

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

Deborah Gefang*

Department of Economics
University of Leicester
UK

November 12, 2007

Abstract

This paper proposes a Bayesian procedure to investigate the purchasing power parity (PPP) utilizing an exponential smooth transition Vector Error Correction Model (VECM). 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 nonlinear 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.

*I would like to thank the helpful comments from Gary Koop, Roberto Leon-Gonzalez, Rodney Strachan and the participants at 6th Annual Advances in Econometrics Conference at LSU. This paper also has benefited from the productive discussions with Roderick McCrorie and Emi Mise. Any remaining errors are the author's responsibility. Financial support from the Department of Economics, University of Leicester is gratefully acknowledged.

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 PPP, multivariate models, especially nonlinear vector error correction mod-

¹Note 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, Kai Li (1999) proposes a system of equations model with hierarchical priors to surmount the problems associated panel data unit root tests.

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

els (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 paper proposes a Bayesian approach to investigate PPP within an exponential smooth transition VECM (ESVECM) framework. Specifically, we follow the Bayesian cointegration space approach introduced by Strachan and Inder (2004) and the Bayesian logistic smooth transition Vector Autoregressive (LSVAR) approach of Gefang and Strachan (2007).³ Our method jointly captures the equilibrium and the presence of nonlinearity in the ESVECM 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 ESVECM. Yet, jagged likelihood functions do not create any particular problems in our 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

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

era. We take account of model uncertainty through Bayesian model selection and Bayesian model averaging. The posterior model probabilities are derived from the Bayes factor approach of Koop and Potter (1999) and the Bayesian information criterion (Schwarz, 1978)(BIC) approximation approach of Kass and Raftery (1995), respectively. 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 the magnitudes of the original shocks.

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

2 The Exponential Smooth Transition VECM

Under the relative PPP, the nominal exchange rates and the 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 the threshold VECM (TVECM) and the ESVECM. 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. While in TVECM or ESVECM, 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 transition function. In TVECM, the regime changes are assumed to be discrete, whereas in the ESVECM, the regimes change smoothly. Since the market force driving PPP adjustment is an aggregated process, following the suggestions of Teräsvirta (1994), we use ESVECM to model the nonlinear convergence towards PPP between two countries. ESVECM 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 the 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 $t=1, \dots, T$ as follows.

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

where $\Delta y_t = y_t - y_{t-1}$. The error term ε_t is a white noise or innovation process, that is $E(\varepsilon_t) = 0$, $E(\varepsilon_s'\varepsilon_t) = \Sigma$ for $s = t$, and $E(\varepsilon_s'\varepsilon_t) = 0$ for $s \neq t$.

The dimensions of Γ_h and Γ_h^z are 3×3 , and the dimensions of β , α' , and α^z are $3 \times r$, where r is the rank of the cointegration vector β . If PPP holds, the value of r should be equal to 1.

In model (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):

$$F(z_t) = 1 - \exp(-\gamma(z_t - c)^2) \quad (2)$$

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; γ is the smooth parameter that governs the speed of the transition process between extreme regimes, with higher values of γ implying faster transition.

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 $z_t - c \rightarrow \pm\infty$. As convention, we define $F(z_t) = 0$ and $F(z_t) = 1$ corresponding to the middle and the outer regimes, respectively. In the middle regime, model (1) becomes a linear VECM, with the adjustment process governed by (α, ξ, Γ_h) ; while in the outer regime, model (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 the PPP adjustment is determined

⁴Note 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.

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 (1) allows a set of models which vary in the rank of the cointegration vector (from 0 to 3), the order of the autoregressive process, the lag length of the transition variable, and the presence of the nonlinearity.

2.1 The Likelihood Function

For notational convenience, we can re-write model (1) as

$$\Delta y_t = x_{1,t-1}\beta\alpha + x_{2,t}\Phi + F(z_t)(x_{1,t-1}\beta\alpha^z + x_{2,t}\Phi^z) + \varepsilon_t \quad (3)$$

where $x_{1,t-1} = y_{t-1}$, $x_{2,t} = (d_t, \Delta y_{t-1}, \dots, \Delta y_{t-p})$, $\Phi = (\xi', \Gamma'_1, \dots, \Gamma'_p)'$, $\Phi^z = (\xi^{z'}, \Gamma_1^{z'}, \dots, \Gamma_p^{z'})'$.

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

$$X_0 = X_1\beta\alpha + X_2\Phi + F^z X_1\beta\alpha^z + F^z X_2\Phi^z + E = XB + E \quad (4)$$

The likelihood function of model (4) is following.

$$L(y|\Sigma, B, \beta) \propto |\Sigma|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2} \text{tr} \Sigma^{-1} E' E\right\} \quad (5)$$

Vectorizing model (4), we transform model (1) into

$$x_0 = xb + e \quad (6)$$

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

$$\begin{aligned} \text{tr} \Sigma^{-1} E' E &= e' (\Sigma^{-1} \otimes I_T) e \\ &= s^2 + (b - \hat{b})' V^{-1} (b - \hat{b}) \end{aligned} \quad (7)$$

where $s^2 = x_0' M_v x_0$, $M_v = \Sigma^{-1} \otimes [I_T - X(X'X)^{-1}X']$, $\hat{b} = [I_n \otimes (X'X)^{-1}X']x_0$ and $V = \Sigma \otimes (X'X)^{-1}$. Thus, the likelihood function in equation (5) can be re-written as

$$L(y|\Sigma, B, \beta) \propto |\Sigma|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2} [s^2 + (b - \hat{b})' V^{-1} (b - \hat{b})]\right\} \quad (8)$$

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 β 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 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 problem. As well documented in the literature, a linear VECM suffers from both the global and local nonidentification of the cointegration vectors and the 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 β , researchers first normalize β into $\beta = [I_r \ V']'$, then impose uninformative prior on the sub-vector V . However, as argued by Strachan and van Dijk (2004), this approach has an undesirable side-effect that it 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 (1999) for further explanation] becomes more serious in Bayesian context than in the classical framework. Bauwens et al (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 γ as Gamma.

The nonidentification problem faced by the ESVECM model is slightly different. Although the Davies' problem remains relatively the same as in the 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 β , α , α^z and γ sequentially once we choose a way to normalize β .

Following the suggestions in Koop, Strachan, van Dijk and Villani (2006), we elicit the prior of β directly from the cointegration space, in particular, we adopt the approach developed in Strachan and Inder (2004).⁵ However, despite following Strachan and Inder (2004) to elicit uninformative priors for the cointegration space, we set *pseudo*-uninformative priors for the space of β . Our prior differs from that of Strachan and Inder (2004) in two aspects.

⁵To 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.

First, we only explore the situation where the rank of the cointegration vector is 1. Second, to save computing time, we elicit uninformative priors from the cointegration 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.

It is simple to set our *pseudo*-uninformative priors of β based on the uninformative priors proposed in Strachan and Inder (2004). First, we specify the space of the three by one vector β to be uniformly distributed over the two dimensional Grassman manifold $G_{1,2}$.

$$p(\beta) = \frac{1}{c_1^3} \quad (9)$$

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). Since the rank of β is 1, we can use polar coordinates to denote the semi-orthogonal β 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 β in (9) by $\sin(\theta_1)$, the Jacobian of the transformation (Muirhead, 1982). As explained before, we want to restrict the signs of the elements in β instead of allowing

the space of β to move freely. The objective can be easily achieved by restricting the range of θ_1 to be from $\pi/2$ to π , and the range of θ_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 Σ :

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

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; 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 $b = \text{vec}(B)$, $\underline{V} = \Sigma \otimes \eta^{-1} I_k$, $k = 2(r + 1 + np)$, η is the shrinkage prior as proposed by Ni and Sun (2003). As practiced in Koop, Leon-Gonzalez and Strachan (2006), we draw η from the Gibbs sampler. In our case, we set the relatively uninformative prior distribution of η 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 α and α^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 γ 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 γ does not render model (1) in favor of the nonlinear effect. In the empirical work, we set the prior distribution of γ as Gamma(1, 0.0001) to let the data speak.

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 β).

2.3 Posterior Computation

We use full conditional Gibbs sampler for posterior computations. From the priors just elicited and the likelihood function as derived before, we find that the posterior of Σ 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 $\bar{V} = \Sigma \otimes (X'X + \eta I_k)^{-1}$. Note that the posterior distributions of θ_1 , θ_2 , γ and c are not of any standard form. However, the ranges of θ_1 and θ_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 θ_1 and θ_2 . Thus, we can use Griddy Gibbs Sampling introduced in Ritter and Tanner (1992) to draw θ_1 , θ_2 and c within the main Gibbs Sampler.

With respect to γ , 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 η as well. The conditional posterior of η is calculated as

$$p(\eta|B, \Sigma, \gamma, c, Y, X) \propto p(\eta)|\Sigma^{-1} \otimes \eta I_k|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} b' (\Sigma^{-1} \otimes \eta I_k) b \right\} \quad (10)$$

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

The Gibbs Sampling Scheme can be summarized as follows:

1. Initialize $(b, \Sigma, \theta_1, \theta_2, \gamma)$;
2. Draw $\Sigma|b, \theta_1, \theta_2, \gamma$ from $IW(E'E, T)$;
3. Draw $b|\Sigma, \theta_1, \theta_2, \gamma$ from $N(b, \overline{V}_b)$;
4. Draw $\theta_1, \theta_2|\Sigma, b, \gamma$ numerically by Griddy Gibbs;
5. Draw $\gamma|\Sigma, b, \theta_1, \theta_2$, through Metropolis-Hastings method;
6. Draw c numerically;
7. Repeat steps 2 to 6 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 two approaches, namely the Bayes factor approach of Koop and Potter (1999) and the BIC approximation approach of Kass and Raftery (1995).

As explained in Koop and Potter (1999), by penalizing parameter rich models, the Bayes factor approach can resolve the over fitting problems that generally exist in nonlinear models. Following Koop and Potter (1999) and Koop, Leon-Gonzalez and Strachan (2006), we use Savage-Dickey density ratio (Verdinelli and Wasserman, 1995) to compute the Bayes factors comparing every restricted model nested within the general model (1) with the general model itself. Using this information, we back up 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 α^z and Φ^z are equal to zero. Likewise, the restricted linear VAR model with neither the cointegration nor the nonlinear effect occurs when we impose all the elements of α , α^z and Φ^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 :

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

where the restrictions are $b_i = 0$. Note that this method penalize parameter

⁶It 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 $\alpha = 0$.

rich models as explained in Koop and Potter (1999).

With regards to the second approach, as illustrated by Kass and Raftery (1995), using BIC to approximate posterior odds ratio can not produce correct values even for very large samples. However, the approach is appealing in many aspects. For example, it can produce reasonable results in large samples, and it is easy to implement [see Kass and Raftery (1995) and Koop, Potter and Strachan (2005) for further illustrations]. Similar to the procedure in the Bayes factor approach, we integrate BIC calculation into the Gibbs sampler, then back up the posterior model probabilities by comparing the BIC of each possible model with that of a common base model.

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 the 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. For all series, 1973:1 is set as the base period.

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 Diag-

nostic measure introduced by Geweke (1992). We use the MATLAB program from LeSage’s Econometrics Toolbox (LeSage, 1999) for the diagnostic. Information on the parameter space is presented in tables 1-3. 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 the corresponding elements in the cointegration vector are reported.

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 (ESVAR) and 36 ESVECM models] for each country pair. 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 4 summarizes the total posterior probabilities of the models based on the two model comparison methods. Based on the Bayes factor approach, in all cases, ESVECM models receive over 90% of the posterior model probabilities, which provides strong evidence that PPP holds, and the adjustment process towards PPP is nonlinear. When employing the BIC approach, we observe that the ESVECM models are not as dominant as in the Bayes factor approach. However, since the ESVECM and the ESVAR models jointly

account for more than 80% of the posterior model probabilities across all samples, the existence of PPP and the nonlinear effects triggered by the deviations from PPP are also evident in the BIC approach. Overall, both the Bayes factor and the BIC put most probability on nonlinear models, suggesting it is improper to model the interrelationship among the nominal interest rate and the domestic and foreign price levels in a linear framework.

It may also be illuminating to look into the support for the VECM and the VAR models in the linear context. Using the Bayes factor, we find the linear VECM models are more favored over the linear VAR models in all countries, especially in the cases of US-GER and US-UK. However, the results from the BIC indicate that in all cases, the linear VECM and the linear VAR models receive relatively the same amount of posterior model probabilities, which would cast doubt over the validity of PPP in our sample if we were to choose to neglect the nonlinear effects in the first instance.

To provide information on the degree of model uncertainty, in table 4 we also report the total posterior mass of the top 20 models. Observe that there is a great deal of model uncertainty when we resort to the BIC to calculate the posterior model probabilities.

Table 5 contains results of the sum of the posterior probabilities of the ESVECM models distinguished by the transition variables for each country pair. Using BIC approach, we find that in all cases, the posterior probabilities received by each transition indicator are relatively the same. However, the model comparison results derived from the Bayes factors show that the transition indicator 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 6-8. In all cases, observe that except for a few exceptions, all the top 20 models chosen by both the Bayes factor and the BIC are nonlinear. Observe that there is substantial model uncertainty when the BIC is employed. As a matter of fact, there is little evidence in favor of any single model. Turning to the results calculated from the Bayes factor, we also find model uncertainty, nonetheless less obvious than what we find when using the BIC approach.

When we employ the Bayes factor, 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 the 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-DEU, 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 US-UK. 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%.

Given the deferent degrees of model uncertainty found via the Bayes factor and the BIC, we report the smooth transition functions of the most probable model based on the Bayes factor in figure 1, and the smooth transition functions derived by Bayesian model averaging (BMA) based on the BIC in figure 2. 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.

In figure 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 smooth transition functions, 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. For comparison, we plot the BMA based smooth transition functions in figure 2. Observe that the corresponding graphs in figures 1-2 display a relatively similar pattern. However, since BMA based on the BIC gives roughly equal weights across different models, the range of the transition functions becomes narrower in figure 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 GIP for a shock, v_t , and a history, ω_{t-1} as follows

$$GIP(n, v_t, \omega_{t-1}) = E[P_{t+n|v_t, \omega_{t-1}}] - E[P_{t+n|\omega_{t-1}}] \quad (11)$$

where n is the time horizon. By averaging out the future shocks, in (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 the model uncertainties.

To examine the impulse response functions of PPP, we allow a shock amounting to ± 0.01 and ± 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, 12 different shocks and two methods for model comparison (based on the Bayes factor and the BIC). For brevity, we only present the impulse response functions of PPP for the most preferred models in figures 3-5.⁷ Inspecting the impulse response functions, we have two main findings.

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

⁷The full set of graphs are available on request.

- 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 the PPP adjustment.⁸ As shown in our study, the impacts on PPP 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.⁹

4 Conclusion

In this paper, we introduce a Bayesian approach for estimating an ESVECM 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 long run PPP holds, and the adjustment to PPP is a nonlinear process with the regime changes governed by the magnitude of the deviations. Furthermore, our research casts doubt over

⁸The 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).

⁹Chortareas 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.

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

References

- [1] Bartlett, M. S. (1957), A comment on D. V. Lindley's statistical paradox, *Biometrika*, 44, 533-34.
- [2] 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*, 20, 379-99.
- [3] Bauwens, L., M. Lubrano and J-F. Richard (1999), Bayesian inference in dynamic econometric models, Oxford University Press: New York.
- [4] Cheung, Y. W. and K. S. Lai (1994), Mean reversion in real exchange rates, *Economics Letters*, 46(3),251-56.
- [5] Chib, S. and E. Greenberg (1995), Understanding the Metropolis-Hastings algorithm, *The American Statistician* 49(4), 327-35.
- [6] 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.
- [7] Davies, R. B. (1977), Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* 74, 33-43.
- [8] Dumas, B. (1992), Dynamic equilibrium in a spatially separated world, *Review of Financial Studies*, 5(2), 153-80.
- [9] Edison, H. J. and E. O. Fisher (1991), A long-run view of the European monetary system, *Journal of International Money and Finance*, 10, 53-70.
- [10] 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.
- [11] Geweke, J. (1992), Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, in Bernardo, J.,

- Berger, J., Dawid, A. and Smith, A. (*eds.*) Bayesian Statistics, 4, 641-49, Oxford: Clarendon Press.
- [12] James, A. T. (1954), Normal multivariate analysis and the orthogonal group, *Analysis of Mathematical Statistics*, 25, 40-75.
- [13] Kapetanios, G., Y. Shin and A. Snell (2006), Testing for cointegration in nonlinear smooth transition error correction models, *Econometric Theory*, 22, 279-303.
- [14] Kass, R. and A. Raftery (1995), Bayes Factors, *Journal of the American Statistical Association*, 90, 773-95.
- [15] 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.
- [16] Koop, G., M. H. Pesaran and S. M. Potter (1996), Impulse Response Analysis in Nonlinear Multivariate Models, *Journal of Econometrics*, 74, 491-99.
- [17] Koop, G. and S. M. Potter (1999), Bayes factors and nonlinearity: Evidence from economic time series, *Journal of Econometrics*, 88, 251-81.
- [18] 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.
- [19] 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
- [20] LeSage, J. (1999), Applied econometrics using MATLAB. <http://www.spatial econometrics.com/>.

- [21] Li, K. (1999), Testing symmetry and proportionality in PPP: a panel-data approach, *Journal of Business and Economic Statistics*, 17, 409-18.
- [22] 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*, 37(2), 361-69.
- [23] Lopez, C. and D. H. Papell (2006), Convergence to Purchasing Power Parity at the Commencement of the Euro, *Review of International Economics*, 14, 1-16.
- [24] Lothian, J. R. (1997), Multi-country evidence on the behaviour of purchasing power parity under the current float, *Journal of International Money and Finance*, 16, 19-35.
- [25] 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*, 104(3), 488-509.
- [26] Lubrano, M. (1999a), Bayesian Analysis of Nonlinear Time Series Models with a Threshold, In: *Nonlinear Econometric Modelling*, Cambridge University Press, Cambridge.
- [27] Lubrano, M. (1999b), Smooth Transition GARCH Models: A Bayesian Perspective. Universite Aix-Marseille III G.R.E.Q.A.M. 99a49.
- [28] Mark, N. C. (1990), Real and nominal exchange rates in the long run: an empirical investigation, *Journal of International Economics*, 28, 115-36.
- [29] 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*, 43, 933-48.
- [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*, 105(4), 862-79.

- [31] Muirhead, R. J. (1982), Aspects of multivariate statistical theory, Wiley: New York.
- [32] Ni, S. X. and D. Sun (2003), Noninformative Priors and Frequentist Risks of Bayesian Estimators of Vector-Autoregressive Models, *Journal of Econometrics*, 115, 159-97.
- [33] O'Connell, P.G.(1998), The overvaluation of Purchasing Power Parity, *Journal International Economics*, 44, 1-19.
- [34] O'Hagan, A. (1995), Fractional Bayes factors for model comparison, *Journal of the Royal Statistical Society*, B 57, 99-138.
- [35] Peel, D. A. and I. A. Venetis (2005), Smooth Transition Models and Arbitrage Consistency , *Economica*, 72, 413-30.
- [36] Potter, S. (1994), Nonlinear impulse response functions, Department of Economics working paper (University of California, Los Angeles, CA).
- [37] 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*, 87, 861-68.
- [38] Sarno, L. and M. P. Taylor(1998), Real exchange rate under the recent float: Unequivocal evidence of mean reversion, *Economic Letters*, 60, 131-37.
- [39] 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*, 23, 1-25.
- [40] 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*, 50, 1309-19.
- [41] Seo, B. (2004), Testing for nonlinear adjustment in smooth transition vector error correction models, Econometric Society, Far Eastern Meetings.

- [42] Seo, M. (2006), Bootstrap testing for the null of no cointegration in a threshold vector error correction model, *Journal of Econometrics*, 143, 129-50.
- [43] Strachan, R. W. (2003), Valid bayesian estimation of the cointegrating error correction model, *Journal of Business and Economic Statistics*, 21(1), 185-95.
- [44] Strachan, R. W. and B. Inder (2004), Bayesian analysis of the error correction model, *Journal of Econometrics*, 123, 307-25.
- [45] Strachan, R. W. and H. K. van Dijk (2004), Bayesian model selection with an uninformative prior, Keele Economics Research Papers KERP 2004/01, Centre for Economic Research, Keele University.
- [46] 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.
- [47] Sugita, K. (2006), Bayesian analysis of Markov switching vector error correction model, Discussion Paper /2006/13, Graduate school of economics, Hitotsubashi University.
- [48] Schwarz, G. (1978), Estimating the dimension of a model, *Annals of Statistics*, 6:2, 461-64.
- [49] Teräsvirta, T. (1994), Specification, estimation and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association* , 89(425), 208-18.
- [50] Verdinelli, I. and L. Wasserman (1995), Computing Bayes factors using a generalization of the Savage-Dickey density ratio, *Journal of the American Statistical Association*, 90, 614-18.
- [51] Villani, M. (2005), Bayesian reference analysis of cointegration, *Econometric Theory*, 21, 326-57.

- [52] Zellner, A. (1971), *An introduction to Bayesian inference in econometrics*, John Wiley and Sons: New York.

Table 1: Parameters (a)

US - CAN						
	d=1	d=2	d=3	d=4	d=5	d=6
γ	12.6520 <i>3.5105</i>	13.7470 <i>4.1453</i>	13.3550 <i>4.0240</i>	15.5270 <i>5.2090</i>	13.8080 <i>4.0288</i>	13.9260 <i>3.4595</i>
θ_1	2.3150 <i>0.0537</i>	2.0431 <i>0.0865</i>	2.0648 <i>0.0685</i>	1.9085 <i>0.0790</i>	1.9747 <i>0.0930</i>	1.9319 <i>0.0373</i>
θ_2	1.4372 <i>0.0805</i>	1.5365 <i>0.0246</i>	1.4857 <i>0.0519</i>	1.5337 <i>0.0316</i>	1.5171 <i>0.0463</i>	1.5283 <i>0.0309</i>
c	-0.2820 <i>0.4604</i>	1.3063 <i>0.5750</i>	1.3917 <i>0.4554</i>	2.2738 <i>0.5565</i>	1.8870 <i>0.5909</i>	2.1457 <i>0.2492</i>
η	2.2817 <i>1.1317</i>	2.1475 <i>1.0957</i>	2.0392 <i>1.1172</i>	2.0671 <i>1.0442</i>	1.9965 <i>1.0906</i>	1.8482 <i>0.9917</i>
β_1	0.7291	0.8900	0.8773	0.9429	0.9182	0.9346
β_2	0.0980	0.0305	0.0749	0.0350	0.0493	0.0398
β_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
γ	30.4830 <i>4.8129</i>	28.7710 <i>4.3462</i>	0.1131 <i>0.3682</i>	25.9230 <i>6.5391</i>	0.0339 <i>0.1182</i>	0.0382 <i>0.1242</i>
θ_1	3.0893 <i>0.0758</i>	3.0240 <i>0.0498</i>	2.2722 <i>0.0480</i>	3.0322 <i>0.0954</i>	2.2540 <i>0.0846</i>	2.3632 <i>0.0698</i>
θ_2	1.0376 <i>0.4789</i>	1.0875 <i>0.3591</i>	0.4227 <i>0.1585</i>	1.0655 <i>0.4284</i>	0.5105 <i>0.1852</i>	0.6029 <i>0.1697</i>
c	-4.9709 <i>0.4376</i>	-4.4999 <i>0.3262</i>	1.7649 <i>0.4055</i>	-4.5704 <i>0.5818</i>	1.9835 <i>0.6248</i>	1.1213 <i>0.6242</i>
η	0.0857 <i>0.0514</i>	0.1187 <i>0.0595</i>	0.0008 <i>0.0002</i>	0.1298 <i>0.0700</i>	0.0008 <i>0.0002</i>	0.0008 <i>0.0002</i>
β_1	0.0450	0.1039	0.3134	0.0956	0.3790	0.3981
β_2	0.0266	0.0545	0.6967	0.0529	0.6767	0.5783
β_3	-0.9986	-0.9931	-0.6453	-0.9940	-0.6313	-0.7121

Notes: Standard deviations are in italics.

Table 2: Parameters (b)

US - GER						
	d=1	d=2	d=3	d=4	d=5	d=6
γ	0.0136 <i>0.0653</i>	0.9400 <i>1.8586</i>	0.6267 <i>1.1395</i>	9.0952 <i>1.6984</i>	0.0517 <i>0.1453</i>	14.4090 <i>2.9474</i>
θ_1	1.8754 <i>0.0499</i>	1.8200 <i>0.0609</i>	1.9436 <i>0.0114</i>	1.8318 <i>0.1280</i>	1.9217 <i>0.0355</i>	1.9702 <i>0.0033</i>
θ_2	0.0186 <i>0.0109</i>	0.0180 <i>0.0112</i>	0.0130 <i>0.0057</i>	0.0365 <i>0.0347</i>	0.0139 <i>0.0074</i>	0.0384 <i>0.0255</i>
c	3.2665 <i>0.3401</i>	3.6215 <i>0.3841</i>	2.7841 <i>0.0835</i>	3.5507 <i>0.8344</i>	2.9426 <i>0.2554</i>	2.6839 <i>0.1183</i>
η	0.1382 <i>0.0797</i>	0.2079 <i>0.1278</i>	0.0683 <i>0.0324</i>	0.3126 <i>0.3172</i>	0.0883 <i>0.0591</i>	0.0952 <i>0.0494</i>
β_1	0.0177	0.0174	0.0121	0.0353	0.0131	0.0353
β_2	0.9538	0.9690	0.9312	0.9655	0.9390	0.9206
β_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
γ	4.3500 <i>2.9948</i>	7.8474 <i>3.8444</i>	42.2050 <i>8.1170</i>	6.4511 <i>2.2292</i>	30.5660 <i>2.0644</i>	3.4249 <i>1.7186</i>
θ_1	2.4866 <i>0.0076</i>	2.4649 <i>0.0033</i>	3.0858 <i>0.0451</i>	2.5294 <i>0.0048</i>	3.0539 <i>0.0820</i>	2.4769 <i>0.0187</i>
θ_2	0.0615 <i>0.0590</i>	0.0487 <i>0.0409</i>	0.8980 <i>0.4882</i>	0.0248 <i>0.0158</i>	1.1982 <i>0.3517</i>	0.0339 <i>0.0362</i>
c	-0.3737 <i>0.0732</i>	-0.2265 <i>0.1011</i>	-4.8580 <i>0.3030</i>	-0.8063 <i>0.0546</i>	-4.6900 <i>0.5221</i>	-0.3258 <i>0.1963</i>
η	0.0012 <i>0.0025</i>	0.0020 <i>0.0033</i>	0.0924 <i>0.2001</i>	0.0007 <i>0.0020</i>	0.0959 <i>0.1975</i>	0.0016 <i>0.0024</i>
β_1	0.0374	0.0305	0.0436	0.0143	0.0816	0.0209
β_2	0.6080	0.6255	0.0347	0.5745	0.0319	0.6165
β_3	-0.7931	-0.7797	-0.9985	-0.8184	-0.9962	-0.7871

Notes: See notes in Table 1.

Table 3: Parameters (c)

US - JAP						
	d=1	d=2	d=3	d=4	d=5	d=6
γ	22.7700 <i>32.2510</i>	34.7190 <i>7.3199</i>	85.1300 <i>59.0440</i>	15.2390 <i>6.5107</i>	1.5974 <i>2.9569</i>	48.3000 <i>11.0160</i>
θ_1	2.6538 <i>0.0856</i>	1.9233 <i>0.0206</i>	1.8639 <i>0.0517</i>	2.1151 <i>0.0019</i>	2.3533 <i>0.1240</i>	1.8933 <i>0.0497</i>
θ_2	0.7510 <i>0.2628</i>	1.5155 <i>0.0460</i>	0.9572 <i>0.6698</i>	0.0763 <i>0.0147</i>	0.3485 <i>0.2272</i>	1.5201 <i>0.0568</i>
c	-1.4580 <i>0.6592</i>	3.0393 <i>0.2016</i>	3.7788 <i>0.5806</i>	2.1029 <i>0.0610</i>	0.7404 <i>0.7972</i>	3.3569 <i>0.4702</i>
η	0.1938 <i>0.2085</i>	0.0096 <i>0.0033</i>	0.1523 <i>0.2084</i>	0.0029 <i>0.0010</i>	0.0647 <i>0.2062</i>	0.0259 <i>0.0202</i>
β_1	0.3198	0.9371	0.7827	0.0652	0.2422	0.9472
β_2	0.3426	0.0519	0.5513	0.8530	0.6665	0.0481
β_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
γ	12.6260 <i>5.0581</i>	16.8620 <i>6.9726</i>	17.0540 <i>5.2787</i>	15.2000 <i>5.2428</i>	14.2380 <i>4.5599</i>	14.5250 <i>6.1151</i>
θ_1	2.0099 <i>0.2439</i>	2.3977 <i>0.0288</i>	2.3327 <i>0.0596</i>	2.3090 <i>0.0645</i>	1.7875 <i>0.0891</i>	2.0869 <i>0.2538</i>
θ_2	1.2821 <i>0.2369</i>	0.8151 <i>0.0489</i>	0.9159 <i>0.0901</i>	0.9517 <i>0.0895</i>	1.5146 <i>0.0741</i>	1.2118 <i>0.2544</i>
c	2.4621 <i>0.8302</i>	0.7374 <i>0.2304</i>	1.1670 <i>0.3802</i>	1.3126 <i>0.3856</i>	3.1515 <i>0.2692</i>	2.1100 <i>0.9630</i>
η	0.9275 <i>0.4872</i>	0.5892 <i>0.3166</i>	0.6727 <i>0.3288</i>	0.6477 <i>0.3234</i>	1.2082 <i>0.5066</i>	0.8681 <i>0.4877</i>
β_1	0.8677	0.4929	0.5739	0.6024	0.9751	0.8143
β_2	0.2577	0.4644	0.4407	0.4293	0.0549	0.3055
β_3	-0.4251	-0.7358	-0.6903	-0.6729	-0.2150	-0.4935

Notes: See notes in Table 1.

Table 4: Summarized Posterior Model Probabilities

Model	US - CAN		US - FRA		US - GER		US - ITA		US - JAP		US - UK	
	BF	BIC	BF	BIC	BF	BIC	BF	BIC	BF	BIC	BF	BIC
ESVECM	0.9451	0.4286	0.9997	0.4363	0.9442	0.4209	0.9179	0.4334	0.9989	0.4324	0.9917	0.4279
ESVAR	0.0549	0.4250	0.0003	0.4171	0.0548	0.4337	0.0821	0.4221	0.0011	0.4208	0.0083	0.4265
Top 20 models	0.9911	0.2413	1.0000	0.2478	0.9898	0.2438	1.0000	0.2446	1.0000	0.2433	0.9934	0.2414
VECM/VAR*	8.0570	1.0063	4.4462	1.0007	10108.5000	0.9714	0.7355	0.9774	3.1522	0.9998	118.7717	1.0030

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

Table 5: Posterior Model Probabilities of the Transition Variables

Model	US - CAN		US - FRA		US - GER		US - ITA		US - JAP		US - UK	
	BF	BIC	BF	BIC	BF	BIC	BF	BIC	BF	BIC	BF	BIC
d=1	0.0000	0.1667	0.0040	0.1643	0.0020	0.1667	0.0000	0.1677	0.0000	0.1674	0.0002	0.1666
d=2	0.0000	0.1667	0.9960	0.1647	0.0021	0.1666	0.0000	0.1681	0.0000	0.1651	0.0237	0.1666
d=3	0.0000	0.1667	0.0000	0.1687	0.0020	0.1669	0.2501	0.1641	0.0000	0.1659	0.3377	0.1666
d=4	0.0117	0.1666	0.0000	0.1649	0.1228	0.1664	0.0000	0.1681	0.0000	0.1677	0.6385	0.1667
d=5	0.1058	0.1667	0.0000	0.1687	0.0020	0.1668	0.7499	0.1641	0.0000	0.1682	0.0000	0.1667
d=6	0.8825	0.1666	0.0000	0.1687	0.8691	0.1665	0.0000	0.1680	1.0000	0.1658	0.0000	0.1668

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

Table 6: Top 20 Most Preferred Models (a)

		US - CAN				US - FRA			
No.	Bayes factor	BIC		Bayes factor		BIC		Bayes factor	
1	$ESVECM(6)_{d=6}$	0.1415	$ESVECM(1)_{d=1}$	0.0122	$ESVECM(6)_{d=2}$	0.8478	$ESVECM(1)_{d=1}$	0.0126	
2	$ESVECM(5)_{d=6}$	0.1413	$ESVECM(1)_{d=2}$	0.0122	$ESVECM(3)_{d=2}$	0.0417	$ESVECM(1)_{d=2}$	0.0126	
3	$ESVECM(4)_{d=6}$	0.1402	$ESVECM(1)_{d=3}$	0.0122	$ESVECM(2)_{d=2}$	0.0350	$ESVECM(1)_{d=4}$	0.0126	
4	$ESVECM(2)_{d=6}$	0.1395	$ESVECM(1)_{d=6}$	0.0122	$ESVECM(4)_{d=2}$	0.0266	$ESVAR(6)_{d=5}$	0.0125	
5	$ESVECM(3)_{d=6}$	0.1390	$ESVECM(1)_{d=5}$	0.0122	$ESVECM(1)_{d=2}$	0.0226	$ESVAR(6)_{d=6}$	0.0125	
6	$ESVECM(1)_{d=6}$	0.1356	$ESVECM(1)_{d=4}$	0.0122	$ESVECM(5)_{d=2}$	0.0220	$ESVECM(2)_{d=1}$	0.0125	
7	$ESVAR(6)_{d=6}$	0.0377	$VECM(1)$	0.0121	$ESVECM(6)_{d=1}$	0.0035	$ESVAR(6)_{d=3}$	0.0125	
8	$ESVECM(5)_{d=5}$	0.0164	$ESVECM(2)_{d=1}$	0.0121	$ESVECM(3)_{d=1}$	0.0002	$ESVECM(2)_{d=2}$	0.0124	
9	$ESVECM(5)_{d=5}$	0.0164	$ESVECM(2)_{d=2}$	0.0121	$ESVECM(4)_{d=1}$	0.0001	$ESVECM(2)_{d=4}$	0.0124	
10	$ESVECM(4)_{d=5}$	0.0163	$ESVECM(2)_{d=3}$	0.0121	$ESVECM(5)_{d=1}$	0.0001	$ESVAR(3)_{d=3}$	0.0123	
11	$ESVECM(2)_{d=5}$	0.0162	$ESVECM(2)_{d=6}$	0.0121	$ESVAR(5)_{d=2}$	0.0001	$ESVAR(3)_{d=6}$	0.0123	
12	$ESVECM(3)_{d=5}$	0.0162	$ESVECM(2)_{d=5}$	0.0121	$ESVAR(4)_{d=2}$	0.0001	$ESVAR(3)_{d=5}$	0.0123	
13	$ESVECM(1)_{d=5}$	0.0158	$ESVECM(2)_{d=4}$	0.0121	$ESVECM(2)_{d=1}$	0.0001	$ESVECM(3)_{d=1}$	0.0123	
14	$ESVAR(1)_{d=6}$	0.0042	$VECM(2)$	0.0120	$ESVAR(3)_{d=2}$	0.0001	$ESVECM(3)_{d=2}$	0.0123	
15	$ESVAR(1)_{d=5}$	0.0039	$ESVECM(3)_{d=1}$	0.0120	$ESVECM(6)_{d=4}$	0.0000	$ESVECM(3)_{d=4}$	0.0123	
16	$ESVAR(6)_{d=5}$	0.0029	$ESVECM(3)_{d=2}$	0.0120	$ESVECM(1)_{d=1}$	0.0000	$ESVAR(4)_{d=3}$	0.0123	
17	$ESVAR(5)_{d=6}$	0.0026	$ESVECM(3)_{d=3}$	0.0120	$ESVAR(2)_{d=2}$	0.0000	$ESVAR(4)_{d=5}$	0.0123	
18	$ESVECM(6)_{d=4}$	0.0018	$ESVECM(3)_{d=6}$	0.0120	$ESVAR(6)_{d=2}$	0.0000	$ESVAR(4)_{d=6}$	0.0123	
19	$ESVECM(5)_{d=4}$	0.0018	$ESVECM(3)_{d=5}$	0.0120	$ESVECM(3)_{d=4}$	0.0000	$ESVAR(2)_{d=3}$	0.0123	
20	$ESVECM(4)_{d=4}$	0.0018	$ESVECM(3)_{d=4}$	0.0119	$ESVECM(4)_{d=4}$	0.0000	$ESVAR(1)_{d=6}$	0.0122	

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

Table 7: Top 20 Most Preferred Models (b)

No.	US - GER			US - ITA				
	Bayes factor	BIC	Bayes factor	BIC	BIC			
1	$ESV ECM(6)_{d=6}$	0.2128	$ESV AR(1)_{d=3}$	0.0123	$ESV ECM(2)_{d=5}$	0.1600	$ESV ECM(1)_{d=3}$	0.0124
2	$ESV ECM(3)_{d=6}$	0.1824	$ESV AR(1)_{d=5}$	0.0123	$ESV ECM(6)_{d=5}$	0.1470	$ESV ECM(1)_{d=5}$	0.0124
3	$ESV ECM(4)_{d=6}$	0.1468	$ESV AR(1)_{d=1}$	0.0123	$ESV ECM(3)_{d=5}$	0.1287	$ESV ECM(2)_{d=3}$	0.0123
4	$ESV ECM(2)_{d=6}$	0.1074	$ESV AR(1)_{d=2}$	0.0123	$ESV ECM(4)_{d=5}$	0.1134	$ESV ECM(2)_{d=5}$	0.0123
5	$ESV ECM(1)_{d=6}$	0.1071	$ESV AR(2)_{d=3}$	0.0122	$ESV ECM(5)_{d=5}$	0.0734	$ESV AR(1)_{d=4}$	0.0122
6	$ESV ECM(5)_{d=6}$	0.0817	$ESV AR(2)_{d=5}$	0.0122	$ESV ECM(1)_{d=5}$	0.0708	$ESV AR(1)_{d=6}$	0.0122
7	$ESV AR(1)_{d=6}$	0.0211	$ESV AR(1)_{d=6}$	0.0122	$ESV ECM(6)_{d=3}$	0.0583	$ESV ECM(1)_{d=2}$	0.0122
8	$ESV ECM(6)_{d=4}$	0.0201	$ESV AR(2)_{d=1}$	0.0122	$ESV ECM(4)_{d=3}$	0.0529	$ESV ECM(1)_{d=6}$	0.0122
9	$ESV ECM(4)_{d=4}$	0.0178	$ESV AR(1)_{d=4}$	0.0122	$ESV ECM(3)_{d=3}$	0.0395	$ESV AR(1)_{d=2}$	0.0122
10	$ESV ECM(3)_{d=4}$	0.0176	$ESV AR(2)_{d=2}$	0.0122	$ESV ECM(2)_{d=3}$	0.0331	$ESV ECM(1)_{d=1}$	0.0122
11	$ESV ECM(5)_{d=4}$	0.0159	$ESV ECM(1)_{d=3}$	0.0121	$ESV ECM(5)_{d=3}$	0.0313	$ESV AR(2)_{d=4}$	0.0122
12	$ESV ECM(1)_{d=4}$	0.0153	$ESV AR(3)_{d=3}$	0.0121	$ESV AR(2)_{d=5}$	0.0185	$ESV ECM(1)_{d=4}$	0.0122
13	$ESV ECM(2)_{d=4}$	0.0152	$VAR(1)$	0.0121	$ESV AR(3)_{d=5}$	0.0152	$ESV AR(2)_{d=2}$	0.0122
14	$ESV AR(1)_{d=4}$	0.0127	$ESV ECM(1)_{d=6}$	0.0121	$ESV AR(4)_{d=5}$	0.0136	$ESV AR(3)_{d=4}$	0.0122
15	$ESV AR(6)_{d=6}$	0.0037	$ESV ECM(1)_{d=5}$	0.0121	$ESV ECM(1)_{d=3}$	0.0096	$ESV AR(2)_{d=6}$	0.0122
16	$ESV AR(3)_{d=6}$	0.0033	$ESV AR(3)_{d=5}$	0.0121	$ESV AR(5)_{d=5}$	0.0093	$ESV AR(3)_{d=2}$	0.0122
17	$ESV AR(4)_{d=6}$	0.0030	$ESV ECM(1)_{d=1}$	0.0121	$ESV AR(4)_{d=3}$	0.0086	$ESV ECM(2)_{d=2}$	0.0122
18	$ESV AR(4)_{d=4}$	0.0020	$ESV AR(2)_{d=6}$	0.0121	$ESV AR(3)_{d=3}$	0.0062	$ESV ECM(2)_{d=6}$	0.0122
19	$ESV AR(2)_{d=6}$	0.0019	$ESV ECM(1)_{d=4}$	0.0121	$ESV AR(4)_{d=3}$	0.0056	$ESV AR(3)_{d=6}$	0.0122
20	$ESV AR(3)_{d=4}$	0.0018	$ESV AR(3)_{d=1}$	0.0121	$ESV AR(2)_{d=3}$	0.0049	$ESV ECM(2)_{d=1}$	0.0122

Notes: See notes in Table 6.

Table 8: Top 20 Most Preferred Models (c)

No.	US - JAP			US - UK				
	Bayes factor	BIC	Bayes factor	BIC	BIC			
1	$ESVECM(3)_{d=6}$	0.2838	$ESVECM(1)_{d=1}$	0.0123	$ESVECM(1)_{d=4}$	0.1293	$ESVECM(1)_{d=2}$	0.0122
2	$ESVECM(2)_{d=6}$	0.2083	$ESVECM(1)_{d=6}$	0.0123	$ESVECM(2)_{d=4}$	0.1235	$ESVECM(1)_{d=4}$	0.0122
3	$ESVECM(6)_{d=6}$	0.2045	$ESVECM(1)_{d=3}$	0.0123	$ESVECM(6)_{d=4}$	0.1049	$ESVECM(1)_{d=3}$	0.0122
4	$ESVECM(4)_{d=6}$	0.1473	$ESVECM(1)_{d=2}$	0.0122	$ESVECM(3)_{d=4}$	0.1044	$ESVECM(1)_{d=6}$	0.0122
5	$ESVECM(5)_{d=6}$	0.0989	$ESVAR(1)_{d=5}$	0.0122	$ESVECM(4)_{d=4}$	0.0897	$ESVECM(1)_{d=1}$	0.0122
6	$ESVECM(1)_{d=6}$	0.0562	$ESVECM(1)_{d=5}$	0.0122	$ESVECM(5)_{d=4}$	0.0828	$ESVECM(1)_{d=5}$	0.0122
7	$ESVAR(3)_{d=6}$	0.0004	$ESVECM(1)_{d=4}$	0.0122	$ESVECM(1)_{d=3}$	0.0617	$VECM(2)$	0.0121
8	$ESVAR(2)_{d=6}$	0.0003	$ESVAR(2)_{d=5}$	0.0122	$ESVECM(6)_{d=3}$	0.0613	$ESVECM(2)_{d=2}$	0.0121
9	$ESVAR(6)_{d=6}$	0.0002	$ESVECM(2)_{d=1}$	0.0122	$ESVECM(2)_{d=3}$	0.0575	$ESVECM(2)_{d=4}$	0.0121
10	$ESVAR(4)_{d=6}$	0.0002	$ESVECM(2)_{d=3}$	0.0122	$ESVECM(3)_{d=3}$	0.0552	$ESVECM(2)_{d=3}$	0.0121
11	$ESVAR(5)_{d=6}$	0.0000	$ESVECM(2)_{d=6}$	0.0121	$ESVECM(4)_{d=3}$	0.0500	$ESVECM(2)_{d=6}$	0.0121
12	$ESVECM(6)_{d=2}$	0.0000	$ESVAR(3)_{d=56}$	0.0121	$ESVECM(5)_{d=3}$	0.0486	$ESVECM(2)_{d=1}$	0.0121
13	$ESVECM(5)_{d=2}$	0.0000	$ESVECM(2)_{d=2}$	0.0121	$ESVECM(1)_{d=2}$	0.0046	$ESVECM(2)_{d=5}$	0.0121
14	$ESVECM(6)_{d=5}$	0.0000	$ESVAR(5)_{d=5}$	0.0121	$ESVECM(6)_{d=2}$	0.0040	$ESVAR(1)_{d=6}$	0.0120
15	$ESVECM(4)_{d=5}$	0.0000	$ESVAR(4)_{d=5}$	0.0121	$ESVECM(2)_{d=2}$	0.0039	$ESVAR(1)_{d=5}$	0.0120
16	$ESVECM(5)_{d=5}$	0.0000	$ESVAR(6)_{d=5}$	0.0121	$ESVECM(3)_{d=2}$	0.0038	$ESVAR(1)_{d=4}$	0.0120
17	$ESVECM(3)_{d=5}$	0.0000	$ESVECM(2)_{d=5}$	0.0121	$ESVECM(4)_{d=2}$	0.0033	$ESVAR(1)_{d=1}$	0.0120
18	$ESVAR(6)_{d=2}$	0.0000	$ESVECM(2)_{d=4}$	0.0121	$ESVECM(5)_{d=2}$	0.0030	$ESVAR(1)_{d=3}$	0.0120
19	$ESVAR(5)_{d=5}$	0.0000	$VECM(1)$	0.0121	$ESVAR(2)_{d=4}$	0.0010	$ESVECM(2)_{d=5}$	0.0120
20	$ESVECM(2)_{d=5}$	0.0000	$ESVECM(3)_{d=2}$	0.0121	$ESVAR(3)_{d=4}$	0.0009	$VECM(3)$	0.0120

Notes: See notes in Table 6.

Figure 1
Smooth Transition Functions (most preferred models)

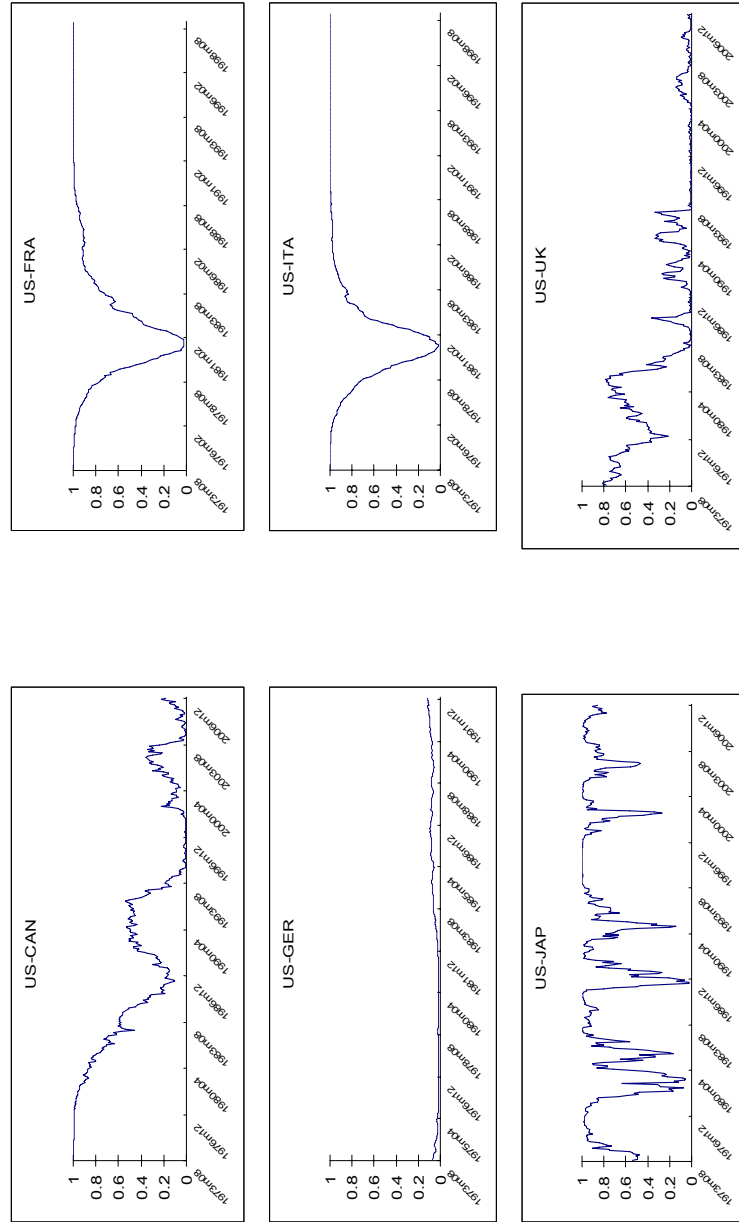


Figure 2
Smooth Transition Functions (BMA results)

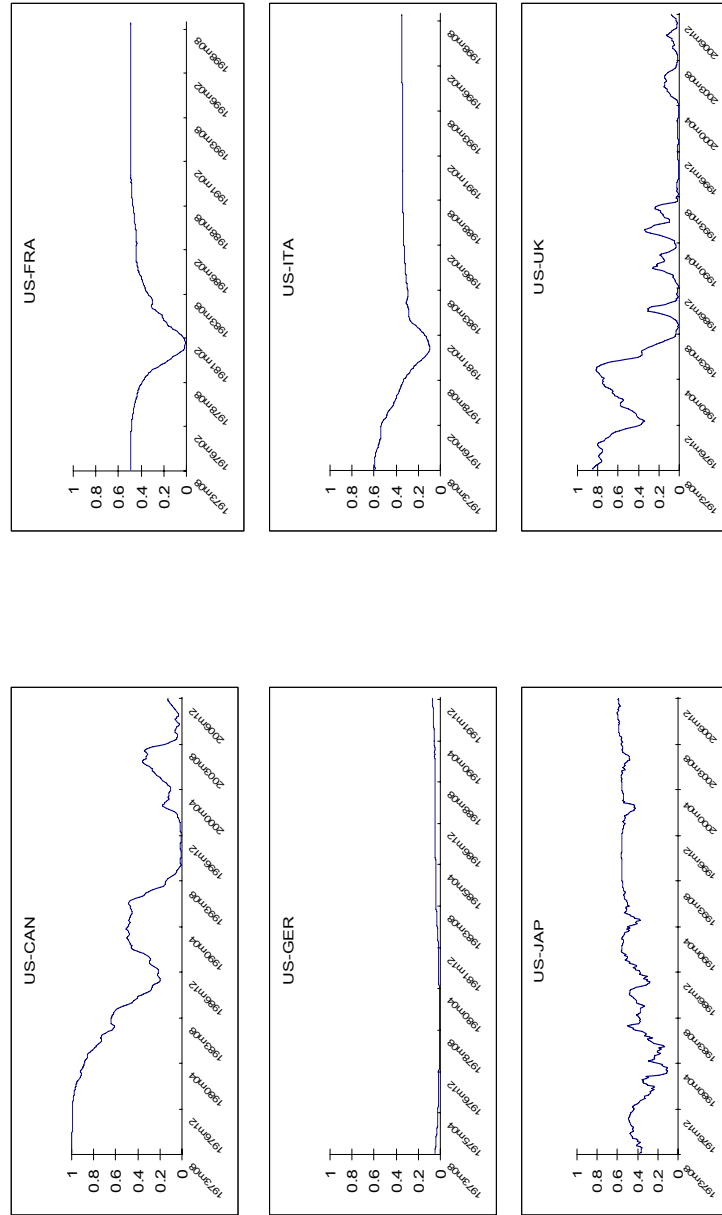


Figure 3
General Impulse Response Functions (a)

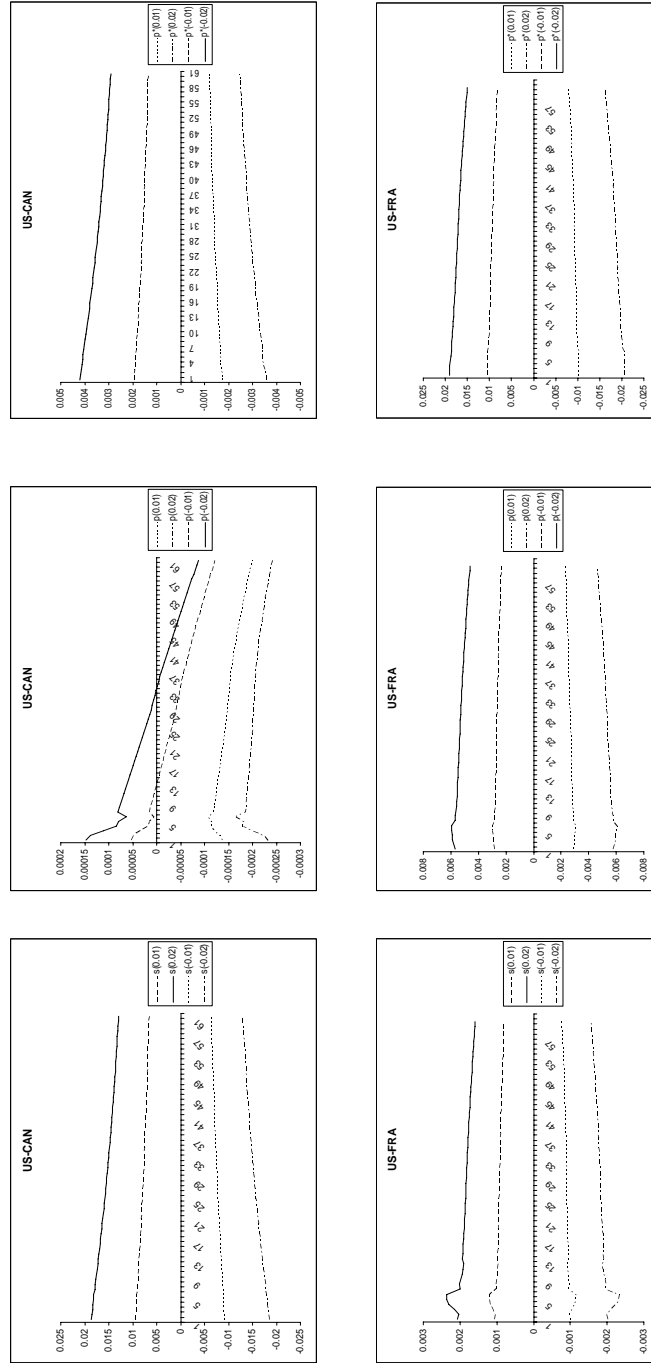


Figure 4
General Impulse Response Functions (b)

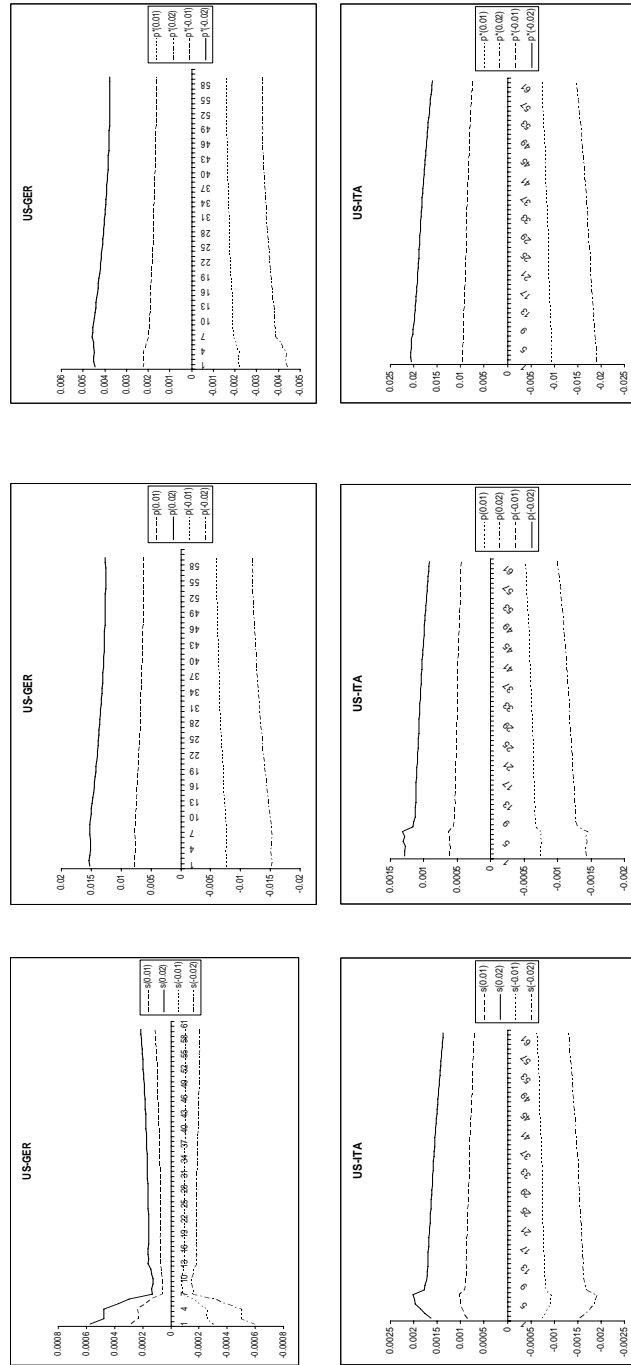


Figure 5
General Impulse Response Functions (c)

