# Three Essays In Macroeconomics and Monetary Economics Using Bayesian Multivariate Smooth Transition Approaches

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Fang Ge

Department of Economics

University of Leicester

August 2008

To: Hong

# Abstract

The first essay introduces a Bayesian logistic smooth transition vector autoregression (LSTVAR) approach to investigating the impact of international business cycles on the UK economy. We find that the British business cycle is asymmetrically influenced by growth in the US, France and Germany. Overall, positive and negative shocks generating in the US or France affect the UK in the same directions as the shock. However, a shock emanating from Germany always exerts negative cumulative effects on the UK. Further, a positive shock arising from Germany adversely affects the UK output growth more than a negative shock of the same size.

The second essay proposes a Bayesian method to investigating the purchasing power parity (PPP) utilizing an exponential smooth transition vector error correction model (ESTVECM). Employing a simple Gibbs sampler, we jointly estimate the cointegrating relationship along with the nonlinearities caused by the departures from the long run equilibrium. By allowing for symmetric regime changes, we provide strong evidence that PPP holds between the US and each of the remaining G7 countries. The model we employed implies that the dynamics of the PPP deviations can be rather complex, which is attested to by the impulse response analysis.

The final essay proposes a Bayesian approach to exploring money-output causality within a logistic smooth transition vector error correction framework (LSTVECM). Our empirical results provide substantial evidence that the postwar US money-output relationship is nonlinear, with regime changes mainly governed by the lagged inflation rates. More importantly, we obtain strong support for long-run non-causality and nonlinear Granger-causality from money to output. Furthermore, our impulse response analysis reveals that a shock to money appears to have a negative aggregate impact on real output over the next fifty years, which calls for more caution when using money as a policy instrument.

# Acknowledgements

I would like to express my deep and sincere gratitude to all those who gave me the possibility to complete this thesis.

First of all I would like to thank my supervisors on Bayesian econometrics, Rodney Strachan and Roberto Leon-Gonzalez. They have offered invaluable guidance and continuous support throughout my PhD study. In particular, I gratefully acknowledge Roberto for spending time to advise this thesis despite his own busy teaching and research commitments at National Graduate Institute for Policy Studies in Tokyo.

I am extremely grateful to Gary Koop. I was first introduced to the beauty of Bayesian analysis in his advanced econometrics class during my MSc study in 2005. From then on, Gary has become my enormous source of knowledge and somewhat an unofficial mentor. His detailed comments on the second and third chapters have significantly improved this thesis. On top of that, Gary has set a great example as teacher and researcher for me to follow.

I would also like to express my deepest gratitude to my other supervisors: Ali al-Nowaihi, Wojciech Charemza, Roderick McCrorie and Emi Mise, who gave valuable advice and comments in different stages. Moreover, I would like to thank Stephen Hall, whose suggestions have tremendously influenced the directions of the third chapter.

My warm thanks also go to staff members at the Department of Economics, computer center and library. Without their consistent assistance, this thesis would not have been possible.

As always, I am also deeply indebted to my family and friends for their unvarying love and support.

Finally, financial supports from the Department of Economics and University of Leicester are gratefully acknowledged.

Leicester, August 2008
Deborah Gefang

# Declaration

An earlier draft of Chapter 1 entitled "Asymmetric Impacts of International Business Cycles on the UK - A Bayesian LSTVAR Approach" has been presented at the 15th Annual Symposium of the Society for Nonlinear Dynamics and Economics (SNDE), Paris, France, 2007.

Chapter 2 entitled "Investigating nonlinear purchasing power parity during the post-Bretton Woods era - a Bayesian exponential smooth transition VECM approach" has been accepted for publication in Advances in Econometrics, Volume 23: Bayesian Econometrics, 2008, Elsevier.

Chapter 3 entitled "Money-output Causality Revisited: A Bayesian Logistic Smooth Transition VECM Perspective" has been accepted for presentation at the 15th Annual Conference of the Multinational Finance Society, Florida, USA, 2008 and the 25th Symposium on Banking and Monetary Economics, Luxembourg, 2008.

# Contents

# Summary of Thesis

Although linear models play a fundamental role in econometrics and economic theory, it is widely accepted that the relationships between economic variables are generally nonlinear [eg. Goldfeld and Quandt (1972), Granger and Teräsvirta (1993)]. In the literature, a wealth of alternative nonlinear models have been proposed to capture the nonlinearities in macroeconomics and monetary economics. Among the time series models allowing for two or more regimes in the data generating process, the three types of most commonly used models are the Markov regime switching model [Hamilton (1989)], the threshold autoregressive model [Tong (1978)], and the smooth threshold autoregressive model [Teräsvirta (1994)].

In recent years, smooth transition models have become popular for they suggest smooth, rather than discrete, adjustment mechanisms in regime changes [see for example, Teräsvirta (1994), Lubrano (1999b), van Dijk, Teräsvirta and Franses (2002)]. However, as noted by Osborn, Perez and Sensier (2005), a vast majority of empirical studies utilizing smooth transition models are conducted in the univariate contexts, and the attempts in

the multivariate systems are relatively rear. Noteworthily, we find that there is a lack of Bayesian literature in multivariate smooth transition models. The main theme of this thesis, consisting of three essays, is to introduce Bayesian multivariate smooth transition approaches to investigating three important topics in macroeconomics and monetary economics.

Compared with the available classical estimation techniques which often require multiple steps and Taylor expansions, our Bayesian method can jointly estimate the autoregressive coefficients and the nuisance parameters in the transition function simultaneously. Therefore, our approach is less susceptible to the issues associated with sequential testing and our finite sample inference is exact. The class of nonlinear models we consider may be subject to the criticism of being too parameter rich. However, as we use Bayes Factors for model selection and model averaging we effectively overcome this issue as this approach rewards more parsimonious models.[1] Finally, as far as we are aware, our Bayesian methods developed in the second chapter are among the first in the literature to identify the existence of the long-run cointegrating relationship among endogenous variables where the nonlinear effect is triggered by the deviations from the long-run equilibrium.

Business cycle linkages between countries have been remaining the focus of public interest. In recent years, nonlinear multivariate models have been widely adopted as researchers become more concerned with the regime shifts in economic interdependencies across countries [Smith and Summers (2005),

[1]See Koop and Potter (1999a, 1999b) for further explanations.

Artis, Galvao and Marcellino (2007), Chen and Shen (2007)]. The first essay introduces a Bayesian LSTVAR approach to investigating the impact of international business cycles on the UK economy. Overall, we find that the UK economy is sensitive to the fluctuations of international business cycles in a asymmetric form. Our results suggest that linear models misspecify the form of the relationship and would result in systematic errors in analysis and policy making due to the presence of substantial nonlinear effects. Furthermore, the negative effects on the UK growth rate exerted by Germany is of intrinsic importance to policy makers. Although we make no attempt here to provide an economic explanation for our empirical findings, our results warrant a closer investigation in business cycle linkages, especially how macroeconomic shocks propagate through the transmission channels such as trade, monetary policy and financial markets.

Given its importance in open economy macro modeling and policy advice, the validity of the purchasing power parity (PPP) over the post-Bretton Woods era has been the subject of intensive study in the literature. Inspired by the transaction cost theory [Dumas (1992), Sercu, Uppal and van Hulle (1995)], the second essay proposes a Bayesian procedure to investigate PPP utilizing an ESTVECM model. Different from the available approaches typically used in the literature, our methods jointly estimate the cointegrating relationship along with the nonlinearities caused by the departures from the long run equilibrium. By allowing for nonlinear regime changes in the interrelationship among the nominal exchange rates and domestic and foreign prices, we provide strong evidence that PPP holds between the US and each of the remaining G7 countries. The model we employed implies that the dynamics of the PPP deviations can be rather complex, which is attested to by the impulse response analysis.

Many researchers use vector error correction model (VECM) to explore the money-output relationship due to the framework′s advantages in capturing both the long-run and short-run dynamics [see, e.g., Johansen (1992), Garratt, Koop, Mise and Vahey (2007)]. However, given the increasing awareness of the importance of possible regime shifts in money-output relationship [see, e.g., Lutekepohl, Terasvirta and Wolters (1999), Escribano (2004)], we have good reasons to use nonlinear VECM to investigate whether money is causal for output. Yet, among the vast literature focusing on the causal effect from money to output, according to our knowledge, only Rothman, van Dijk and Frances (2001) apply a nonlinear multivariate VECM framework to study the money-output relationship. The final essay proposes a Bayesian approach to exploring money-output causality within a LSTVECM context. Our empirical results provide substantial evidence that the postwar US money-output relationship is nonlinear, with regime changes mainly governed by the lagged inflation rates. More importantly, we obtain strong support for long-run non-causality and nonlinear Granger-causality from money to output. Furthermore, our impulse response analysis reveals that an increase in money supply appears to have a negative aggregate impact on real output in the long term. This result is of special importance

when central banks consider injecting money into the market to ease the current credit crunch .

# Chapter 1

# Impacts of International Business Cycles on the UK — a Bayesian Smooth Transition VAR Approach

**Acknowledgements**: I am most grateful to Rodney Strachan for the invaluable supervision. Also, I would like to thank the participants in the SNDE 2007 for their helpful comments on an earlier draft of this chapter. Any remaining errors are the author's responsibility.

## 1.1 Introduction

The study of international business cycle linkages is of special importance to macroeconomic policy research. Numerous studies have sought to identify a common business cycle across countries [see for instance, Artis and Zhang (1997), Wynne and Koo (2000), Inklaar and Haan (2001)]. In recent years, nonlinear multivariate models have become more popular among researchers for such models can effectively capture the cross-country asymmetric inter-dependencies [Smith and Summers (2005), Artis, Galvâo and Marcellino (2007), Chen and Shen (2007), to mention a few].

The present chapter examines the impacts of international business cycles on the UK economy within the framework of a logistic smooth transition vector autoregression (LSTVAR) model. In particular, we attempt to characterize the behaviour of the UK output growth under the influence of the booms and busts in the US, France, and Germany, respectively.

Business cycle linkages between the UK and the three afore mentioned countries have been examined previously by, for example, Artis and Zhang (1997), Inklaar and Haan (2001), and Perez, Osborn and Artis (2006). However, most of the literature focuses on exploring the business cycles synchronization rather than investigating the propagation of different types of shocks (such as positive and negative or large and small shock) across countries. Although the US' effects on the UK economy are investigated in several studies [for example, Artis, Krolzig and Toro (2004), Osborn, Perez and Sensier (2005), Artis, Galvâo and Marcellino (2007)], to our best knowledge, no evidence on how France and Germany, the two largest continental European economies, influence the UK business cycles has been documented, except for Artis, Galvâo and Marcellino (2007), which look into Germany's impact on the UK business cycles at one point.

Our approach for the LSTVAR estimation is Bayesian. In particular, we extend the Bayesian technique in estimating the univariate smooth transition models introduced in Lubrano (1999a, 1999b) into a multivariate form. Compared with the available classical estimation techniques which often require multiple steps and Taylor expansions, our Bayesian method can jointly estimate the autoregressive coefficients and the nuisance parameters in the transition function in one stage. Therefore, our approach is less susceptible to the sequential testing and inaccurate approximations problems. Furthermore, considering that nonlinear models are generally subject to the criticism of being too parameter rich, we resort to Bayes Factors for model selection and model averaging to reward more parsimonious models.[1]

Our results provide strong evidence of asymmetry in the bivariate relationship across the three country pairs. For all cases, LSTVAR models receive overwhelming support over the linear models. Additionally, we find that business cycles in the US, the UK and Germany play important roles in leading regimes changes, while the changes in France output would not

[1] As discussed by Koop and Potter (1999a, 1999b), Bayes Factors include an automatic penalty for more complex models.

cause salient nonlinear effect.

Impulse response analysis implies that features of the impact from the three countries are quite different. Among the three countries, the US' impact is the most persistent. Observe that the effects from France or Germany die out in relatively five years, while with a much clearer cyclical pattern, the impacts of the US growth shocks are still evident after nine years. It is not surprising to observe that the shocks from the US and France would affect the UK in the same direction. However, different from Artis, Galvâo and Marcellino (2007), we find that both the expansion and recession of Germany would thwart the UK output growth. Most strikingly, we find that a boom in Germany brings more negative effects to the UK's economy than a bust in Germany.

Overall, we find that the UK's economy is sensitive to the fluctuations of international business cycles in a asymmetric form. Our research nonetheless suggests that relying on linear models would result in systematic mistakes in analysis and policy making due to the presence of substantial nonlinear effect. Furthermore, it goes without saying that pernicious effects on the UK growth rate exerted by Germany is of intrinsic importance to policy makers.

The rest of the chapter is structured as follows. Section 2 introduces the LSTVAR model and Bayesian inferences. Section 3 presents empirical results. Section 4 concludes.

## 1.2 Logistic Smooth Transition VAR Model

The vector autoregressive model (VAR) has proven very successful in modeling endogenous relationships among macroeconomic variables without imposing restrictions that may be 'incredible' in the sense of Sims (1972, 1980). We therefore model the pairwise business cycle linkages in a reduced form VAR based on two considerations. First, VAR is ideally suited to the analysis of endogenously determined processes where dynamics are important but where we have little or no clear economic structure. Second, VAR provides an atheoretical framework for analysis and allows very rich dynamics.[2] Considering the possible presence of nonlinearities in the cross-country business cycle linkages, we model the annual growth rates of the two countries of concern in a bivariate LSTVAR system introduced by Weise (1999).

Let $y_t = (y_{1,t}, y_{2,t})$, where $y_{1,t}$ is the annual real GDP growth rate of the country other than the UK (the US, France or Germany), $y_{2,t}$ is the British annual real GDP growth rate. For time $t$=*1,...,T*, the cyclical linkages between the UK and another country can be expressed in the nonlinear autoregressive process of order $p$ as follows.

$$y_t = \Phi + \Sigma_{h=1}^{p} \Gamma_h y_{t-h} + F(z_t) \left[\Phi^z + \Sigma_{h=1}^{p} \Gamma_h^z y_{t-h}\right] + \varepsilon_t, \tag{1.1}$$

where $\varepsilon_t$ is a white noise process, that is $E(\varepsilon_t) = 0$, $E(\varepsilon_s' \varepsilon_t) = \Sigma$ for $s = t$,

[2]Many studies of co-movements of business cycles among the main industrial countries use VAR for modeling the interrelationships, see for example, Norrbin and Schlagenhauf (1996), Helbling and Bayoumi (2003).

and $E(\varepsilon_s'\varepsilon_t) = 0$ for $s \neq t$.

The regime changes are assumed to be captured by the first order logistic smooth transition function defined by the transition variable $z_t$

$$F(z_t) = \left[1 + exp\left\{-\gamma\left(z_t - c\right)/\sigma\right\}\right]^{-1} \tag{1.2}$$

In function (1.2), the parameter $\gamma$ (which is non-negative) determines the speed of the smooth transition. We can see that when $\gamma \to \infty$, the transition function becomes a Dirac function, then model (1.1) becomes a two-regime threshold VAR model along the lines of Tong (1983). When $\gamma = 0$, the logistic function becomes a constant (equal to 0.5), and the nonlinear model (1.1) collapses into a linear VAR($p$). The parameter $c$ is the threshold around which the dynamics of the model change. The value for the parameter $\sigma$ is chosen by the researcher and could reasonably be set to one. However, if we set $\sigma$ equal to the standard deviation of the process $z_t$, this effectively normalizes $\gamma$ such that we can give $\gamma$ 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 $\gamma$.

The principle underlying the LSTVAR is that as $z_t$ increases, moving from well below some threshold $c$ to well above this threshold, the dynamics of the vector process $y_t$ changes from one regime to another. That is, if $z_t$ is very low - i.e., well into what we will call the *lower* regime for nominal purposes

- then the process $y_t$ may be generated by the VAR model as follows.

$$y_t = \Phi + \Sigma_{h=1}^{p} \Gamma_h y_{t-h} + \varepsilon_t \tag{1.3}$$

However, when $z_t$ is very high - i.e., well into what we will call the *upper* regime - then the process $y_t$ may be generated by the VAR given by

$$y_t = \Phi^1 + \Sigma_{h=1}^{p} \Gamma_h^1 y_{t-h} + \varepsilon_t \tag{1.4}$$

The transition between these two regimes is smooth and governed by a smooth function of $z_t$ denoted by $F(z_t)$. The value of $F(z_t)$ is bounded by *0* and *1*. $F(z_t) = 0$ when $z_t$ is very low, and $F(z_t) = 1$ when $z_t$ is very high.

Thus we may express the full process as

$$y_t = (1 - F(z_t)) \left[\Phi + \Sigma_{h=1}^{p} \Gamma_h y_{t-h}\right] + F(z_t) \left[\Phi^1 + \Sigma_{h=1}^{p} \Gamma_h^1 y_{t-h}\right] + \varepsilon_t \tag{1.5}$$

which can equivalently be written as model (1.1).

Observe that model (1.1) encompasses a set of models distinguished by the choice of the transition variable, the order of the autoregressive process, and whether there exist nonlinear effects.

### 1.2.1 Likelihood Function

For notation convenience, we set $x_t = (1, y_{t-1}, ..., y_{t-p})$, and $x_t^\theta = [x_t \;\; F(z_t)x_t]$. Next we stack the vectors over $t$ as $Y = (y_1^{'}, y_2^{'}, ..., y_T^{'})'$, $X^\theta = (X_1^{\theta\prime}, X_2^{\theta\prime}, ..., X_T^{\theta\prime})'$,

$B = (\Phi, \Gamma_1, ..., \Gamma_p, \Phi^z, \Gamma_1^z, ..., \Gamma_p^z)'$, and $E = (\varepsilon_1^{'}, \varepsilon_2^{'}, ..., \varepsilon_T^{'})'$.

Now we can write model (1.1) in a more compact form as

$$Y = X^\theta B + E \tag{1.6}$$

where the dimensions of Y and E are $(T \times 2)$, the dimension of $X^\theta$ is $(T \times k)$, and the dimension of B is $2k$, with $k = 2(1 + 2p)$.

Given the assumptions on the error terms, the likelihood function of the model can be expressed as

$$L(B, \Sigma, \gamma, c) \propto |\Sigma|^{-T/2} exp \left\{ -\frac{1}{2} tr \Sigma^{-1} E'E \right\} \tag{1.7}$$

Using standard algebraic results, it is possible to show that

$$E'E = S + (B - \widehat{B})' X^{\theta'} X^\theta (B - \widehat{B})$$

where $\widehat{B} = (X^{\theta'} X^\theta)^{-1} X^{\theta'} Y$, and $S = (Y - X^\theta \widehat{B})'(Y - X^\theta \widehat{B})$. Thus, the likelihood function can then be rewritten as

$$L(B, \Sigma, \gamma, c) \propto |\Sigma|^{-T/2} exp \left\{ -\frac{1}{2} tr S \Sigma^{-1} - \frac{1}{2} tr (B - \widehat{B})' X'^\theta X^\theta (B - \widehat{B}) \Sigma^{-1} \right\} \tag{1.8}$$

Vectorizing model (1.6), we can transform model (1.1) into

$$y = x^\theta b + e, \tag{1.9}$$

where $y = vec(Y)$, $b = vec(B)$, $x^{\theta} = I_n \otimes X^{\theta}$, and $e = vec(E)$.

Now, using the relationship between the trace function and the vectorizing operation, we can write the term in the exponent of (1.7) as

$$tr\Sigma^{-1}E'E = e'(\Sigma^{-1} \otimes I_T)e = s^2 + (b - \widehat{b})'V^{-1}(b - \widehat{b}) \tag{1.10}$$

where $s^2 = y'M_V y, M_V = \Sigma^{-1} \otimes \left(I_T - X^{\theta}(X^{\theta'}X^{\theta})^{-1}X^{\theta'}\right), \widehat{b} = vec(\widehat{B})$ and $V = \Sigma \otimes (X^{\theta'}X^{\theta})^{-1}$.

Hence, the likelihood function in (1.7) can also be written as

$$L(b, \Sigma, \gamma, c) \propto |\Sigma|^{-T/2} exp\left\{-\frac{1}{2}\left[s^2 + (b - \widehat{b})'V^{-1}(b - \widehat{b})\right]\right\} \tag{1.11}$$

which has a more familiar Normal form for vector $b$.

### 1.2.2 Priors

In setting the values for the priors we take into account a number of considerations. It is apparent that the LSTVAR model is highly parameterized and the degree of parameterizations influences the quality of inference in finite samples. Priors that are tight around zero (i.e., very informative) tend to improve estimation in VARs [Ni and Sun (2003)]. Also, we use Bayes Factors for inference on models. As discussed in Strachan and van Dijk (2004b), the Bayes factors are functions of the prior normalizing constants and so the prior settings can have a strong influence on the posterior model weights.

Generally, less informative priors will tend to penalize more highly parameterized models. A final consideration is that we have little understanding of the behaviour of economic growth beyond anecdotal evidence and how it can be reasonably modeled. Thus, we face a potential conflict between our desire to specify uninformative priors for a large number of parameters, and priors that are informative which would improve the efficiency of estimation. Furthermore, we do not want to completely avoid or prefer the use of large models *a priori*. Taking into account these considerations, we elicit the priors as follows.

To start with, we assume all models to be *a priori* equally likely. Next, following Zellner (1971), we specify a standard Jeffreys prior for $\Sigma$ as

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

We plan to compute posterior probabilities for model inference. For these probabilities to be well defined, the priors for any parameters that change dimensions, i.e. $b$, must be proper [see Bartlett (1957) and Strachan and Van Dijk (2004b) for further discussion]. Hence, we assume the prior for $b$ is Normal with zero mean and covariance matrix $\underline{V} = \eta^{-1} I_{nk}$, where $\eta$ is a shrinkage prior distributed as Gamma with mean $\underline{\mu_{\eta}}$, and degrees of freedom $\underline{\nu_{\eta}}$. Note that the prior variance for $b$ depends on $\eta$. Large values of $\eta$ imply greater shrinkage towards zero which will tend to reduce the expected frequentist risk of the estimator. However, smaller values of $\eta$ will imply a less informative prior. To allow prior for $b$ that is relatively uninformative, but still allow for a degree of shrinkage, we specify the prior of $\eta$ distributed as $G(10, 0.001)$, where 10 is the mean, and 0.001 is the degree of freedom.

As explained in Bauwens, Lubrano and Richard (1999), at the point where $\gamma = 0$, the smooth transition function in (1.2) becomes a constant and, as a consequence, elements of $b$ become unidentified. While when $\gamma \to \infty$, under a flat prior for $\gamma$, the posterior is not integrable. Hence, following the suggestion of Lubrano (1999a, 1999b), we exclude *a priori* the point $\gamma = 0$ from the support of $\gamma$. Specifically, we assume the prior of $\gamma$ is a Gamma distribution with mean $\underline{\mu_\gamma}$ and degree of freedom $\underline{\nu_\gamma}$. Note that although the prior for $\gamma$ excludes zero, as the prior for $b$ is centered on zero, this restriction does not bias in favor of asymmetry. We define the prior mean of $\gamma$ as 1, in line with the starting values of grid search in most of the classical works [see, for example, Öcal and Osborn (2000) and Sensier, Osborn and Öcal (2002)], while our assumption that the degree of freedom of the prior Gamma distribution is 0.001 is for minimizing the prior's influence on posterior computations.

In the end, we assume the prior of the location parameter $c$ as uniformly distributed between the upper and lower limits of the middle 80% of the observed transition variables.

### 1.2.3 Posteriors Computations

We use Gibbs Sampling to compute the outputs from the posteriors. Conditional upon $\gamma$, $c$, and $\eta$, the model is linear. Thus the conditional posterior distributions of $\Sigma$ and $b$ and are of standard forms. Combining the likelihood function (1.7) and the priors, we obtain the conditional posterior distribution for $\Sigma$ as an inverted Wishart with scale matrix $E'E$ and degrees of freedom $T$, and the conditional posterior distribution for the vector $b$ as Normal with mean $\bar{b}$ and variance $\overline{V}$, where $\overline{V} = (V^{-1} + \eta I_{nk})^{-1}$, and $\bar{b} = \overline{V} V^{-1} \widehat{b}$.

To obtain the conditional posterior for $\eta$, we combine the prior and the likelihood to obtain the expression

$$p(\eta|b, \Sigma, \gamma, c, y, x) \propto \eta^{\frac{\underline{\nu_\eta}+nk-2}{2}} exp(-\frac{\eta \underline{\nu_\eta}}{2\underline{\mu_\eta}} - \frac{1}{2} b'b\eta) \tag{1.12}$$

Thus with a Gamma prior, the conditional posterior distribution of $\eta$ is Gamma with degrees of freedom $\overline{\nu_\eta} = nk + \underline{\nu_\eta}$, and mean$\overline{\mu_\eta} = \frac{\overline{\nu_\eta}\underline{\mu_\eta}}{\underline{\nu_\eta}+\underline{\mu_\eta}b'b}$.

The posterior distributions for the remaining parameters, $\gamma$ and $c$, have nonstandard forms. However, we can use Metropolis-Hastings algorithms [Chib and Greenberg (1995)] within Gibbs to estimate $\gamma$, and the Griddy Gibbs sampler [Ritter and Tanner(1992)] to estimate $c$.

The Gibbs sampling scheme for our posterior computation, therefore, takes the following form.

1. Initialize $(b, \Sigma, \gamma, c, \eta) = (b^0, \Sigma^0, \gamma^0, c^0, \eta^0)$;
2. Draw $\Sigma|b, \gamma, c, \eta$ from $IW(E'E, T)$;

3. Draw $b|\Sigma, \gamma, c, \eta$ from $N(\bar{b}, \overline{V})$;
4. Draw $\gamma|b, \Sigma, c, \eta$ through Metropolis-Hastings method;
5. Draw $c|b, \Sigma, \gamma, \eta$ numerically by Griddy Gibbs;
6. Draw $\eta|b, \Sigma, \gamma, c$ from $G(\overline{\mu_\eta}, \overline{\nu_\eta})$;
7. Repeat step 2 to 6 for a suitable number of replications.

To avoid the draws from Metropolis-Hastings simulator getting stuck in a local mode, we try different starting values for the sampler.

### 1.2.4 Posterior Model Probabilities

There has been a great deal of work on the theories of business cycles and even on the asymmetries observed in business cycles. However, there are relatively fewer formal theories on the nonlinear effects in international business cycle linkages. Thus we have little guidance on how to specify the model prior to introducing the data. Further, notwithstanding the few studies that do exist, we do not wish at this stage of the research to impose any restrictions implied by particular theories. Our interest is on the existence of the linkages and the form of the asymmetries. These concerns were important motivations for considering LSTVAR models. However, we also have reason to expect that the real data generating process might be nonlinear, yet we do not wish to exclude the possibility that the model is linear. A linear model may prove more robust if the asymmetric effect is trivial. Thus, we include the standard linear VAR in our model set. Furthermore, we can not confidently pre-specify the driving force of the asymmetric dynamics (if there is any) nor predetermine the duration of the dynamics, so we allow for a range of specifications of $z_t$ and lag lengths $p$.

Bayesian methods provide us a formal method for 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 Bayesian Model Averaging (also known as the Bayesian Model Pooling technique in the literature[3]). 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.

The Bayes Factor for comparing one model to a second model where each model is parameterized by $\zeta = (\zeta_1, \zeta_2)$ and $\psi$ respectively, is

$$B_{12} = \frac{\int \ell(\zeta)p(\zeta)d(\zeta)}{\int \ell(\psi)p(\psi)d(\psi)},$$

where $\ell(.)$ is the likelihood function and $p(.)$ is the prior density of the parameters for each model.

If the second model nests within the first at the point $\psi = \zeta_1$ and $\zeta_2 = \zeta^*$, then, subject to further conditions, we can compute the Bayes factor $B_{12}$ via the Savage-Dickey density ratio [see, for example, Koop and Potter (1999a), Koop, Leon-Gonzales and Strachan (2006) for further discussion in this class

[3] We would like to thank Mateusz Pipień for pointing this out.

of models]. For the simple example discussed here, the Savage-Dickey density ratio is:

$$B_{12} = \frac{p(\zeta_2 = \zeta^* | Y)}{p(\zeta_2 = \zeta^*)},$$

where the numerator is the marginal posterior density of $\zeta_2$ for the unrestricted model evaluated at the point $\zeta_2 = \zeta^*$, and the denominator is the prior density of $\zeta_2$ also evaluated at the point $\zeta_2 = \zeta^*$.

Since the conditional posterior of $b$ is normal, it is easy to incorporate the estimation of the numerator of the Savage-Dickey density ratio in the Gibbs sampler. As to the denominator of the Savage-Dickey density ratio, using the properties of the Gamma distribution and the Normal distribution, we derive the marginal prior for a sub-vector of $b$ evaluated at zeros as

$$\{(\frac{\underline{\mu}_\eta}{\pi \underline{\nu}_\eta})^{\omega/2} \Gamma(\frac{\omega + \underline{\nu}_\eta}{2})\}/[\Gamma(\frac{\underline{\nu}_\eta}{2})]$$

where $\Gamma(.)$ is the Gamma function, and $\omega$ is the number of elements in $b$ being restricted to be zero.

A simple restriction in our application to choose is the point where all lag coefficients are zero, i.e., $\Gamma_h = \Gamma_h^z = 0$, at which point we have the model with $p = 0$. 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.

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 Bayesian Model Averaging and weight features of interest across different models using posterior model probabilities [as suggested by Leamer (1978)].

## 1.3 Empirical Application

The data we used are quarterly observations of real GDP for the UK, the US, France and Germany over the period of 1970:Q1-2004:Q4. All series are taken from Datastream. For all cases, the first quarter of 1970 is set as the base time for index purposes. We construct the annual growth rates by taking the fourth-difference of log real GDP index.[4]

The growth rates for the four countries are plotted in figure 1.1. Note that all the series are stationary and free from seasonal components. The average annual growth rates for the sample period are: 2.34% for the UK, 3.08% for the US, 2.49% for France and 2% for Germany. The correlations between the annual growth rate for the UK and that of the US, France and Germany are 0.5941, 0.3606 and 0.3693, respectively. Note that the dynamics of recessions are quite different from those of expansions, a phenomenon which might imply the presence of asymmetry.

For all countries, we assume the maximum order of the unrestricted bivariate LSTVAR is 4. Although the driving force of the asymmetry can be

[4]The jump in German data due to the reunification in 1991 has been corrected.

any exogenous or endogenous variables of concern, following the convention, we simply choose a specific lag of the observed growth rate from our selected countries as the transition variable. However, instead of picking a plausible lagged growth rate from a particular country, we allow $z_t$ to be any of the 16 observations of the lagged (from 1-4) annual growth rates for the UK, the US, France or Germany. Note that this specification allows for the driving force of the regimes to be generated within or beyond the two countries being examined under the bivariate VAR. As we allow the order of the VAR to vary from one to four, then for each of the three bilateral relationships we consider a total of 68 models.[5]

### 1.3.1 Posterior Evidence on Alternative Models

We calculate Bayesian posterior model probabilities from the Bayes Factors comparing the nested models to the unrestricted LSTVAR(4, $z_t$) models.[6] The Gibbs Sampler for each of the unrestricted LSTVAR (4, $z_t$) model 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 diagnostic.

The posterior probabilities for the top 10 models evaluated at Bayes Fac-

[5] The total number of models is calculated as 4 (maximum order of the nonlinear VAR)$\times$ 4 (choices for $z_t$)$\times$ 4 (lags of $z_t$) + 4 (the number of linear VAR models)= 68.

[6] Where the order of the model is 4, and the transition variable $z^t$ equals to $US_{t-1}, US_{t-2}, US_{t-3}, US_{t-4}, FR_{t-1}, FR_{t-2}, FR_{t-3}, FR_{t-4}, UK_{t-1}, UK_{t-2}, UK_{t-3}, UK_{t-4}, GER_{t-1}, GER_{t-2}, GER_{t-3}, GER_{t-4}$, respectively.

tor are reported in table 1.1. As we calculate posterior model probabilities with relatively uninformative priors, we would expect this to reward parsimony and, as such, penalize the nonlinear models. However, there is little overall evidence for linear models (which, for a given lag length, is the most parsimonious model). This reinforces the evidence in favor of asymmetry in the bilateral business cycle linkages between the UK and each of the other three countries.

Posterior model probabilities reveal that model uncertainty is not a significant issue in this data. For France and the UK, we find the bivariate relationship can be jointly captured by LSTVAR(4, $US_{t-2}$) and LSTVAR(4, $UK_{t-2}$), with posterior probabilities 52.34% and 36.99%, respectively. While model comparison results involving the US and Germany show that a single model receives substantial posterior support in each case. For US-UK, LSTVAR(4, $UK_{t-4}$) accounts for 90.38% of the posterior probability. For Germany-UK, the posterior model probability of LSTVAR(4, $GER_{t-3}$) is 92.68%.

We observe four interesting findings from our model comparison results. First, the US growth rates play a leading role in triggering the regime changes for France-UK and a non-negligible role in causing the nonlinear effects for Germany-UK. Second, the regime changes are governed by the UK business cycles in the case of US-UK. Third, Germany's economic performance is important for the regime changes in all cases, in particular, it plays a deterministic role in the case of Germany-UK. Finally, we find that the role of

France's growth rate in triggering the regime changes is nearly negligible in all cases. Observe that even though for France-UK, the nonlinear effects are mainly determined by the growth rates of the US and the UK.

It is hard to explain the parameters in such big nonlinear models. Yet, we present the estimated UK equations for the three most preferred models in table 1.2, for the smooth transition functions and the impulse response analysis we are going to report are based on these results.

To better understand the form of the asymmetric affect, we plot the graphs of the time profile of $F(z_t)$ and the corresponding transition functions over the range of $z_t$ for the three most probable models in figures 1.2 - 1.3. For comparison, we also report the time profiles of $F(z_t)$ derived from Bayesian Model Averaging in figure 1.4. Observe that for US-UK, the dynamics of the regime changes remains to be between the upper and lower regimes, for France-UK, the model is most often in the upper regimes, while for Germany-UK, more abrupt regime changes can be spotted. From these figures, we can see that the regime changes are rather smooth in all the three cases. Thus, it is improper to model the nonlinear effects using functions that only allow for abrupt changes.

### 1.3.2 Impulse Response Analysis

The nonlinear LSTVAR allows for asymmetries in the behaviour of the business cycle linkages. Thus the model provides richer inference on the possible response paths that account for both the nature of the shocks and the current economic environment. In analyzing the response of the UK economy to the foreign shocks we are interested in how the economy responds taking into account the magnitude of the shock, whether the shock is positive or negative and whether UK growth is negative or positive at the time of the shock. For example, it would seem natural to expect that the response to a positive growth shock from the US, say, will have a different effect upon UK's growth if the UK is currently growing quickly than if the UK is in a recession.

As discussed in, *inter alia*, Potter (1995), Koop, Pesaran and Potter (1996), Koop and Potter (2000), impulse response functions of nonlinear models are history- and shock- dependent. This contrasts with the traditional impulse response analysis in a linear VAR in which positive and negative shocks are treated symmetrically and independent of the current state of the business cycle. Thus, the traditional methods of computing impulse responses are unable to inform us on nonlinearities in responses [see Koop, Pesaran and Potter (1996) for detailed discussions]. We therefore follow these earlier papers and use generalized impulse response functions (GIRF)[7] to measure the effect of a shock on the asymmetric system.

Following Koop, Pesaran and Potter (1996), we examine the GIRF where we have a shock $\upsilon_t$ and a history $\omega_{t-1}$ which is defined as follows

$$GI_y(n, \upsilon_t, \omega_{t-1}) = E[y_{t+n}|\upsilon_t, \omega_{t-1}] - E[y_{t+n}|\omega_{t-1}] \tag{1.13}$$

[7]The term impulse response functions, if without any specific description, also refers to general impulse response functions hereafter.

where $n$ is the number of periods into the future after the time $t$.

The definition in (1.13) is the expected response path where the expectation is taken with respect to the distribution of all future shocks, the distribution of the parameters and, if model averaging is employed, with respect to the posterior distribution of the models. That is, the impulse response is the expected deviation of $y_{t+n}$ subject to the shock $\upsilon_t$ from the expected value of $y_{t+n}$ without fixed future shocks and conditional only upon the history at time $t$, $\omega_{t-1}$.

Estimation of the GIRF for a specific model with given parameters is detailed in the literature mentioned above. Here, we only outline how we achieve an estimate that is not conditional upon any parameter values.

We wish to calculate the GIRF for a given shock $\upsilon_t$ and history $\omega_{t-1}$. Assume we have the $i^{th}$ draw from the Gibbs sampler of the parameters in the model which we will denote by $\theta^{(i)}$. For each draw we compute $GI_y(n, \upsilon_t, \omega_{t-1}|\theta^{(i)})$ which is simply (1.13) for a given value of the parameters. Next assume we have $N$ draws of $\theta^{(i)}$ where $i = 1, ..., N$. Then we can compute an estimate of (1.13) from by

$$\widehat{GI}_y(n, \upsilon_t, \omega_{t-1}) = \frac{1}{N}\Sigma_{i=1}^{N} GI_y(n, \upsilon_t, \omega_{t-1}|\theta^{(i)}).$$

By drawing randomly from histories and averaging across these, we are able to obtain an estimate of $GI_y(n, \upsilon_t)$ which is not conditional upon the current state of the economy. Furthermore, we report the estimates of

$GI_y(n, \upsilon_t, \omega_{t-1})$ conditional upon some special $\omega_{t-1}$ since we believe these paths may differ for different histories. To be specific, we are interested in whether the path of $GI_y(n, \upsilon_t, \omega_{t-1})$ differs when the UK economy exhibits a positive growth in comparison to a negative growth.

Finally, we report the estimated path of $GI_y(n, \upsilon_t, \omega_{t-1})$ when the shock $\upsilon_t$ is a negative one/two standard deviations shock to the US, France or German economy, as well as when $\upsilon_t$ is a positive one/two standard deviations shock to the US, France or German economy. In the estimation of the posterior distributions of these functions, we found that outliers distorted the posterior means of the GIRFs in some cases. Therefore, we report the median of the GIRFs instead of the mean.[8]

Graphs of the median estimates of the GIRFs for the most preferred model and the BMA results, respectively, are plotted in figures 1.5 - 1.10. In each figure, we use six graphs to examine general impulses from different dimensions. In the upper panel of the figure, we display the impact on the UK growth of positive and negative shocks from the other country but where we have averaged across all the UK histories. The middle panel of the figure shows the same response of UK growth but the path is conditional upon the UK's economy being in expansion at the time of the shock. The lower panel of the figure presents the corresponding effects when UK's economy is in contraction at the time of the shock.

[8]The mean of the GIRFs with the outliers being dropped share the similar pattern with the median results. Graphs depicting mean values of GIRF are available upon request.

An inspection of all the graphs reveals that the GIRFs plots for the most preferred model and that of the BMA results appear to be similar for all the three country pairs, which is in consistent with the model comparison result earlier reported.

Observing the GIRFs for US-UK plotted in figures 1.5 - 1.6, we see that the impact of a US shock on the UK is in all cases prominent for the first seven to eight quarters, after which there remain much smaller cyclical effects. Finally, the impulse responses die out in about nine years. It is seen that the cumulative effect of a positive US shock will increase the UK's output growth rate, while the cumulative effect of a negative shock from the US will decrease the UK's output growth rate.

With respects to France-UK, from figures 1.7 - 1.8, we can see that while there are strong immediate positive and negative responses to shocks of the same sign, the cyclical effect is much less pronounced than in the case of US-UK. Observe that much of the impact takes place in the first six quarters after the shock. Afterwards, only some smaller cyclical effect remains for another nine quarters. Overall, the impact from France dies out in five years. Similar to that of US-UK, we find a positive shock emanating from France would boost the UK economy, and a negative shock from France would offset the UK's growth.

By visual inspection, we can hardly find any nonlinearities in the GIRFs for US-UK and France-UK. First, the graphs for positive shocks appear to mirror the graphs for negative shocks. Second, the impacts of shocks of differing magnitude seem to have proportionate effects. Third, it looks like that the dynamics of the impulse responses is independent of the status of the UK's economy when the shocks hit.

Noticeable nonlinearities in impulse response functions are observed in the case of Germany-UK. Observing figures 1.9 - 1.10, we find the paths of the responses will not just differ given the sign and the magnitude of the shock, but also given the current state of the UK economy. Surprisingly, we find that the cumulative effect of any type of innovations in Germany is to slow down the UK economy. More strikingly, we find a positive shock from Germany brings more negative effect to the UK output growth than a negative shock. For a given status of the UK economy when the shock from Germany happens, we can order the shocks by gravity for negatively affecting the UK growth rate. We find, in descending order of severity, that it is the large positive shock, the small positive shock, the large negative shock and the small negative shock. Finally, we observe that when the UK economy is in recession when the shock happens, the overall setting back effect from Germany is less than when the UK's economy is in expansion.

## 1.4 Conclusions

In this chapter, we investigate bivariate relationships between the UK and three main industrial countries - the US, France, and Germany - within the framework of a LSTVAR model. We employ Bayesian methods to develop an approach to model estimation and evaluation.

The estimation results show that the UK's business cycles are asymmetrically influenced by the other three countries. Overall, it would seem that the UK benefits from positive shocks emanating from the US and France, while suffers from negative shocks from these two countries. However, we also observe that Germany always play a pernicious role in the UK's economy. More strikingly, we find that a boom in Germany would bring more negative impact on the UK than a bust.

As a purely atheoretical study, this chapter only describes the behaviour of the linkages between the UK and each of the other three countries. For a better understanding of the forms and sources of these linkages, further investigations (for examples, on the transmission channels) which are beyond our current research are called for.

Table 1.1: Top 10 models

| No. | US - UK | | FR - UK | | GER - UK | |
|---|---|---|---|---|---|---|
| 1 | $LSTVAR(4)_{UK_{t-4}}$ | 0.9038 | $LSTVAR(4)_{US_{t-2}}$ | 0.5239 | $LSTVAR(4)_{GER_{t-3}}$ | 0.9268 |
| 2 | $LSTVAR(4)_{GER_{t-3}}$ | 0.0465 | $LSTVAR(4)_{UK_{t-2}}$ | 0.3699 | $LSTVAR(4)_{US_{t-1}}$ | 0.0501 |
| 3 | $LSTVAR(4)_{UK_{t-3}}$ | 0.0292 | $LSTVAR(4)_{FR_{t-2}}$ | 0.0225 | $LSTVAR(4)_{GER_{t-1}}$ | 0.0126 |
| 4 | $LSTVAR(4)_{GER_{t-4}}$ | 0.0147 | $LSTVAR(4)_{US_{t-1}}$ | 0.0189 | $LSTVAR(4)_{GER_{t-2}}$ | 0.0017 |
| 5 | $LSTVAR(4)_{US_{t-1}}$ | 0.0032 | $LSTVAR(4)_{FR_{t-3}}$ | 0.0162 | $LSTVAR(4)_{US_{t-2}}$ | 0.0012 |
| 6 | $LSTVAR(4)_{UK_{t-1}}$ | 0.0015 | $LSTVAR(4)_{FR_{t-1}}$ | 0.0162 | $LSTVAR(4)_{UK_{t-1}}$ | 0.0011 |
| 7 | $LSTVAR(4)_{US_{t-4}}$ | 0.0007 | $LSTVAR(4)_{GER_{t-4}}$ | 0.0113 | $LSTVAR(4)_{UK_{t-4}}$ | 0.0009 |
| 8 | $LSTVAR(4)_{UK_{t-2}}$ | 0.0001 | $LSTVAR(4)_{UK_{t-3}}$ | 0.0088 | $LSTVAR(4)_{US_{t-4}}$ | 0.0008 |
| 9 | $LSTVAR(4)_{GER_{t-23}}$ | 0.0001 | $LSTVAR(4)_{FR_{t-4}}$ | 0.0061 | $LSTVAR(4)_{FR_{t-1}}$ | 0.0008 |
| 10 | $LSTVAR(4)_{US_{t-2}}$ | 0.0000 | $LSTVAR(4)_{US_{t-3}}$ | 0.0020 | $LSTVAR(4)_{UK_{t-3}}$ | 0.0007 |

Notes:

Numbers in parentheses are the orders of VAR. The subscript denotes the transition variable.

Table 1.2: Estimated parameters for the most preferred models

| | US-UK | FR-UK | GER-UK |
|---|---|---|---|
| $\eta$ | 8.9416 (3.4356) | 17.7540 (6.0606) | 15.6150 (5.1681) |
| $c$ | 0.0353 (0.0103) | 0.0097 (0.0056) | 0.0378 (0.0050) |
| $\gamma$ | 0.9886 (0.3812) | 2.7632 (1.4184) | 6.5763 (3.7615) |
| lower regime | | | |
| $\Phi$ | 0.0058 (0.0244) | 0.0183 (0.0080) | 0.0138 (0.0035) |
| $\Gamma_{1,1}$ | 1.0682 (0.1833) | 0.5870 (0.2198) | 0.8401 (0.1487) |
| $\Gamma_{2,1}$ | 0.7063 (0.1913) | 0.5491 (0.1743) | 0.5352 (0.1419) |
| $\Gamma_{1,2}$ | 0.0867 (0.2200) | 0.0757 (0.2246) | 0.1233 (0.1686) |
| $\Gamma_{2,2}$ | 0.0250 (0.1954) | -0.0275 (0.1684) | 0.0343 (0.1551) |
| $\Gamma_{1,3}$ | -0.2249 (0.2223) | -0.4639 (0.2379) | -0.1322 (0.3041) |
| $\Gamma_{2,3}$ | 0.3403 (0.1975) | 0.3824 (0.1697) | 0.3026 (0.1480) |
| $\Gamma_{1,4}$ | -0.1364 (0.1948) | -0.3274 (0.2209) | -0.3537 (0.1477) |
| $\Gamma_{2,4}$ | 0.0748 (0.3669) | -0.1476 (0.1539) | -0.0572 (0.1305) |
| upper regime | | | |
| $\Phi^1$ | 0.0280 (0.0724) | -0.0010 (0.0105) | 0.0499 (0.0404) |
| $\Gamma^1_{1,1}$ | 0.1948 (0.3747) | 0.4036 (0.2478) | -0.2431 (0.2850) |
| $\Gamma^1_{2,1}$ | -0.1739 (0.3531) | 0.2292 (0.2082) | 0.0808 (0.2565) |
| $\Gamma^1_{1,2}$ | 0.0603 (0.4005) | 0.1724 (0.2599) | 0.0359 (0.2918) |
| $\Gamma^1_{2,2}$ | -0.3365 (0.3568) | -0.0735 (0.2237) | 0.2225 (0.2631) |
| $\Gamma^1_{1,3}$ | -0.0867 (0.4138) | 0.0456 (0.2570) | -0.2639 (0.3687) |
| $\Gamma^1_{2,3}$ | 0.0886 (0.3877) | -0.2466 (0.2300) | -0.0758 (0.2437) |
| $\Gamma^1_{1,4}$ | -0.1776 (0.4054) | 0.3180 (0.2502) | -0.5306 (0.2856) |
| $\Gamma^1_{2,4}$ | -0.9271 (0.4002) | -0.0724 (0.2038) | -0.6705 (0.2598) |

Notes:

*Standard errors are in parenthesis.

** The first subscript indicates the country, where 1 denotes the country other than the UK, 2 denotes UK. The second subscript denotes the lag length of the variable.

***The superscript 1 indicates the parameter is of the upper regime.

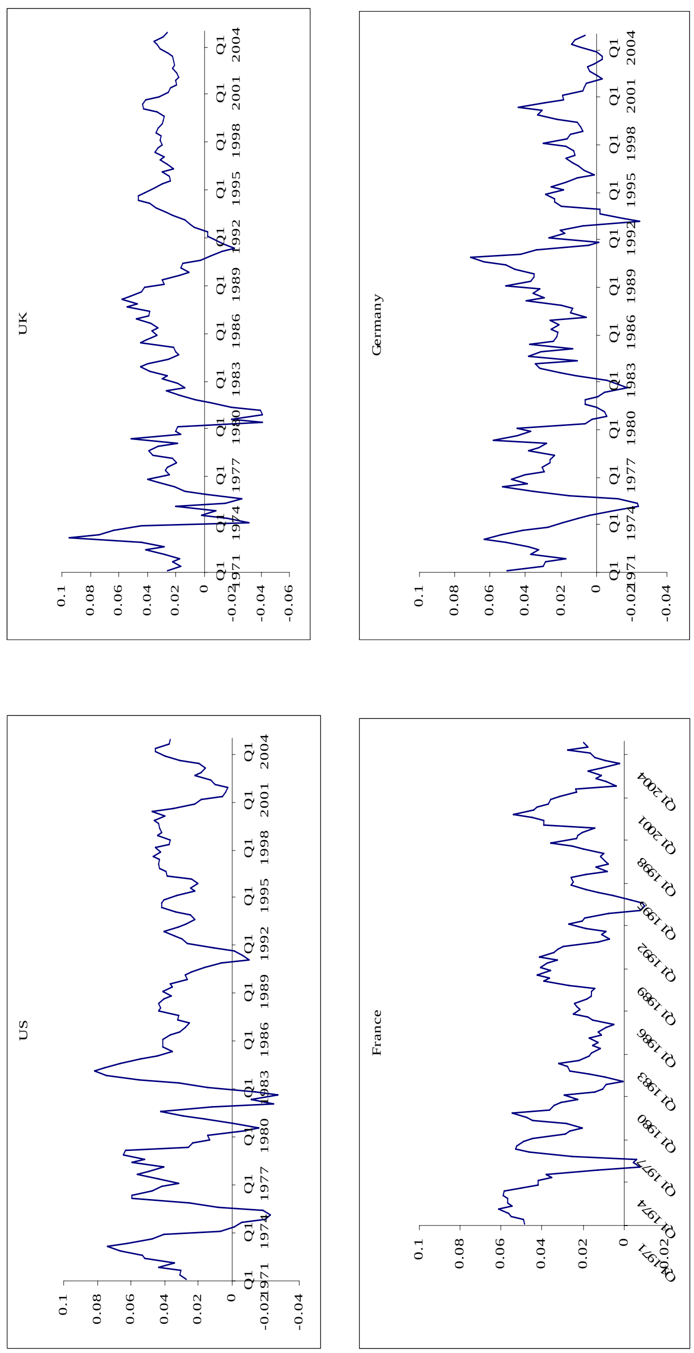
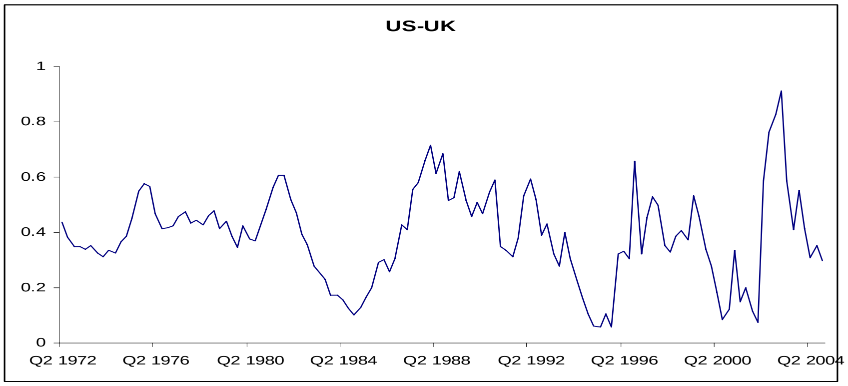
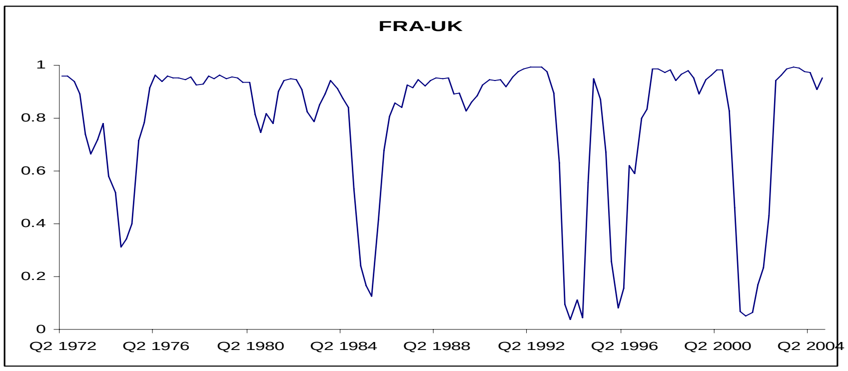



Figure 1.1
Annual Growth Rates

Figure 1.2
Time Profiles of Smooth Transition Functions ___ Most Preferred Models



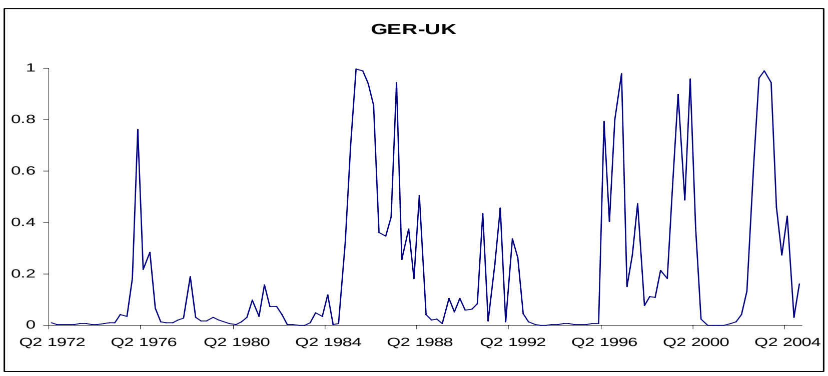



Figure 1.3
Smooth Transition Functions

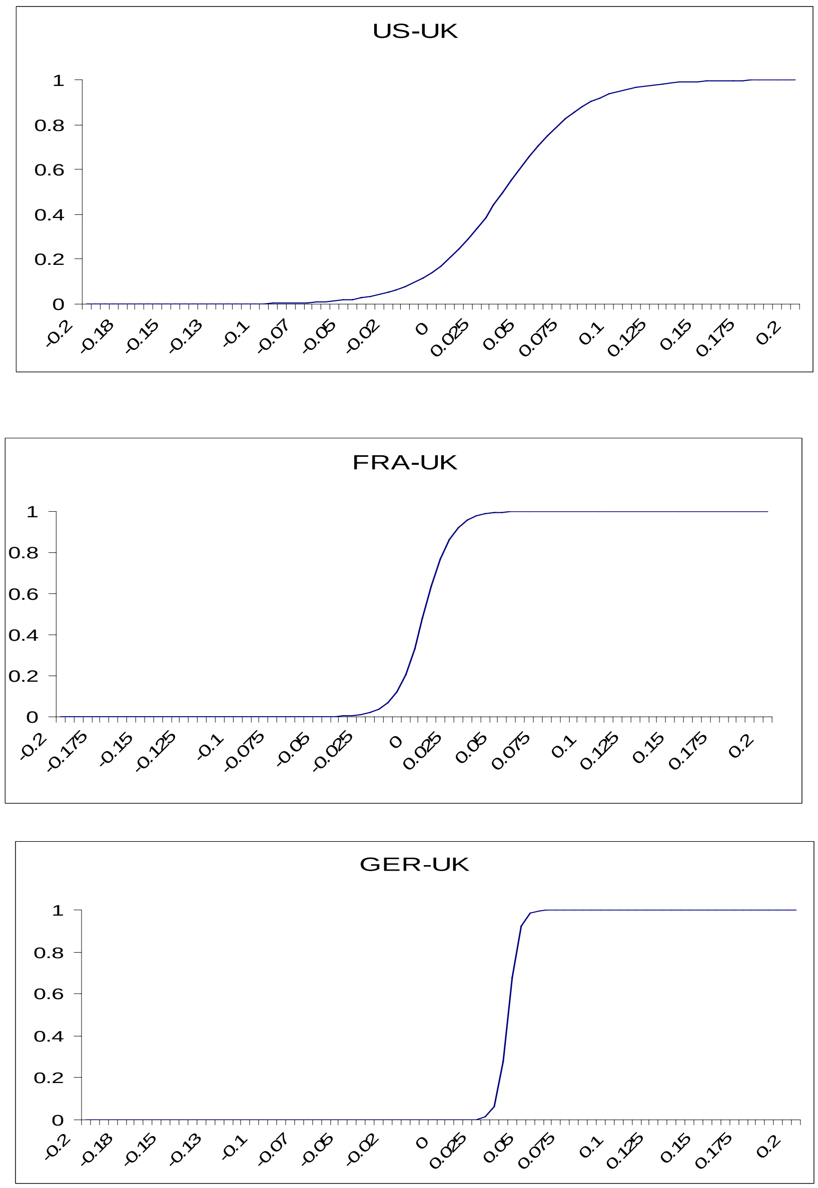


Figure 1.4
Time Profiles of Smooth Transition Functions __ BMA results

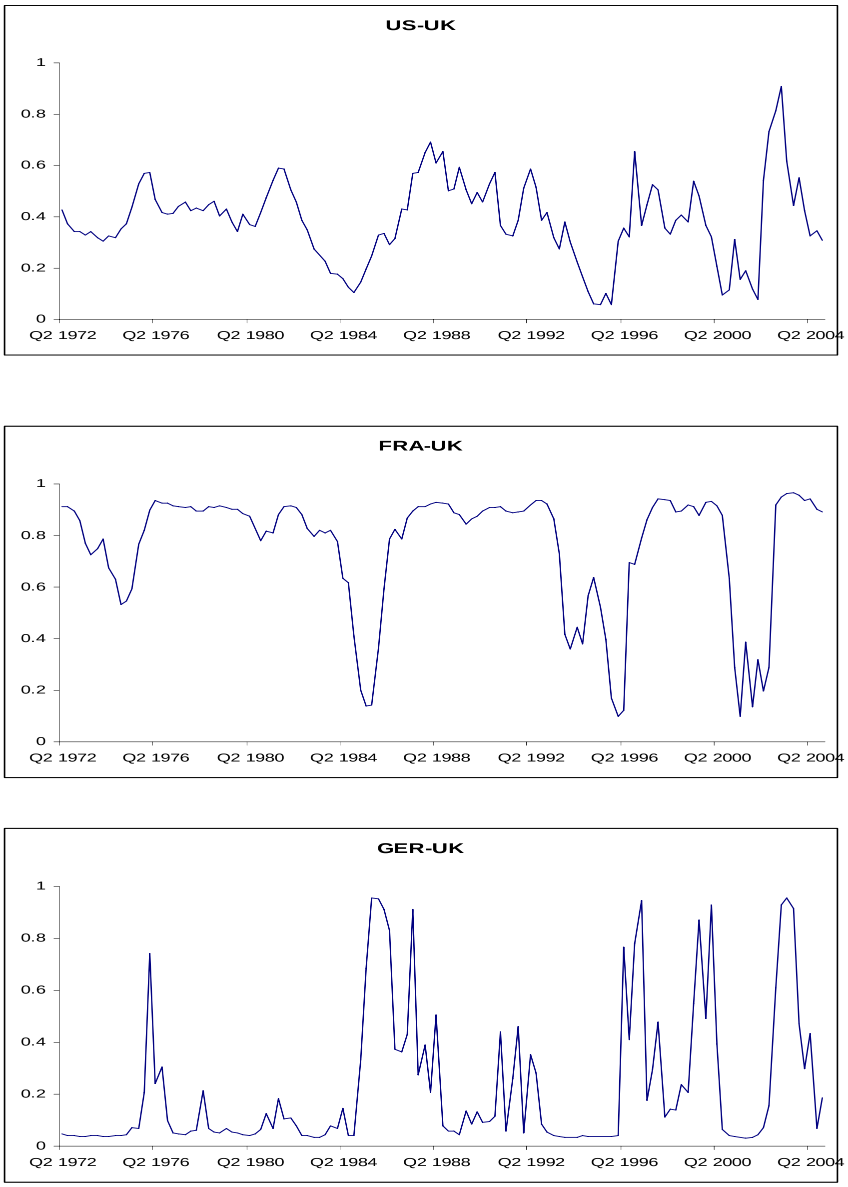
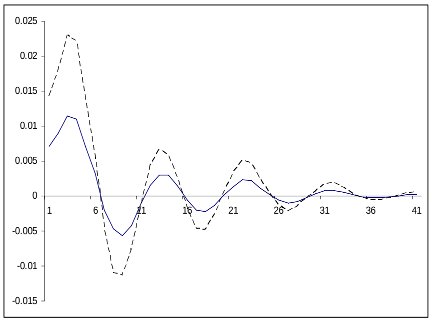
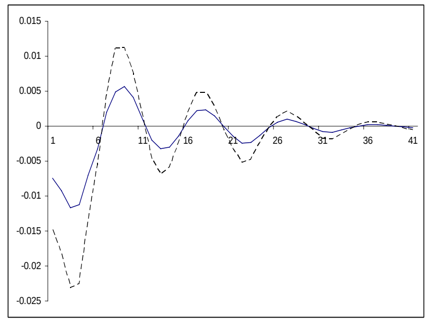


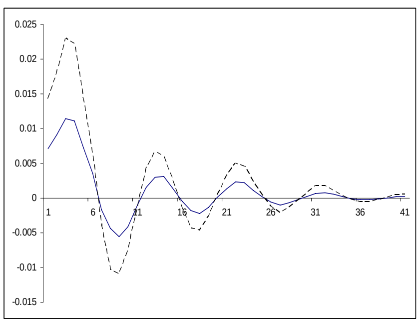
Figure 1.5
General Impulse Response Functions ___ Most Preferred Models



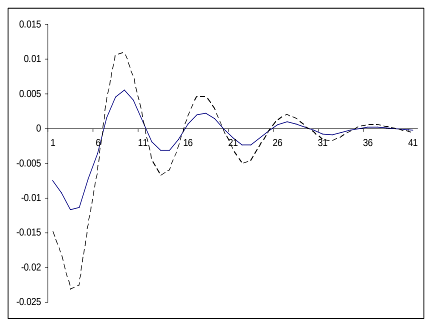

**Impacts of US' Positive Shocks**



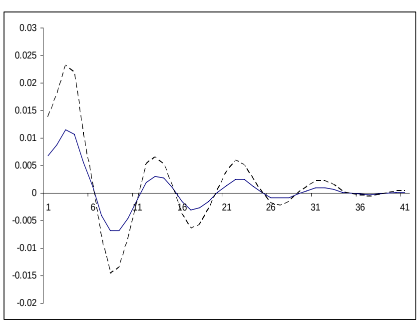

**Impacts of US' Negative Shocks**



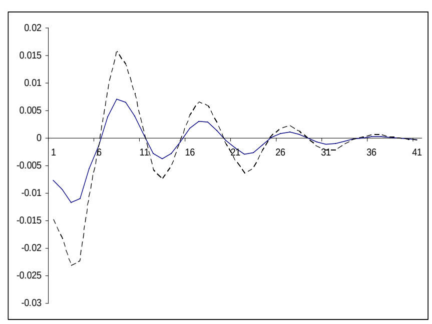

**Impacts of US' Positive Shocks When UK is in Expansion**




**Impacts of US' Negative Shocks When UK is in Expansion**



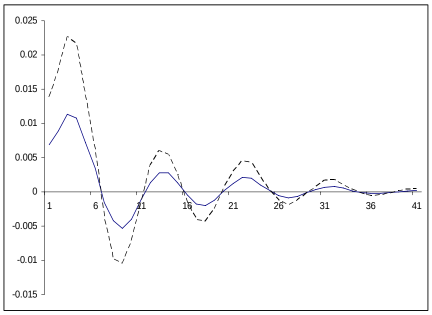

**Impacts of US' Positive Shocks When UK is in Recession**



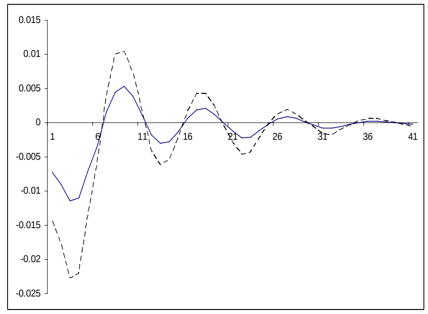

**Impacts of US' Negative Shocks When UK is in Recession**

Notes:
Solid line is for the impulse response function when the shock equal to the standard deviation of the US growth rates. Dashed line is for the impulse response function when the shock equal to two times the standard deviation of the US growth rates.

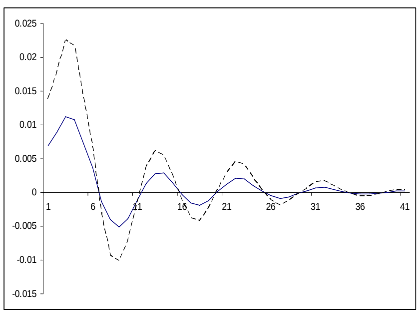
Figure 1.6
General Impulse Response Functions ___ BMA



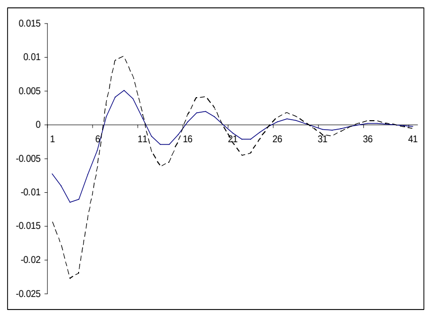

**Impacts of US' Positive Shocks**



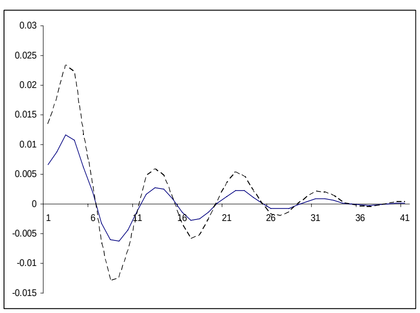

**Impacts of US' Negative Shocks**



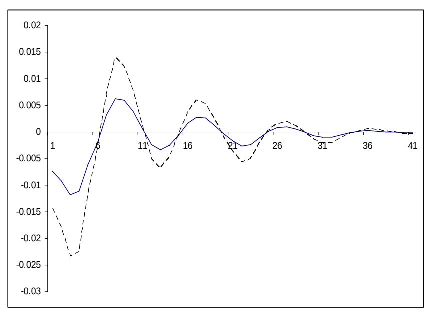

**Impacts of US' Positive Shocks When UK is in Expansion**




**Impacts of US' Negative Shocks When UK is in Expansion**



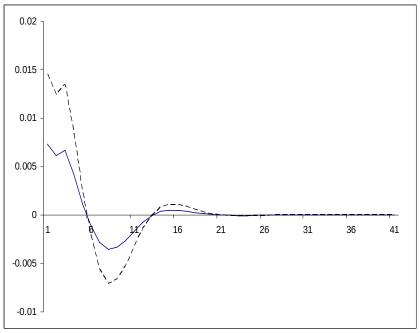

**Impacts of US' Positive Shocks When UK is in Recession**



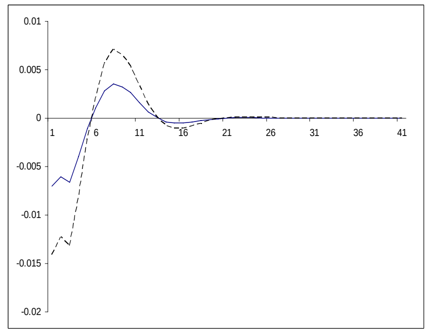

**Impacts of US' Negative Shocks When UK is in Recession**

Notes:
See notes in figure 1.5.

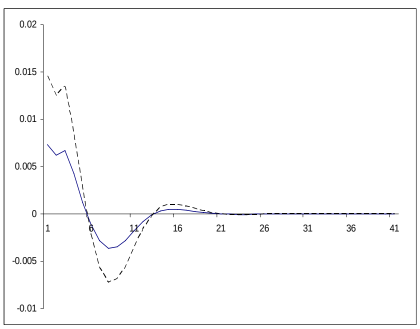
Figure 1.7
General Impulse Response Functions ___ Most Preferred Model



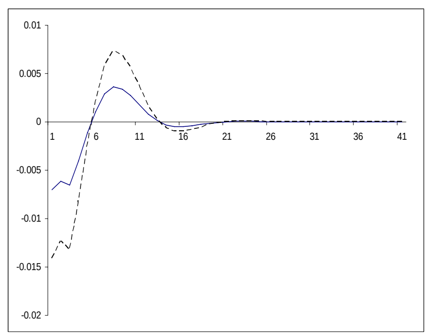

**Impacts of France's Positive Shocks**



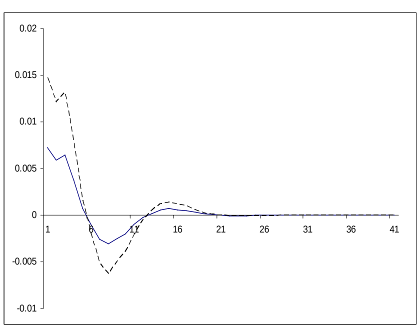

**Impacts of France's Negative Shocks**



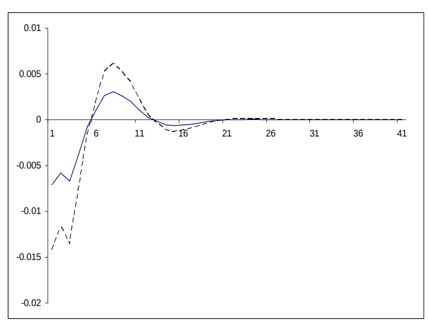

**Impacts of France's Positive Shocks When UK is in Expansion**




**Impacts of France's Negative Shocks When UK is in Expansion**



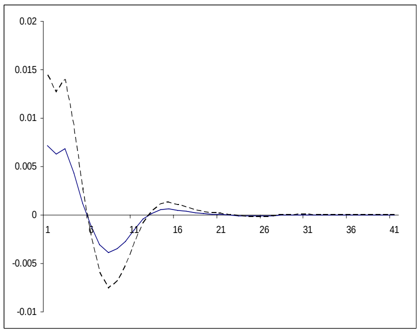

**Impacts of France's Positive Shocks When UK is in Recession**



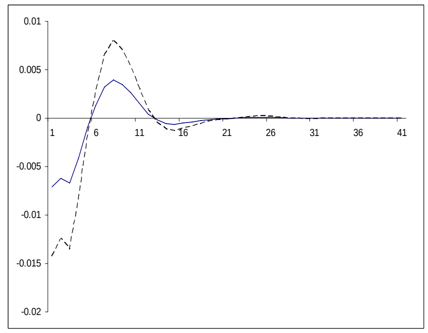

**Impacts of France's Negative Shocks When UK is in Recession**

Notes:
Solid line is for the impulse response function when the shock equal to the standard deviation of France's growth rates. Dashed line is for the impulse response function when the shock equal to two times the standard deviation of France's growth rates.

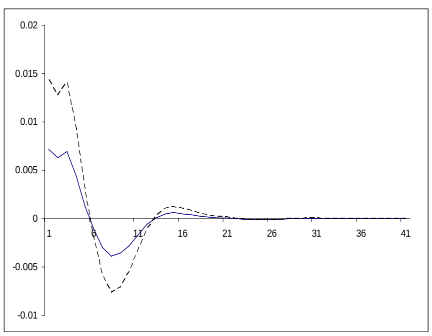
Figure 1.8
General Impulse Response Functions ___ BMA



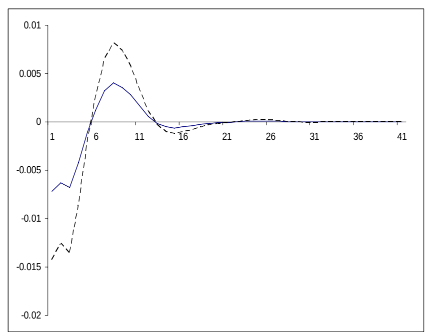

**Impacts of France's Positive Shocks**



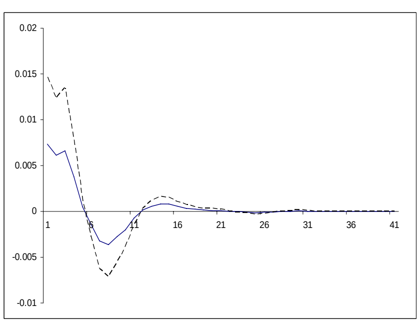

**Impacts of France's Negative Shocks**



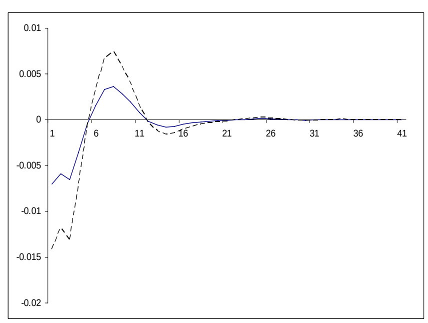

**Impacts of France's Positive Shocks When UK is in Expansion**




**Impacts of France's Negative Shocks When UK is in Expansion**



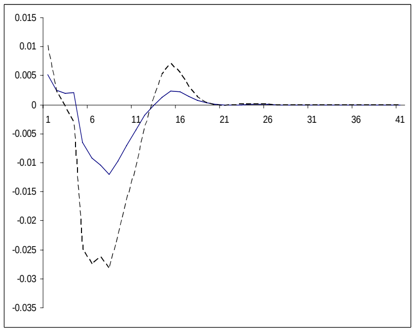

**Impacts of France's Positive Shocks When UK is in Recession**



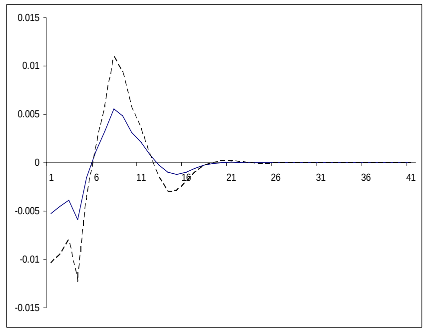

**Impacts of France's Negative Shocks When UK is in Recession**

Notes:
See notes in figure 1.7.

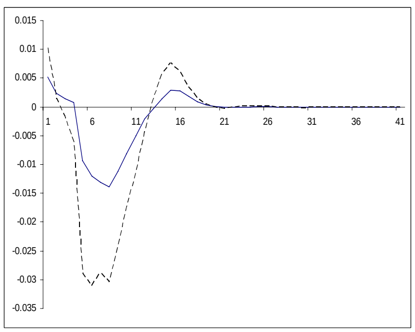
Figure 1.9
General Impulse Response Functions ___ Most Preferred Model



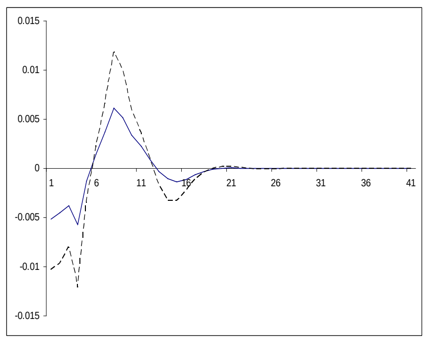

**Impacts of Germany's Positive Shocks**



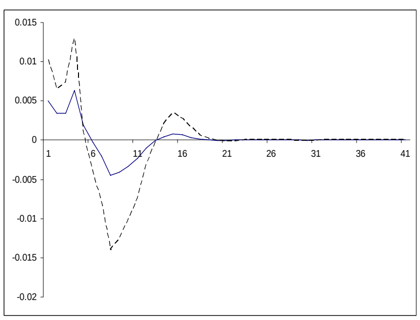

**Impacts of Germany's Negative Shocks**



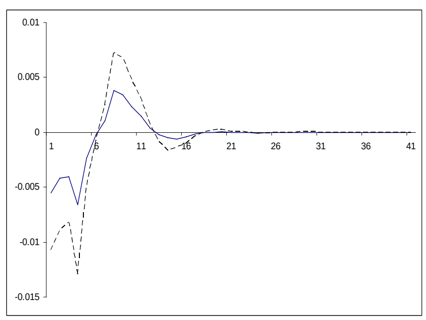

**Impacts of Germany's Positive Shocks When UK is in Expansion**




**Impacts of Germany's Negative Shocks When UK is in Expansion**



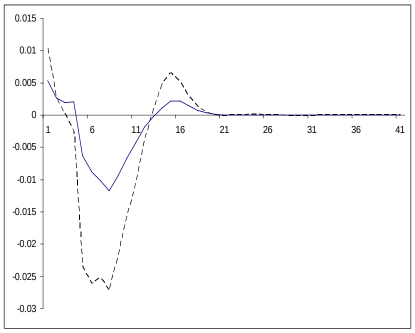

**Impacts of Germany's Positive Shocks When UK is in Recession**



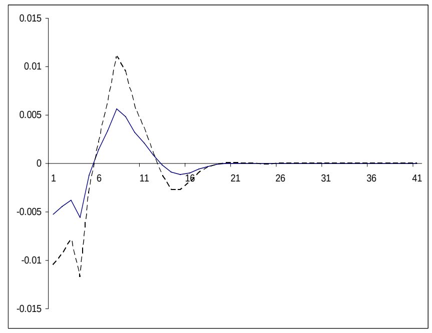

**Impacts of Germany's Negative Shocks When UK is in Recession**

Notes:
Solid line is for the impulse response function when the shock equal to the standard deviation of Germany's growth rates.
Dashed line is for the impulse response function when the shock equal to two times the standard deviation of Germany's growth rates.

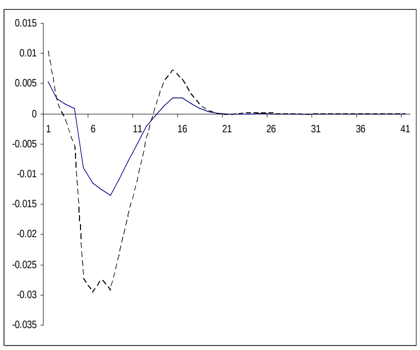
Figure 1.10
General Impulse Response Functions ___BMA



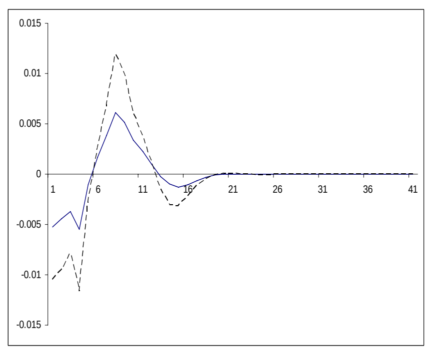

**Impacts of Germany's Positive Shocks**



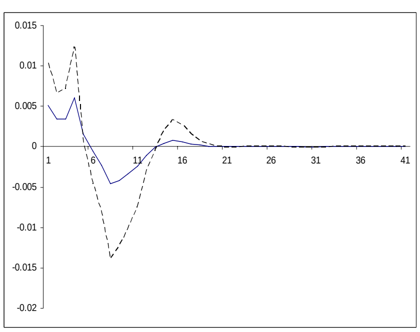

**Impacts of Germany's Negative Shocks**



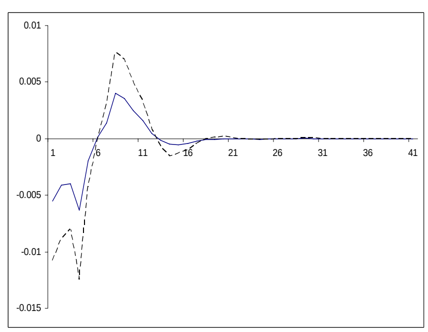

**Impacts of Germany's Positive Shocks When UK is in Expansion**




**Impacts of Germany's Negative Shocks When UK is in Expansion**




**Impacts of Germany's Positive Shocks When UK is in Recession**




**Impacts of Germany's Negative Shocks When UK is in Recession**

Notes:
See notes in figure 1. 9.

# Chapter 2

# Investigating nonlinear purchasing power parity during the post-Bretton Woods era — a Bayesian exponential smooth transition VECM approach

**Acknowledgements:** I thank an anonymous referee, Gary Koop, Roberto Leon-Gonzalez, Mateusz Pipień, Rodney Strachan and the participants at 6th Annual Advances in Econometrics Conference at LSU for their helpful comments and suggestions. This chapter also has benefited from the productive discussions with Roderick McCrorie and Emi Mise. Any remaining errors are the author's responsibility.

## 2.1 Introduction

Given its importance in open economy macro modeling and policy advice, the validity of PPP over the post-Bretton Woods era has been the subject of intensive study in the literature. Employing unit root tests or cointegration tests in a linear framework, earlier work generally fails to confirm the presence of PPP over the modern floating exchange rate regime [e.g., Meese and Rogoff (1988), Edison and Fisher (1991), Mark (1990)]. Inspired by the theoretical arguments that emphasize the role of the transaction cost as proposed by Dumas (1992) and Sercu, Uppal and van Hulle (1995), among others, recent studies turn to analyze whether PPP adjustment follows a nonlinear process. This research has led to evidence in favor of relative PPP [e.g., Michael, Nobay and Peel(1997), Baum, Barkoulas and Caglayan (2001), Sarno, Taylor and Chowdhury (2004), Peel and Venetis (2005)].[1]

The majority of the literature modeling PPP in a nonlinear framework uses univariate models. In these models, the variable of concern is the real exchange rate which is calculated by imposing a cointegrating vector on the nominal exchange rates and the foreign and domestic price levels.[2] However, given the interrelationships among the three variables that constitute

[1] Note that the research adopting a panel data framework [e.g., Lothian (1997), Lopez and Papell (2006)] usually finds support for PPP in the real exchange rates under the recent floating exchange rate regime. However, the panel data approach is not free from controversies [e.g. O'Connell (1998), Sarno and Taylor (1998)]. In Bayesian framework, Li (1999) proposes a system of equations model with hierarchical priors to surmount the problems associated panel data unit root tests.

[2] Generally, the imposed cointegrating vector is either in accord with the strict version of PPP or is pre-estimated through a linear VECM.

PPP, multivariate models, especially nonlinear vector error correction models (VECM), can be more effective in capturing both the long run and short run dynamics of PPP adjustment. Perhaps the reason why researchers have not followed this route is due to 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), Seo (2006), Kapetanios, Shin and Snell (2006) for the latest developments in the nonlinear VECM tests].

This chapter proposes a Bayesian approach to investigate PPP within an exponential smooth transition VECM (ESTVECM) framework. Specifically, we follow the Bayesian cointegration space approach introduced by Strachan and Inder (2004) and the Bayesian logistic smooth transition Vector Autoregressive (LSTVAR) approach of Gefang and Strachan (2007).[3] Our method jointly captures the equilibrium and the presence of nonlinearity in the ESTVECM in a single step. Compared with the available classical estimation techniques which often require multiple steps and Taylor expansions, our approach is less susceptible to the sequential testing and inaccurate approximations 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 ESTVECM. Yet, jagged likelihood functions do not create any particular problems in our

[3]Their approach is based on the univariate smooth transition model estimation technique introduced by Bauwens, Lubrano and Richard (1999).

Gibbs sampling scheme.

In our empirical investigation, we analyze the validity of PPP between the US and the remaining six G7 countries over the post-Bretton Woods era. We take account of model uncertainty through Bayesian model selection and Bayesian model averaging. Following Koop and Potter (1999a), We use Bayes factors derived from the Savage-Dickey density ratio (SDDR) to calculate the posterior model probabilities. Here, different models are distinguished by the presence of the cointegration relationship, the order of the model, whether there exist nonlinear effects, and the transition variables which trigger the regime changes. Our estimation results strongly support that PPP holds, while the dynamics of the adjustment process to PPP is nonlinear. Furthermore, our results from the general impulse response functions show that the dynamics of the misalignment from PPP is determined by the sources and magnitudes of the original shocks.

The rest of the chapter is structured as follows. Section two introduces the ESTVECM model and Bayesian inferences. Section three reports empirical results. Section four concludes.

## 2.2 Exponential Smooth Transition VECM

Under the relative PPP, the nominal exchange rates and domestic and foreign prices should follow a cointegration relationship. However, as argued by Dumas (1992), among others, due to the presence of the transaction cost, the adjustment towards PPP should follow a nonlinear process, where small deviations from PPP are left uncorrected for the profit is not large enough to cover the transaction costs.

In the multivariate framework, this type of nonlinear adjustment can be captured by a threshold VECM (TVECM) or an ESTVECM. In a VECM the adjustment process induced by deviations from the long run equilibrium is a linear function of the magnitude of the deviations from that long run equilibrium. In contrast, in a TVECM or an ESTVECM, the dynamics of the adjustment process change across regimes, and the driving force of the regime changes is governed by the observed deviations from the equilibrium through the transaction function. In a TVECM, the regime changes are assumed to be discrete, whereas in an ESTVECM, the regimes change smoothly. Since the market force driving PPP adjustment is an aggregated process, following the suggestions of Teräsvirta (1994), we use an ESTVECM to model the nonlinear convergence towards PPP between two countries. ESTVECM appears to have another attractive property for it allows the same dynamics of regime changes for deviations above and below the equilibrium level.

Let $y_t = [s_t \;\; p_t \;\; p_t^*]$, where $s_t$, $p_t$, and $p_t^*$ are the logarithms of the foreign price of the domestic currency and respective domestic and foreign price levels. Assuming the cointegration relationships are common across different regimes, we model PPP in the exponential smooth transition VECM for t=1,...,T as follows.

$$\begin{aligned}\triangle y_t =& y_{t-1}\beta\alpha + d_t\xi + \Sigma_{h=1}^{p}\triangle y_{t-h}\Gamma_h \\ &+ F(z_t)(y_{t-1}\beta\alpha^z + d_t\xi^z + \Sigma_{h=1}^{p}\triangle y_{t-h}\Gamma_h^z) + \varepsilon_t\end{aligned} \tag{2.1}$$

where $\triangle y_t = y_t - y_{t-1}$. The error term $\varepsilon_t$ 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 $s \neq t$. The dimensions of $\Gamma_h$ and $\Gamma_h^z$ are $3 \times 3$, and the dimensions of $\beta$, $\alpha'$, and $\alpha^{z'}$ are $3 \times r$, where $r$ is the rank of the cointegration space, with $r$ can be either 0 or 1. If PPP holds, the value of $r$ should be equal to 1.

In model (2.1), the regime changes are assumed to be caused by a past deviation from the equilibrium relationship, and the dynamics of the regime changes is captured by the symmetric **U** shaped exponential smooth transition function proposed by Teräsvirta (1994):

$$F(z_t) = 1 - exp(-\gamma(z_t - c)^2) \tag{2.2}$$

where the transition variable $z_t = y_{t-d}\beta$ is the cointegrating combination among $s$, $p$, and $p^*$ at period $t - d$;[4] $c$ is the equilibrium level of the cointegrating relationship, also the threshold around which the regime changes; $\gamma$ is the smooth parameter that governs the speed of the transition process between extreme regimes, with higher values of $\gamma$ implying faster transition.

[4]Note that the driving force of the regime changes can be any exogenous or endogenous variables of concern. In this study, we only examine the nonlinear effects caused by the misalignments from PPP.

The transition function $F(z_t)$ is bounded by 0 and 1. It is seen that $F(z_t) = 0$ when $z_t - c = 0$, and $F(z_t) = 1$ when $z_t - c \to \pm\infty$. As convention, we define $F(z_t) = 0$ and $F(z_t) = 1$ corresponding to the middle and outer regimes, respectively. In the middle regime, model (2.1) becomes a linear VECM, with the adjustment process governed by $(\alpha, \xi, \Gamma_h)$; while in the outer regime, model (2.1) becomes a different linear VECM, where the dynamics of the model are determined by $(\alpha + \alpha^z, \xi + \xi^z, \Gamma_h + \Gamma_h^z)$. Between the two extreme regimes, the speed of PPP adjustment is determined by the deviations from the equilibrium. For small deviations from PPP, the model is more dependent on the parameters of the middle regime. Once the deviations get larger, the adjustment process will be more influenced by the parameters in the outer regime.

Finally note that equation (2.1) allows a set of models which vary in the rank of the cointegration vector (0 or 1), the order of the autoregressive process, the lag length of the transition variable, and the presence of the nonlinearity.

### 2.2.1 The Likelihood Function

For notational convenience, we can re-write model (2.1) as

$$\triangle y_t = x_{1,t-1}\beta\alpha + x_{2,t}\Phi + F(z_t)(x_{1,t-1}\beta\alpha^z + x_{2,t}\Phi^z) + \varepsilon_t \qquad (2.3)$$

where $x_{1,t-1} = y_{t-1}$, $x_{2,t} = (d_t, \triangle y_{t-1}, ..., \triangle y_{t-p})$, $\Phi = (\xi', \Gamma'_1, ..., \Gamma'_p)'$, $\Phi^z = (\xi^{z'}, \Gamma^{z'}_1, ..., \Gamma^{z'}_p)'$.

To simplify the notation, we first define the $T \times n$ matrix $X_0 = (\triangle y'_1, \triangle y'_2, ..., \triangle y'_T)'$ and $T \times 2(r + 2 + np)$ matrix $X = (X_1\beta \;\; X_2 \;\; F^z X_1\beta \;\; F^z X_2)$, where $X_1 = (x'_{1,1}, \;\; x'_{1,2}, \;\; ..., \;\; x'_{1,T})'$, $X_2 = (x'_{2,1}, \;\; x'_{2,2}, \;\; ..., \;\; x'_{2,T})'$, and $F^z = diag(F(z_1), F(z_2), ..., F(z_T))$. Next, we set $B = (\alpha' \;\; \Phi' \;\; \alpha^{z'} \;\; \Phi^{z'})'$. Finally, we stack the error terms $\varepsilon_t$ in the $T \times n$ matrix E, where $E = (\varepsilon'_1, \varepsilon'_2, ..., \varepsilon'_T)'$. Hence, model (2.1) can be written as

$$X_0 = X_1\beta\alpha + X_2\Phi + F^z X_1\beta\alpha^z + F^z X_2\Phi^z + E = XB + E \tag{2.4}$$

The likelihood function of model (2.4) is following.

$$L(y|\Sigma, B, \beta) \propto |\Sigma|^{-\frac{T}{2}} exp\{-\frac{1}{2} tr\Sigma^{-1}E'E\} \tag{2.5}$$

Vectorizing model (2.4), we transform model (2.1) into

$$x_0 = xb + e \tag{2.6}$$

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

$$\begin{aligned} tr\Sigma^{-1}E'E &= e'(\Sigma^{-1} \otimes I_T)e \\ &= s^2 + (b - \widehat{b})'V^{-1}(b - \widehat{b}) \end{aligned} \tag{2.7}$$

where $s^2 = x_0' M_v x_0$, $M_v = \Sigma^{-1} \otimes [I_T - X(X'X)^{-1}X']$, $\widehat{b} = [I_n \otimes (X'X)^{-1}X']x_0$ and $V = \Sigma \otimes (X'X)^{-1}$. Thus, the likelihood function in equation (2.5) can be re-written as

$$L(y|\Sigma, B, \beta) \propto |\Sigma|^{-\frac{T}{2}} exp\{-\frac{1}{2}[s^2 + (b - \widehat{b})'V^{-1}(b - \widehat{b})]\} \qquad (2.8)$$

### 2.2.2 Priors

Although the strict version of PPP states that the combination $s_t + p_t - p^*$ should be stationary, there is no theoretical guidelines to specify the values of $\beta$ in the cointegration relationship for the relative PPP. Furthermore, it is impossible to impose meaningful informative priors for the coefficients of the long run/short run adjustment in the VECM nor for the parameter that indicates the speed of regime changes in the transition function. Therefore, we use the 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, *a priori*, equally likely.

Before eliciting our priors of the parameters, it is worthwhile to stress the identification problems in our model setting. Note that both a linear VECM and a simple smooth transition VAR model suffer from the identification problems. As well documented in the literature, a linear VECM suffers from both the global and local nonidentification of the cointegration vectors and parameters corresponding to the long-run adjustment. In Bayesian literature, a great effort has been made to surmount this problem. In earlier research, to set uninformative prior for the cointegration vector $\beta$, researchers first normalize $\beta$ into $\beta = [I_r \;\; V']'$, then impose uninformative prior on the sub-vector $V$. However, as argued by Strachan and van Dijk (2004a), this approach has an undesirable side-effect that it favors 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 [e.g. Strachan (2003), Strachan and Inder (2004), Villani (2005)]. As pointed out by Koop, Strachan, van Dijk and Villani (2006) in their survey on the Bayesian approaches to cointegration, 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 regards to the smooth transition part of the model, as explained by 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 Bayesian context than in the classical framework. Bauwens, Lubrano and Richard (1999) and Lubrano (1999a, 1999b) introduce a number of prior settings to solve the problem. Following Gefang and Strachan (2007), we tackle this problem by simply setting the prior distribution of $\gamma$ as Gamma.

The nonidentification problem faced by the ESTVECM model is slightly different. Although the Davies' problem remains relatively the same as in a smooth transition VAR, the problem in identifying the cointegration vector and its adjustment parameters is subject to the additional influence from the transition parameter. Here the cointegration vector comes forth in three combinations, namely $\beta\alpha$, $\beta\gamma$ and $\beta\alpha^z$. However, this difference does not render the identification problem more complicated than what we have to deal with in a single linear VECM and a LSTVAR. As long as we can rule out the possibility that $\gamma = 0$, we can identify $\beta$, $\alpha$, $\alpha^z$ and $\gamma$ sequentially once we choose a way to normalize $\beta$.

Following the arguments of Koop, Strachan, van Dijk and Villani (2006), we elicit the prior of $\beta$ indirectly from the prior expressed upon the cointegration space.[5] While we adopt the general approach developed in Strachan and Inder (2004), we diverge in two aspects important for this application. First, as is standard in the cointegration analysis of PPP (as there is considerable empirical evidence and theoretical support for this restriction), we only consider a single equilibrium cointegrating relationship. Second, we restrict ourselves to the economically justifiable region of the cointegrating space where the signs of the elements in the cointegration vector are [+ + -] or [- - +]. Therefore, our method does not explore if there are any other long run equilibrium relationships different from PPP. Having restricted the support to this region, the prior is otherwise uninformative on this space.

Specifically, we set the prior of $\beta$ as following. First, we specify the

[5]To our knowledge, in literature, only Sugita (2006) applies the Strachan and Inder (2004) methodology in defining the prior density for cointegrating vector in a nonlinear VECM. In his model the regimes changes are assumed to follow a Markov switching process.

space of the three by one vector $\beta$ to be uniformly distributed over the two dimensional Grassman manifold $G_{1,2}$.

$$p(\beta) = \frac{1}{c_1^3} \tag{2.9}$$

where $c_1^3 = \int_{G_{1,2}} dg_1^3$ is a constant, which is the volume of the compact space $G_{1,2}$ [James (1954), Muirhead (1982)]. Next, we restrict $\beta'\beta = 1$ for the purpose of identification as the normalization method does not distort the distribution of the cointegration space (see Strachan and Inder, 2004 for further explanation). Thus, we can use polar coordinates to denote the semi-orthogonal $\beta$ as follows:

$$\beta = [sin(\theta_1)sin(\theta_2) \quad sin(\theta_1)cos(\theta_2) \quad cos(\theta_1)]'$$

To describe the uniform distribution of the cointegration space in polar coordinates, we multiply the uninformative prior of the space of $\beta$ in (2.9) by $sin(\theta_1)$, the Jacobian of the transformation from rectangular coordinates to polar coordinates [Muirhead (1982), p55]. As explained before, we want to restrict the signs of the elements in $\beta$ instead of allowing the space of $\beta$ to move freely. The objective can be easily achieved by restricting the range of $\theta_1$ to be from $\pi/2$ to $\pi$, and the range of $\theta_2$ to be from 0 to $\pi/2$.

With regards to the variance covariance matrix of the error terms, following Zellner (1971), we set standard diffuse prior for $\Sigma$:

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

.

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 Bayes factors well defined, we are not allowed to set flat prior for $b$ [see Bartlett (1957), O'Hagan (1995) for details]. Therefore, following Strachan and van Dijk (2006), we set weakly informative conditional proper prior for $b$ as:

$$P(b|\Sigma, \theta_1, \theta_2, \gamma, c, M_\omega) \propto N(0, \underline{V})$$

where $b = vec(B)$, $\underline{V} = \Sigma \otimes \eta^{-1} I_k$, $k = 2(r+1+np)$, $\eta$ is the shrinkage prior as proposed by Ni and Sun (2003). As practiced in Koop, Leon-Gonzalez and Strachan (2006), we draw $\eta$ from the Gibbs sampler. In our case, we set the relatively uninformative prior distribution of $\eta$ as Gamma with mean $\underline{\mu_\eta}$, and degrees of freedom $\underline{\nu_\eta}$, where $\underline{\mu_\eta}$=10, $\underline{\nu_\eta}$=0.0001. Note that in our prior setting, the conditional weakly informative priors for $\alpha$ and $\alpha^z$ are the same, which are normal with zero mean and covariance matrix $\Sigma \otimes (\beta' \eta I_3 \beta)^{-1}$.

To avoid the Davies' problem in the nuisance parameter space, following Gefang and Strachan (2007), we set the prior distribution for $\gamma$ as Gamma, which exclude *a priori* the point $\gamma = 0$ from the integration range. Since the nonlinear part of $b$ can still be a vector of 0s as $\gamma > 0$, the prior specification of $\gamma$ does not render model (2.1) in favor of the nonlinear effect. In the empirical work, we set the prior distribution of $\gamma$ as Gamma(1, 0.0001), which is relatively uninformative.

Finally, to interpret our results more sensibly, we elicit the conditional prior of $c$ as uniformly distributed between the upper and lower limits of the middle 80% of the transition variables (which, in our case, is the product of $[s_{t-d} \;\; p_{t-d} \;\; p^*_{t-d}]$ and the cointegrating vector $\beta$). Note that the bounds of the support for $c$ are both data dependent and dependent upon $\beta$.

### 2.2.3 Posterior Computation

We use full conditional Gibbs sampler for posterior computations. From the priors just elicited and likelihood function derived in Section 2.1, we find that the posterior of $\Sigma$ is Inverted Wishart (IW) with scale matrix $E'E$, and the degrees of freedom $T$, while the conditional posterior of $b$ is Normal with mean $\bar{b} = vec[(X'X + \eta I_k)^{-1}X'X_0]$ and covariance matrix $\overline{V} = \Sigma \otimes (X'X + \eta I_k)^{-1}$. Note that the posterior distributions of $\theta_1$, $\theta_2$, $\gamma$ and $c$ are not of any standard form. However, the ranges of $\theta_1$ and $\theta_2$ are restricted as explained in the previous section, and in each run of the Gibbs sampler, the range of $c$ can be predetermined based on the current draws of $\theta_1$ and $\theta_2$. Thus, we can use Griddy Gibbs Sampling introduced in Ritter and Tanner (1992) to draw $\theta_1$, $\theta_2$ and $c$ within the main Gibbs Sampler. With respect to $\gamma$, we resort to Metropolis-Hastings algorithms [Chib and

Greenberg (1995)] within Gibbs for the estimation. In order to carry out all the aforementioned posterior analysis, we need to know the posterior of $\eta$ as well. The conditional posterior of $\eta$ is calculated as

$$p(\eta|B,\Sigma,\gamma,c,Y,X) \propto p(\eta)|\Sigma^{-1}\otimes\eta I_k|^{\frac{1}{2}} exp\left\{-\frac{1}{2}b'(\Sigma^{-1}\otimes\eta I_k)b\right\} \quad (2.10)$$

which indicates that the conditional posterior of $\eta$ is distributed as Gamma with the mean $\overline{\mu_\eta} = \frac{\overline{\nu_\eta}\underline{\mu_\eta}}{\underline{\nu_\eta}+\underline{\mu_\eta}tr(B'B\Sigma^{-1})}$, and the degrees of freedom $\overline{\nu_\eta} = nk + \underline{\nu_\eta}$.

The Gibbs Sampling Scheme can be summarized as follows:

1. Initialize $(b,\Sigma,\theta_1,\theta_2,\gamma,c,\eta)$;
2. Draw $\Sigma|b,\theta_1,\theta_2,\gamma,c,\eta$ from $IW(E'E,T)$;
3. Draw $b|\Sigma,\theta_1,\theta_2,\gamma,c,\eta$ from $N(\bar{b},\overline{V}_b)$;
4. Draw $\theta_1,\theta_2|\Sigma,b,\gamma,c,\eta$ numerically by Griddy Gibbs;
5. Draw $\gamma|\Sigma,b,\theta_1,\theta_2,c,\eta$ through Metropolis-Hastings method;
6. Draw $c|b,\Sigma,\theta_1,\theta_2,\gamma,\eta$ numerically by Griddy Gibbs;
7. Draw $\eta|b,\Sigma,\theta_1,\theta_2,\gamma,c$ from $G(\overline{\mu_\eta},\overline{\nu_\eta})$;
8. Repeat steps 2 to 7 for a suitable number of replications.

In case the draws from Metropolis-Hastings simulator get stuck in a local mode, we try different starting values for the sampler.

One of the main concerns of our study is to examine the posterior probabilities of different possible models and trace the effects of cointegration and nonlinearity. For this purpose, we resort to the SDDR approach of Koop and Potter (1999a) to calculate the Bayes factors.

As explained in Koop and Potter (1999a), by penalizing parameter rich models, using Bayes factors to calculate posterior odds ratio can resolve the over fitting problems that generally exist in nonlinear models. Following Koop and Potter (1999a) and Koop, Leon-Gonzalez and Strachan (2006), we use SDDR to compute Bayes factors comparing every restricted model nested within the general model (2.1) with the general model itself. Using this information, we back out the posterior model probabilities for each country pair through a base model (e.g. the model where all the parameters in $b$ are restricted to be zero). Note that the restricted linear VECM model occurs when all the elements of $\alpha^z$ and $\Phi^z$ are equal to zero. Likewise, the restricted linear VAR model with neither the cointgration nor the nonlinear effect occurs when we impose all the elements of $\alpha$, $\alpha^z$ and $\Phi^z$ to be equal to zeros.[6] Hence, we can use the conditional posterior distribution and the conditional priors of $b$ to compute the Bayes factor for the restricted model $M_1$ (nested in model $M_2$) versus the unrestricted model $M_2$ using the SDDR

[6]It is important to stress that as explained by Koop, Leon-Gonzalez and Strachan (2006), in the linear VECM model, the rank of the cointegration relationship equal to zero if and only if $\alpha = 0$.

which is given by the expression:

$$B_{1,2} = \frac{Pr(M_1|y)}{Pr(M_2|y)} = \frac{p(b|M_2, y)|_{b_i=0}}{p(b|M_2)|_{b_i=0}}$$

where the restrictions are $b_i = 0$. Note that this method penalizes parameter rich models as explained in Koop and Potter (1999a).

## 2.3 Empirical Results

In this section, we investigate whether PPP holds between the US and the other six G7 countries—Canada (CAN), France (FRA), Germany (GER), Italy (ITA), Japan (JAP), and the UK. In all cases, the US is considered the foreign country. We extract monthly nominal exchange rates and consumer price index (CPI) series from the International Financial Statistics database. For Canada, Japan and the UK, the data span the period 1973:1 to 2006:12. For France and Italy, the sample period covers from 1973:1 until the fixing of the Euro conversion rate 1998:12. For Germany, we use the former West Germany data running from 1973:1 to 1991:12.

The 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 diagnostic. The parameter estimates are presented in tables 2.1 - 2.3. Given the large amount of parameters being estimated, we only report the estimation results for the cointegration relationship, the threshold, and the smooth variable which indicate the speed of the regime changes. To aid comprehension, both the angles in the polar coordinates and corresponding elements in the cointegration vector are reported. [7]

### 2.3.1 Model Comparison Results

In this section, we report results relating to the posterior model probabilities of 85 different models [namely 1 model with only the error terms, 6 linear VARs, 6 linear VECMs, 36 exponential smooth transition VAR models (ESTVAR) and 36 ESTVECM models] for each country pair.[8] Among these models, both the maximal order of the autoregressive process and the longest lag length of the transition indicator are allowed to be 6. We assume the 85 models are exhaustive and mutually independent.

Table 2.4 summarizes the total posterior probabilities of the models. In all cases, ESTVECM models receive over 90% of the posterior model probabilities, which provides strong eveidence that PPP holds, and the adjustment process towards PPP is nonlinear.This finding suggests that it is improper to model the interrelationship among the nominal interest rate and domestic

[7] We identify the parameters by normalizing $\beta'\beta = 1$; Linear identification can be achieved by first dividing the reported $\beta$ by one of its element that of concern, then transform the reported $\gamma$ and $c$ accordingly.

[8] Note that by imposing restrictions on the long-run adjustment parameters $\alpha$ and $\alpha^z$ in the unrestricted model, the linear and nonlinear VAR models we considered are in differences.

and foreign price levels in a linear framework.

It may also be illuminating to look into the support for the VECM and VAR models in the linear context. It is seen that except for the case of US-ITA, the linear VECM models are more favored over linear VAR models in all countries pairs.

Table 2.5 contains results of the sum of the posterior probabilities of the ESTVECM models distinguished by the transition variables for each country pair. The model comparison results show that the transition indicators with longer lag lengths are generally preferred over the shorter ones in modeling the nonlinear effects. Given the time lags between the contract and settlement in international trade, this result is not surprising. However, in the case of US-FRA, it turns out that the most preferred lag length of the transition indicator is 2, and it receives nearly 100% of the posterior mass.

To shed more light on the properties of the posterior probabilities, we report the individual top 20 models for each country pairs in tables 2.6 - 2.7. Observe that, for all country pairs, the top 20 models account for more than 99% of the total posterior mass. However, the degree of model uncertainties are rather different across country pairs. In the case of US-FRA, with the single most preferred model obtains 84.78% of the posterior model probabilities, more than 99% of the posterior model probabilities are taken by the top six models. While in others cases, although a great majority of the posterior mass is also taken by the top six models, the posterior model probabilities tend to spread across the six models more evenly. For example, in the case of US-CAN, each of the top six models accounts for relatively 14% of the posterior model probabilities; In the case of US-GER, the posterior model probabilities of the top six models range from 8.17% to 21.27%. The most obvious case of model uncertainty can be found in the US-UK pair. For US-UK, 96.89% of the posterior mass is scattered across twelve models, with their posterior models probabilities ranging from 4.85% to 12.93%.

We report the time profiles of the smooth transition functions of the most probable models in figure 2.1. Observe that throughout the years, in the cases of US-CAN and US-UK, the dynamics of regime changes is gradually switching from the outer regimes towards the middle regimes. In the cases of US-FRA and US-ITA, we observe U-shaped time profiles, with the former hit the middle regime in November, 1980, and the latter hit the middle regime in August, 1980. In the case of US-GER, the dynamics of the PPP adjustment remains very close to the middle regime. In contrast, for US-JAP, the dynamics of the regime changes is in the outer regime during most of the time. The graphs show that the regime switching processes are rather smooth for all cases, thus it is improper to adopt an abrupt function to model the nonlinear effects.

### 2.3.2 Impulse Response Analysis

It is acknowledged that the impulse response functions of the nonlinear models are history- and shock- dependent [e.g. Potter (1994), Koop, Pesaran and Potter (1996)]. We use the generalized impulse response function proposed in Koop, Pesaran and Potter (1996) to examine the effect of a shock on the PPP relationship. In particular, we examine the generalized impulse response functions of $GI_P$ for a shock, $\upsilon_t$, and a history, $\omega_{t-1}$ as follows

$$GI_P(n, \upsilon_t, \omega_{t-1}) = E[P_{t+n}|\upsilon_t, \omega_{t-1}] - E[P_{t+n}|\omega_{t-1}] \tag{2.11}$$

where $n$ is the time horizon. By averaging out the future shocks, in (2.11), we treat the impulse responses as an average of what might happen given what has happened. Using Bayesian approach, we calculate the generalized impulse responses by averaging out the history uncertainties, the future uncertainties, the parameter uncertainties and model uncertainties.

To examine the impulse response functions of the cointegrating PPP combination, we allow a shock amounting to $\pm 0.01$ and $\pm 0.02$, respectively, to hit each of the three variables (namely $s_t$, $p_t$ and $p_t^*$). The time horizon of the impulse responses is set to 60 months. Note that for each country pair, we have 85 models and 12 different shocks for model comparison. For brevity, we only present the impulse response functions of the PPP combinations for the most preferred models in figures 2.2 - 2.4. Inspecting the impulse response functions, we have two main findings.

i. The dynamics of PPP deviations are determined by the sources and magnitudes of the initial shocks that hit $s_t$, $p_t$ and $p_t^*$.

ii. Deviations from PPP are mean-reverting in the next 5 years in all cases except for when shocks are originated from Canada's price levels.

However, the types of convergence processes are rather different across different cases.

Our findings in the impulse response functions of PPP might shed some light on the discussions regarding the half life of PPP adjustment.[9] As shown in our study, the impacts on PPP relationship varies with the sources and magnitudes of the initial shocks hitting $s_t$, $p_t$ and $p_t^*$. In the cointegrating context, an amount of deviation from PPP can be traced to a myriad combinations of initial shocks that hit $s_t$, $p_t$ and $p_t^*$. Hence, we suggest that any assertions on the speed of PPP convergence which neglect the causes of the deviation can be misleading.[10]

## 2.4 Conclusion

In this chapter, we introduce a Bayesian approach to estimating an ESTVECM model to investigate whether purchasing power parity holds between the US and the other six G7 countries. The model comparison results are in accord with the theoretical assertion that in the long run PPP holds, and the adjustment to PPP is a nonlinear process with the regime changes governed by the magnitude of deviations from the long-run PPP equilibrium. Furthermore, our research casts doubt over the practice of estimating the half life of PPP

[9] The half life estimates has been extensively used in the literature to indicate the speed of PPP adjustment on real exchange rates [e.g. Cheung and Lai (1994), Lothian and Taylor (1996), Lopez, Murray and Papell (2005)].

[10] Chortareas and Kapetanios (2005) also claim that using the half life measure to analyze PPP adjustment might be problematic. However, their reasoning are different from ours.

deviations. The analysis of the impulse response functions show that the mean-reverting process of the PPP misalignment can be rather complex.

Table 2.1: Parameters (a)

| | d=1 | d=2 | d=3 | d=4 | d=5 | d=6 |
|---|---|---|---|---|---|---|
| **US - CAN** | | | | | | |
| $\gamma$ | 12.6520 | 13.7470 | 13.3550 | 15.5270 | 13.8080 | 13.9260 |
| | *3.5105* | *4.1453* | *4.0240* | *5.2090* | *4.0288* | *3.4595* |
| $\theta_1$ | 2.3150 | 2.0431 | 2.0648 | 1.9085 | 1.9747 | 1.9319 |
| | *0.0537* | *0.0865* | *0.0685* | *0.0790* | *0.0930* | *0.0373* |
| $\theta_2$ | 1.4372 | 1.5365 | 1.4857 | 1.5337 | 1.5171 | 1.5283 |
| | *0.0805* | *0.0246* | *0.0519* | *0.0316* | *0.0463* | *0.0309* |
| $c$ | -0.2820 | 1.3063 | 1.3917 | 2.2738 | 1.8870 | 2.1457 |
| | *0.4604* | *0.5750* | *0.4554* | *0.5565* | *0.5909* | *0.2492* |
| $\beta_1$ | 0.7291 | 0.8900 | 0.8773 | 0.9429 | 0.9182 | 0.9346 |
| $\beta_2$ | 0.0980 | 0.0305 | 0.0749 | 0.0350 | 0.0493 | 0.0398 |
| $\beta_3$ | -0.6774 | -0.4550 | -0.4742 | -0.3313 | -0.3930 | -0.3534 |
| **US - FRA** | | | | | | |
| | d=1 | d=2 | d=3 | d=4 | d=5 | d=6 |
| $\gamma$ | 30.4830 | 28.7710 | 0.1131 | 25.9230 | 0.0339 | 0.0382 |
| | *4.8129* | *4.3462* | *0.3682* | *6.5391* | *0.1182* | *0.1242* |
| $\theta_1$ | 3.0893 | 3.0240 | 2.2722 | 3.0322 | 2.2540 | 2.3632 |
| | *0.0758* | *0.0498* | *0.0480* | *0.0954* | *0.0846* | *0.0698* |
| $\theta_2$ | 1.0376 | 1.0875 | 0.4227 | 1.0655 | 0.5105 | 0.6029 |
| | *0.4789* | *0.3591* | *0.1585* | *0.4284* | *0.1852* | *0.1697* |
| $c$ | -4.9709 | -4.4999 | 1.7649 | -4.5704 | 1.9835 | 1.1213 |
| | *0.4376* | *0.3262* | *0.4055* | *0.5818* | *0.6248* | *0.6242* |
| $\beta_1$ | 0.0450 | 0.1039 | 0.3134 | 0.0956 | 0.3790 | 0.3981 |
| $\beta_2$ | 0.0266 | 0.0545 | 0.6967 | 0.0529 | 0.6767 | 0.5783 |
| $\beta_3$ | -0.9986 | -0.9931 | -0.6453 | -0.9940 | -0.6313 | -0.7121 |

Notes: Standard deviations are in italics.

Table 2.2: Parameters (b)

| | d=1 | d=2 | d=3 | d=4 | d=5 | d=6 |
|---|---|---|---|---|---|---|
| **US - GER** | | | | | | |
| $\gamma$ | 0.0136 | 0.9400 | 0.6267 | 9.0952 | 0.0517 | 14.4090 |
| | *0.0653* | *1.8586* | *1.1395* | *1.6984* | *0.1453* | *2.9474* |
| $\theta_1$ | 1.8754 | 1.8200 | 1.9436 | 1.8318 | 1.9217 | 1.9702 |
| | *0.0499* | *0.0609* | *0.0114* | *0.1280* | *0.0355* | *0.0033* |
| $\theta_2$ | 0.0186 | 0.0180 | 0.0130 | 0.0365 | 0.0139 | 0.0384 |
| | *0.0109* | *0.0112* | *0.0057* | *0.0347* | *0.0074* | *0.0255* |
| $c$ | 3.2665 | 3.6215 | 2.7841 | 3.5507 | 2.9426 | 2.6839 |
| | *0.3401* | *0.3841* | *0.0835* | *0.8344* | *0.2554* | *0.1183* |
| $\beta_1$ | 0.0177 | 0.0174 | 0.0121 | 0.0353 | 0.0131 | 0.0353 |
| $\beta_2$ | 0.9538 | 0.9690 | 0.9312 | 0.9655 | 0.9390 | 0.9206 |
| $\beta_3$ | -0.2999 | -0.2466 | -0.3642 | -0.2580 | -0.3438 | -0.3889 |
| **US - ITA** | | | | | | |
| | d=1 | d=2 | d=3 | d=4 | d=5 | d=6 |
| $\gamma$ | 4.3500 | 7.8474 | 42.2050 | 6.4511 | 30.5660 | 3.4249 |
| | *2.9948* | *3.8444* | *8.1170* | *2.2292* | *2.0644* | *1.7186* |
| $\theta_1$ | 2.4866 | 2.4649 | 3.0858 | 2.5294 | 3.0539 | 2.4769 |
| | *0.0076* | *0.0033* | *0.0451* | *0.0048* | *0.0820* | *0.0187* |
| $\theta_2$ | 0.0615 | 0.0487 | 0.8980 | 0.0248 | 1.1982 | 0.0339 |
| | *0.0590* | *0.0409* | *0.4882* | *0.0158* | *0.3517* | 0.0362 |
| $c$ | -0.3737 | -0.2265 | -4.8580 | -0.8063 | -4.6900 | -0.3258 |
| | *0.0732* | *0.1011* | *0.3030* | *0.0546* | *0.5221* | *0.1963* |
| $\beta_1$ | 0.0374 | 0.0305 | 0.0436 | 0.0143 | 0.0816 | 0.0209 |
| $\beta_2$ | 0.6080 | 0.6255 | 0.0347 | 0.5745 | 0.0319 | 0.6165 |
| $\beta_3$ | -0.7931 | -0.7797 | -0.9985 | -0.8184 | -0.9962 | -0.7871 |

Notes: Standard deviations are in italics.

Table 2.3: Parameters (c)

| | | | US - JAP | | | |
|---|---|---|---|---|---|---|
| | d=1 | d=2 | d=3 | d=4 | d=5 | d=6 |
| $\gamma$ | 22.7700 | 34.7190 | 85.1300 | 15.2390 | 1.5974 | 48.3000 |
| | *32.2510* | *7.3199* | *59.0440* | *6.5107* | *2.9569* | *11.0160* |
| $\theta_1$ | 2.6538 | 1.9233 | 1.8639 | 2.1151 | 2.3533 | 1.8933 |
| | *0.0856* | *0.0206* | *0.0517* | *0.0019* | *0.1240* | *0.0497* |
| $\theta_2$ | 0.7510 | 1.5155 | 0.9572 | 0.0763 | 0.3485 | 1.5201 |
| | *0.2628* | *0.0460* | *0.6698* | *0.0147* | *0.2272* | *0.0568* |
| $c$ | -1.4580 | 3.0393 | 3.7788 | 2.1029 | 0.7404 | 3.3569 |
| | *0.6592* | *0.2016* | *0.5806* | *0.0610* | *0.7972* | *0.4702* |
| $\beta_1$ | 0.3198 | 0.9371 | 0.7827 | 0.0652 | 0.2422 | 0.9472 |
| $\beta_2$ | 0.3426 | 0.0519 | 0.5513 | 0.8530 | 0.6665 | 0.0481 |
| $\beta_3$ | -0.8834 | -0.3452 | -0.2889 | -0.5178 | -0.7051 | -0.3170 |
| | | | US - UK | | | |
| | d=1 | d=2 | d=3 | d=4 | d=5 | d=6 |
| $\gamma$ | 12.6260 | 16.8620 | 17.0540 | 15.2000 | 14.2380 | 14.5250 |
| | *5.0581* | *6.9726* | *5.2787* | *5.2428* | *4.5599* | *6.1151* |
| $\theta_1$ | 2.0099 | 2.3977 | 2.3327 | 2.3090 | 1.7875 | 2.0869 |
| | *0.2439* | *0.0288* | *0.0596* | *0.0645* | *0.0891* | *0.2538* |
| $\theta_2$ | 1.2821 | 0.8151 | 0.9159 | 0.9517 | 1.5146 | 1.2118 |
| | *0.2369* | *0.0489* | *0.0901* | *0.0895* | *0.0741* | *0.2544* |
| $c$ | 2.4621 | 0.7374 | 1.1670 | 1.3126 | 3.1515 | 2.1100 |
| | *0.8302* | *0.2304* | *0.3802* | *0.3856* | *0.2692* | *0.9630* |
| $\beta_1$ | 0.8677 | 0.4929 | 0.5739 | 0.6024 | 0.9751 | 0.8143 |
| $\beta_2$ | 0.2577 | 0.4644 | 0.4407 | 0.4293 | 0.0549 | 0.3055 |
| $\beta_3$ | -0.4251 | -0.7358 | -0.6903 | -0.6729 | -0.2150 | -0.4935 |

Notes: Standard deviations are in italics.

Table 2.4: Summarized Posterior Model Probabilities

| Model | US - CAN | US - FRA | US - GER | US - ITA | US - JAP | US - UK |
|---|---|---|---|---|---|---|
| ESVECM | 0.9451 | 0.9997 | 0.9442 | 0.9179 | 0.9989 | 0.9917 |
| ESVAR | 0.0549 | 0.0003 | 0.0548 | 0.0821 | 0.0011 | 0.0083 |
| Top 20 models | 0.9911 | 1.0000 | 0.9898 | 1.0000 | 1.0000 | 0.9934 |
| VECM/VAR* | 8.0570 | 4.4462 | 10108.5000 | 0.7355 | 3.1522 | 118.7717 |

Notes: * reports the ratio of the total posterior probabilities of the linear VECM models to that of the linear VAR models.

Table 2.5: Posterior Model Probabilities of the Transition Variables

| Model | US - CAN | US - FRA | US - GER | US - ITA | US - JAP | US - UK |
|---|---|---|---|---|---|---|
| p=1 | 0.0000 | 0.0040 | 0.0020 | 0.0000 | 0.0000 | 0.0002 |
| p=2 | 0.0000 | 0.9960 | 0.0021 | 0.0000 | 0.0000 | 0.0237 |
| p=3 | 0.0000 | 0.0000 | 0.0020 | 0.2501 | 0.0000 | 0.3377 |
| p=4 | 0.0117 | 0.0000 | 0.1228 | 0.0000 | 0.0000 | 0.6385 |
| p=5 | 0.1058 | 0.0000 | 0.0020 | 0.7499 | 0.0000 | 0.0000 |
| p=6 | 0.8825 | 0.0000 | 0.8691 | 0.0000 | 1.0000 | 0.0000 |

Notes: $d$ is the lag length of the transition variable.

Table 2.6: Top 20 Most Preferred Models (a)

| No. | US - CAN | | US - FRA | | US - GER | |
|---|---|---|---|---|---|---|
| 1 | $ESVECM(6)_{p=6}$ | 0.1415 | $ESVECM(6)_{p=2}$ | 0.8478 | $ESVECM(6)_{p=6}$ | 0.2128 |
| 2 | $ESVECM(5)_{p=6}$ | 0.1413 | $ESVECM(3)_{p=2}$ | 0.0417 | $ESVECM(3)_{P=6}$ | 0.1824 |
| 3 | $ESVECM(4)_{p=6}$ | 0.1402 | $ESVECM(2)_{p=2}$ | 0.0350 | $ESVECM(4)_{p=6}$ | 0.1468 |
| 4 | $ESVECM(2)_{p=6}$ | 0.1395 | $ESVECM(4)_{p=2}$ | 0.0266 | $ESVECM(2)_{p=6}$ | 0.1074 |
| 5 | $ESVECM(3)_{p=6}$ | 0.1390 | $ESVECM(1)_{p=2}$ | 0.0226 | $ESVECM(1)_{P=6}$ | 0.1071 |
| 6 | $ESVECM(1)_{p=6}$ | 0.1356 | $ESVECM(5)_{p=2}$ | 0.0220 | $ESVECM(5)_{p=6}$ | 0.0817 |
| 7 | $ESVAR(6)_{p=6}$ | 0.0377 | $ESVECM(6)_{p=1}$ | 0.0035 | $ESVAR(1)_{p=6}$ | 0.0211 |
| 8 | $ESVECM(5)_{p=5}$ | 0.0164 | $ESVECM(3)_{P=1}$ | 0.0002 | $ESVECM(6)_{p=4}$ | 0.0201 |
| 9 | $ESVECMb(5)_{p=5}$ | 0.0164 | $ESVECM(4)_{P=1}$ | 0.0001 | $ESVECM(4)_{p=4}$ | 0.0178 |
| 10 | $ESVECM(4)_{p=5}$ | 0.0163 | $ESVECM(5)_{P=1}$ | 0.0001 | $ESVECM(3)_{P=4}$ | 0.0176 |
| 11 | $ESVECM(2)_{p=5}$ | 0.0162 | $ESVAR(5)_{p=2}$ | 0.0001 | $ESVECM(5)_{p=4}$ | 0.0159 |
| 12 | $ESVECM(3)_{p=5}$ | 0.0162 | $ESVAR(4)_{p=2}$ | 0.0001 | $ESVECM(1)_{P=4}$ | 0.0153 |
| 13 | $ESVECM(1)_{p=5}$ | 0.0158 | $ESVECM(2)_{p=1}$ | 0.0001 | $ESVECM(2)_{p=4}$ | 0.0152 |
| 14 | $ESVAR(1)_{p=6}$ | 0.0042 | $ESVAR(3)_{p=2}$ | 0.0001 | $ESVAR(1)_{p=4}$ | 0.0127 |
| 15 | $ESVAR(1)_{p=5}$ | 0.0039 | $ESVECM(6)_{p=4}$ | 0.0000 | $ESVAR(6)_{p=6}$ | 0.0037 |
| 16 | $ESVAR(6)_{p=5}$ | 0.0029 | $ESVECM(1)_{P=1}$ | 0.0000 | $ESVAR(3)_{p=6}$ | 0.0033 |
| 17 | $ESVAR(5)_{p=6}$ | 0.0026 | $ESVAR(2)_{p=2}$ | 0.0000 | $ESVAR(4)_{p=6}$ | 0.0030 |
| 18 | $ESVECM(6)_{p=4}$ | 0.0018 | $ESVAR(6)_{p=2}$ | 0.0000 | $ESVAR(4)_{p=4}$ | 0.0020 |
| 19 | $ESVECM(5)_{p=4}$ | 0.0018 | $ESVECM(3)_{p=4}$ | 0.0000 | $ESVAR(2)_{p=6}$ | 0.0019 |
| 20 | $ESVECM(4)_{p=4}$ | 0.0018 | $ESVECM(4)_{p=4}$ | 0.0000 | $ESVAR(3)_{p=4}$ | 0.0018 |

Notes: The order of the model is in parenthesis, and the subscript $d$ denotes the lag length of the transition variable.

Table 2.7: Top 20 Most Preferred Models (b)

| No. | US - ITA | | US - JAP | | US - UK | |
|---|---|---|---|---|---|---|
| 1 | $ESVECM(2)_{p=5}$ | 0.1600 | $ESVECM(3)_{p=6}$ | 0.2838 | $ESVECM(1)_{p=4}$ | 0.1293 |
| 2 | $ESVECM(6)_{p=5}$ | 0.1470 | $ESVECM(2)_{p=6}$ | 0.2083 | $ESVECM(2)_{p=4}$ | 0.1235 |
| 3 | $ESVECM(3)_{p=5}$ | 0.1287 | $ESVECM(6)_{p=6}$ | 0.2045 | $ESVECM(6)_{p=4}$ | 0.1049 |
| 4 | $ESVECM(4)_{p=5}$ | 0.1134 | $ESVECM(4)_{p=6}$ | 0.1473 | $ESVECM(3)_{p=4}$ | 0.1044 |
| 5 | $ESVECM(5)_{p=5}$ | 0.0734 | $ESVECM(5)_{p=6}$ | 0.0989 | $ESVECM(4)_{p=4}$ | 0.0897 |
| 6 | $ESVECM(1)_{p=5}$ | 0.0708 | $ESVECM(1)_{p=6}$ | 0.0562 | $ESVECM(5)_{p=4}$ | 0.0828 |
| 7 | $ESVECM(6)_{p=3}$ | 0.0583 | $ESVAR(3)_{p=6}$ | 0.0004 | $ESVECM(1)_{p=3}$ | 0.0617 |
| 8 | $ESVECM(4)_{p=3}$ | 0.0529 | $ESVAR(2)_{p=6}$ | 0.0003 | $ESVECM(6)_{p=3}$ | 0.0613 |
| 9 | $ESVECM(3)_{p=3}$ | 0.0395 | $ESVAR(6)_{p=6}$ | 0.0002 | $ESVECM(2)_{p=3}$ | 0.0575 |
| 10 | $ESVECM(2)_{p=3}$ | 0.0331 | $ESVAR(4)_{p=6}$ | 0.0002 | $ESVECM(3)_{p=3}$ | 0.0552 |
| 11 | $ESVECM(5)_{p=3}$ | 0.0313 | $ESVAR(5)_{p=6}$ | 0.0000 | $ESVECM(4)_{p=3}$ | 0.0500 |
| 12 | $ESVAR(2)_{p=5}$ | 0.0185 | $ESVECM(6)_{p=2}$ | 0.0000 | $ESVECM(5)_{p=3}$ | 0.0486 |
| 13 | $ESVAR(3)_{p=5}$ | 0.0152 | $ESVECM(5)_{p=2}$ | 0.0000 | $ESVECM(1)_{p=2}$ | 0.0046 |
| 14 | $ESVAR(4)_{p=5}$ | 0.0136 | $ESVECM(6)_{p=5}$ | 0.0000 | $ESVECM(6)_{p=2}$ | 0.0040 |
| 15 | $ESVECM(1)_{p=3}$ | 0.0096 | $ESVECM(4)_{p=5}$ | 0.0000 | $ESVECM(2)_{p=2}$ | 0.0039 |
| 16 | $ESVAR(5)_{p=5}$ | 0.0093 | $ESVECM(5)_{p=5}$ | 0.0000 | $ESVECM(3)_{p=2}$ | 0.0038 |
| 17 | $ESVAR(4)_{p=3}$ | 0.0086 | $ESVECM(3)_{p=5}$ | 0.0000 | $ESVECM(4)_{p=2}$ | 0.0033 |
| 18 | $ESVAR(3)_{p=3}$ | 0.0062 | $ESVAR(6)_{p=2}$ | 0.0000 | $ESVECM(5)_{p=2}$ | 0.0030 |
| 19 | $ESVAR(4)_{p=3}$ | 0.0056 | $ESVAR(5)_{p=5}$ | 0.0000 | $ESVAR(2)_{p=4}$ | 0.0010 |
| 20 | $ESVAR(2)_{p=3}$ | 0.0049 | $ESVECM(2)_{p=5}$ | 0.0000 | $ESVAR(3)_{p=4}$ | 0.0009 |

Notes: The order of the model is in parenthesis, and the subscript $d$ denotes the lag length of the transition variable.

Figure 2.1
Smooth Transition Functions (most preferred models)

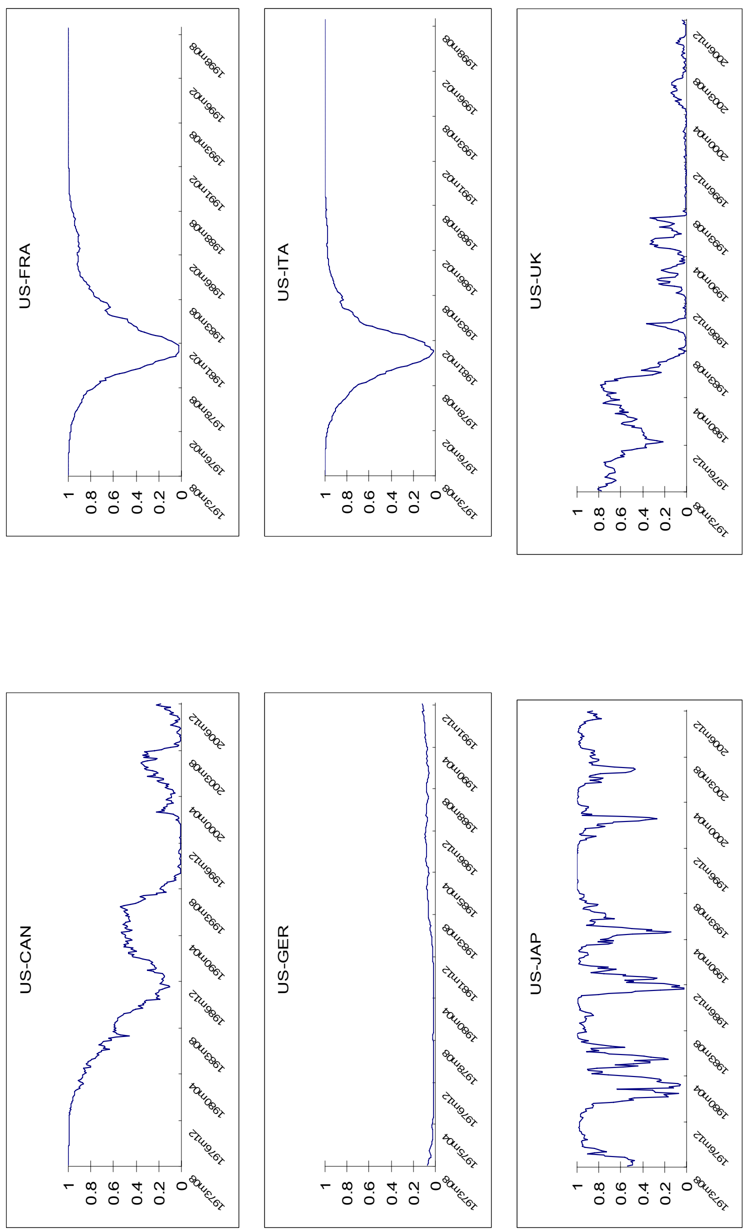


Figure 2.2
General Impulse Response Functions (a)

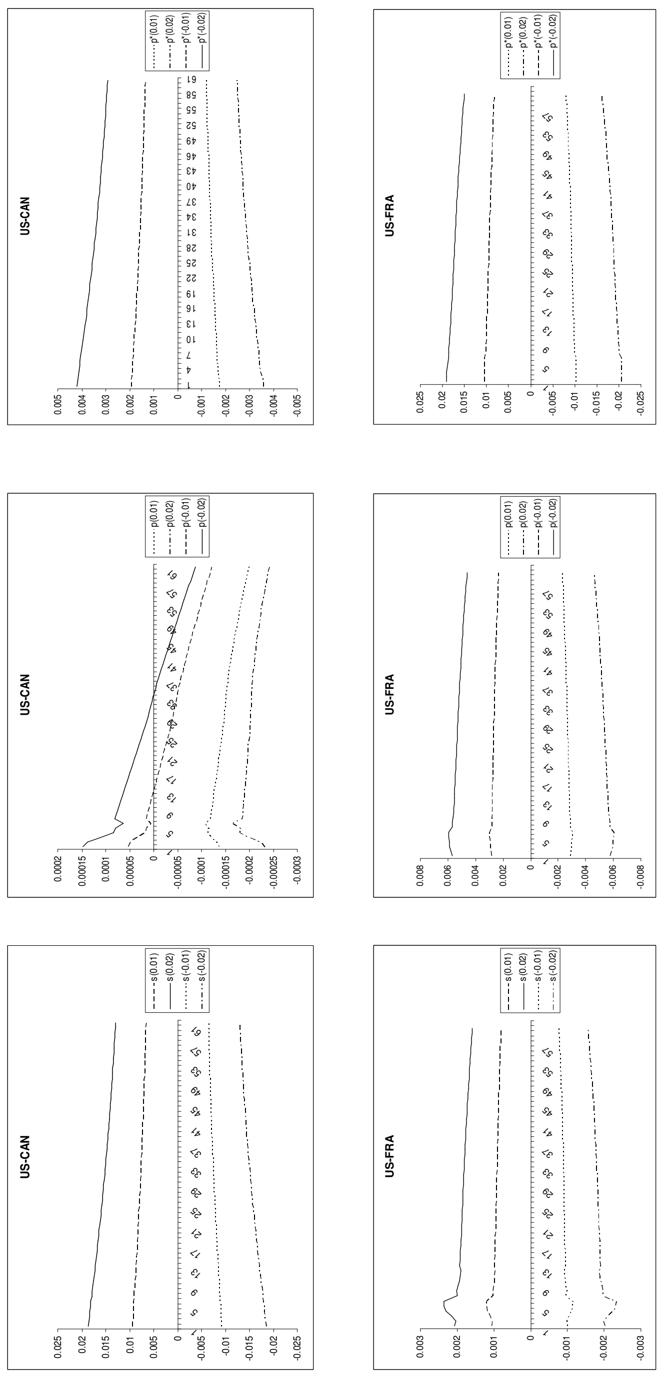


Figure 2.3
General Impulse Response Functions (b)

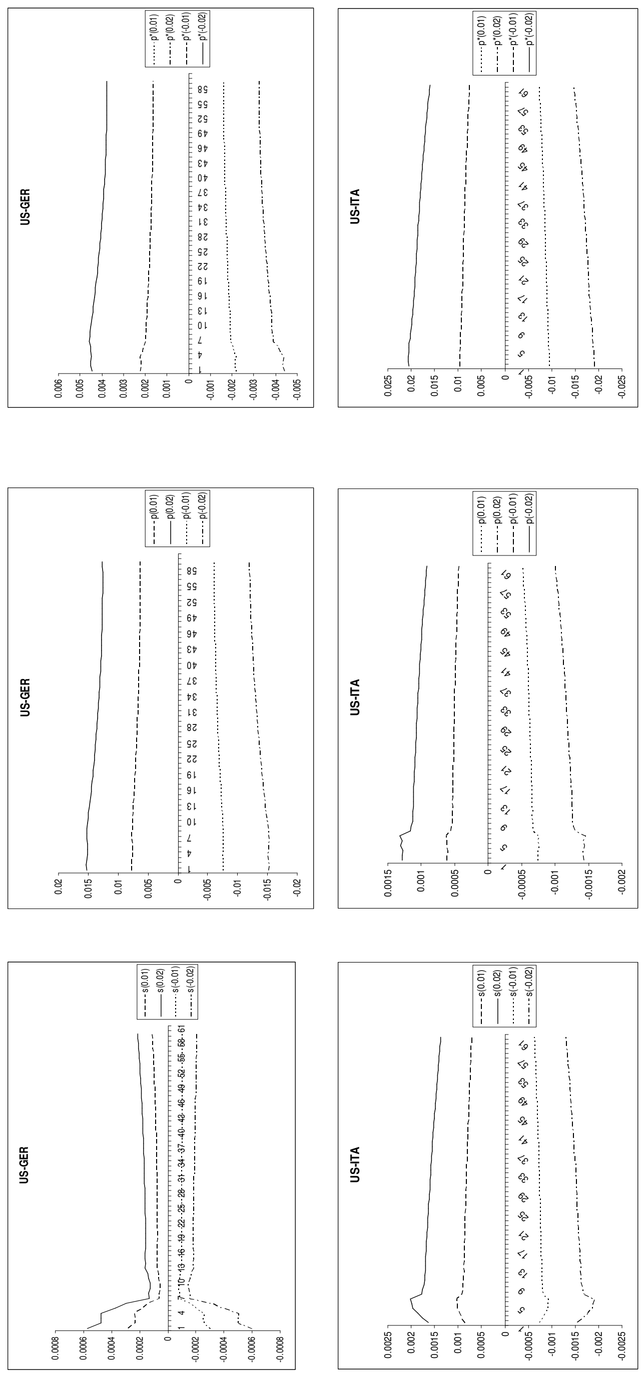
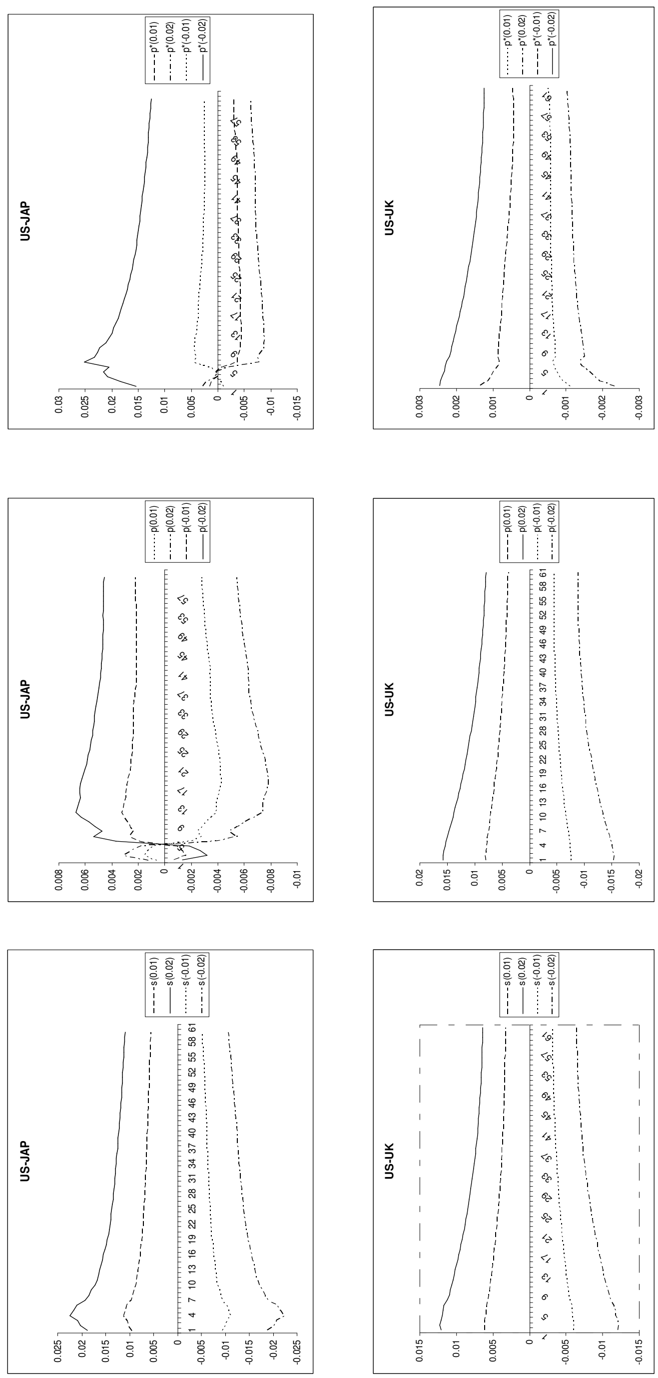


Figure 2.4
General Impulse Response Functions (c)



# Chapter 3

# Money-output causality revisited – A Bayesian logistic smooth transition VECM perspective

**Acknowledgements**: I am indebted for the constructive comments to Gary Koop and Roberto Leon-Gonzalez. This chapter has also benefited from the helpful discussions with Stephen Hall, Roderick McCrorie, Emi Mise and Rodney Strachan. Any remaining errors are the author's responsibility.

## 3.1 Introduction

From the late 1980s through the early 2000s, with the prevalence of studies based on Taylor rule [Taylor (1993)], the role of money (monetary base or monetary aggregates) had been deemphasized in much research on monetary policy and macroeconomic modeling [see, e.g., Barro (1989), Taylor (1999), Clarida, Galí and Gertler (2000)]. However, there has been a renewed interest in the effect of money in recent years. 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 any other policy instruments. Moreover, we find money reemerges as an important variable of concern in a number of most recent empirical work [for instance, Wang and Wen (2005), Sims and Zha (2006), Hill (2007), to mention a few].[1]

This chapter contributes to the discussion on whether money matters by revisiting an old topic: the causal effects from money to output in the postwar US data.[2] However, the current research 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

[1]As of the time of writing, Federal Reserve, the European Central Bank, Bank of England and central banks from Canada and Switzerland jointly announced cash injection plans to lessen the credit squeeze triggered by the sub-prime mortgages losses. Although the consequence of the intervention is yet to know, this unprecedent operation clearly implies that money remains a vital instrument for monetary policy.

[2]The money-output relationship has been intensively investigated in the literature. However, there is much less consensus about how money affects output [see, e.g. Sims (1972, 1980), Stock and Watson (1989), King and Watson (1997), Coe and Nason (2004)].

(LSTVECM) incorporating cointegration of an unknown form. Second, we develop a simple Bayesian approach to investigating the causal effects from money to output.

Single-equation logistic smooth transition error correction models have been widely used in the literature to capture the possible nonlinear money-output relationship [Lütekepohl, Teräsvirta and Wolters (1999), Teräsvirta and Eliasson (2001), Escribano (2004), Haug and Tam (2007), to mention a few]. However, considering the interplay between endogenously determined money, interest rates and the ultimate policy targets output and inflation, we believe LSTVECM can be more effective in capturing both the long run and short run dynamics in the linkages among all the variables. Perhaps the reason why researchers have not followed this route is due to 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), Seo (2006), Kapetanios, Shin and Snell (2006) for details]. In the literature, only Rothman, van Dijk and Frances (2001) apply a multivariate LSTVECM framework which is closest to us to study the money-output relationship.[3] Yet, Rothman, van Dijk and Frances (2001) pre-impose a theory based long-run cointegrating relationship in their estimation. While recognizing that the actual money-output interrelation is rather complex, unlike Rothman, van Dijk and Frances (2001), we let both

[3]Rothman, van Dijk and Frances (2001) test Granger causality from money to output in a classical context involving rolling window forecasting.

the cointegration rank and cointegrating vectors to be determined by the data.

Our estimation technique is Bayesian. Specifically, we extend the Bayesian cointegration space approach introduced in Strachan and Inder (2004) and collapsed Gibbs sampler developed in Koop, Leon-Gonzalez and Strachan (2005) into the nonlinear framework. Our method jointly captures the long-run cointegrating relationship and presence of nonlinearity in the LSTVECM in a single step. Compared with the available classical estimation techniques which often require multiple steps and Taylor expansions, our 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 LSTVECM. Yet, jagged likelihood functions do not create any particular problems in our Gibbs sampling scheme.

Considering that the large model we employed might be subject to the criticism of being too parameter rich, we use Bayes Factors for model comparison in order to reward more parsimonious models.[4] Alternative models are specified by placing zero restrictions on certain parameters of the unrestricted LSTVECM. Our approach to examining whether money long-run causes output is in spirit similar to that in Hall and Wickens (1993), Hall

[4]Bayes Factors include an automatic penalty for more complex models (see Koop and Potter, (1999a, 1999b) for details).

and Milne (1994) and Granger and Lin (1995). With respect to the Granger causality test from money to output, aside from considering if money directly enters the output equation as described in Rothman, van Dijk and Frances (2001), we look into whether money indirectly affects output through the channels of price and interest rate.

An important finding of our study is that the postwar US money-output relationship is nonlinear, with the regime shifting mainly driven by the lagged inflation rates. In terms of triggering regime changes, compared with the key role played by inflation rates, the role of lagged annual growth rates of output is less important, while the roles played by changes in oil prices, money and interest rates are nearly negligible. However, it is worth stressing that, in our study, nonlinear models consistently outperform linear models.

We find substantial evidence that money does not long-run cause output in the postwar US data. Additionally, consistent with the in-sample testing results in Rothman, van Dijk and Frances (2001), our studies show that money is nonlinearly Granger-causal for output. The impulse response analysis shows that the dynamic paths of output given a shock to money is rather complex. Most strikingly, we find that the accumulated effect of a shock to money is negative on real output in the next 50 years, regardless of the size and sign of the initial shock. This result calls for a word of caution when using money as a policy instrument.

The outline of this chapter is as follows. Section 2 describes the model and the Bayesian estimation technique. Section 3 reports the empirical results.

Section 4 concludes.

## 3.2 LSTVECM Model and Bayesian Inference

Following a majority of empirical work [for example, Lütekepohl, Teräsvirta and Wolters (1999), Rothman, van Dijk and Frances (2001)], we investigate the money-output relationship in a system of output, money, prices and interest rates.

We use the monthly US data spanning from 1959:1 to 2006:12. The data are obtained from the database of Federal Reserve Bank of St. Louis. Various measures of output, money, prices and interest rates are used in the literature. In this chapter, we adopt the seasonally adjusted industrial production index ($i_t$), 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 ($r_t$) for the measures of output, money, prices and interest rates, respectively. All variables are in logarithms except for interest rates which are in percent.

To catch the possible regime changes in US monetary policy, we model the interrelationship among output, money, prices and interest rates in a LSTVECM.[5] Let $y_t = [i_t \quad m_t \quad p_t \quad r_t]$, the LSTVECM of the $1{\times}4$ vector time series process $y_t$, t=1,...,T, conditioning on the p observations t= -p+1,...,0,

[5] The possible regime changes in US monetary policy have been well documented in the literature [see, e.g., Weise (1999), Clarida, Galí and Gertler (2000), Leeper and Zha (2003)].

can be specified as

$$\begin{aligned}\triangle y_t =& y_{t-1}\beta\alpha + \xi + \Sigma_{h=1}^{p}\triangle y_{t-h}\Gamma_h \\ &+ F(z_t)(y_{t-1}\beta^z\alpha^z + \xi^z + \Sigma_{h=1}^{p}\triangle y_{t-h}\Gamma_h^z) + \varepsilon_t\end{aligned} \tag{3.1}$$

$\varepsilon_t$ is a Gaussian white noise process where $E(\varepsilon_t) = 0$, $E(\varepsilon_s'\varepsilon_t) = \Sigma$ for $s = t$, and $E(\varepsilon_s'\varepsilon_t) = 0$ for $s \neq t$. Note that $\triangle y_t = y_t - y_{t-1}$. The dimensions of $\Gamma_h$ and $\Gamma_h^z$ are $n \times n$, and the dimensions of $\beta$, $\alpha'$, $\beta^z$, and $\alpha^{z'}$ are $n \times r$. Since we are using monthly data, without loss of generality, we set $p = 6$.

In model (3.1), 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):

$$F(z_t) = \{1 + exp[-\gamma(z_t - c)/\sigma]\}^{-1} \tag{3.2}$$

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 Bayesian approach, we are able to search over large numbers of choices for $z_t$ (or average over them). In this chapter, following Rothman, van Dijk and Frances (2001), we set $z_t$ to be the lagged annual growth rates of output, the lagged annual growth rates of money, the lagged annual inflation rates, the lagged annual changes in interest rates and the lagged annual growth rates in oil prices, respectively.[6] In particular, we allow the lag length of the

[6]Rothman, van Dijk and Frances (2001) point out that using annual growth rates

transition variables to vary from 1 to 6.

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 (3.2), the smoothing parameter $\gamma$ (which is non-negative) determines the speed of the smooth transition. Observe that when $\gamma \to \infty$, the transition function becomes a Dirac function, then model (3.1) becomes a two-regime threshold VECM model along the lines of Tong (1983). When $\gamma = 0$, the logistic function becomes a constant (equal to 0.5), and the nonlinear model (3.1) 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 $\sigma$ is chosen by the researcher and could reasonably be set to one. In this study, we set $\sigma$ equal to the standard deviation of the process $z_t$. This effectively normalizes $\gamma$ such that we can give $\gamma$ 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 $\gamma$.

Observe that model (3.1) encompasses a set of models distinguished by the number of the long-run equilibrium relationships, the cointegrating vectors, the order of the autoregressive process, the existence of the nonlinear effects, the choice of the transition variable, and whether Granger non-causality or long-run non-causality from money to output is imposed.

---

instead of monthly changes as plausible transition variables is in accord with the commonly accepted perception that the regimes in the money-output relationship are quite persistent.

### 3.2.1 Likelihood Function

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

To start with, restricting $\beta$ and $\beta^z$ to be semi-orthogonal, we write (3.1) as

$$\triangle y_t = x_{1,t-1}\beta\alpha + x_{2,t}\Phi + F(z_t)(x_{1,t-1}\beta^z\alpha^z + x_{2,t}\Phi^z) + \varepsilon_t \tag{3.3}$$

where $x_{1,t-1} = y_{t-1}$, $x_{2,t} = (1, \triangle y_{t-1}, ..., \triangle y_{t-p})$, $\Phi = (\xi', \Gamma_1', ..., \Gamma_p')'$, $\Phi^z = (\xi^{z'}, \Gamma_1^{z'}, ..., \Gamma_p^{z'})'$. To simplify the notation, we then define the $T \times n$ matrix $X_0 = (\triangle y_1', \triangle y_2', ..., \triangle y_T')'$ and the $T \times 2(r+1+np)$ matrix $X = (X_1\beta \quad X_2 \quad F^z X_1 \beta^z \quad F^z X_2)$, where $X_1 = (x_{1,1}', \quad x_{1,2}', \quad ..., \quad x_{1,T}')'$, $X_2 = (x_{2,1}', \quad x_{2,2}', \quad ..., \quad x_{2,T}')'$, and $F^z = diag(F(z_1), F(z_2), ..., F(z_T))$. Next, we set $B = (\alpha' \quad \Phi' \quad \alpha^{z'} \quad \Phi^{z'})'$, and stack the error terms $\varepsilon_t$ in the $T \times n$ matrix E, where $E = (\varepsilon_1', \varepsilon_2', ..., \varepsilon_T')'$. Finally, we rewrite model (3.1) as

$$X_0 = X_1\beta\alpha + X_2\Phi + F^z X_1\beta^z\alpha^z + F^z X_2 \Phi^z + E = XB + E \tag{3.4}$$

It is seen that the likelihood function of (3.4) is

$$L(y|\Sigma, B, \beta, \beta^z, \gamma, c) \propto |\Sigma|^{-\frac{T}{2}} exp\{-\frac{1}{2} tr \Sigma^{-1} E'E\} \tag{3.5}$$

Vectorizing (3.4), model (3.1) can be transformed into

$$x_0 = xb + e \tag{3.6}$$

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

Given that

$$\begin{aligned} tr\Sigma^{-1}E'E &= e'(\Sigma^{-1} \otimes I_T)e \\ &= s^2 + (b - \widehat{b})'V^{-1}(b - \widehat{b}) \end{aligned} \tag{3.7}$$

where $s^2 = x_0'M_v x_0$, $M_v = \Sigma^{-1} \otimes [I_T - X(X'X)^{-1}X']$, $\widehat{b} = [I_n \otimes (X'X)^{-1}X']x_0$ and $V = \Sigma \otimes (X'X)^{-1}$. The likelihood (3.5) can be written as

$$L(y|\Sigma, B, \beta, \beta^z, \gamma, c) \propto |\Sigma|^{-\frac{T}{2}} exp\{-\frac{1}{2}[s^2 + (b - \widehat{b})'V^{-1}(b - \widehat{b})]\} \tag{3.8}$$

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 matrices $\kappa$ and $\kappa^z$ of rank $r$, 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 $\kappa = (\alpha\alpha')^{\frac{1}{2}} = (\beta^{*'}\beta^*)^{\frac{1}{2}}$, and $\kappa^z = (\alpha^z\alpha^{z'})^{\frac{1}{2}} = (\beta^{z*'}\beta^{z*})^{\frac{1}{2}}$, we find $\alpha^{*'}$ and $\alpha^{z*'}$ are semi-orthogonal if $\beta$ and $\beta^z$ are semi-orthogonal. Therefore, we can reexpress equation (3.4) as

$$
\begin{aligned}
X_0 - X_2\Phi - F^z X_2 \Phi^z &= X_1 \beta \alpha + F^z X_1 \beta^z \alpha^z + E \\
&= X_1 \beta^* \alpha^* + F^z X_1 \beta^{*z} \alpha^{*z} + E
\end{aligned} \tag{3.9}
$$

Setting $\widetilde{x_0} = vec(X_0 - X_2\Phi - F^z X_2 \Phi^z)$, $\widetilde{x} = [\alpha^{*'} \otimes X_1 \;\; \alpha^{*z'} \otimes F^z X_1]$, $\widetilde{b} = [vec(\beta^*)' \;\; vec(\beta^{*z})']'$, we find equation (3.9) can be written as

$$
\widetilde{x_0} = \widetilde{x}\widetilde{b} + e \tag{3.10}
$$

where the dimension of $\widetilde{x_0}$ is $Tn \times 1$, the dimension of $\widetilde{x}$ is $Tn \times 2nr$, and the dimension of $\widetilde{b}$ is $2nr \times 1$.

Thus, we find the second likelihood representation from (3.10) is

$$
L(y|\Sigma, B, \beta, \beta^z, \gamma, c) \propto |\Sigma|^{-\frac{T}{2}} exp\{-\frac{1}{2}[s^2_{\beta^*} + (b_{\beta^*} - \widehat{b_{\beta^*}})' V^{-1}_{\beta^*} (b_{\beta^*} - \widehat{b_{\beta^*}})]\} \tag{3.11}
$$

where $s^2_{\beta^*} = (\widetilde{x_0} - \tilde{x}\widehat{b_{\beta^*}})'(\Sigma^{-1} \otimes I_T)(\widetilde{x_0} - \tilde{x}\widehat{b_{\beta^*}})$, $\widehat{b_{\beta^*}} = (\tilde{x}'\widetilde{x})^{-1}\tilde{x}'\widetilde{x_0}$, $V^{-1}_{\beta^*} = \tilde{x}'(\Sigma^{-1} \otimes I_T)\tilde{x}$.

### 3.2.2 Priors

Although the most commonly elicited quantity money demand equation indicates that the velocity of money is stationary [see, e.g., Rothman, van Dijk and Frances (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, e.g., Ambler (1989), Friedman and Kuttner (1992), Swanson (1998)]. Furthermore, it is impossible to impose meaningful informative priors for the coefficients of the long run/short run adjustment in the VECM or for parameters that indicates 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 logistic smooth transition VAR model (LSTVAR) suffer from identification problems.

As well documented in the literature, a linear VECM suffers from both the global and local nonidentifications 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 $\beta$, researchers first normalize $\beta$ into $\beta = [I_r \quad V']'$, then impose uninformative prior on the sub-vector $V$. However, as argued by Strachan and van Dijk (2004a), this approach has an undesirable side-effect that it favors 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, e.g., Strachan and Inder (2004), Villani (2005)]. As noted in Koop, Strachan, van Dijk and Villani (2006), 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 to 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 classical framework. Bauwens, Lubrano and Richard (1999) and Lubrano (1999a, 1999b) introduce a number of prior settings to solve the problem. Following Gefang and Strachan (2007), we tackle this problem by simply setting the prior distribution of $\gamma$ as Gamma.

The nonidentification problem faced by the LSTVECM is slightly different. Although the Davies' problem remains relatively the same as in the LSTVAR, 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 LSTVAR. As long as we can rule out the possibility that $\gamma = 0$, we can identify $\beta$, $\beta^z$, $\alpha$ and $\alpha^z$ sequentially once we choose a way to normalize $\beta$ and $\beta^z$.

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 $\Sigma$.

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

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 $b = vec(B)$, $k = 2(r + 1 + np)$. $\eta$ is the shrinkage prior as proposed by Ni and Sun (2003). As practiced in Koop, Leon-Gonzalez and Strachan (2006), we draw $\eta$ from the Gibbs sampler. In our case, we set the relatively uninformative prior distribution of $\eta$ as Gamma with mean $\underline{\mu_\eta}$, and degrees of freedom $\underline{\nu_\eta}$, where $\underline{\mu_\eta}$=10, $\underline{\nu_\eta}$=0.001.

Following the arguments of Koop, Strachan, van Dijk and Villani (2006), we elicit the uninformative prior of $\beta$ and $\beta^z$ indirectly from the prior expressed upon the cointegration space. In particular, following Strachan and Inder (2004), for $r \in (0, 4)$, we specify $\beta'\beta = I_r$ and $\beta^{z'}\beta^z = I_r$ for the purpose of normalization.[7] Moreover, in line with Koop, Leon-Gonzalez and Strachan (2005), we set the prior for $b_{\beta^*}$ as $p(b_{\beta^*}|\eta) \sim N(0, \eta^{-1} I_{2nr})$ in order to obtain a Normal form for the posterior.

To avoid the Davies' problem in the nuisance parameter space, following Lubrano (1999a, 1999b) and Gefang and Strachan (2007), we set the prior distribution for $\gamma$ as Gamma, which exclude *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 $\gamma$ does not render model (3.1) in favor of the nonlinear effect. In empirical work, we use Gamma(1,0.001) to allow the data information to dominate the prior of $\gamma$.

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.

### 3.2.3 Posterior Computation

Using the priors just identified and the likelihood functions in (3.5) and (3.11), we obtain the full conditional posteriors as follows.

Conditional on $\beta$, $\beta^z$, $\gamma$, $c$, and $b$, the posterior of $\Sigma$ 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 $\overline{b} = \overline{V}_b V^{-1} \hat{b}$ and covariance matrix $\overline{V}_b = \Sigma \otimes (X'X + \eta I_k)^{-1}$. Conditional on $\Sigma$, $b$, $\gamma$, and $c$,

[7]Note that the priors over the cointegration spaces of $\beta$ and $\beta^z$ are proper. See James (1954), Strachan and Inder (2004) for further explanation on the uniform distribution of the cointegration space.

the posterior of $b_{\beta^*}$ is Normal with mean $\bar{b}_{\beta^*} = \overline{V}_{\beta^*} V_{\beta^*}^{-1} \widehat{b_{\beta^*}}$ and covariance matrix $\overline{V}_{\beta^*} = [V_{\beta^*}^{-1} + \eta I_{nr}]^{-1}$.

To obtain the conditional posterior for $\eta$, we combine the prior and likelihood to obtain the expression

$$p(\eta|b, \Sigma, \gamma, c, y, x) \propto \eta^{\frac{\underline{\nu_\eta}+nk-2}{2}} exp(-\frac{\eta \underline{\nu_\eta}}{2\underline{\mu_\eta}} - \frac{1}{2} b' b \eta) \tag{3.12}$$

Thus with a Gamma prior, the conditional posterior distribution of $\eta$ is Gamma with degrees of freedom $\overline{\nu_\eta} = nk + \underline{\nu_\eta}$, and mean $\overline{\mu_\eta} = \frac{\overline{\nu_\eta}\underline{\mu_\eta}}{\underline{\nu_\eta}+\underline{\mu_\eta} b' b}$.

The posterior distributions for the remaining parameters, $\gamma$ and $c$, have nonstandard forms. However, we can use Metropolis-Hastings algorithm [Chib and Greenberg (1995)] within Gibbs to estimate $\gamma$, and the Griddy Gibbs sampler [Ritter and Tanner (1992)] to estimate $c$.

Following Koop, Leon-Gonzalez and Strachan (2005), we construct the collapsed Gibbs sampler as following.

1. Initialize $(b, \Sigma, b_\beta, \gamma, c)$;

2. Draw $\Sigma|b, b_\beta, \gamma, c$ from $IW(E'E, T)$;

3. Draw $b|\Sigma, b_\beta, \gamma, c$ from $N(\bar{b}, \overline{V}_b)$;

4. Calculate $\alpha^* = (\alpha\alpha')^{-\frac{1}{2}}\alpha$, $\alpha^{z*} = (\alpha^z \alpha^{z'})^{-\frac{1}{2}}\alpha^z$;

5. Create $\widetilde{x_0}$;

6. Draw $b_{\beta^*}|\Sigma, b, \gamma, c, \widetilde{x_0}$ from $N(\bar{b}_{\beta^*}, \overline{V}_{\beta^*})$;

7. construct $\kappa = (\beta^{*'}\beta^*)^{\frac{1}{2}}$, calculate $\beta = \beta^*\kappa^{-1}$. Construct $\alpha = \kappa\alpha^*$. Use the same procedure to derive $\beta^z$ and $\alpha^z$;

8. Draw $\gamma|\Sigma, b, b_\beta, c$ using M-H algorithm;

9. Draw $c|\Sigma, b, b_\beta, \gamma$ using Griddy-Gibbs sampler;

10. Repeat steps 2 to 9 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 relationships, the lag length of the autoregressive process, the existence of the nonlinear effects, and the transition variable triggering regime changes.

Similar to Rothman, van Dijk and Frances (2001), we specify that if money does not Granger-cause output, the lagged money variables do not enter the equation for output, and money can not be identified as the transition variable triggering regime changes. Moreover, enlightened by Hill (2007), we define that if money does not Granger-cause output, the lagged money does not enter the equations for price and interest rate.[8] In terms of long-run causality, following Hall and Wickens (1993), Hall and Milne (1994) and Granger and Lin (1995), we specify that if money does not appear in any cointegration relationships which enter the output equation, money is not

[8] As explained in Hill (2007), the situation that A causes B and B causes C implies A eventually causes C.

long-run causal (or weakly causal) for output.[9]

Bayesian methods provide us 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 Bayesian Model Averaging. 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.

The Bayes Factor for comparing one model to a second model where each model is parameterized by $\zeta = (\zeta_1, \zeta_2)$ and $\psi$ respectively, is

$$B_{12} = \frac{\int \ell(\zeta)p(\zeta)d(\zeta)}{\int \ell(\psi)p(\psi)d(\psi)},$$

where $\ell(.)$ is the likelihood function and $p(.)$ is the prior density of the parameters for each model.

If the second model nests within the first at the point $\zeta_2 = \zeta^*$, then, subject to further conditions, we can compute the Bayes factor $B_{12}$ via the Savage-Dickey density ratio [see, for example, Koop and Potter (1999a), Koop, Leon-Gonzalez and Strachan (2006) for further discussion in this class of models]. For the simple example discussed here, the Savage-Dickey density

[9] See Hall and Wickens (1993), Hall and Milne (1994) and Granger and Lin (1995) for details.

ratio is:

$$B_{12} = \frac{p(\zeta_2 = \zeta^*|Y)}{p(\zeta_2 = \zeta^*)},$$

where the numerator is the marginal posterior density of $\zeta_2$ for the unrestricted model evaluated at the point $\zeta_2 = \zeta^*$, and the denominator is the prior density of $\zeta_2$ also evaluated at the point $\zeta_2 = \zeta^*$.

Since the conditional posterior of $b$ is Normal, it is easy to incorporate the estimation of the numerator of the Savage-Dickey density ratio in the Gibbs sampler. As to the denominator of the Savage-Dickey density ratio, using the properties of the Gamma and Normal distributions, we derive the marginal prior for a sub-vector of $b$ evaluated at zeros as

$$\{(\frac{\underline{\mu}_\eta}{\pi\underline{\nu}_\eta})^{\omega/2}\Gamma(\frac{\omega+\underline{\nu}_\eta}{2})\}/[\Gamma(\frac{\underline{\nu}_\eta}{2})]$$

where $\Gamma(.)$ is the Gamma function, and $\omega$ is the number of elements in $b$ restricted to be zeros.

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 lag coefficients are zero, i.e., $\Gamma_h = \Gamma_h^z = 0$, at which point we have the model with $p = 0$. 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.

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 Bayesian Model Averaging and weight features of interest across different models using posterior model probabilities [as suggested by Leamer (1978)].

## 3.3 Empirical Results

In empirical work, we allow the cointegration rank of the unrestricted model (3.1) to vary from 1 to 3.[10] For unrestricted models with a specific cointegration rank, we allow for 5 types of possible transition variables to trigger the regime changes, namely the lagged annual output growth, the lagged annual money growth, the lagged annual inflation rates, the lagged annual changes in interest rates and lagged annual growth rates in oil prices, respectively. Among these models, both the maximal order of the autoregressive process and longest lag length of the transition indicator are allowed to be 6. In total, we investigate the causal effects from money to output in the postwar US data by estimating 90 unrestricted LSTVECM models.

Altogether, we run 90 Gibbs sampling schemes. 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

[10]We don't 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.

introduced by Geweke (1992). We use the MATLAB program from LeSage's Econometrics Toolbox [LeSage (1999)] for the diagnostic.

### 3.3.1 Model Comparison Results

In this section, we report the results relating to the posterior model probabilities associated with a set of 2766 possible models nested in the original 90 unrestricted models.[11] Assuming the 2766 models are mutually exclusive, in calculating the Bayes Factors, we have each of the 2766 models receive an *a priori* equal weight.

We find compelling evidence that money does not long-run cause output in the postwar US data. First, assuming all the 2766 models are mutually exclusive and exhaustive, we find money long-run non-causality models jointly account for 95.16% of the posterior mass.[12] Second, assuming all models nested in the unrestricted models with the same number of cointegration ranks are mutually exclusive and exhaustive, we observe that money long-run non-causality models are predominant in all the three cases. Specifically, for models nested in the unrestricted LSTVECM models with only one cointegration relationship, money long-run non-causality models jointly account

[11] Altogether, we examine 66 linear models and 2700 nonlinear models. Namely 6 linear VARs, 6 linear VARs with money Granger non-causality restriction, 18 linear VECMs, 18 linear VECMs with money Granger non-causality restriction, 18 linear VECMs with money long-run non-causality restriction, 540 nonlinear VARs, 540 nonlinear VARs with money Granger non-causality restriction, 540 nonlinear VECMs, 540 nonlinear VECMs with money Granger non-causality restriction, and 540 nonlinear VECMs with money long-run non-causality restriction.

[12] In the remainder of the chapter, we use money long-run non-causality model to indicate the restricted model where money does not long-run cause output.

for 96.06% of the posterior probabilities; for models nested in the unrestricted LSTVECM models with two cointegration relationships, overall, money long-run non-causality models receive 97.68% of the posterior mass; for models nested in the unrestricted LSTVECM models with three stationary cointegration relationships, money long-run non-causality models altogether get 95.16% of the posterior probability. Finally, if we assume models nested in each of the 90 unrestricted LSTVECM models are mutually exclusive and exhaustive, we find that in each cases, money long-run non-causality models are constantly overwhelmingly supported over other types of models.[13]

Assuming models nested within LSTVECMs with the same number of cointegration ranks (from 1 to 3) to be exhaustive, we reports the top 10 models with the highest posterior model probabilities in table 1. Note that the top 10 models of all the 2766 models are exactly the same as the top 10 models nested in the LSTVECM models with three cointegration relationships, for nonlinear models of rank 3 get nearly 100% of the posterior mass among all the 2766 models. Table 1 reinforces the substantial support for money long-run non-causality models. It is worth noting that the most preferred models for all cases are nonlinear money long-run non-causality models. Another interesting finding is that there is no pronounced model uncertainty if we focus on all the 2766 possible models or a subset of models nested within the unrestricted LSTVECM models with 2 or 3 stochastic

[13]The model comparison results for models nested in each of the 90 unrestricted LSTVECM models are available upon request.

trends. Yet, more evidence of model uncertainty emerges if we pre-impose the cointegration rank of the unrestricted models to be 1. Finally, note that the most preferred model among all the possible 2766 models is the restricted money long-run non-causality LSTVECM of rank 3, order 6, and with lagged 2 inflation rates as the transition indicator.

Overall, we find little support for models indicating money is not Granger-causal for output. The posterior mass for all models (linear types of VAR, VECM models and nonlinear types of LSTVAR, LSTVECM models) with zero restrictions on the lagged money in the equations for output, price and interest rates is nearly negligible. Furthermore, observing that the total posterior model probability associated with the unrestricted LSTVECMs and the restricted money long-run non-causality LSTVECMs is almost 100%, we find that money nonlinearly Granger-causes output, which is in accord with the in-sample evidence in Rothman, van Dijk and Frances (2001).

Given the substantial support for nonlinear models, it is interesting to examine which transition variable plays a more important role in triggering regime changes. Examining all the possible nonlinear models, we find that lagged annual inflation rates consistently predominate over other candidate transition variables in driving regime changes. All together, nonlinear models with lagged inflation rates as transition variables receive 89.17% of the posterior mass. The next important triggers for regime changes are lagged annual output growth rates. Note that nonlinear models with regime shifting governed by the lagged annual output growth rates account for 10.83% of the posterior mass. Last, compared with lagged inflation and output growth, lagged changes in money, interest rates and oil prices play trivial roles in triggering regime changes.

To highlight the nonlinear feature of the interrelationship among money, output, prices and interest rates, in figure 3.1 we plot the values of the logistic smooth transition function over time for the most preferred model chosen from all the 2766 candidate models.[14] Observe that although the plot is quite volatile, the values of the transition function are almost always bounded by 0.4 and 0.6 throughout the time. This result implies that the regime changes in the postwar US money-output relationship are quite modest, which is in line with the findings of Primiceri (2005) and Sargent, Williams and Zha (2006). However, given the compelling support for nonlinear models over linear models, it is worth stressing that we find it improper to model the post-war US money-output relationship in linear models.

Table 2 contains the estimates of the cointegration vectors and transition parameters for the most preferred models nested in the unrestricted LSTVECM models of rank 1, 2, 3, respectively. Recall that the most preferred model among the whole set of 2766 candidate models is exactly the same most preferred model selected from all the possible models nested in the unrestricted LSTVECM models of rank 3.

To aid in interpretation, in table 2, we normalize the cointegration vectors on output and money, respectively. Assuming that the cointegration rank is

[14]The whole set of the time profiles of the transition functions are available upon request.

1, we find the parameters for output, money and price levels appear to have reasonable economic interpretations. For example, inflation brings about (nominal) higher output level. Yet, it is not so straightforward to explain why effects of interest rates are quite different between the lower and upper regimes. Focusing on the model with 2 cointegration relationships, we find that in each regime, the first cointegrating vectors could be said to correspond to the theory of the (log) quasi-velocity of money as defined in Rothman, van Dijk and Frances (2001), while it is hard to find an economic theory to explain the second long-run equilibrium relationship. For the most preferred model among all the possible 2766 models (or the most preferred model among all the models nested in LSTVECMs of rank 3), we find it even more difficult to find a theory-based explanation for the long-run cointegrating interrelationships. Yet, it is clear that there are enormous differences in the cointegration vectors between the upper and lower regimes.

The estimated values of the smoothing parameter $\gamma$ presented in table 2 are relatively small. With the speed of the transition determined by $\gamma$, small value of $\gamma$ indicates that the transition between regimes is rather smooth. As to the estimated value of $c$, recall that for all cases, the transition variable is the lagged inflation rates. In our sample, the mean of inflation rates is 0.0352. Given the threshold $c$ is greater than 0.05 for each cases, it is seen that the upper regimes only become active when the transition variable is very large.

Finally, it is illuminating to look into the model comparison results in the linear framework. Assuming the 66 linear models are exhaustive, we find that the unrestricted linear VECM of rank 3 and order 6 receives nearly 100% of the posterior mass. Thus, models denoting long-run money non-causality are no longer supported in the linear frameworks. Furthermore, we find unrestricted VECM of order 6 dominates money long-run non-causality models when we pre-specify the rank of the cointegration space to be 1 or 2. Nevertheless, these results prove that ignoring nonlinear effects can lead to quite misleading conclusions, such as money is long-run causal for output.

### 3.3.2 Impulse Response Analysis

To shed further light on the causal effects from money to output, we analyze the impulse responses of output given a shock to money. The nonlinear LSTVECM allows for asymmetries in the behaviour of the money-output linkages. In this study, we are interested in two types of asymmetric effects. First, whether positive and negative shocks to money have unbalanced effects on real 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 [e.g. Potter (1994), Koop, Pesaran and Potter (1996)]. We use the generalized impulse response function proposed in Koop, Pesaran and Potter (1996) to examine the response of output to a money shock. In particular, we examine the generalized impulse response functions of $GI_P$ for a shock, $\upsilon_t$, and a history, $\omega_{t-1}$ as follows

$$GI_P(n, \upsilon_t, \omega_{t-1}) = E[P_{t+n}|\upsilon_t, \omega_{t-1}] - E[P_{t+n}|\omega_{t-1}] \tag{3.13}$$

where $n$ is the time horizon. By averaging out the future shocks, in (3.13), we treat the impulse responses as an average of what might happen given what has happened. Using Bayesian approach, we calculate the generalized impulse responses by averaging out the history uncertainties, future uncertainties and parameter uncertainties.

In each replication of the Gibbs Sampler after the initial burning runs, we calculate the generalized impulse response functions for all the alternative models as follows.

1. Randomly draw an $\omega_{t-1}$ in the observed sample as the history.

2. For a pre-specified shock that hits money, randomly draw from $\Sigma$ the corresponding shocks hitting the other three variables at time $t$.

3. Set the maximum horizon as $n$ and randomly sample $n+1$ four by one vectors of innovations from $\Sigma$.

4. Calculate the expected realizations of output using the shocks calculated in step 2 and the last $n$ innovations in step 3.

5. Calculate the shock-independent expected realizations of output using all the $n+1$ innovations in step 3.

6. Take the difference of the results from step 4 and step 5 to generate the impulse responses of output for the current draw.

At the end of the Gibbs sampling scheme, we derive the generalized impulse response functions for each possible model by integrating out all the parameter uncertainties. Note that if there is a great deal of model uncertainty, we can also average across models to derive the impacts of money on output weighting by the posterior model probabilities.

We set the magnitudes of the initial shocks amounting to $\pm 1$ and $\pm 2$ times the standard deviation of monthly money growth rates, namely $\pm 1$ and $\pm 2$ units of shocks. The time horizon of the impulse responses is set as 600 months (50 years). Given the large number of models and four different shocks, we only present the impulse response functions for the most preferred model among all the 2766 models in figures 3.2 - 3.3. For comparison, both the impulse response functions of output (nominal output) and real output are provided. The following observations are noteworthy in figures 3.2 - 3.3.

1. Positive and negative money shocks of the same magnitude appear to have asymmetric affects on both nominal output 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.

2. Impacts on both nominal output and real output appear to vary disproportionately with the size of the shock to money.

3. Impacts on nominal output appear to steadily increase in the same direction of the initial shocks in the first 10 years. After that, the impact responses become more volatile.

4. Compared with the responses of nominal output, the impact responses of real output to a money shock are rather volatile. More strikingly, the total effect on real output appears to be negative in the next 50 years after a shock to money, regardless of the size and sign of the shock.

## 3.4 Conclusion

This chapter investigates the causal effects from money to output using postwar US data. We develop a Bayesian approach to catch the interrelationship among money, output, prices and interest rates in a LSTVECM model. Different from similar nonlinear modeling method in the literature, we jointly estimate the cointegration relationship and nonlinear effects in a single step without pre-imposing any theory based restrictions.

Our model comparison results indicate that the postwar US money-output relationship is nonlinear. Yet, we find that the transition between regimes is rather smooth, and it is improper to use any abrupt transition framework to model the money-output linkage. Through model comparison, we find substantial evidence in favor of money long-run non-causality for output. In addition, we find little evidence against Granger causality from money to output. More precisely, our result strongly support that money nonlinearly

Granger-causes output during the postwar period in the US.

Our impulse response analysis sheds further light on the nonlinear causal effects from money to output. An important finding is that although a positive money shock can increase nominal output, we have to be cautious in using money as a policy instrument, for it appears that a shock to money will have negative cumulative effects on real output over the next fifty years, regardless of the size and sign of the shock.

Table 3.1: Most preferred models

| | Rank1 | | Rank2 | | Rank3/All 2766 models | |
|---|---|---|---|---|---|---|
| No. | Model | Posterior Mass | Model | Posterior Mass | Model | Posterior Mass |
| 1 | $LSTVECMb(1,3,5,6)$ | 0.4409 | $LSTVECMb(2,3,5,6)$ | 0.9606 | $LSTVECMb(3,3,2,6)$ | 0.8771 |
| 2 | $LSTVECMb(1,1,4,6)$ | 0.2098 | $LSTVECM(2,3,5,6)$ | 0.0217 | $LSTVECM(3,1,1,6)$ | 0.0737 |
| 3 | $LSTVECMb(1,3,1,6)$ | 0.1689 | $LSTVECMb(2,1,5,6)$ | 0.0150 | $LSTVECM(3,1,1,6)$ | 0.0342 |
| 4 | $LSTVECMb(1,3,6,6)$ | 0.1028 | $LSTVECM(2,1,5,6)$ | 0.0015 | $LSTVECM(3,3,2,6)$ | 0.0141 |
| 5 | $LSTVECM(1,1,4,6)$ | 0.0301 | $LSTVECMb(2,3,2,6)$ | 0.0010 | $LSTVECMb(3,3,1,6)$ | 0.0006 |
| 6 | $LSTVECMb(1,5,4,6)$ | 0.0266 | $LSTVECMb(2,1,3,6)$ | 0.0002 | $LSTVECMb(3,1,3,6)$ | 0.0003 |
| 7 | $LSTVECMb(1,2,4,6)$ | 0.0037 | $LSTVECMb(2,3,6,6)$ | 0.0000 | $LSTVECM(3,1,3,6)$ | 0.0001 |
| 8 | $LSTVECM(1,3,5,6)$ | 0.0030 | $LSTVECMb(2,3,2,6)$ | 0.0000 | $LSTVECM(3,1,3,6)$ | 0.0000 |
| 9 | $LSTVECMb(1,5,5,6)$ | 0.0026 | $LSTVECM(2,1,3,6)$ | 0.0000 | $LSTVECM(3,1,4,6)$ | 0.0000 |
| 10 | $LSTVECM(1,5,4,6)$ | 0.0021 | $LSTVECMb(2,4,2,6)$ | 0.0000 | $LSTVECM(3,3,1,6)$ | 0.0000 |

Notes:

*i*. The columns labeled Rank 1 report the top 10 most preferred models and their corresponding posterior probabilities when we assume all possible models nested within the LSTVECM models with only one cointegration relationship are exhaustive. Same procedure applies to columns labeled Rank 2 and Rank 3.

*ii*. LSTVECMb indicates the LSTVECM model is restricted so money is long-run non-causal for output.

*iii*. In parenthesis, the first subscript indicates the rank of the unrestricted model; The second subscript indicates the transition variable causing regime changes, where 1,2,3,4 and 5 denote annual output growth, annual money growth, annual inflation, annual growth in interest rates, and annual growth in oil prices; The third subscript indicates the lag length of the transition variable; The fourth subscript indicates the order of the model.

Table 3.2: Cointegration vectors and smooth transition parameters

| | Rank1 | Rank2 | Rank3 |
|---|---|---|---|
| Lower Regime | $i$-0.2031$m$-0.2855$p$+0.0950$r$ | $i$-1.1168$m$-0.2555$p$<br>$m$+0.2149$p$+0.2829$r$ | $i$+0.1923$m$<br>$m$+1.8885$r$<br>$m$-10.5341$p$ |
| Upper Regime | $i$-0.4150$m$-0.6940$p$-0.0986$r$ | $i$-0.4831$m$-0.7252$p$<br>$m$+1.0513$p$+0.7710$r$ | $i$-0.1417$m$<br>$m$+0.8494$r$<br>$m$-0.8133$p$ |
| $\gamma$ | 0.4068 (*0.0001 2.2946*)* | 0.1957 (*0.0002 1.2811*) | 0.1826 (*0.0000 1.2678*) |
| $c$ | 0.0518 (*0.0301*)** | 0.0519 (*0.0301*) | 0.0513 (*0.0302*) |

Notes:

* The 95% highest posterior density intervals (HPDIs) for $\gamma$s are reported in parenthesis.

** Standard deviations of $c$ are reported in parenthesis.

Figure 3.1
Smooth transition function

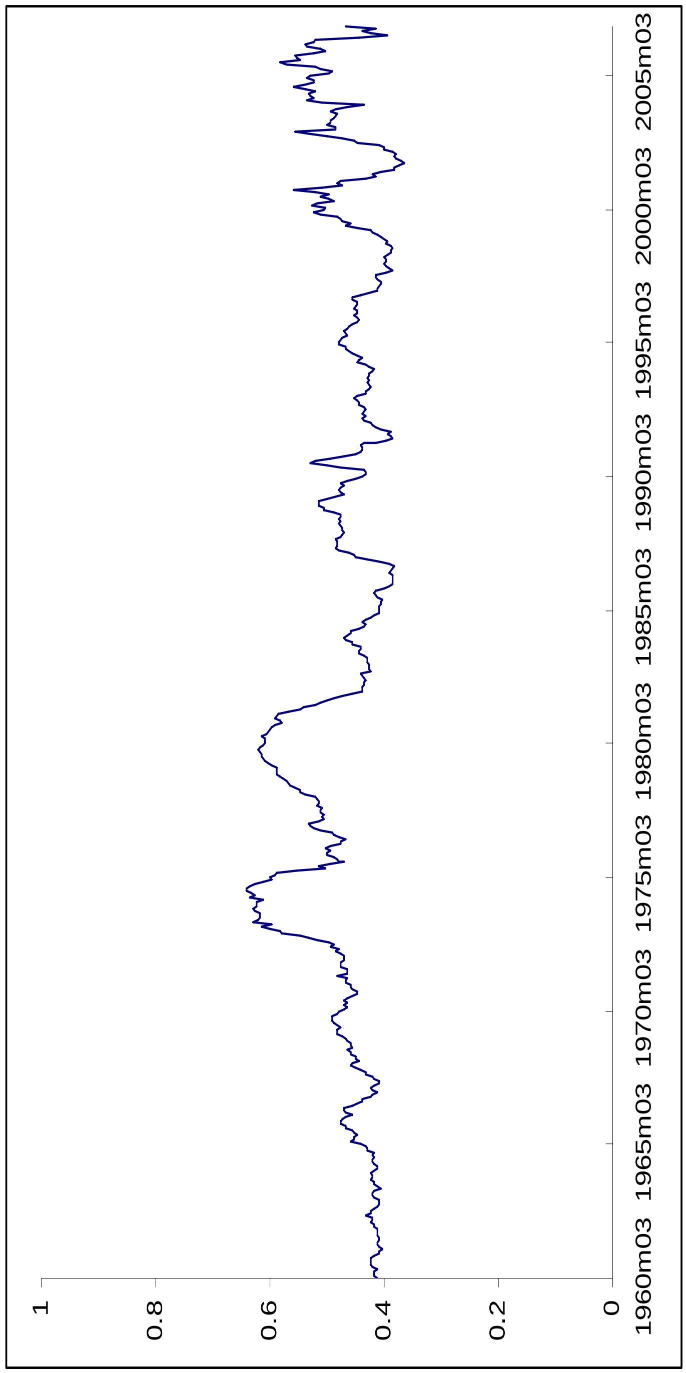
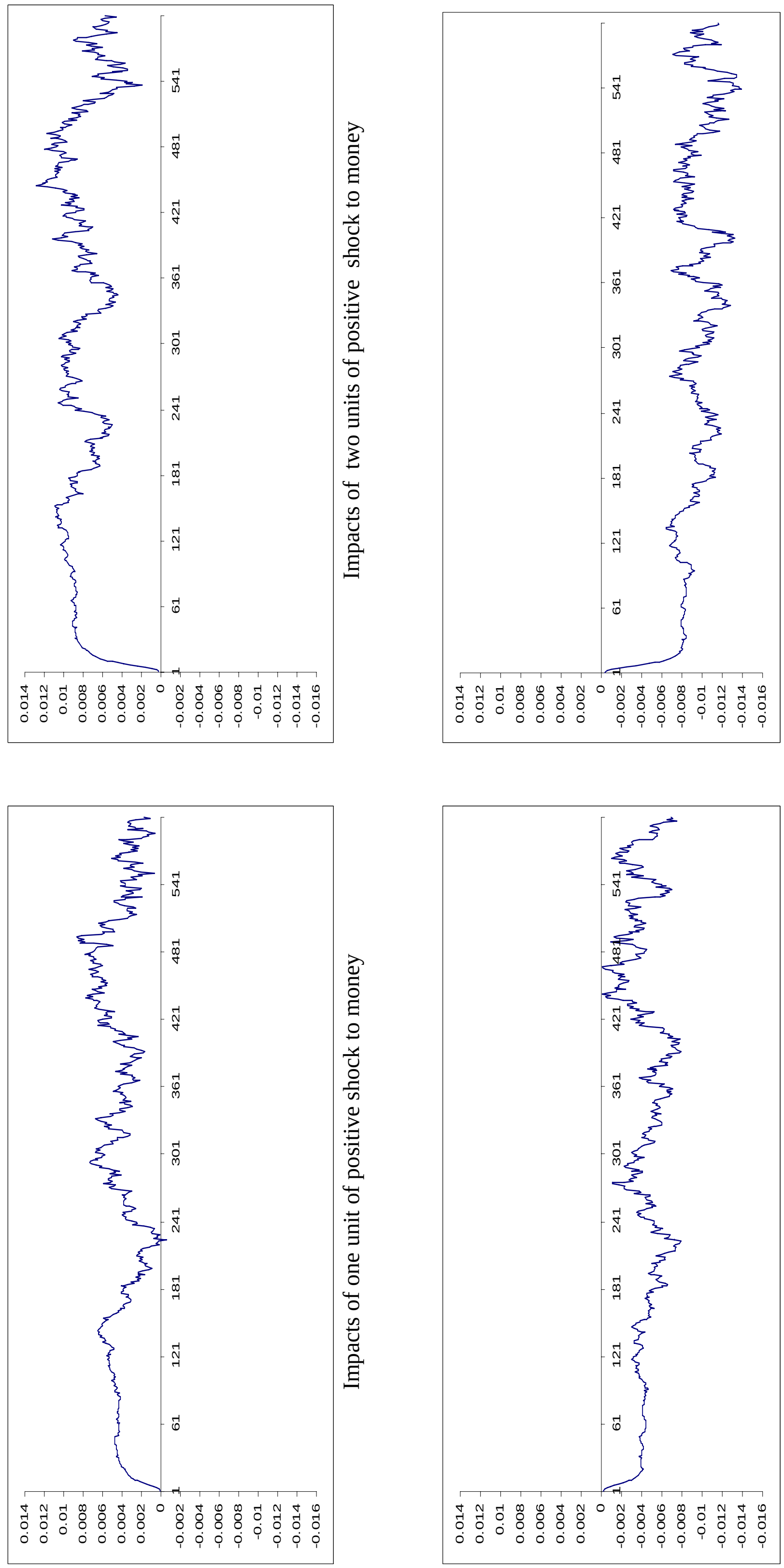


Figure 3.2
**Impulse response functions of (nominal) output**



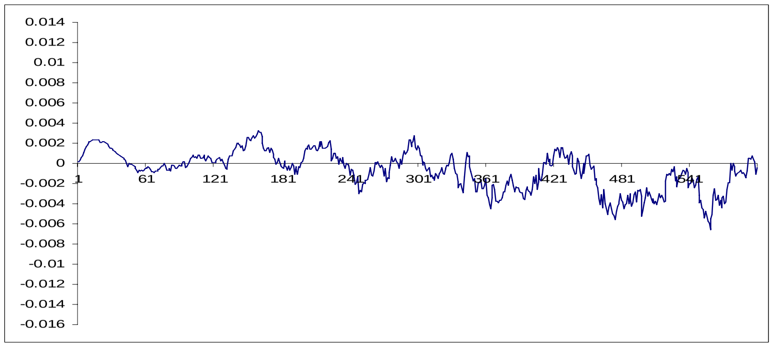

Impacts of one unit of positive shock to money

Impacts of two units of positive shock to money

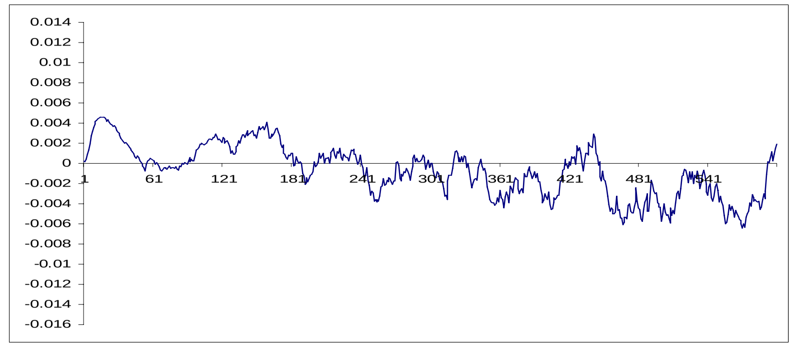
Impacts of one unit of negative shock to money

Impacts of two units of negative shock to money

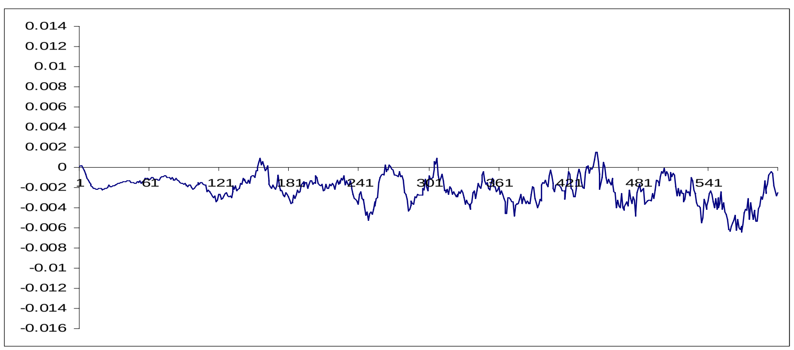
Figure 3.3
**Impulse response functions of real output**



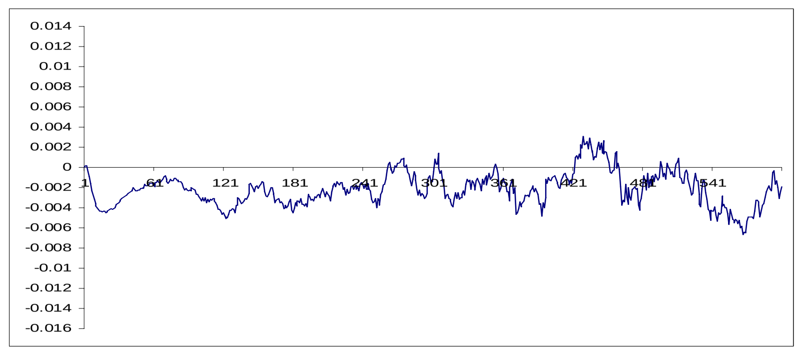
Impacts of one unit of positive shock to money



Impacts of two units of positive shock to money



Impacts of one unit of negative shock to money



Impacts of two units of negative shock to money

# List of References


Ambler, S. (1989). 'Does Money Matter in Canada? Evidence from a Vector Error Correction Model', *The Review of Economics and Statistics*, Vol. 71(4), pp. 651-58.

Artis, M., A. B. Galvâo and M. Marcellino (2007). 'The transmission mechanism in a changing world', *Journal of Applied Econometrcis*, Vol. 22, pp. 39-61.

Artis, M., H-M. Krolzig and J. Toro (2004). 'The European business Cycle', *Oxford Economic Papers*, Vol. 56, pp. 1-44.

Artis, M. and W. Zhang (1997). 'International business cycles and the ERM: is there a European business Cycle?' *International Journal of Finance and Economics*, Vol. 2, pp. 1-16.

Barro, R. J. (1989). 'Interest-rate targeting', *Journal of Monetary Economics* Vol. 23(1), pp. 3-30.

Bartlett, M, S. (1957). 'A comment on D. V. Lindley's statistical paradox', *Biometrika*, Vol. 44, pp. 533-34.

Baum, C. F., J. T. Barkoulas and M. Caglayan (2001). 'Nonlinear adjustment to purchasing power parity in the post-Bretton Woods era', *Journal of International Money and Finance*, Vol. 20, pp. 379-99.

Bauwens, L., M. Lubrano and J-F. Richard (1999). 'Bayesian inference in dynamic econometric models', Oxford University Press: New York.

Chen, S-W. and C-H. Shen (2007). 'A sneeze in the US, a cough in Japan, but pneumonia in Taiwan? An application of the Markov-Switching vector autoregressive model', *Economic Modelling*, Vol. 48, pp. 1-48.

Cheung, Y. W. and K. S. Lai (1994). 'Mean reversion in real exchange rates', *Economics Letters*, Vol. 46(3), pp. 251-56.

Chib, S. and E. Greenberg (1995). 'Understanding the Metropolis-Hastings algorithm', *The American Statistician* Vol. 49(4), pp. 327-35.

Chortareas G. and G. Kapetanios (2005). 'How Puzzling is the PPP Puzzle? An Alternative Half-Life Measure of convergence to PPP', Money Macro and Finance (MMF) Research Group Conference 2005-36.

Clarida, R., J. Galí and M. Gertler (2000). 'Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory', *Quarterly Journal of Economics*, Vol. 115(1), pp. 147-80.

Coe, P. J. and J. M. Nason (2004). 'Long-run monetary neutrality and long-horizon regressions', *Journal of Applied Econometrics*, Vol. 19(3), pp. 355-73.

Davies, R. B. (1977). 'Hypothesis testing when a nuisance parameter is present only under the alternative', *Biometrika*, Vol. 74, pp. 33-43.

Duca, J. V. and D. D. VanHoose (2004). 'Recent Developments in Understanding the Demand for Money', *Journal of Economics and Business*, Vol. 56, pp. 247-72.

Dumas, B. (1992). 'Dynamic equilibrium in a spatially separated world', *Review of Financial Studies*, Vol. 5(2), pp. 153-80.

Edison, H. J. and E. O. Fisher (1991). 'A long-run view of the European monetary system', *Journal of International Money and Finance*, Vol. 10, pp. 53-70.

Escribano, A. (2004). 'Nonlinear Error Correction: The Case of Money Demand in the United Kingdom 1878-2000', *Macroeconomic Dynamics*, Vol. 8(1), pp. 76-116.

Friedman, B. M and K. N. Kuttner (1992). 'Money, Income, Prices, and Interest Rates', *American Economic Review*, Vol. 82(3), pp. 472-92.

Garratt, A., G. Koop, E. Mise and S. P. Vahey (2007). 'Real-time Prediction with UK Monetary Aggregates in the Presence of Model Uncertainty', *http://www.ems.bbk.ac.uk/faculty/garratt/index html.*

Gefang, D. and R. W. Strachan (2007). 'Asymmetric Impacts of International Business Cycles on the UK - a Bayesian LSTVAR Approach', 15th Annual Symposium of the Society for Nonlinear Dynamics and Econometrics, Paris.

Geweke, J. (1992). 'Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments', in Bernodo, J., Berger, J., Dawid, A. and Smith, A. *(eds.)* Bayesian Statistics, Vol. 4, pp. 641-49, Oxford: Clarendon Press.

Goldfeld, S. M. and R. E. Quandt (1972). 'Nonlinear methods in econometrics', North-Holland Publ. Co.: Amsterdam and London.

Granger, C. and J-L. Lin (1995). 'Causality in the Long Run', *Econometric Theory*, Vol. 11(3), pp. 530-36.

Granger, C. and T. Teräsvirta, (1993). 'Modelling Nonlinear Economic Relationships', Oxford University Press: New York.

Hall, S. G. and A. Milne (1994). 'The Relevance of P-Star Analysis to UK Monetary Policy', *Economic Journal*, Vol. 104(424), pp. 597-604.

Hall, S.G. and M. R. Wickens (1993). 'Causality in Integrated Systems', D.P. no. 27C93, Centre for Economic Forecasting, London Business School.

Hamilton, J. D. (1989). 'A New Approach to the Economic Analysis of

Nonstationary Time Series and the Business Cycle', *Econometrica*, Vol. 57, pp. 357-384.

Haug, A. A. and J. Tam (2007). 'A Closer Look at Long-Run U.S. Money Demand: Linear or Nonlinear Error-Correction With M0, M1, or M2?' *Economic Inquiry*, Vol. 45(2), pp. 363-76.

Helbling, T. F. and T. A. Bayoumi (2003). 'Are They All in the Same Boat? The 2000-2001 Growth Slowdown and the G-7 Business Cycle Linkages', Manuscript, IMF.

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(4), pp. 747-65.

Inklaar, R. and J. de Haan (2001). 'Is There Really a European Business Cycle?: A Comment' , *Oxford Economic Papers*, Vol. 53, pp. 215-20.

James, A. T. (1954). 'Normal multivariate analysis and the orthogonal group', *Analysis of Mathematical Statistics*, Vol. 25, pp. 40-75.

Johansen, S. (1992). 'Testing weak exogeneity and the order of cointegration in UK money demand data', *Journal of Policy Modeling*, Vol. 14(3), pp. 313-34.

Kapetanios, G., Y. Shin and A. Snell (2006). 'Testing for cointegration in nonlinear smooth transition error correction models', *Econometric Theory*, Vol. 22, pp. 279-303.

King, R. G. and M. W. Watson (1997). 'Testing long-run neutrality', *Economic Quarterly*, Federal Reserve Bank of Richmond, Vol. 83, pp. 69-101.

Koop, G., R. Leon-Gonzalez and R. W. Strachan (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 Apr 2006.

Koop, G., R. Leon-Gonzalez and R. W. Strachan (2006). 'Bayesian inference in a cointegration panel data model', Discussion Papers in Economics 06/2, Department of Economics, University of Leicester.

Koop, G., M. H. Pesaran and S. M. Potter (1996). 'Impulse Response Analysis in Nonlinear Multivariate Models', *Journal of Econometrics*, Vol. 74, pp. 491-99.

Koop, G. and S. M. Potter (1999a). 'Bayes factors and nonlinearity: Evidence from economic time series', *Journal of Econometrics*, Vol. 88, pp. 251-81.

Koop, G. and S. M. Potter (1999b). 'Dynamic Asymmetries in US Unemployment', *Journal of Business and Economic Statistics*, Vol. 17, pp. 298-313.

Koop, G. and S. M. Potter (2000). 'The Vector Floor and Ceiling Model', Working Paper No 04/15, Department of Ecomomics, University of

Leicester.

Koop, G., S. M. Potter and R. W. Strachan (2005). 'Re-examining the consumption-wealth relationship: the role of model uncertainty', Discussion Papers in Economics 05/3, Department of Economics, University of Leicester.

Koop, G., R. W. Strachan, H. van Dijk and M. Villani (2006). 'Bayesian Approaches to Cointegration', in T. Mills and K. Patterson (*eds*) The Palgrave Handbook of Econometrics, Volume 1: Theoretical Econometrics, Palgrave-Macmillan: Basingstoke.

Leamer, E. E. (1978). 'Specification Searches', John Wiley, New York.

Leeper, E. M. and T. Zha (2003). 'Modest policy interventions', *Journal of Monetary Economics*, Vol. 50(8), pp. 1673-1700.

LeSage, J. (1999). 'Applied econometrics using MATLAB', *http://www.spatialeconometrics.com/.*

Li, K. (1999). 'Testing symmetry and propotionality in PPP: a panel-data approach', *Journal of Business and Economic Statistics*, Vol. 17, pp. 409-18.

Lopez, C., C. J. Murray and D. H. Papell (2005). 'State of the Art Unit Root Tests and Purchasing Power Parity', *Journal of Money, Credit and Banking*, Vol. 37(2), pp. 361-69.

Lopez, C. and D. H. Papell (2006). 'Convergence to Purchasing Power Parity at the Commencement of the Euro', *Review of International Economics*, Vol. 14, pp. 1-16.

Lothian, J. R. (1997). 'Multi-counry evidence on the behaviour of purchasing power parity under the current float', *Journal of International Money and Finance*, Vol. 16, pp. 19-35.

Lothian, J., R. and M. P. Taylor (1996). 'Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries', *Journal of Political Economy*, Vol. 104(3), pp. 488-509.

Lubrano, M. (1999a). 'Bayesian Analysis of Nonlinear Time Series Models with a Threshold', *Nonlinear Econometric Modelling*, Cambridge University Press, Cambridge.

Lubrano, M. (1999b). 'Smooth Transition GARCH Models: A Bayesian Perspective', Universite Aix-Marseille III G.R.E.Q.A.M. 99a49.

Lütkepohl, H., T. Teräsvirta and J. Wolters (1999). 'Investigating Stability and Linearity of a German M1 Money Demand Function', *Journal of Applied Econometrics*, Vol. 14(5), pp. 511-25.

Mark, N. C. (1990). 'Real and nominal exchange rates in the long run: an empirical investigation', *Journal of International Economics*, Vol. 28, pp. 115-36.

Meese, R. A. and K. S. Rogoff (1988). 'Was it real? The exchange-rate interest differential relation over the modern floating rate period', *Journal of Finance*, Vol. 43, pp. 933-48.

Meltzer, A. H. (2001). 'The Transmission Process', in: Deutsche Bundesbank (Hrsg.), The Monetary Transmission Process - Recent Developments and Lessons for Europe, Basingstoke, Palgrave, pp. 112-30.

Michael, P., A. R. Nobay and D. A. Peel (1997). 'Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation', *Journal of Political Economy*, Vol. 105(4), pp. 862-79.

Muirhead, R. J. (1982). 'Aspects of multivariate statistical theory', Wiley: New York.

Nelson, E. (2002). 'Direct effects of base money on aggregate demand: theory and evidence', *Journal of Monetary Economics*, Vol. 49(4), pp. 687-708.

Nelson, E. (2003). 'The future of monetary aggregates in monetary policy analysis', *Journal of Monetary Economics*, Vol. 50(5), pp. 1029-59.

Ni, S. X. and D. Sun (2003). 'Noninformative Priors and Frequentist Risks of Bayesian Estimators of Vector-Autoregressive Models', *Journal of Econometrics*, Vol. 115, pp. 159-97.

Norrbin, S. C. and D. E. Schlagenhauf (1996). 'The role of International

Factors in the Business Cycle: A Multi-country Study', *Journal of International Economics*, Vol. 40, pp. 85-104.

O'Connell, P.G. (1998). 'The overvaluation of Purchasing Power Parity', *Journal International Economics*, Vol. 44, pp. 1-19.

O'Hagan, A. (1995). 'Fractional Bayes factors for model comparison', *Journal of the Royal Statistical Society*, B 57, pp. 99-138.

Osborn, D. R., P. J. Perez and M. Sensier (2005). 'Business Cycle Linkages for the G7 Countries: Does the US Lead the World?', Disscussion Paper 50, Center for Growth and Business Cycle Research, University of Manchester.

Öcal, N. and D. R. Osborn (2000). 'Business Cycle Non-linearities in UK Consumption and Production', *Journal of Applied Econometrics*, Vol. 15, pp. 27-43.

Peel, D. A. and I. A. Venetis (2005). 'Smooth Transition Models and Arbitrage Consistency ', *Economica*, Vol. 72, pp. 413-30.

Perez, P., D. Osborn and M. Artis (2006). 'The International Business Cycle in a Changing World: Volatility and the Propagation of Shocks in the G-7', *Open Economies Review*, Springer, Vol. 17(3), pp. 255-279.

Potter, S. (1994). 'Nonlinear impulse response functions', Department of Economics working paper (University of California, Los Angeles, CA).

Potter, S. (1995). 'A Nonlinear Approach to U.S. GNP', *Journal of Applied Econometrics*, Vol. 10, pp. 109-25.

Primiceri, G. (2005). 'Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy', NBER Working Papers 11147.

Ritter, C. and M. A. Tanner (1992). 'Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler', *Journal of the American Statistical Association*, Vol. 87, pp. 861-68.

Rothman, P., D. van Dijk and P. H. Franses (2001). 'Multivariate Star Analysis of Money-Output Relationship', *Macroeconomic Dynamics*, Cambridge University Press, Vol. 5(4), pp. 506-32.

Sargent, T., N. Williams and T. Zha (2006). 'Shocks and Government Beliefs: The Rise and Fall of American Inflation', *American Economic Review*, Vol. 96(4), pp. 1193-224.

Sarno, L. and M. P. Taylor(1998). 'Real exchange rate under the recent float: Unequivacal evidence of mean reversion', *Economic Letters*, Vol. 60, pp. 131-37.

Sarno, L., M. P. Taylor and I. Chowdhury (2004). 'Nonlinear dynamics in deviations from the law of one price: a broad-based empirical study', *Journal of International Money and Finance*, Vol. 23, pp. 1-25.

Sensier, M., Osborn, D. R. and N. Öcal (2002). 'Asymmetric Interest Rate

Effects for the UK Real Economy', *Oxford Bulletin of Economics and Statistics*, Vol. 64, pp.315-39.

Seo, B. (2004). 'Testing for nonlinear adjustment in smooth transition vector error correction models, Econometric Society', 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. 129-50.

Sercu, P., R. Uppal and C. Van Hulle (1995). 'The exchange rate in the presence of transaction costs: implications for tests of purchasing power parity', *Journal of Finance*, Vol. 50, pp. 1309-19.

Sims, C. (1972). 'Money, Income and Causality', *American Economic Review*, Vol. 62, pp. 540-52.

Sims, C. (1980). 'Macroeconomics and Reality', *Econometrica*, Vol. 48, pp. 1-48.

Sims, C. A. and T. Zha (2006). 'Were There Regime Switches in U.S. Monetary Policy?' *American Economic Review*, Vol. 96(1), pp. 54-81.

Smith, P. A. and P. M. Summers (2005). 'How will do Markov switching models describe actual business cycles? The case of synchronization', *Journal of Applied Econometrics*, Vol. 20, pp. 253-74.

Stock, J. H. and M. W. Watson (1989). 'Interpreting the Evidence on Money-Income Causality', *Journal of Econometrics*, Vol. 40, pp. 161-81.

Strachan, R. W. (2003). 'Valid bayesian estimation of the cointegrating error correction model', *Journal of Business and Economic Statistics*, Vol. 21(1), pp. 185-95.

Strachan, R. W. and B. Inder (2004). 'Bayesian analysis of the error correction model', *Journal of Econometrics*, Vol. 123, pp. 307-25.

Strachan, R. W. and H. K. van Dijk (2004a). 'Bayesian model selection with an uninformative prior', Keele Economics Research Papers KERP 2004/01, Centre for Economic Research, Keele University.

Strachan, R. W. and H. K. van Dijk (2004b). 'Exceptions to Bartlett's Paradox', Keele Economics Research Papers KERP 2004/03, Centre for Economic Research, Keele University.

Strachan, R. W. and H. K. van Dijk (2006). 'Model Uncertainty and Bayesian Model Averaging in Vector Autoregressive Processes', Discussion Papers in Economics 06/5, Department of Economics, University of Leicester.

Sugita, K. (2006). 'Bayesian analysis of Markov switching vector error correction model', Discussion Paper /2006/13, Graduate school of economics, Hitotsubashi University.

Swanson, N. R. (1998). 'Money and output viewed through a rolling window', *Journal of Monetary Economics*, Vol. 41(3), pp. 455-74.

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(425), pp. 208-18.

Teräsvirta, T. and A-C. Eliasson (2001). 'Non-linear error correction and the UK demand for broad money, 1878-1993', *Journal of Applied Econometrics*, Vol. 16(3), pp. 277-88.

Tong, H. (1978). 'On a threshold model', in C. H. Chen (*eds*) Pattern Recognition and Signal Processing, Amsterdam: Sijhoff & Noordhoff.

Tong, H. (1983). 'Threshold Models in Non-linear Time Series Analysis', Springer-Verlag, New York.

van Dijk, D., T. Teräsvirta and P.H. Franses (2002). 'Smooth Transition Autoregressive Models - A Survey of Recent Developments', *Econometric Reviews*, Vol. 21(1), pp. 1-47.

Villani, M. (2005). 'Bayesian reference analysis of cointegration', *Econometric Theory*, Vol. 21, pp. 326-57.

Wang, P. and Y. Wen (2005). 'Endogenous money or sticky prices?comment on monetary non-neutrality and inflation dynamics', *Journal of Economic Dynamics and Control*, Vol. 29, pp. 1361-83.

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.

Wynne, M. and J. Koo (2000). 'Business cycles under monetary union: A comparison of the EU and the US', *Economica*, Vol. 67, pp. 347-74.

Zellner, A. (1971). 'An introduction to Bayesian inference in econometrics', John Wiley and Sons: New York.