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.
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-

\(^1\)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, Kai 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.
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

\(^3\)Their 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 equi-
librium through the transaction 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. ES-
VECM appears to have another attractive property for it allows the same
dynamics of regime changes for deviations above and below the equilibrium
level.

Let  \( y_t = [s_t \quad p_t \quad p^*_t] \), 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,...,T as follows.

\[
\Delta y_t = y_{t-1} \beta \alpha + d_t \xi + \sum_{h=1}^{p} \Delta y_{t-h} \Gamma_h \\
+ F(z_t)(y_{t-1} \beta \alpha^z + d_t \xi^z + \sum_{h=1}^{p} \Delta y_{t-h} \Gamma_h^z) + \varepsilon_t \tag{1}
\]

where \( \Delta y_t = y_t - y_{t-1} \). The error term \( \varepsilon_t \) 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 $\Gamma_h$ 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 vector $\beta$. 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):

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

where the transition variable $z_t = y_{t-d}\beta$ is the cointegrating combination among $s$, $p$, and $p^*$ at period $t - d$; $c$ 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.

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 the outer regimes, respectively. In the middle regime, model (1) becomes a linear VECM, with the adjustment process governed by $(\alpha, \xi, \Gamma_h)$; 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

---

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.
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 $$x_{1,t-1} = y_{t-1}$$, $$x_{2,t} = (d_t, \Delta y_{t-1}, \ldots, \Delta y_{t-p})$$, $$\Phi = (\xi', \Gamma_1', \ldots, \Gamma_p')$$, $$\Phi z = (\xi z', \Gamma_1 z', \ldots, \Gamma_p z')'$$.

To simplify the notation, we first define the $$T \times n$$ matrix $$X_0 = (\Delta y_1', \Delta y_2', \ldots, \Delta y_T')'$$ 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 $$X_1 = (x_{1,1}', \ x_{1,2}', \ldots, \ x_{1,T}')'$$, $$X_2 = (x_{2,1}', \ x_{2,2}', \ldots, \ x_{2,T}')'$$, and $$F^z = diag(F(z_1), F(z_2), \ldots, F(z_T))$$. 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 $$E = (\varepsilon_1', \varepsilon_2', \ldots, \varepsilon_T')'$$. 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.

\[
L(y|\Sigma, B, \beta) \propto |\Sigma|^{-\frac{r}{2}} \exp\left\{-\frac{1}{2} tr \Sigma^{-1} E'E\right\}
\]  

(5)

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

\[
x_0 = xb + e
\]  

(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,

\[
tr \Sigma^{-1} E'E = e'(\Sigma^{-1} \otimes I_T)e
\]

(7)

\[
= s^2 + (b - \hat{b})' V^{-1} (b - \hat{b})
\]

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

\[
L(y|\Sigma, B, \beta) \propto |\Sigma|^{-\frac{r}{2}} \exp\left\{-\frac{1}{2} [s^2 + (b - \hat{b})' V^{-1} (b - \hat{b})]\right\}
\]  

(8)

2.2 Priors

Although the strict version of PPP states that the combination \(s_t + p_t - p^*_t\) 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 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 $\beta$, 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 (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 $\gamma$ 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 $\beta \alpha^2$. 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 $\beta$, $\alpha$, $\alpha^2$ and $\gamma$ sequentially once we choose a way to normalize $\beta$.

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

\footnote{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.}
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 \(\beta\) based on the uninformative priors proposed in Strachan and Inder (2004). First, we specify the space of the three by one vector \(\beta\) to be uniformly distributed over the two dimensional Grassmann manifold \(G_{1,2}\).

\[
p(\beta) = \frac{1}{c_1^3}
\]  

(9)

where \(c_1^3 = \int_{G_{1,2}} dq_1^3\) 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 \(\beta\) is 1, we can use polar coordinates to denote the semi-orthogonal \(\beta\) as follows:

\[
\beta = \begin{bmatrix} sin(\theta_1)sin(\theta_2) & sin(\theta_1)cos(\theta_2) & cos(\theta_1) \end{bmatrix}'
\]

To describe the uniform distribution of the cointegration space in polar coordinates, we multiply the uninformative prior of the space of \(\beta\) 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 \(\beta\) 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 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:

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

where $b = vec(B)$, $V = \Sigma \otimes \eta^{-1}I_k$, $k = 2(r + 1 + np)$, $\eta$ 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 $\mu_\eta$, and degrees of freedom $\nu_\eta$, where $\mu_\eta=10$, $\nu_\eta=0.0001$. 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 $\Sigma \otimes (\beta'\eta I_3 \beta)^{-1}$. 

12
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 (1) in favor of the nonlinear effect. In the empirical work, we set the prior distribution of $\gamma$ 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 $[s_{t-d} \ p_{t-d} \ p^*_t \ d_{t-d}]$ and the cointegrating vector $\beta$).

### 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 $\Sigma$ 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 $\bar{b} = \text{vec}[(X'X + \eta I_k)^{-1}X'X_0]$ and covariance matrix $\Sigma = \Sigma \otimes (X'X + \eta I_k)^{-1}$. 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

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

(10)

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

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 $N(b, \nabla_b)$;
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 $\alpha^z$ and $\Phi^z$ 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 $\alpha, \alpha^z$ and $\Phi^z$ to be equal to zeros.\footnote{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 $\alpha = 0$.} 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
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 $GI_P$ for a shock, $\nu_t$, and a history, $\omega_{t-1}$ as follows

$$GI_P(n, \nu_t, \omega_{t-1}) = E[P_{t+n}|\nu_t, \omega_{t-1}] - E[P_{t+n}|\omega_{t-1}]$$  (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 $\pm 0.01$ 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, 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.\textsuperscript{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 $p_t^*$.

\textsuperscript{7}The 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

---

\(^8\)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).

\(^9\)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.
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


[7] Davies, R. B. (1977), Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika 74, 33-43.


Table 1: Parameters (a)

<table>
<thead>
<tr>
<th></th>
<th>US - CAN</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d=1</td>
<td>d=2</td>
<td>d=3</td>
<td>d=4</td>
<td>d=5</td>
<td>d=6</td>
</tr>
<tr>
<td></td>
<td>3.5105</td>
<td>4.1453</td>
<td>4.0240</td>
<td>5.2090</td>
<td>4.0288</td>
<td>3.4595</td>
</tr>
<tr>
<td>θ₁</td>
<td>2.3150</td>
<td>2.0431</td>
<td>2.0648</td>
<td>1.9085</td>
<td>1.9747</td>
<td>1.9319</td>
</tr>
<tr>
<td></td>
<td>0.0537</td>
<td>0.0865</td>
<td>0.0685</td>
<td>0.0790</td>
<td>0.0930</td>
<td>0.0373</td>
</tr>
<tr>
<td>θ₂</td>
<td>1.4372</td>
<td>1.5365</td>
<td>1.4857</td>
<td>1.5337</td>
<td>1.5171</td>
<td>1.5283</td>
</tr>
<tr>
<td></td>
<td>0.0805</td>
<td>0.0246</td>
<td>0.0519</td>
<td>0.0316</td>
<td>0.0463</td>
<td>0.0309</td>
</tr>
<tr>
<td>c</td>
<td>-0.2820</td>
<td>1.3063</td>
<td>1.3917</td>
<td>2.2738</td>
<td>1.8870</td>
<td>2.1457</td>
</tr>
<tr>
<td></td>
<td>0.4604</td>
<td>0.5750</td>
<td>0.4554</td>
<td>0.5565</td>
<td>0.5909</td>
<td>0.2492</td>
</tr>
<tr>
<td>η</td>
<td>2.2817</td>
<td>2.1475</td>
<td>2.0392</td>
<td>2.0671</td>
<td>1.9965</td>
<td>1.8482</td>
</tr>
<tr>
<td></td>
<td>1.1317</td>
<td>1.0957</td>
<td>1.1172</td>
<td>1.0442</td>
<td>1.0906</td>
<td>0.9917</td>
</tr>
<tr>
<td>β₁</td>
<td>0.7291</td>
<td>0.8900</td>
<td>0.8773</td>
<td>0.9429</td>
<td>0.9182</td>
<td>0.9346</td>
</tr>
<tr>
<td></td>
<td>0.0980</td>
<td>0.0305</td>
<td>0.0749</td>
<td>0.0350</td>
<td>0.0493</td>
<td>0.0398</td>
</tr>
<tr>
<td>β₂</td>
<td>-0.6774</td>
<td>-0.4550</td>
<td>-0.4742</td>
<td>-0.3313</td>
<td>-0.3930</td>
<td>-0.3534</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>US - FRA</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d=1</td>
<td>d=2</td>
<td>d=3</td>
<td>d=4</td>
<td>d=5</td>
<td>d=6</td>
</tr>
<tr>
<td>γ</td>
<td>30.4830</td>
<td>28.7710</td>
<td>0.1131</td>
<td>25.9230</td>
<td>0.0339</td>
<td>0.0382</td>
</tr>
<tr>
<td></td>
<td>4.8129</td>
<td>4.3462</td>
<td>0.3682</td>
<td>6.5391</td>
<td>0.1182</td>
<td>0.1242</td>
</tr>
<tr>
<td>θ₁</td>
<td>3.0893</td>
<td>3.0240</td>
<td>2.2722</td>
<td>3.0322</td>
<td>2.2540</td>
<td>2.3632</td>
</tr>
<tr>
<td></td>
<td>0.0758</td>
<td>0.0498</td>
<td>0.0480</td>
<td>0.0954</td>
<td>0.0846</td>
<td>0.0698</td>
</tr>
<tr>
<td>θ₂</td>
<td>1.0376</td>
<td>1.0875</td>
<td>0.4227</td>
<td>1.0655</td>
<td>0.5105</td>
<td>0.6029</td>
</tr>
<tr>
<td></td>
<td>0.4789</td>
<td>0.3591</td>
<td>0.1585</td>
<td>0.4284</td>
<td>0.1852</td>
<td>0.1697</td>
</tr>
<tr>
<td>c</td>
<td>-4.9709</td>
<td>-4.4999</td>
<td>1.7649</td>
<td>-4.5704</td>
<td>1.9835</td>
<td>1.1213</td>
</tr>
<tr>
<td></td>
<td>0.4376</td>
<td>0.3262</td>
<td>0.4055</td>
<td>0.5818</td>
<td>0.6248</td>
<td>0.6242</td>
</tr>
<tr>
<td>η</td>
<td>0.0857</td>
<td>0.1187</td>
<td>0.0008</td>
<td>0.1298</td>
<td>0.0008</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>0.0514</td>
<td>0.0595</td>
<td>0.0002</td>
<td>0.0700</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td>β₁</td>
<td>0.0450</td>
<td>0.1039</td>
<td>0.3134</td>
<td>0.0956</td>
<td>0.3790</td>
<td>0.3981</td>
</tr>
<tr>
<td></td>
<td>0.0266</td>
<td>0.0545</td>
<td>0.6967</td>
<td>0.0529</td>
<td>0.6767</td>
<td>0.5783</td>
</tr>
<tr>
<td>β₂</td>
<td>-0.9986</td>
<td>-0.9931</td>
<td>-0.6453</td>
<td>-0.9940</td>
<td>-0.6313</td>
<td>-0.7121</td>
</tr>
</tbody>
</table>

Notes: Standard deviations are in italics.
Table 2: Parameters (b)

<table>
<thead>
<tr>
<th></th>
<th>US - GER</th>
<th></th>
<th>US - ITA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d=1</td>
<td>d=2</td>
<td>d=3</td>
</tr>
<tr>
<td>γ</td>
<td>0.0136</td>
<td>0.9400</td>
<td>0.6267</td>
</tr>
<tr>
<td></td>
<td>0.0653</td>
<td>1.8586</td>
<td>1.1395</td>
</tr>
<tr>
<td>θ₁</td>
<td>1.8754</td>
<td>1.8200</td>
<td>1.9436</td>
</tr>
<tr>
<td></td>
<td>0.0499</td>
<td>0.0609</td>
<td>0.0114</td>
</tr>
<tr>
<td>θ₂</td>
<td>0.0186</td>
<td>0.0180</td>
<td>0.0130</td>
</tr>
<tr>
<td>c</td>
<td>3.2665</td>
<td>3.6215</td>
<td>2.7841</td>
</tr>
<tr>
<td></td>
<td>0.1382</td>
<td>0.2079</td>
<td>0.0683</td>
</tr>
<tr>
<td>η</td>
<td>0.0177</td>
<td>0.0174</td>
<td>0.0121</td>
</tr>
<tr>
<td>β₁</td>
<td>0.9538</td>
<td>0.9690</td>
<td>0.9312</td>
</tr>
<tr>
<td>β₂</td>
<td>-0.2999</td>
<td>-0.2466</td>
<td>-0.3642</td>
</tr>
<tr>
<td>β₃</td>
<td>-0.0012</td>
<td>-0.0020</td>
<td>-0.0924</td>
</tr>
<tr>
<td>η</td>
<td>-0.0374</td>
<td>-0.0305</td>
<td>-0.0436</td>
</tr>
<tr>
<td>β₃</td>
<td>-0.7931</td>
<td>-0.7797</td>
<td>-0.9985</td>
</tr>
</tbody>
</table>

Notes: See notes in Table 1.
<table>
<thead>
<tr>
<th></th>
<th>US - JAP</th>
<th></th>
<th></th>
<th>US - UK</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>d=1</td>
<td>d=2</td>
<td>d=3</td>
<td>d=4</td>
<td>d=5</td>
<td>d=6</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>22.7700</td>
<td>34.7190</td>
<td>85.1300</td>
<td>15.2390</td>
<td>1.5974</td>
<td>48.3000</td>
</tr>
<tr>
<td></td>
<td>32.2510</td>
<td>7.3199</td>
<td>59.0440</td>
<td>6.5107</td>
<td>2.9569</td>
<td>11.0160</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>2.6538</td>
<td>1.9233</td>
<td>1.8639</td>
<td>2.1151</td>
<td>2.3533</td>
<td>1.8933</td>
</tr>
<tr>
<td></td>
<td>0.0856</td>
<td>0.0306</td>
<td>0.0517</td>
<td>0.0019</td>
<td>0.1240</td>
<td>0.0497</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.7510</td>
<td>1.5155</td>
<td>0.9572</td>
<td>0.0763</td>
<td>0.3485</td>
<td>1.5201</td>
</tr>
<tr>
<td></td>
<td>0.2628</td>
<td>0.0460</td>
<td>0.6698</td>
<td>0.0147</td>
<td>0.2272</td>
<td>0.0568</td>
</tr>
<tr>
<td>c</td>
<td>-1.4580</td>
<td>3.0393</td>
<td>3.7788</td>
<td>2.1029</td>
<td>0.7404</td>
<td>3.3569</td>
</tr>
<tr>
<td></td>
<td>0.6592</td>
<td>0.2016</td>
<td>0.5806</td>
<td>0.0610</td>
<td>0.7972</td>
<td>0.4702</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.1938</td>
<td>0.0096</td>
<td>0.1523</td>
<td>0.0029</td>
<td>0.0647</td>
<td>0.0259</td>
</tr>
<tr>
<td></td>
<td>0.2085</td>
<td>0.0033</td>
<td>0.2084</td>
<td>0.0010</td>
<td>0.2062</td>
<td>0.0202</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.3198</td>
<td>0.9371</td>
<td>0.7827</td>
<td>0.0652</td>
<td>0.2422</td>
<td>0.9472</td>
</tr>
<tr>
<td></td>
<td>0.3426</td>
<td>0.0519</td>
<td>0.5513</td>
<td>0.8530</td>
<td>0.6665</td>
<td>0.0481</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.8834</td>
<td>-0.3452</td>
<td>-0.2889</td>
<td>-0.5178</td>
<td>-0.7051</td>
<td>-0.3170</td>
</tr>
<tr>
<td></td>
<td>0.9275</td>
<td>0.5892</td>
<td>0.6727</td>
<td>0.6477</td>
<td>1.2082</td>
<td>0.8681</td>
</tr>
<tr>
<td>Notes: See notes in Table 1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Summarized Posterior Model Probabilities

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>BF</td>
<td>BIC</td>
<td>BF</td>
<td>BIC</td>
<td>BF</td>
<td>BIC</td>
<td>BF</td>
<td>BIC</td>
<td>BF</td>
<td>BIC</td>
<td>BF</td>
<td>BIC</td>
<td>BF</td>
</tr>
<tr>
<td>ESVECM</td>
<td>0.9451</td>
<td>0.4286</td>
<td>0.9997</td>
<td>0.4363</td>
<td>0.9442</td>
<td>0.4209</td>
<td>0.9179</td>
<td>0.4334</td>
<td>0.9989</td>
<td>0.4324</td>
<td>0.9917</td>
<td>0.4279</td>
</tr>
<tr>
<td>ESVAR</td>
<td>0.0549</td>
<td>0.4250</td>
<td>0.0003</td>
<td>0.4171</td>
<td>0.0548</td>
<td>0.4337</td>
<td>0.0821</td>
<td>0.4221</td>
<td>0.0011</td>
<td>0.4208</td>
<td>0.0083</td>
<td>0.4265</td>
</tr>
<tr>
<td>Top 20 models</td>
<td>0.9911</td>
<td>0.2413</td>
<td>1.0000</td>
<td>0.2478</td>
<td>0.9898</td>
<td>0.2438</td>
<td>1.0000</td>
<td>0.2446</td>
<td>1.0000</td>
<td>0.2433</td>
<td>0.9934</td>
<td>0.2414</td>
</tr>
<tr>
<td>VECM/VAR*</td>
<td>8.0570</td>
<td>1.0063</td>
<td>4.4462</td>
<td>1.0007</td>
<td>10108.5000</td>
<td>0.9714</td>
<td>0.7355</td>
<td>0.9774</td>
<td>3.1522</td>
<td>0.9998</td>
<td>118.7717</td>
<td>1.0030</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>d=1</td>
<td>BF 0.0000</td>
<td>BIC 0.1667</td>
<td>BF 0.0040</td>
<td>BIC 0.1643</td>
<td>BF 0.0000</td>
<td>BIC 0.1677</td>
</tr>
<tr>
<td>d=2</td>
<td>BF 0.0000</td>
<td>BIC 0.1667</td>
<td>BF 0.9960</td>
<td>BIC 0.1647</td>
<td>BF 0.0021</td>
<td>BIC 0.1666</td>
</tr>
<tr>
<td>d=3</td>
<td>BF 0.0000</td>
<td>BIC 0.1667</td>
<td>BF 0.0000</td>
<td>BIC 0.1687</td>
<td>BF 0.0020</td>
<td>BIC 0.1669</td>
</tr>
<tr>
<td>d=4</td>
<td>BF 0.0117</td>
<td>BIC 0.1666</td>
<td>BF 0.0000</td>
<td>BIC 0.1649</td>
<td>BF 0.1228</td>
<td>BIC 0.1664</td>
</tr>
<tr>
<td>d=5</td>
<td>BF 0.1058</td>
<td>BIC 0.1667</td>
<td>BF 0.0000</td>
<td>BIC 0.1687</td>
<td>BF 0.0020</td>
<td>BIC 0.1668</td>
</tr>
<tr>
<td>d=6</td>
<td>BF 0.8825</td>
<td>BIC 0.1666</td>
<td>BF 0.0000</td>
<td>BIC 0.1687</td>
<td>BF 0.8691</td>
<td>BIC 0.1665</td>
</tr>
</tbody>
</table>

Notes: d is the lag length of the transition variable.
Table 6: Top 20 Most Preferred Models (a)

<table>
<thead>
<tr>
<th>No.</th>
<th>US - CAN</th>
<th>US - FRA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bayes factor</td>
<td>BIC</td>
</tr>
<tr>
<td>1</td>
<td>ESV ECM(6) d=6</td>
<td>0.1415</td>
</tr>
<tr>
<td>2</td>
<td>ESV ECM(5) d=6</td>
<td>0.1413</td>
</tr>
<tr>
<td>3</td>
<td>ESV ECM(4) d=6</td>
<td>0.1402</td>
</tr>
<tr>
<td>4</td>
<td>ESV ECM(2) d=6</td>
<td>0.1395</td>
</tr>
<tr>
<td>5</td>
<td>ESV ECM(3) d=6</td>
<td>0.1390</td>
</tr>
<tr>
<td>6</td>
<td>ESV ECM(1) d=6</td>
<td>0.1356</td>
</tr>
<tr>
<td>7</td>
<td>ESVAR(6) d=6</td>
<td>0.0377</td>
</tr>
<tr>
<td>8</td>
<td>ESV ECM(5) d=5</td>
<td>0.0164</td>
</tr>
<tr>
<td>9</td>
<td>ESV ECM b(5) d=5</td>
<td>0.0164</td>
</tr>
<tr>
<td>10</td>
<td>ESV ECM(4) d=5</td>
<td>0.0163</td>
</tr>
<tr>
<td>11</td>
<td>ESV ECM(2) d=5</td>
<td>0.0162</td>
</tr>
<tr>
<td>12</td>
<td>ESV ECM(3) d=5</td>
<td>0.0162</td>
</tr>
<tr>
<td>13</td>
<td>ESV ECM(1) d=5</td>
<td>0.0158</td>
</tr>
<tr>
<td>14</td>
<td>ESVAR(1) d=6</td>
<td>0.0042</td>
</tr>
<tr>
<td>15</td>
<td>ESVAR(1) d=5</td>
<td>0.0039</td>
</tr>
<tr>
<td>16</td>
<td>ESVAR(6) d=5</td>
<td>0.0029</td>
</tr>
<tr>
<td>17</td>
<td>ESVAR(5) d=6</td>
<td>0.0026</td>
</tr>
<tr>
<td>18</td>
<td>ESV ECM(6) d=4</td>
<td>0.0018</td>
</tr>
<tr>
<td>19</td>
<td>ESV ECM(5) d=4</td>
<td>0.0018</td>
</tr>
<tr>
<td>20</td>
<td>ESV ECM(4) d=4</td>
<td>0.0018</td>
</tr>
</tbody>
</table>

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)

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

Notes: See notes in Table 6.
Table 8: Top 20 Most Preferred Models (c)

<table>
<thead>
<tr>
<th>No.</th>
<th>US - JAP</th>
<th>US - UK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bayes factor</td>
<td>BIC</td>
</tr>
<tr>
<td>1</td>
<td>ESV ECM(3)</td>
<td>0.2838</td>
</tr>
<tr>
<td>2</td>
<td>ESV ECM(2)</td>
<td>0.2083</td>
</tr>
<tr>
<td>3</td>
<td>ESV ECM(6)</td>
<td>0.2045</td>
</tr>
<tr>
<td>4</td>
<td>ESV ECM(4)</td>
<td>0.1473</td>
</tr>
<tr>
<td>5</td>
<td>ESV ECM(1)</td>
<td>0.0989</td>
</tr>
<tr>
<td>6</td>
<td>ESV ECM(1)</td>
<td>0.0562</td>
</tr>
<tr>
<td>7</td>
<td>ESV ECM(6)</td>
<td>0.0004</td>
</tr>
<tr>
<td>8</td>
<td>ESV ECM(2)</td>
<td>0.0003</td>
</tr>
<tr>
<td>9</td>
<td>ESV ECM(6)</td>
<td>0.0002</td>
</tr>
<tr>
<td>10</td>
<td>ESV ECM(4)</td>
<td>0.0002</td>
</tr>
<tr>
<td>11</td>
<td>ESV ECM(5)</td>
<td>0.0000</td>
</tr>
<tr>
<td>12</td>
<td>ESV ECM(6)</td>
<td>0.0000</td>
</tr>
<tr>
<td>13</td>
<td>ESV ECM(5)</td>
<td>0.0000</td>
</tr>
<tr>
<td>14</td>
<td>ESV ECM(6)</td>
<td>0.0000</td>
</tr>
<tr>
<td>15</td>
<td>ESV ECM(4)</td>
<td>0.0000</td>
</tr>
<tr>
<td>16</td>
<td>ESV ECM(5)</td>
<td>0.0000</td>
</tr>
<tr>
<td>17</td>
<td>ESV ECM(3)</td>
<td>0.0000</td>
</tr>
<tr>
<td>18</td>
<td>ESV ECM(6)</td>
<td>0.0000</td>
</tr>
<tr>
<td>19</td>
<td>ESV ECM(5)</td>
<td>0.0000</td>
</tr>
<tr>
<td>20</td>
<td>ESV ECM(2)</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

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)

US-CAN

-0.025
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02
0.025
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
s(0.01)
s(0.02)
s(-0.01)
s(-0.02)

US-FRA

-0.003
-0.002
-0.001
0
0.001
0.002
0.003
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
s(0.01)
s(0.02)
s(-0.01)
s(-0.02)

US-CAN

-0.008
-0.006
-0.004
-0.002
0
0.002
0.004
0.006
0.008
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
p(0.01)
p(0.02)
p(-0.01)
p(-0.02)

US-FRA

-0.025
-0.02
-0.015
-0.01
-0.005
0
0.005
0.01
0.015
0.02
0.025
1
5
9
13
17
21
25
29
33
37
41
45
49
53
57
p*(0.01)
p*(0.02)
p*(-0.01)
p*(-0.02)
Figure 4
General Impulse Response Functions (b)

US-GER
-0.0008
-0.0006
-0.0004
-0.0002
0
0.0002
0.0004
0.0006
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52 55 58 61
s(0.01)
s(0.02)
s(-0.01)
s(-0.02)

US-ITA
-0.0025
-0.002
-0.0015
-0.001
-0.0005
0
0.0005
0.001
0.0015
0.002
0.0025
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61
p(0.01)
p(0.02)
p(-0.01)
p(-0.02)

p*(0.01)
p*(0.02)
p*(-0.01)
p*(-0.02)
Figure 5
General Impulse Response Functions (c)