00
2	Restricted VAR, money does not Granger cause output	6	0.05
3	Unrestricted STVAR, money Granger causes output	108	0.01
4	Restricted STVAR, money does not Granger cause output	102	2.64
5	Unrestricted VECM, money Granger causes output	18	3.83
6	Restricted VECM, money does not Granger cause output	18	0.00
7	Restricted VECM, money is not long-run causal to output	18	0.00
8	Unrestricted STVECM, money nonlinearly Granger causes output	972	87.44
9	Restricted STVECM, money does not Granger cause output	918	4.99
10	Restricted STVECM, money does not long-run cause output	972	1.03

Notes: STVAR, smooth transition VAR model; VECM, vector error correction model; STVECM, smooth transition VECM.

sidered in Table 3. In contrast to the model uncertainty just described, it is seen that 99.99% of the posterior mass for the transition indicators are centred on three candidates, namely the price level, the monthly output growth rate and the monthly interest rate growth rate. Among them, output growth, receiving 45.79% of the posterior mass, can be identified as the most important state variable driving the switching between regimes, followed by the price level, which receives 42.41% of the posterior probabilities. It is noteworthy to find that money (the monthly money growth) only accounts for 0.01% of the posterior probabilities. Thus, although money cannot be ruled out as a candidate transition variable that may cause regime changes, it is obvious that money's role in triggering regime switching is rather limited. Another finding that might be of some importance is that there is little evidence that time ordering can be identified as a transition indicator.

An important finding from Table 2 is that different forms of VECM models receive 97.29% of the posterior mass. Thus, we find it necessary to report the posterior supports for the cointegration ranks of the upper and lower regimes. As outlined in Table 4, there is

TABLE 3
Posterior probabilities for different candidate transition variables

<i>f</i>	Description	<i>Pr(f data)</i> (%)
1	$z_t = y_{t-1}$	0.00
2	$z_t = m_{t-1}$	0.00
3	$z_t = p_{t-1}$	42.41
4	$z_t = i_{t-1}$	0.00
5	$z_t = \Delta^1 y_{t-1}$	45.79
6	$z_t = \Delta^1 m_{t-1}$	0.01
7	$z_t = \Delta^1 p_{t-1}$	0.00
8	$z_t = \Delta^1 i_{t-1}$	11.79
9	$z_t = \Delta^4 y_{t-1}$	0.00
10	$z_t = \Delta^4 m_{t-1}$	0.00
11	$z_t = \Delta^4 p_{t-1}$	0.00
12	$z_t = \Delta^4 i_{t-1}$	0.00
13	$z_t = \Delta^{12} y_{t-1}$	0.00
14	$z_t = \Delta^{12} m_{t-1}$	0.00
15	$z_t = \Delta^{12} p_{t-1}$	0.00
16	$z_t = \Delta^{12} i_{t-1}$	0.00
17	$z_t = y_{t-1} - m_{t-1} - p_{t-1}$	0.00
18	$z_t = t$	0.00

Note: Here we only focus on nonlinear models.

TABLE 4
Posterior probabilities for different ranks

<i>r</i>	Description	<i>Pr(r data)</i> (%)
1	$r_1 = 1, r_2 = 1$	3.90
2	$r_1 = 2, r_2 = 2$	16.84
3	$r_1 = 3, r_2 = 3$	3.04
4	$r_1 = 1, r_2 = 2$	23.68
5	$r_1 = 1, r_2 = 3$	34.39
6	$r_1 = 2, r_2 = 1$	4.01
7	$r_1 = 2, r_2 = 3$	7.69
8	$r_1 = 3, r_2 = 1$	3.37
9	$r_1 = 3, r_2 = 2$	3.07

Note: Here we only consider the cointegration models.

a certain degree of uncertainty about the combinations of r_1 and r_2 . However, we find that models assuming the cointegration rank changes smoothly from the lower to upper regimes receive far more support than models assuming the cointegration rank remains unchanged over time. In particular, we notice that the models assuming $r_1 = 1$ at the lower regime and $r_2 = 2$ and $r_2 = 3$ at the upper regimes together account for 58.07% of the posterior mass. Given that there is also a 16.84% chance that $r_1 = r_2 = 2$, we conjecture that it is highly likely that the cointegration ranks of the upper regimes are greater than 1.

To highlight the nonlinear feature of the interrelationship among money, output, prices and interest rates, in Figure 1 we plot the values of the smooth transition function over

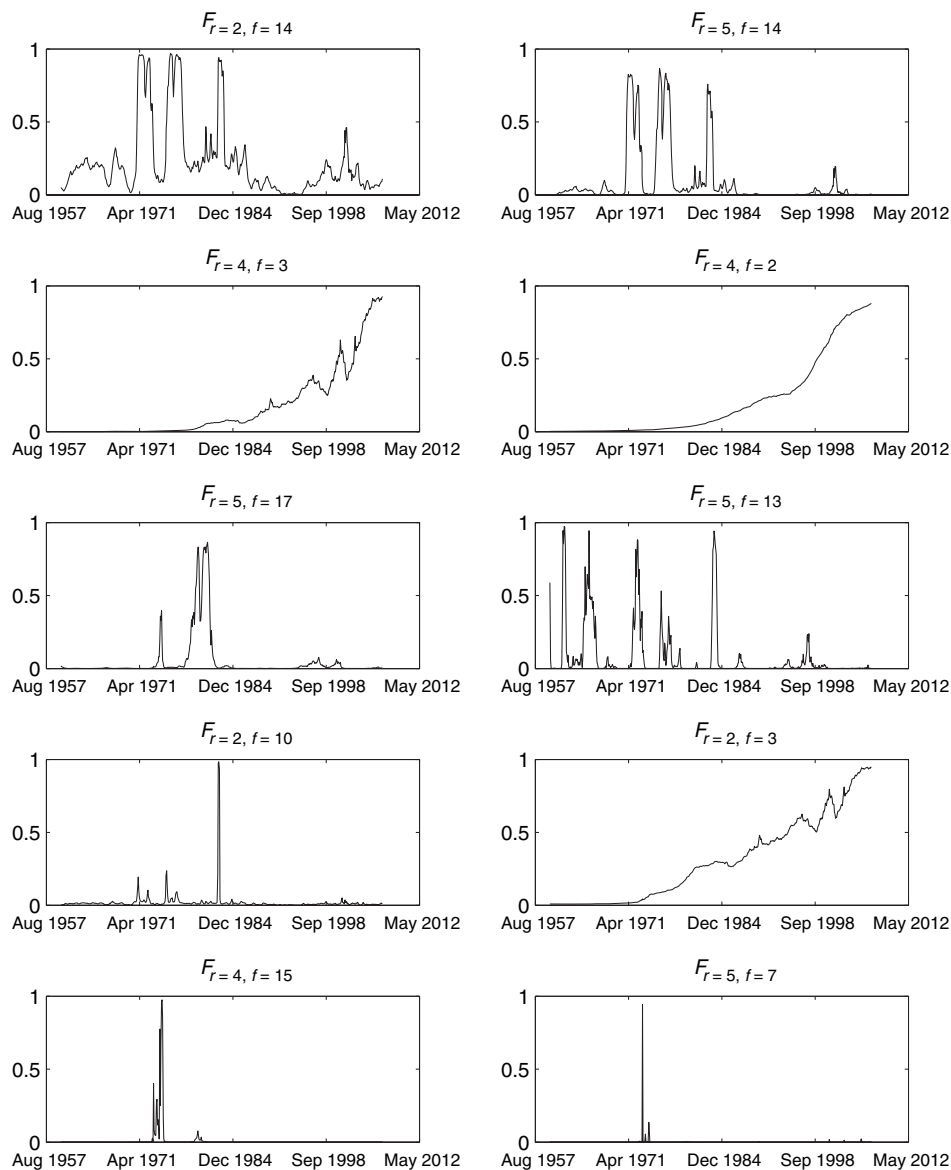


Figure 1. Smooth transition functions

time for the top 10 most preferred nonlinear models. Although in a few cases the regime changes may only look like several blips, a key information conveyed in Figure 1 is that regimes are changing smoothly over time. In fact, we find the same message when we examine the whole set of the smooth transition functions plotted against time. Thus, we have good reason to argue that it might be misleading to use discrete regime switching models to capture the nonlinear effect in the post war US money-output linkage.

Since the STVECM we adopted can be subject to the criticism of being too parameter rich, it is illuminating to examine the posterior support for the different lag lengths

TABLE 5
 Posterior probabilities for different time lags

p	1	2	3	4	5	6
$Pr(p data)(\%)$	55.49	43.43	0.06	0.03	0.00	0.99

presented in Table 5. Observe that $p = 1$ and $p = 2$ receive 55.49% and 43.43% of the posterior support, respectively, and the support for the rest of the lag lengths is very small. It appears that, adopting STVECM, we can use less than four autoregressive lags, that is required in a standard linear VECM, to better capture the dynamics of the system.

Impulse response analysis

To shed further light on the causal effects from money to output, we analyse the impulse responses of output given an unexpected money shock. We are interested in finding two types of asymmetric effects which usually characterize the nonlinear models. First, whether positive and negative shocks to money have unbalanced effects on output. Second, whether big and small money shocks have disproportionate effects.

It is acknowledged that the impulse response functions of the nonlinear models are history and shock dependent (see Potter, 1994; Koop, Pesaran and Potter, 1996). We use the generalized impulse response function proposed in Koop *et al.* (1996) to examine the response of output to an unexpected money shock. Using the Bayesian approach, we calculate the generalized impulse responses by averaging out the history uncertainties, future uncertainties, parameter uncertainties and model uncertainties.

We set the magnitudes of the initial shocks amounting to ± 1 and ± 2 times the standard deviation of monthly money growth rates, and name them ± 1 unit and ± 2 units of shocks, respectively. The time horizon of the impulse responses is set as 120 months (10 years). For comparison, we present the impulse response functions of the nominal output in Figure 2 and that of the real output in Figure 3. Each figure contains four graphs. The impulse response functions caused by the unexpected positive money shocks are plotted in the upper graphs, and those attributable to the negative shocks are plotted in the lower graphs. Also, the impacts given ± 1 units of the money shocks are depicted in the left panels and the impacts given ± 2 units of the money shocks are depicted in the right panels.

Observing these figures, we find that the following findings are noteworthy.

- (i) Positive and negative money shocks of the same magnitude appear to have asymmetric affects on both the nominal and real output. Observe that the time path of the impulse responses to positive shocks never mirror that of the impulse responses to negative shocks.
- (ii) Impacts on both the nominal and real output appear to vary disproportionately with the size of the unexpected shock to money.
- (iii) A $+1$ unit money shock seems to do more harm than good to output: the shock exerts a sustained downwards impact on nominal output growth after boosting it for less than 12 months. Furthermore, exactly during that short period when

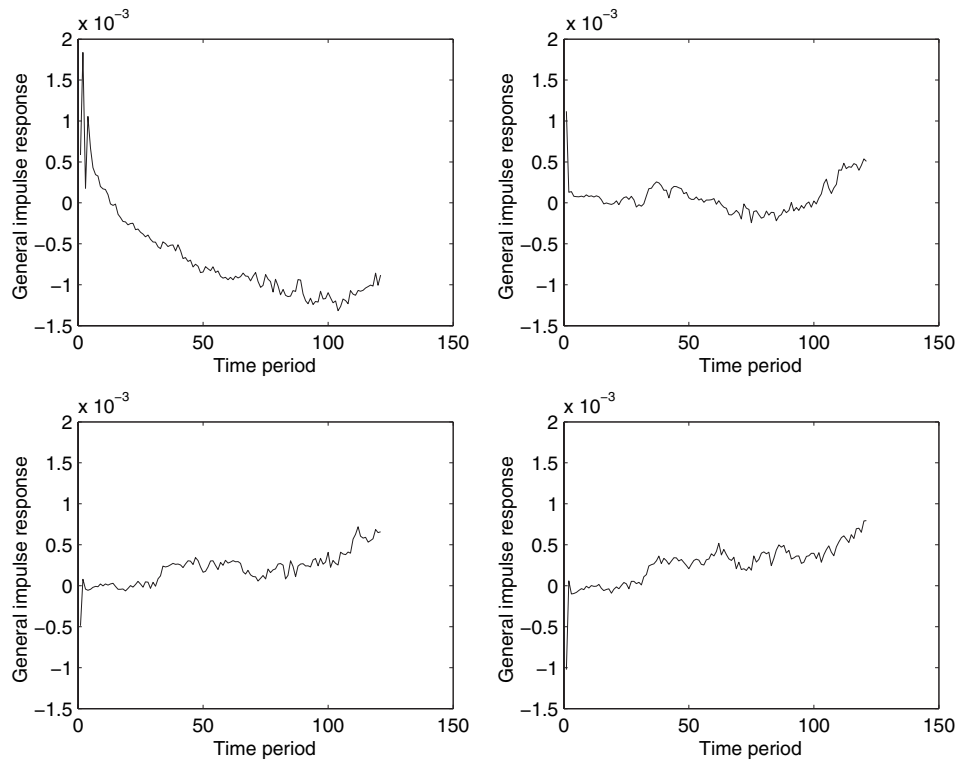


Figure 2. Impacts on nominal output

nominal output is growing faster, the same money shock actually causes a quick sharp drop in real output growth.

- (iv) A positive money shock amounting to two standard deviations of money growth seems to bring about more positive impacts on both nominal and real output growth than the other three types of shocks.
- (v) A negative shock on money seems to bring about a prolonged positive nominal output growth after its initial downward effect on output growth, yet the shock does not necessarily have the same effect on the real output growth.

Overall, the impulse response functions presented in Figures 2 and 3 appear to be in accord with our model comparison results that the nonlinear models account for an overwhelming posterior support. As a result, it seems that the most sensible implication of the impulse response analysis is that, depending on its size and sign, an unexpected money shock can have rather complicated impact on output.

IV. Conclusion

This article investigates the causal effects from money to output using postwar US data. We develop a Bayesian approach to capture the interrelationship among money, output, prices and interest rates using a STVECM model. Different from similar nonlinear modelling methods in the literature, we jointly estimate the cointegration relationships

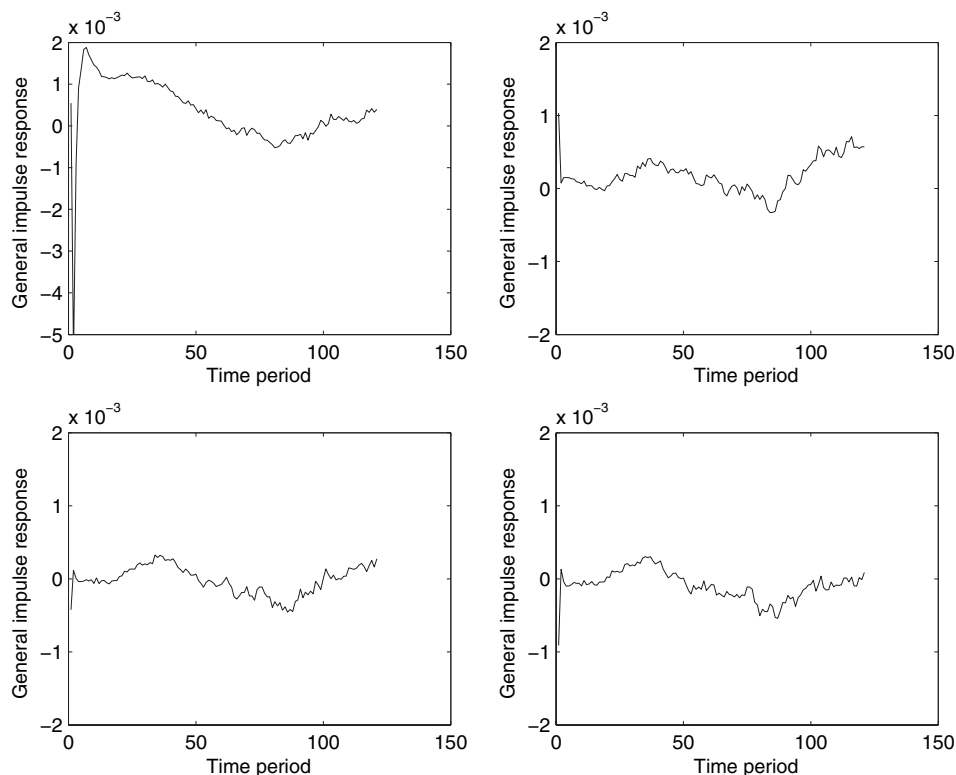


Figure 3. Impacts on real output

and the nonlinear effects in a single step without pre-imposing any theory-based restrictions. By treating models as random variables, we make comparisons across a large number of models distinguished by different important features.

There is a high degree of model uncertainty leaping out of the model comparison exercise. Yet, both the model comparison and impulse response analysis provide substantial evidence that the postwar US money-output relationship is nonlinear. Among a variety of possible transition indicators which might trigger regime changes, it is seen that output growth and price levels jointly receive an overwhelming posterior support, while money's role in driving regime changes is quite small. Overall, STVECM models implying that money is nonlinearly Granger causal to output account for a substantial posterior mass. However, there is also some evidence for models indicating that money is not Granger causal or long-run causal to output.

Final Manuscript Received: December 2010

References

- Ambler, S. (1989). 'Does money matter in Canada? Evidence from a vector error correction model', *The Review of Economics and Statistics*, Vol. 71, pp. 651–658.

- Aslanidis, N. and Kouretas, G. P. (2005). 'Testing for two-regime threshold cointegration in the parallel and official markets for foreign currency in Greece', *Economic Modelling*, Vol. 22, pp. 665–682.
- Bartlett, M. S. (1957). 'A comment on D. V. Lindley's statistical paradox', *Biometrika*, Vol. 44, pp. 533–534.
- Bauwens, L., Lubrano, M. and Richard, J.-F. (1999). *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, New York.
- Bec, F. and Rahbek, A. (2004). 'Vector equilibrium correction models with nonlinear discontinuous adjustments', *Econometrics Journal*, Vol. 7, pp. 628–651.
- Bec, F., Rahbek, A. and Shephard, N. (2008). 'The ACR model: a multivariate dynamic mixture autoregression', *Oxford Bulletin of Economics and Statistics*, Vol. 70, pp. 583–618.
- Chib, S. and Greenberg, E. (1995). 'Understanding the Metropolis–Hastings algorithm', *The American Statistician*, Vol. 49, pp. 327–335.
- Clarida, R., Gali, J. and Gertler, M. (2000). 'Monetary policy rules and macroeconomic stability: evidence and some theory', *Quarterly Journal of Economics*, Vol. 115, pp. 147–180.
- Davies, R. B. (1977). 'Hypothesis testing when a nuisance parameter is present only under the alternative', *Biometrika*, Vol. 74, pp. 33–43.
- Dolado, J., Pedrero, R. M. and Ruge-Murcia, F. J. (2004). 'Nonlinear monetary policy rules: some new evidence for the US', *Studies in Nonlinear Dynamics and Econometrics*, Vol. 8, Article 2.
- Duca, J. V. and VanHoose, D. D. (2004). 'Recent developments in understanding the demand for money', *Journal of Economics and Business*, Vol. 56, pp. 247–272.
- Escribano, A. (2004). 'Nonlinear error correction: the case of money demand in the United Kingdom 1878–2000', *Macroeconomic Dynamics*, Vol. 8, pp. 76–116.
- Fong, P., Li, W. and Wong, C. (2007). 'On a mixture vector autoregressive model', *The Canadian Journal of Statistics*, Vol. 35, pp. 135–150.
- Francis, N. and Owyang, M. T. (2005). 'Monetary policy in a Markov-switching vector error-correction model', *Journal of Business and Economic Statistics*, Vol. 23, pp. 305–313.
- Friedman, B. M. and Kuttner, K. N. (1992). 'Money, income, prices, and interest rates', *American Economic Review*, Vol. 82, pp. 472–492.
- Gefang, D. and Strachan, R. W. (2010). 'Nonlinear impacts of international business cycles on the UK – a Bayesian smooth transition VAR approach', *Studies in Nonlinear Dynamics and Econometrics*, Vol. 14, Article 2.
- Geweke, J. (1992). 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', in Bernardo J., Berger J., Dawid A. and Smith A. (eds), *Bayesian Statistics, Vol. 4*, Clarendon Press, pp. 641–649.
- Granger, C. and Lin, J.-L. (1995). 'Causality in the long run', *Econometric Theory*, Vol. 11, pp. 530–536.
- Granger, C. and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*, Oxford University Press, New York.
- Hall, S. G. and Milne, A. (1994). 'The relevance of P-Star analysis to UK monetary policy', *Economic Journal*, Vol. 104, pp. 597–604.
- Hall, S. G. and Wickens, M. R. (1993). *Causality in Integrated Systems*, D.P. No. 27C93, Centre for Economic Forecasting, London Business School.
- Hansen, B. and Seo, B. (2002). 'Testing for two-regime threshold cointegration in vector error-correction models', *Journal of Econometrics*, Vol. 110, pp. 293–318.
- Haug, A. A. and Tam, J. (2007). 'A closer look at long-run US, money demand: linear or nonlinear error-correction with M0, M1, or M2?' *Economic Inquiry*, Vol. 45, pp. 363–376.
- Hill, J. B. (2007). 'Efficient tests of long-run causation in trivariate VAR processes with a rolling window study of the money-income relationship', *Journal of Applied Econometrics*, Vol. 22, pp. 747–765.
- James, A. T. (1954). 'Normal multivariate analysis and the orthogonal group', *Analysis of Mathematical Statistics*, Vol. 25, pp. 40–75.
- Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
- Kapetanios, G., Shin, Y. and Snell, A. (2006). 'Testing for cointegration in nonlinear smooth transition error correction models', *Econometric Theory*, Vol. 22, pp. 279–303.

- King, R. G. and Watson, M. W. (1997). 'Testing long-run neutrality', *Economic Quarterly*, Federal Reserve Bank of Richmond, Vol. 83, pp. 69–101.
- Koop, G. and Potter, S. M. (1999a). 'Bayes factors and nonlinearity: evidence from economic time series', *Journal of Econometrics*, Vol. 88, pp. 251–281.
- Koop, G. and Potter, S. M. (1999b). 'Dynamic asymmetries in US unemployment', *Journal of Business and Economic Statistics*, Vol. 17, pp. 298–313.
- Koop, G., Pesaran, M. H. and Potter, S. M. (1996). 'Impulse response analysis in nonlinear multivariate models', *Journal of Econometrics*, Vol. 74, pp. 491–499.
- Koop, G., Leon-Gonzalez, R. and Strachan, R. W. (2005). *Efficient Posterior Simulation for Cointegrated Models with Priors on the Cointegration Space*, Discussion Papers in Economics 05/13, Department of Economics, University of Leicester, revised April 2006.
- Koop, G., Leon-Gonzalez, R. and Strachan, R. W. (2006a). *Bayesian Inference in a Cointegration Panel Data Model*, Discussion Papers in Economics 06/2, Department of Economics, University of Leicester.
- Koop, G., Strachan, R. W., van Dijk, H. and Villani, M. (2006b). 'Bayesian approaches to cointegration', in Mills T. and Patterson K. (eds), *The Palgrave Handbook of Econometrics, Volume 1: Theoretical Econometrics*, Palgrave-Macmillan, Basingstoke.
- Kristensen, D. and Rahbek, A. (2010). 'Likelihood-based inference in non-linear error-correction models', *Journal of Econometrics*, Vol. 158, pp. 78–94.
- Leamer, E. E. (1978). *Specification Searches*, John Wiley, New York.
- LeSage, J. (1999). 'Applied econometrics using MATLAB'. Available at <http://www.spatial-econometrics.com/html/mbook.pdf> (accessed September, 2007).
- Lubrano, M. (1999a). 'Bayesian analysis of nonlinear time series models with a threshold', in Barnett W., Hendry D., Hylleberg S., Terasvirta T., Tjostheim D. and Wartz A. (eds), *Nonlinear Econometric Modeling in Time Series*, Cambridge University Press, Cambridge, pp. 79–118.
- Lubrano, M. (1999b). 'Smooth transition GARCH models: a Bayesian perspective', Universite Aix-Marseille III G.R.E.Q.A.M. 99a49.
- Lütkepohl, H., Teräsvirta, T. and Wolters, J. (1999). 'Investigating stability and linearity of a German M1 money demand function', *Journal of Applied Econometrics*, Vol. 14, pp. 511–525.
- Meltzer, A. H. (2001). 'The transmission process', in Deutsche Bundesbank (eds), *The Monetary Transmission Process – Recent Developments and Lessons for Europe*, Palgrave, Basingstoke, pp. 112–130.
- Nelson, E. (2002). 'Direct effects of base money on aggregate demand: theory and evidence', *Journal of Monetary Economics*, Vol. 49, pp. 687–708.
- Nelson, E. (2003). 'The future of monetary aggregates in monetary policy analysis', *Journal of Monetary Economics*, Vol. 50, pp. 1029–1059.
- Ni, S. X. and Sun, D. (2003). 'Noninformative priors and frequentist risks of Bayesian estimators of vector-autoregressive models', *Journal of Econometrics*, Vol. 115, pp. 159–197.
- O'Hagan, A. (1995). 'Fractional Bayes factors for model comparison', *Journal of the Royal Statistical Society*, Vol. 57, pp. 99–138.
- Potter, S. (1994). *Nonlinear Impulse Response Functions*, Department of Economics Working Paper, University of California, Los Angeles, CA.
- Ritter, C. and Tanner, M. A. (1992). 'Facilitating the Gibbs sampler: the Gibbs stopper and the Griddy–Gibbs sampler', *Journal of the American Statistical Association*, Vol. 87, pp. 861–868.
- Rothman, P., van Dijk, D. and Franses, P. H. (2001). 'Multivariate star analysis of money-output relationship', *Macroeconomic Dynamics*, Vol. 5, pp. 506–532.
- Saikkonen, P. (2005). 'Stability results for nonlinear error correction models', *Journal of Econometrics*, Vol. 127, pp. 69–81.
- Saikkonen, P. (2008). 'Stability of regime switching error correction models under linear cointegration', *Econometric Theory*, Vol. 24, pp. 294–318.
- Seo, B. (2004). 'Testing for nonlinear adjustment in smooth transition vector error correction models', No. 749, Economic Society 2004 Far Eastern Meetings.
- Seo, M. (2006). 'Bootstrap testing for the null of no cointegration in a threshold vector error correction model', *Journal of Econometrics*, Vol. 143, pp. 129–150.
- Sims, C. (1972). 'Money, income and causality', *American Economic Review*, Vol. 62, pp. 540–552.

- Sims, C. (1980). 'Macroeconomics and reality', *Econometrica*, Vol. 48, pp. 1–48.
- Sims, C. A. and Zha, T. (2006). 'Were there regime switches in U.S. monetary policy?' *American Economic Review*, Vol. 96, pp. 54–81.
- Sinharay, S. and Stern, H. S. (2002). 'On the sensitivity of Bayes factors to the prior distributions', *The American Statistician*, Vol. 56, pp. 196–201.
- Stock, J. H. and Watson, M. W. (1989). 'Interpreting the evidence on money-income causality', *Journal of Econometrics*, Vol. 40, pp. 161–181.
- Strachan, R. W. (2003). 'Valid Bayesian estimation of the cointegrating error correction model', *Journal of Business and Economic Statistics*, Vol. 21, pp. 185–195.
- Strachan, R. W. and van Dijk, H. K. (2004). *Bayesian Model Selection with an Uninformative Prior*, Keele Economics Research Papers KERP 2004/01, Centre for Economic Research, Keele University.
- Strachan, R. W. and van Dijk, H. K. (2006). *Model Uncertainty and Bayesian Model Averaging in Vector Autoregressive Processes*, Discussion Papers in Economics 06/5, Department of Economics, University of Leicester.
- Strachan, R. W. and Inder, B. (2004). 'Bayesian analysis of the error correction model', *Journal of Econometrics*, Vol. 123, pp. 307–325.
- Sugita, K. (2008). 'Bayesian analysis of a Markov switching temporal cointegration model', *Japan and the World Economy*, Vol. 20, pp. 257–274.
- Swanson, N. R. (1998). 'Money and output viewed through a rolling window', *Journal of Monetary Economics*, Vol. 41, pp. 455–474.
- Taylor, J. B. (1993). 'Discretion versus policy rules in practice', *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 195–214.
- Taylor, J. B. (ed.) (1999). *Monetary Policy Rules*, University of Chicago Press, Chicago.
- Teräsvirta, T. (1994). 'Specification, estimation and evaluation of smooth transition autoregressive models', *Journal of the American Statistical Association*, Vol. 89, pp. 208–218.
- Teräsvirta, T. and Eliasson, A.-C. (2001). 'Non-linear error correction and the UK demand for broad money, 1878–1993', *Journal of Applied Econometrics*, Vol. 16, pp. 277–288.
- Tong, H. (1983). *Threshold Models in Non-linear Time Series Analysis*, Springer-Verlag, New York.
- Verdinelli, I. and Wasserman, L. (1995). 'Computing Bayes factors using a generalization of the Savage–Dickey density ratio', *Journal of the American Statistical Association*, Vol. 90, pp. 614–618.
- Villani, M. (2005). 'Bayesian reference analysis of cointegration', *Econometric Theory*, Vol. 21, pp. 326–357.
- Wang, P. and Wen, Y. (2005). 'Endogenous money or sticky prices? Comment on monetary non-neutrality and inflation dynamics', *Journal of Economic Dynamics and Control*, Vol. 29, pp. 1361–1383.
- Weise, C. L. (1999). 'The asymmetric effects of monetary policy: a nonlinear vector autoregression approach', *Journal of Money Credit and Banking*, Vol. 31, pp. 85–108.
- Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.

Appendix A: Bayesian econometric method

Likelihood function

Koop *et al.* (2005) develop an efficient collapsed Gibbs sampler for the VECM estimation in linear contexts, which provides great computational advantages over conventional methods. To incorporate the collapsed Gibbs sampler into our posterior simulation algorithm, following Koop *et al.* (2005), we obtain two representations of the likelihood.

To start with, restricting β and β^z to be semi-orthogonal such that $\beta'\beta = I_{r_1}$ and $\beta^z'\beta^z = I_{r_2}$, we write equation (1) as:

$$\Delta x_t = \omega_{1,t-1}\beta\alpha + \omega_{2,t}\Phi + F(z_t)(\omega_{1,t-1}\beta^z\alpha^z + \omega_{2,t}\Phi^z) + \varepsilon_t, \quad (5)$$

where $\omega_{1,t-1} = x_{t-1}$, $\omega_{2,t} = (\Delta x_{t-1}, \dots, \Delta x_{t-p}, d_t)$, $\Phi = (\Gamma'_1, \dots, \Gamma'_p, \zeta')'$, $\Phi^z = (\Gamma^z'_1, \dots, \Gamma^z'_p, \zeta^z)'$. To simplify the notation, we then define $X_0 = (\Delta x'_1, \Delta x'_2, \dots, \Delta x'_T)'$ and $X = (X_1\beta, X_2,$

$F^z X_1 \beta^z, F^z X_2$), where $X_1 = (\omega'_{1,1}, \omega'_{1,2}, \dots, \omega'_{1,T})'$, $X_2 = (\omega'_{2,1}, \omega'_{2,2}, \dots, \omega'_{2,T})'$ and $F^z = \text{diag}(F(z_1), F(z_2), \dots, F(z_T))$. Next, we set $B = (\alpha', \Phi', \alpha^z', \Phi^z)'$ and $E = (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_T)'$. Finally, we rewrite model (1) as:

$$X_0 = X_1 \beta \alpha + X_2 \Phi + F^z X_1 \beta^z \alpha^z + F^z X_2 \Phi^z + E = XB + E. \quad (6)$$

It is seen that the likelihood function of equation (6) is:

$$L(x | \Sigma, B, \beta, \beta^z, \gamma, c) \propto |\Sigma|^{-(T/2)} \exp \left\{ -\frac{1}{2} \text{tr} \Sigma^{-1} E' E \right\}. \quad (7)$$

Vectorizing equation (6), model (1) can be transformed into

$$x_0 = xb + e, \quad (8)$$

where $x_0 = \text{vec}(X_0)$, $x = I_n \otimes X$, $b = \text{vec}(B)$, and $e = \text{vec}(E)$. Note that $E(ee') = V_e = \Sigma \otimes I_T$. Given that

$$\begin{aligned} \text{tr} \Sigma^{-1} E' E &= e' (\Sigma^{-1} \otimes I_T) e \\ &= s^2 + (b - \hat{b})' V^{-1} (b - \hat{b}), \end{aligned} \quad (9)$$

where $s^2 = x'_0 M_v x_0$, $M_v = \Sigma^{-1} \otimes [I_T - X(X'X)^{-1}X']$, $\hat{b} = [I_n \otimes (X'X)^{-1}X']x_0$ and $V = \Sigma \otimes (X'X)^{-1}$. The likelihood (7) can be written as:

$$L(x | \Sigma, B, \beta, \beta^z, \gamma, c) \propto |\Sigma|^{-(T/2)} \exp \left\{ -\frac{1}{2} [s^2 + (b - \hat{b})' V^{-1} (b - \hat{b})] \right\}. \quad (10)$$

Observe that the likelihood of b is normal conditional on all other parameters.

With a normal form for the likelihood of b , we next obtain a normal form for the likelihood of the cointegration vectors.

For any positive definite matrix κ and κ^z of rank r_1 and r_2 , we have $\beta \alpha = \beta \kappa \kappa^{-1} \alpha = \beta^* \alpha^*$ and $\beta^z \alpha^z = \beta^z \kappa^z \kappa^{z(-1)} \alpha^z = \beta^{z*} \alpha^{z*}$, where $\beta^* = \beta \kappa$ and $\alpha^* = \kappa^{-1} \alpha$, $\beta^{z*} = \beta^z \kappa^z$ and $\alpha^{z*} = \kappa^{z(-1)} \alpha^z$. Moreover, restricting $\kappa = (\alpha \alpha')^{(1/2)} = (\beta^{*'} \beta^*)^{(1/2)}$, and $\kappa^z = (\alpha^z \alpha^{z'})^{(1/2)} = (\beta^{z*'} \beta^{z*})^{(1/2)}$, we find $\alpha^* \alpha^{*'}$ and $\alpha^{z*} \alpha^{z*'}$ are semi-orthogonal if β and β^z are semi-orthogonal. Therefore, we can reexpress equation (6) as:

$$\begin{aligned} X_0 - X_2 \Phi - F^z X_2 \Phi^z &= X_1 \beta \alpha + F^z X_1 \beta^z \alpha^z + E \\ &= X_1 \beta^* \alpha^* + F^z X_1 \beta^{z*} \alpha^{z*} + E. \end{aligned} \quad (11)$$

Setting $\tilde{x}_0 = \text{vec}(X_0 - X_2 \Phi - F^z X_2 \Phi^z)$, $\tilde{x} = [\alpha^{*'} \otimes X_1 \alpha^{z*'} \otimes F^z X_1]$, $\tilde{b} = [\text{vec}(\beta^*)' \text{vec}(\beta^{z*})']'$, we find that equation (11) can be written as:

$$\tilde{x}_0 = \tilde{x} \tilde{b} + e. \quad (12)$$

Thus, we find that the second likelihood representation from equation (12) is:

$$L(x | \Sigma, B, \beta, \beta^z, \gamma, c) \propto |\Sigma|^{-(T/2)} \exp \left\{ -\frac{1}{2} [s_{\beta^*}^2 + (b_{\beta^*} - \hat{b}_{\beta^*})' V_{\beta^*}^{-1} (b_{\beta^*} - \hat{b}_{\beta^*})] \right\}, \quad (13)$$

where $s_{\beta^*}^2 = (\tilde{x}_0 - \tilde{x} \hat{b}_{\beta^*})' (\Sigma^{-1} \otimes I_T) (\tilde{x}_0 - \tilde{x} \hat{b}_{\beta^*})$, $\hat{b}_{\beta^*} = (\tilde{x}' \tilde{x})^{-1} \tilde{x}' \tilde{x}_0$, $V_{\beta^*}^{-1} = \tilde{x}' (\Sigma^{-1} \otimes I_T) \tilde{x}$.

Priors

Although the most commonly elicited quantity money demand equation indicates that the velocity of money is stationary (see Rothman *et al.*, 2001; Teräsvirta and Eliasson, 2001), empirical work does not rule out the possibility that the number of the long-run cointegration relationships and the cointegration vectors are in fact data based (see Ambler, 1989; Friedman and Kuttner, 1992; Swanson, 1998). Furthermore, it is difficult to impose meaningful informative priors for the coefficients of the long-run/short-run adjustment in the VECM nor for parameters that indicate the speed of regime changes in the transition function. Hence, we use uninformative or weakly informative priors to allow the data information to dominate any prior information. To start with, we assume that all possible models are to be independent and, *a priori*, equally likely.

Before setting our priors for the parameters, it is worthwhile to stress the identification problems in our model setting. Note that both the linear VECM and smooth transition VAR model (STVAR) suffer from identification problems.

As documented in the literature, a linear VECM suffers from both the global and local non-identifications of the cointegration vectors and parameters corresponding to the long-run adjustments. In Bayesian literature, a great effort has been made to surmount this problem. In earlier research, to set uninformative prior for the cointegration vector β , researchers first normalize β into $\beta = [I_r, V']'$, then impose uninformative prior on the sub-vector V . However, as argued by Strachan and van Dijk (2004), this approach has an undesirable side effect that it favours the regions of cointegration space where the imposed linear normalization is actually invalid. In most recent work, researchers have worked on putting uninformative priors on the cointegration space (see Strachan and Inder, 2004; Villani, 2005). As noted in Koop *et al.* (2006b), since only the space of the cointegration vector can be derived from the data, it is better to elicit priors in terms of the cointegration space than in terms of cointegration vectors.

With regard the smooth transition part of the model, as explained in Lubrano (1999a), since Bayesians have to integrate over the whole domain of the smooth parameter, the identification problem that arises from $\gamma = 0$ [the so-called Davies' problem (Davies, 1977), see Koop and Potter (1999a) for further explanation] becomes more serious in the Bayesian context than in a classical framework. Bauwens *et al.* (1999) and Lubrano (1999a,b) introduce a number of prior settings to solve the problem. Following Gefang and Strachan (2008), we tackle this problem by simply setting the prior distribution of γ as Gamma.

The non-identification problem faced by the STVECM is slightly different. Although the Davies' problem remains relatively the same as in the STVAR, the problem in identifying the cointegration vector and its adjustment parameters is subject to the additional influence from the transition parameters. Here the cointegration vectors come forth in two combinations, namely $\beta\alpha$ and $\beta^z\alpha^z$. However, this difference does not render the identification problem more complicated than what we have to deal with in a linear VECM or a STVAR. As long as we can rule out the possibility that $\gamma = 0$, we can identify β , β^z , α and α^z sequentially once we choose a way to normalize β and β^z .

In the rest of the section, we construct prior distributions for all the parameters. With regards to the variance-covariance matrix of the error terms, following Zellner (1971), we set standard diffuse prior for Σ .

$$p(\Sigma) \propto |\Sigma|^{-(n+1)/2}.$$

For the purpose of our research, we need to calculate posterior model probabilities to compare across different possible models. As the dimension of b changes across different model specifications, to have the Bayes factors well defined, we are not allowed to set flat prior for b (see Bartlett, 1957, and O'Hagan, 1995, for details). Therefore, following Strachan and van Dijk (2006), we set weakly informative conditional proper prior for b as:

$$P(b | \Sigma, \beta, \gamma, c, M_\omega) \propto N(0, \eta^{-1}I_k),$$

where $b = \text{vec}(B)$, $k = 2(2 + np) + r_1 + r_2$, η is the shrinkage parameter as proposed by Ni and Sun (2003). As practised in Koop *et al.* (2006a), we draw η from the Gibbs sampler. In our case, we set the prior distribution of η as Gamma with mean $\underline{\mu}_\eta$, and degrees of freedom \underline{v}_η , where $\underline{\mu}_\eta = 3$, $\underline{v}_\eta = 4$. Note that the prior for η renders the prior for b relatively uninformative.

Following the arguments of Koop *et al.* (2006b), we elicit the uninformative prior of β and β^z indirectly from the prior expressed upon the cointegration space. In particular, following Strachan and Inder (2004), we set the cointegration spaces of β and β^z are uniformly distributed on the Grassmann manifolds G_{n,r_1} and G_{n,r_2} , respectively. Using the fact that the uniform distribution on the Stiefel manifold $V_{n,r}$ induces a uniform distribution on the Grassmann manifold $G_{n,r}$, we impose the identification restriction $\beta'\beta = I_{r_1}$ and $\beta^z'\beta^z = I_{r_2}$ which do not restrict the possible cointegration space.⁸ Moreover, in line with Koop *et al.* (2005), we set the prior for b_{β^*} as $p(b_{\beta^*} | \eta) \sim N(0, \eta^{-1}I_{n(r_1+r_2)})$ to obtain a normal form for the posterior.

To avoid the Davies' problem in the nuisance parameter space, following Lubrano (1999a,b) and Gefang and Strachan (2008), we set the prior distribution for γ as Gamma, which excludes *a priori* the point $\gamma = 0$ from the integration range. Since the nonlinear part of b can still be a vector of zeros as $\gamma > 0$, the prior specification of γ does not render model (1) in favour of the nonlinear effect. In empirical work, we use Gamma(1, 0.001) to allow the data information to dominate the prior of γ .

As to the prior of c , to make more sense in the context of economic interpretation, we elicit the conditional prior of c as uniformly distributed between the middle 80% ranges of the transition variables.

Posterior computation

Using the priors just identified and the likelihood functions in equations (10) and (13), we obtain the full conditional posteriors as follows.

Conditional on $\beta, \beta^z, \gamma, c$ and b , the posterior of Σ is inverted Wishart (IW) with scale matrix $E'E$ and degree of freedom T ; conditional on $\Sigma, \beta, \beta^z, \gamma$ and c , the posterior of b is normal with mean $\bar{b} = \bar{V}_b V^{-1} \hat{b}$ and covariance matrix $\bar{V}_b = \Sigma \otimes (X'X + \eta I_k)^{-1}$. Conditional on Σ, b, γ and c , the posterior of b_{β^*} is normal with mean $\bar{b}_{\beta^*} = \bar{V}_{\beta^*} V_{\beta^*}^{-1} \hat{b}_{\beta^*}$ and covariance matrix $\bar{V}_{\beta^*} = [V_{\beta^*}^{-1} + \eta I_{n(r_1+r_2)}]^{-1}$.

⁸Note that the priors over the cointegration spaces of β and β^z are proper. See James (1954), and Strachan and Inder (2004) for further explanations.

To obtain the conditional posterior for η , we combine the prior and likelihood to obtain the expression

$$p(\eta | b, \Sigma, \gamma, c, y, x) \propto \eta^{\overline{v}_\eta + nk - 2)^{1/2}} \exp\left(-\frac{\eta \overline{v}_\eta}{2\overline{\mu}_\eta} - \frac{1}{2} b' b \eta\right). \quad (14)$$

Thus with a Gamma prior, the conditional posterior distribution of η is Gamma with degrees of freedom $\overline{v}_\eta = nk + v_\eta$ and mean $\overline{\mu}_\eta = \overline{v}_\eta \overline{\mu}_\eta / (\overline{v}_\eta + \mu_\eta b' b)$.

The posterior distributions for the remaining parameters, γ and c , have non-standard forms. However, we can use the independent chain Metropolis–Hastings (M–H) algorithm (Chib and Greenberg, 1995) within Gibbs to estimate γ , and the Griddy–Gibbs sampler (Ritter and Tanner, 1992) to estimate c . In practice, we draw γ from a candidate-generating Gamma density. The acceptance rates are roughly 20–40%.

Following Koop *et al.* (2005), we construct the collapsed Gibbs sampler as following:

- (i) Initialize $(b, \Sigma, b_\beta, \gamma, c, \eta)$.
- (ii) Draw $\Sigma | b, b_\beta, \gamma, c, \eta$ from $IW(E' E, T)$.
- (iii) Draw $b | \Sigma, b_\beta, \gamma, c, \eta$ from $N(\overline{b}, \overline{V}_b)$.
- (iv) Calculate $\alpha^* = (\alpha \alpha')^{-(1/2)} \alpha$, $\alpha^{z*} = (\alpha^z \alpha^{z'})^{-(1/2)} \alpha^z$.
- (v) Create \tilde{x}_0 .
- (vi) Draw $b_{\beta^*} | \Sigma, b, \gamma, c, \eta, \tilde{x}_0$ from $N(\overline{b}_{\beta^*}, \overline{V}_{\beta^*})$.
- (vii) Construct $\kappa = (\beta^* \beta^{*'})^{(1/2)}$, calculate $\beta = \beta^* \kappa^{-1}$. Construct $\alpha = \kappa \alpha^*$. Use the same procedure to derive β^z and α^{z*} .
- (viii) Draw $\gamma | \Sigma, b, b_\beta, c, \eta$ using M–H algorithm.
- (ix) Draw $c | \Sigma, b, b_\beta, \gamma, \eta$ using Griddy–Gibbs sampler.
- (x) Draw $\eta | \Sigma, b, b_\beta, \gamma, c$ from $G(\overline{v}_\eta, \overline{\mu}_\eta)$.
- (xi) Repeat steps (ii)–(x) for a suitable number of replications.

We consider a wide range of models to investigate the causal effects from money to output. Alternative models are distinguished by the number of the long-run cointegration relationship, the lag length of the autoregressive process, the existence of the nonlinear effects, the transition variable triggering regime changes and whether the restrictions of non-Granger causality or no long-run causality from money to output is imposed.

Bayesian methods provide us with a formal approach to evaluating the support for alternative models by comparing posterior model probabilities. These posterior probabilities can be used to select the best model for further inference, or to use the information in all or an important subset of the models to obtain an average of the economic object of inference by BMA. The posterior odds ratio – the ratio of the posterior model probabilities – is proportional to the Bayes factor. Once we know the Bayes factors and prior probabilities, we can compute the posterior model probabilities.

We use SDDR to compute Bayes factor for comparing one model with a second model (Verdinelli and Wasserman, 1995). Note that all our restrictions can be finally put on b . Since the conditional posterior of b is Normal, we can incorporate the estimation of the numerators of the SDDRs in the Gibbs sampler. The denominators of the SDDRs are the marginal priors of b , which can be analytically derived outside of the Gibbs loop. The method is detailed in Koop *et al.* (2006a) and the reader is referred there for precise derivations. Note that Bayes factors enable us to derive the posterior probabilities for restricted

models nested in different unrestricted models. A simple restriction in our application to choose is the point where all coefficients in b are zero except for the ones that corresponding to the intercepts, at which point we have the model with no explanatory variables except for the intercept variables. This restricted model is useful as it nests within all models. Once we have the Bayes factor for each model to the zero lag model, via simple algebra we can back out the posterior probabilities for all models.

Appendix B: Prior sensitivity analysis

It is well acknowledged that Bayes factors can be highly sensitive to the priors assumed (see Koop and Potter, 1999a; Sinharay and Stern, 2002). In the prior sensitivity exercise, we leave all the uninformative priors (e.g. the error covariance, the cointegration space and the model space) unchanged, then we compute posterior model probabilities using different priors on η and γ . Overall, we find the model comparison results are relatively insensitive to the prior choices of η and γ . For brevity, in Table B1, we only present the posterior supports for different model types using $\underline{\mu}_\eta = 10$, $\underline{v}_\eta = 0.01$, $\underline{\mu}_\gamma = 10$ and $\underline{v}_\gamma = 0.01$.

TABLE B1
Posterior probabilities for different types of models

M_i	Description	Number of models	$Pr(M_i data)$ (%)
1	Unrestricted VAR, money Granger causes output	6	0.00
2	Restricted VAR, money does not Granger cause output	6	0.06
3	Unrestricted STVAR, money Granger causes output	108	0.01
4	Restricted STVAR, money does not Granger cause output	102	2.58
5	Unrestricted VECM, money Granger causes output	18	3.81
6	Restricted VECM, money does not Granger cause output	18	0.00
7	Restricted VECM, money is not long-run causal to output	18	0.00
8	Unrestricted STVECM, money nonlinearly Granger causes output	972	87.58
9	Restricted STVECM, money does not Granger cause output	918	4.68
10	Restricted STVECM, money does not long-run cause output	972	1.29

Notes: STVAR, smooth transition VAR model; VECM, vector error correction model; STVECM, smooth transition VECM.