THE LONG-RUN U.S./U.K. REAL
EXCHANGE RATE

Charles Engel
Chang-Jin Kim

THE LONG-RUN U.S./U.K. REAL EXCHANGE RATE

ABSTRACT

We investigate the behavior of the long-run U.S./U.K. real exchange rate from 1885 to 1995. Our long-run real exchange rate series is derived from an unobserved components model which divides the real exchange rate into permanent and transitory components. The transitory component is modeled as having variances which switch, according to a Markov-switching process, among low, medium and high variance states. The underlying assumptions of our time-series model are based on an economic theory in which the permanent component represents real influences, while the transitory component represents primarily short-run movements due to nominal exchange rate fluctuations. Because the model is difficult to estimate by standard methods, we describe how the method of Gibbs sampling can handle this model. We find that our long-run real exchange rate series moves similarly to other measures proposed in the literature based on economic models.

Charles Engel
Department of Economics
University of Washington
Seattle, WA 98195
and NBER
cmengel@u.washington.edu

Chang-Jin Kim
Department of Economics
Korea University
Anam-dong, Seongbuk-ku
Seoul 136-701
KOREA
cjkim@kuccnx.korea.ac.kr

The notion that there are large and persistent transitory fluctuations in real exchange rates for industrialized countries is commonplace in the theoretical and empirical literature. One view is that nominal prices adjust slowly to their equilibrium values, while nominal exchange rates move quickly on financial news, so that the real exchange rate diverges considerably from its long-run value for long periods of time.

The behavior of the transitory component of the real exchange rate makes the detection and examination of the permanent component extremely difficult. To understand the behavior of its permanent part, the real exchange rate's activity over long periods of time must be studied. But, if the standard view of the transitory component is correct, then we are confronted with a bewildering variety of short-run movements.

Consider, for example, the factors influencing the U.S./U.K. real exchange rate from 1885 to 1995. Over much of the period prior to 1970, the nominal exchange rate between those two countries was fixed. During the period of fixed nominal exchange rates, the real exchange rate tended to be very stable (with the exception of a few periods of rapid price movements following each of the two World Wars and in the very late 19th century. See Figure 1.) However, during this period, there were occasionally large jumps in the real exchange rate that correspond to nominal revaluations.

We offer a new model that can separate the temporary from the permanent component when there are these types of switches in behavior. We capture the changes in regime that seem to characterize movements of the real exchange rate. Specifically, the model we propose for the U.S./U.K. real exchange rate is one where the permanent component of the real exchange rate follows a simple, homoskedastic random walk over the entire 111 year period, 1885-1995. But, the transitory component

1

switches regimes among periods of low, medium and high variance. While our general method allows for more diverse behavior (switches in regimes for both the permanent and transitory components, and switches in means as well as variances), this relatively simple model appears to do an excellent job characterizing the permanent and transitory components.

In fact, our measure of the two components seems very plausible on economic grounds, even though it was extracted using a pure time-series model. We find that it matches fairly well some measures of the long-run real exchange rate as functions of economic variables that have been proposed in the literature (such as a function of the relative levels of production in the two countries.) There are economic reasons for the a priori restrictions we place on the pure time-series model -- restrictions such as independence of the permanent and transitory shocks, homoskedasticity of the permanent shocks, heteroskedasticity of the transitory shocks, and a constant speed of adjustment -- which may contribute to the economic plausibility of our extracted series.

The model we propose would be quite difficult to estimate using standard procedures. There are unobserved permanent and transitory components, and the transitory component switches between unobserved states. However, we extend a relatively new estimation procedure -- Gibbs sampling -- to estimate our model. It turns out that Gibbs sampling handles a model such as ours easily. Also, the Gibbs sampling methodology allows us to perform in a natural way some diagnostics on the model allowing for our uncertainty about the parameter estimates.

Section 1 lays out the empirical model. Section 2 describes the data and the estimation technique. (A detailed outline of the estimation procedure is contained in the Appendix.) Section 3 presents the results of the estimation. In section 4, we show why we model the real exchange rate as containing a unit root component, even though standard tests for unit roots (such as the Augmented Dickey-Fuller test) reject

2

the null hypothesis of a unit root in our data. We extend the argument of Engel (1996) that these tests are very badly sized in some frequently encountered circumstances. Then in section 5, we compare the permanent component that our model extracts from the data to some measures of the permanent component derived from economic models (such as those in Mark and Choi (1996) and Engel (1995).) We find that our permanent component behaves similarly to some of the model-based measures over the 1960-1995 period. Section 6 concludes.

1. The Model

The real exchange rate q_t is assumed to be comprised of a permanent component, y_t, and a transitory component, x_t:

$$(1) \qquad q_t = y_t + x_t.$$

Our underlying philosophy in modeling the real exchange rate is that the permanent component represents the shocks to tastes and technologies that cause permanent movements in relative price levels. The temporary component mainly captures the slow movement of nominal prices to economic shocks. The restrictions we place on the behavior of y_t and x_t are motivated by this paradigm.

We take y_t to be a simple, homoskedastic random walk:

$$(2) \qquad y_t = y_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2).$$

We model the permanent component as being driftless. We believe the permanent component of the real exchange rate captures the economic forces described in classical optimizing models such as those of Razin (1995), Stockman and Tesar (1995), Asea and Mendoza (1994), Obstfeld and Rogoff (1995) and Brock and Turnovsky (1994). In these models, the real exchange rate represents the price of goods produced in one country relative to the price of goods produced in the other country, and this relative price changes over time as demand and supply change. It seems unlikely that

there would be taste changes or technological changes that would impart drift to the relative price over a period as long as 111 years. Thus, one of the identifying assumptions in our model is that the permanent component has no drift.

The homoskedasticity assumption is not necessitated by either economic reasoning or econometric methodology. Indeed, initially we allowed for the variance of the real exchange rate to change over time according to a Markov-switching process. However, we found that the heteroskedastic version of the model did not significantly improve on the homoskedastic version, so we have reverted to the simpler model. Note, though, that while there are obvious economic reasons for believing that the transitory component has switches in variance regimes (see below), it is not apparent a priori that such switches should occur in the variance of the permanent component. Hence, the finding of homoskedasticity of the v_t is not surprising.

The transitory component is assumed to follow

$$(3) \qquad x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + e_t, \qquad e_t \sim N(0, \sigma^2_{e,t}).$$

The AR(2) assumption is based on the data (that is, higher-order autoregressive terms are statistically insignificant.)

The variance of e_t is assumed to be time-varying. In particular, we assume that it depends on a discrete-valued first-order Markov-switching variable, S_t, (S_t = 1, 2, or 3) which evolves independently of v_t and e_t, according to the following transition probabilities:[1]

$$(4) \qquad \Pr[S_t = j \mid S_{t-1} = i] = p_{ij}; \qquad i,j = 1,2,3; \quad \sum_{j=1}^{3} p_{ij} = 1.$$

Thus, we can write the variance of the transitory component as:

$$(5) \qquad \sigma^2_{e,t} = \sigma^2_1 S_{1t} + \sigma^2_2 S_{2t} + \sigma^2_3 S_{3t}, \qquad \sigma^2_1 < \sigma^2_2 < \sigma^2_3,$$

---

[1] Engel and Hamilton (1990) employed Hamilton's (1989) Markov-switching model to examine switches in variances of exchange rate regimes. Kim (1993) combines variance switching with the unobserved components model.

4

where $S_{kt} = 1$ if $S_t = k$ $(k = 1,2,3)$, and $S_{kt} = 0$ otherwise. Note that by definition, state 1 is the low variance state, state 2 is the medium variance state, and state 3 is the high variance state.

We model the transitory component as switching between these three states because, under our economic interpretation of the transitory component, the variance is likely to be closely related to the variance of the nominal exchange rate. (See, Mussa (1986) and Stockman (1988).) This, in turn, appears to have three states: a very quiescent state in which the nominal exchange rate is fixed (or nearly so); an extremely volatile state during which revaluations and devaluations occur under a "fixed" regime, or when there are extremely rapid price changes in one of the countries; and, an intermediate level of volatility during which nominal rates are floating.

We also assume the shocks v_t and e_t are independent. This assumption is important in helping to distinguish the temporary from the permanent component. It is motivated by our economic interpretation of the two components. Shocks to the permanent component represent taste and technology changes that favor one sector relative to another. These cause prices of some goods to change relative to others, which leads to a change in the relative aggregate price levels of the two countries. On the other hand, we interpret shocks to the transitory component as coming primarily from nominal demand sources, such as monetary shocks. To a first approximation, these aggregate nominal shocks can be thought of as independent of the sectoral real shocks that drive the permanent component.

2. Data and Estimation

The real exchange rate data is constructed from producer price indexes for the U.S. and the U.K., and the nominal dollar/pound exchange rate, monthly from January

5

1885 to November 1995. The data is updated from the data set used by Grilli and Kaminsky (1991). A complete description of the data sources is available in that paper.

The model is estimated using the technique of Gibbs sampling. Because this technique is relatively new, and has not been used much in applications in economics, we will briefly describe the technique here. More complete introductions to Gibbs sampling are available in Albert and Chib (1993), Casella and George (1992), and Gelfand and Smith (1990). Our use of the Gibbs sampler is closely related to that of Albert and Chib, who use it to estimate Hamilton's (1989) autoregressive time-series model with Markov-switching, and to that of Carter and Kohn (1994), who use it to make inferences on unobserved components in a state-space model. Our application extends the Gibbs sampling technique to general state-space models such as those estimated by Kim (1994) using maximum likelihood techniques. (For applications of the unobserved component model with Markov-switching heteroskedasticity, see Kim (1993) and Kim and Kim (1996).)

Gibbs sampling is a technique for generating random variables from a distribution indirectly, without having to calculate the density. For example, suppose we would like to make inferences about two unobserved random variables, P and R, given some data, $\tilde{W}^T$, which is a vector of T observations on some random variable W_t. In particular we would like to obtain the densities, $f(p|\tilde{W}^T)$ and $f(r|\tilde{W}^T)$, because we are interested in, say, the mean and variance of these densities.

The Gibbs sampling technique allows us to generate a sample from $f(p|\tilde{W}^T)$ and $f(r|\tilde{W}^T)$ by sampling from the densities $f(p|R,\tilde{W}^T)$ and $f(r|P,\tilde{W}^T)$. As we shall see, frequently the latter densities are readily available, while $f(p|\tilde{W}^T)$ and $f(r|\tilde{W}^T)$ are not. Begin with an initial guess for a value of R, call it r_0. Then, we can generate a draw, call it p_1, from the density $f(p|r_0,\tilde{W}^T)$. Take this value of p_1, and generate r_1 from $f(r|p_1,\tilde{W}^T)$. Continue this process iteratively to generate the sequence

6

$r_0, p_1, r_1, p_2, r_2, \ldots, p_k, r_k$. These values are realizations of the sequence of random variables $R_0, P_1, R_1, P_2, R_2, \ldots, P_k, R_k$. Under fairly general conditions, the density of P_k converges to $f(p \mid \widetilde{W}^T)$ and R_k to $f(r \mid \widetilde{W}^T)$.

Thus, one might generate K values of p_i and r_i. Then, if K is large enough, it might be reasonable to assume for values of $i > K$, p_i and r_i are drawn from their true distributions, $f(p \mid \widetilde{W}^T)$ and $f(r \mid \widetilde{W}^T)$. So, then, we can generate a sample of P and R drawn from these distributions by generating an additional N draws. These N draws then can be used to calculate the moments of $f(p \mid \widetilde{W}^T)$ and $f(r \mid \widetilde{W}^T)$ that we are interested in.

For the model estimated here, we actually would like to find the density of 3T + 12 random variables, conditional on the vector of observed real exchange rates, $\widetilde{Q}^T$. There are T values of the permanent component, y_t (call this vector $\widetilde{Y}^T$); T values of the transitory component, x_t (call this vector $\widetilde{X}^T$); T values of the state, S_t (call this vector $\widetilde{S}^T$); and twelve parameters, $\widetilde{\theta}$ = $\{\phi_1, \phi_2, \sigma_\nu^2, \sigma_{e,1}^2, \sigma_{e,2}^2, \sigma_{e,3}^2, p_{11}, p_{12}, p_{21}, p_{31}, p_{32}\}$. Notice that the parameters of the model are treated as random variables with prior distributions in the Bayesian context.

While the derivation of 3T + 12 random samples based on appropriate conditional densities is slightly more complicated than the example above with two random variables, the principles are the same. The data on real exchange rates, $\widetilde{Q}^T$ play the role of the observable $\widetilde{W}^T$ in our example above.

The analogs to $f(p \mid R, \widetilde{W}^T)$ and $f(r \mid P, \widetilde{W}^T)$ in our example would be the densities of each row or element of $\widetilde{Z}^T = [\widetilde{Y}^T \ \widetilde{X}^T]$, $\widetilde{S}^T$ and $\widetilde{\theta}$ conditional on $\widetilde{Q}^T$ and on all of the other values of $\widetilde{Z}^T$, $\widetilde{S}^T$ and $\widetilde{\theta}$. The above example was one of "single-move" Gibbs sampling. We actually make use of the "multi-move" Gibbs sampling of Carter and Kohn (1994). In multi-move Gibbs sampling, a Gibbs block can be the whole $\widetilde{Z}^T$ matrix or $\widetilde{S}^T$

vector, as opposed to the single-move Gibbs sampling, in which each row or element of the Z^T matrix or the S^T vector is generated one at a time.

These conditional densities of each Gibbs block of Z^T and S^T, and of each element of $\tilde{\theta}$, are easily derived. For example, conditional on S^T and $\tilde{\theta}$, the model reduces to a standard linear unobserved components model. So $f(Z^T | S^T, \tilde{\theta}, Q^T)$ can be generated as in Carter and Kohn (1994). Conditional on Z^T, the distribution of S^T is independent of the data, Q^T, and just depend on the values of X^T. So, $f(S^T | Z^T, \tilde{\theta}, Q^T) = f(S^T | X^T, \tilde{\theta})$. The problem of finding $f(S^T | X^T, \tilde{\theta})$ was solved by Albert and Chib (1993). The derivation of $f(\tilde{\theta} | Z^T, S^T, Q^T)$ is a straightforward Bayesian exercise. The detailed derivation of each conditional distribution for our Gibbs sampler is given in the Appendix.

In our estimation, the Gibbs sampler appears to converge after about two hundred iterations. We run the Gibbs sampler for eleven thousand observations, and, to be on the safe side, discard the first one thousand. Our distributions of Y^T, X^T and S^T are based on the last ten thousand iterations. For our distribution of $\tilde{\theta}$, we take every fifth observation from the final ten thousand observations (because of potential serial correlation across the iterations.)

3. Estimation Results

Table 1 presents our parameter estimates. The numbers reported are the mean of our Gibbs samples, and the 95 per cent confidence intervals.

There are several interesting things to note about these parameter estimates. First, the variance of the permanent component is relatively small in comparison to the variances of the transitory component in its three states. The y_t variance is about 1.01, which is slightly larger than the variance of the x_t component in its low-variance state (0.81), but much smaller than the variance of the x_t component in its

medium-variance (5.93) and high-variance (24.70) states. This indicates that the real shocks that, according to our paradigm, drive the permanent component tend to be relatively small compared to the nominal shocks that cause temporary deviations in the real exchange rate from its permanent component.

It might seem that we "forced" the transitory component to have the shocks with higher variance, because we allow the transitory component to shift among three different regimes, while we only allow a single variance regime for the permanent shock. But, initially we allowed for two variance regimes for the permanent component (and three for the transitory component). For the permanent component, the variance of those two regimes would be $\sigma_{\nu,1}^2$, and $\sigma_{\nu,2}^2$. We estimate $\sigma_{\nu,2}^2$ by estimating the parameter $h_{\nu,2}$ defined by $\sigma_{\nu,2}^2 = \sigma_{\nu,1}^2(1+h_{\nu,2})$. We found that the 95 per cent confidence interval on $h_{\nu,2}$ was very wide and easily contained zero. On the other hand, examination of Table 1 shows that the confidence intervals for the estimates of the variances of the transitory component in each state do not overlap. This is a conservative way of seeing that the variance is significantly different in the three states. So, the data appear to be telling us that the high-variance regimes for the real exchange rate correspond to high variance regimes for its transitory component.

Next, we note that the low variance and medium variance regimes are extremely persistent. The expected duration of the low-variance state, given by $\frac{1}{1-p_{11}}$, is equal to 74.8 months, or over six years. The expected duration of the medium-variance state, $\frac{1}{1-p_{22}}$, is 38.7 months. The high-variance state is not as persistent, with an expected duration ($= \frac{1}{p_{31}+p_{32}}$) of 6.4 months. We shall see shortly that the medium-variance regime generally corresponds to the post-1973 era of floating exchange rates; the low-variance regime is associated with the periods of fixed nominal exchange rates prior to 1973; and the high-variance regime generally matches either times of exchange rate revaluation prior or periods of rapid inflation in one of the two countries prior to 1973.

The transitory component, x_t, itself is highly persistent. The sum of ϕ_1 and ϕ_2 is 0.987. This implies a half-life of transitory shocks of 55 months. This estimate of the persistence of transitory shocks is very much consistent with other estimates in the literature (see Rogoff (1996).) However, it points out the difficulty alluded to in the introduction to this paper. Because transitory shocks are so persistent, and at times highly volatile, it is difficult to extract the permanent component.

Figure 2 plots the permanent and transitory components of the log of the real exchange rate. In figure 2.1, the permanent component is plotted along with its 95 per cent confidence interval.[2] The real exchange rate here is defined as the relative price of U.K. to U.S. producer goods (that is, the product of the dollar/pound nominal exchange rate and the U.K. producer price index, all divided by the U.S. producer price index.) The diagram shows that the permanent component evolves gradually (low innovation variance), and that there has been a tendency for the permanent component of the relative U.K./U.S. prices to rise since World War II.

Figure 2.2 shows the transitory component, x_t, and its 95 per cent confidence interval. The transitory component is subject to rapid swings, which correspond to the periods of rapid swings in the real exchange rate itself. There are periods of high volatility compared to the permanent component, including a sustained period after 1973.

In figure 3, we plot the probability that the transitory component is in each of states 1, 2 and 3 at each point in time. (These correspond to the smoothed probabilities described in Hamilton (1989)). First, consider Figure 3.2, which plots the probability of being in the medium-variance state. For almost the entire period

---

[2] The obvious method of arriving at the 95 per cent confidence interval would be to take the central 95 per cent of the draws from the Gibbs sampler for each point in time. However, this would require saving an enormous amount of data in memory on the computer. Instead, we calculated the 95% confidence interval as the 1.96 standard deviation band (which necessitated only saving the cumulative sums of y_t and y_t^2.) Given the assumption of Normal distributions, the two methods are asymptotically equivalent.

10

after 1973, the probability of $S_t = 2$ is near unity. We can associate the medium-variance state with the period of floating nominal exchange rates. Interestingly, the other periods in which there is a fairly high probability that $S_t = 2$ also correspond to times in which the dollar/pound nominal exchange rate was floating. As Grilli and Kaminsky (1991) summarize it, there was a period of controlled floating from December 1914 to March 1919, and a regime with a more volatile float from April 1919 to April 1925. During this period, our estimates show that the probability that $S_t = 2$ hovers between 0.25 and 0.70. There is one other period in which the probability that $S_t = 2$ rises above 0.5: in the early 1930s, when the dollar/pound exchange rate returned to floating.

Figure 3.1 shows that for most of the time prior to 1973 (with the exception of those periods just discussed), the probability that we are in the low-variance state $(S_t = 1)$ is very nearly one. Those times in which the real exchange rate is in its low-variance state correspond to times in which the nominal exchange rate was fixed.

Figure 3.3 shows several short periods in which the probability that we are in the high-variance state is high, while most of the time the probability is near zero. We can identify each of these high-variance times with specific historical events. The first occurred in 1898-1899 and again at the end of 1900, when there were large fluctuations in prices in both the U.S. and U.K. These can be associated with the period of the Alaskan Gold Rush. There was a great deal of instability in the money supply during this period, according to to Friedman and Schwartz (1963, p. 145.)

Next, we find that the periods in the early 1920s and the 1930s in which the nominal exchange rate floated and was very volatile are also periods in which there is a high probability that $S_t = 3$. Note that the model assigns some probability that these are periods of $S_t = 2$, but a higher probability that $S_t = 3$. It appears that during floating nominal exchange rate regimes, when the volatility of nominal exchange rates is fairly large (as it was after 1973, in the period 1914 to about 1922 and in

11

1932), the model assigns the real exchange rate to the medium volatility regime. When nominal exchange rates are extremely volatile (as in the 1922-1925 period and most of the period from 1931-1933) the model places a high probability on $S_t = 3$. We should note that during some of these early periods in which there is a high probability that $S_t = 2$ or $S_t = 3$, there is not only nominal exchange rate volatility, but also periods in which nominal prices fluctuate more than they have tended to in the post- World War II era.

The next period in which there is a high probability of the high variance state is late 1939, when Britain devalued the pound. Two other periods in which it is likely that $S_t = 3$ correspond to devaluations of the pound: in 1949 and 1967.

The remaining period of high volatility in the real exchange rate is immediately following World War II, in which there was rapid inflation in the U.S. after price controls were removed.

So, all of the periods of medium and high volatility of the real exchange rate appear to be linked to nominal events -- either periods in which the nominal exchange rate floated, or periods of extraordinary inflation that can be linked to monetary events. This pattern is reminiscent of Mussa's (1986) finding that real exchange rate volatility increased directly with nominal exchange rate volatility for a large number of time periods and currencies.

4. Unit Roots

The Augmented Dickey-Fuller (ADF) test estimates the following regression equation:

(6) $\quad q_t = \beta_0 + \rho q_{t-1} + \beta_1(q_{t-1}-q_{t-2}) + \beta_2(q_{t-2}-q_{t-3}) + ... + \beta_k(q_{t-k}-q_{t-k-1}) + u_t.$

We estimate equation (6) for the real exchange rate data, choosing the lag length k by a data-based method, as advocated by Ng and Perron (1995). We initially estimate the

12

equation with 12 lags, and test the significance of β_{12}. If it is not significantly different from zero at the 95 per cent level, we drop the last lag and reestimate the equation with $k = 11$, etc. We in fact end up with $k = 2$.

Our Dickey-Fuller statistic is 3.007, which just exceeds the five per cent critical value of 2.86. The estimated value of ρ is 0.9877.

If we can reject a unit root, why do we model the real exchange rate as having a permanent component as in section 2?

Engel (1996) argues that the ADF test has very large size biases when the variable being tested is, in fact, the sum of a random walk component and a very persistent and highly volatile transitory component. Engel presents an example in which the real exchange rate is posited to be the sum of a random walk and an AR(1). Both components are assumed to be homoskedastic, but the innovation variance of the stationary component is assumed to be very large compared to that of the random walk. He shows that the stationary representation for the real exchange rate is that the first-differences follow an ARMA(1,1) process. The root of the MA component is near unity in absolute value. The ADF test, which is designed to test against the null of stationary AR processes (i.e., with no MA component), has a very large size bias even when the random walk component accounts for a large fraction of the variance of the real exchange rate change within the sample.

Engel's (1996) examples are all for homoskedastic permanent and stationary components. In this section, we repeat some of Engel's Monte Carlo exercises to assess the size bias of the ADF test under the null hypothesis that the real exchange rate is generated by the model of section 2.

Our first Monte Carlo exercise generates ten thousand draws from the model of section 2, using the parameter estimates reported in Table 1. The program was run in Gauss, using the Gauss random number generators. The lag length of each of the 10,000 ADF tests was chosen by the iterative procedure described above.

The first panel of Table 2 reports the true size of the ADF test for various nominal sizes. For example, if the nominal size is 0.05, we find the test actually rejects nearly 30 per cent of the time (29.51 per cent). One exercise in Engel (1996) is to calibrate his homoskedastic model to one hundred years of U.S./U.K. real exchange rates.[3] Depending on the parameter values chosen, the true size of the ADF test with nominal size of 0.05 ranged from 0.2986 to 0.3566. These numbers are not far off from our finding.

Since the "5 per cent" ADF test actually rejects nearly 30 per cent of the time when the model of section 2 is correct, we should not take the evidence that we just barely reject a unit root at the "5 per cent" level as very strong evidence against the model.

This first Monte Carlo exercise was undertaken assuming that the parameter estimates reported in Table 1 are the true values of the parameters. We also assess the size of the ADF test taking into account our uncertainty about the parameters. To do this, we generate a draw of the parameters from the Gibbs sampler. We then use that parameter draw to generate an artificial real exchange rate series. We run the ADF test on our artificial series. We repeat this procedure ten thousand times. Thus, our ten thousand ADF tests are on series generated from different parameter values each time, with the distribution of parameter values corresponding to the distribution generated by the Gibbs sampler.

The second panel of Table 2 reports those results. It shows that there is still substantial size bias in the ADF tests, though somewhat less than in the first panel which assumed parameter certainty. Thus, a nominal five per cent test has a true size of 19.08 per cent taking into account parameter uncertainty, but 29.51 per cent when we assume the parameters are known.

---

[3] Engel's real exchange rates were based on personal consumption deflators, rather than the PPI real exchange rates used here.

5. Interpretation of Permanent Component

In this section we investigate the nature of our generated permanent component, y_t. We ask whether our series, which comes from a univariate time-series model with a priori restrictions motivated from economic theory, looks like other measures of the permanent real exchange rate that have been proposed in the literature. Specifically, we look at other measures which are based on economic models and use data other than the real exchange rate in their construction.

Five of the alternative permanent series we use for comparison are from Mark and Choi (1996). The sixth comes from Engel (1995).

The series from Mark and Choi are constructed by first positing models of the long-run real exchange rate, y_t, as functions of economic variables. Call the vector of economic variables for model i, w_{it}. Then Mark and Choi estimate the cointegrating vector between q_t and w_{it}. The fitted values from the cointegrating regressions are taken as the measure of the long-run real exchange rate.

The models of the long-run real exchange rate, v_{it} (i = 1,2,...,6), from Mark and Choi that we consider are:

1) v_{1t} is a function of the productivity in U.K. manufacturing relative to productivity in U.S. manufacturing. The notion here is that if the relative productivity in the U.K. increases faster than in the U.S., relative U.K./U.S. prices should decline because the U.K.'s relative costs have declined.

2) v_{2t} is a function of the relative productivities and of the difference in the ratio of government consumption to income in the U.K. and U.S. The government spending variable enters here following Rogoff's (1992) analysis which allows for demand factors to have effects on the long-run real exchange rate under intersectoral factor immobility and poor international capital market integration.

15

3) v_{3t} is a function of per capita output in the U.K. relative to the U.S. One motivation for this might be similar to the motivation for using productivity measures as in (1) above: the faster growing country might be experiencing declining relative costs.

4) v_{4t} is a function of the real interest differential between the U.K. and the U.S. Mark and Choi mention that in Dornbusch (1976), the real exchange rate is related to the real interest differential, though they acknowledge that the relationship is posited to be a short-run one in Dornbusch.

5) v_{5t} is a function of the relative money supplies and relative levels of output in the U.K. and U.S. Here, the motivation is that the effects of monetary shocks on the real exchange rate might indeed be permanent (or, so long-lasting as to be indistinguishable from permanent), so that long-run monetary neutrality is rejected.

(6) Our final formulation, v_{6t}, comes from Engel (1995), who examines various measures of the relative price of non-traded goods. Many models of real exchange rates relate its movement to a relative relative price: the price of non-traded to traded goods in one country relative to the other country. Here, we take as a measure of the price of non-traded to traded goods the consumer price index relative to the producer price index in the U.K. and the U.S. A priori, it seems unlikely that this model will hold up well. The real exchange rate we are examining in this paper is one based on producer prices, which generally price goods which are tradable. So, the relative price of non-traded goods does not seem to be a good candidate for explaining this real exchange rate.[4]

The measures of the long-run real exchange rate from Mark and Choi are monthly from 1960:1 to 1993:11. The relative relative price from Engel is monthly from 1963:1 to 1995:11.

---

[4] In fact, the producer price indexes enter the two relative prices here with opposite signs. Our measure of the real exchange rate, q_t, takes s + ppiuk - ppius, while v_{6t} takes cpiuk - ppiuk - (cpius - ppius), where all these variables are in logs, and s is the log of the nominal exchange rate.

We would like to get an idea of how closely our series for the permanent component, y_t, moves with these other measures. It would be inappropriate to look at something like their correlation coefficient. All series are assumed to be I(1). Instead, we investigate the hypothesis that the series are cointegrated, with cointegrating vector (1,-1). That is, we ask whether they move together in the long-run.

Our series y_t is posited to be a pure random walk. The series based on economic data could have some transitory component as well as the permanent component. So, movements in y_t and these series should not necessarily be closely related in the short run. But, we can test the long-run hypothesis of no cointegration with a known cointegrating vector (1,-1) by implementing the error-correction model test proposed by Zivot (1995). With v_{it} being the economic measures described above (i = 1,2,...,6), we run the following regressions:

(7) $\quad \Delta v_{it} = \alpha + \rho(y_{t-1} - v_{it-1}) + \delta_0 \Delta y_t + \delta_1 \Delta y_{t-1} + ... + \delta_k \Delta y_{t-k} + \gamma_1 \Delta v_{it-1} + \gamma_2 \Delta v_{it-2}$

$\qquad + ... + \gamma_k \Delta v_{it-k}.$

The lag length is chosen by an iterative method similar to the one described above for the ADF test. Zivot demonstrates that the critical values derived in another context by Hansen (1995) are appropriate for the error-correction model of Kremers, Ericsson and Dolado (1992) when the cointegrating vector is known.

The first column of Table 3 reports the slope coefficients from regressions of v_{it} on y_t. If we were interested in estimating the cointegrating vector, these coefficients, β_i, would be an estimate of the cointegrating vector $(1,-\beta_i)$. However, here we present the β_i more to give an intuition of how these series move together over the long run, while we formally test the null that they are not cointegrated with cointegrating vector (1,-1). We see that v_1, v_4 and v_6 do not move with y over the thirty-plus years of data after 1960. The productivity differential measure, the real

interest differential measure, and the non-traded goods price measure are driven by some forces other than the ones driving y_t.

It is notable that β_6 is negative. The relative price of non-traded goods seems particularly unrelated to our measure of the permanent component of real exchange rates. This is consistent with Engel's (1995) observation that this relative price generally accounts for little of the real exchange rate movement at short, medium and long horizons.

On the other hand, β_2, β_3 and β_5 are not too far away from unity. Indeed, with $\beta_3 = 1.132$, it appears that relative income movements (v_{3t}) and y_t move together one-for-one.

The second column of Table 3 reports the test of the null that the the series are not cointegrated with cointegrating vector (1,-1). To be clear, the null can be rejected only when the series are cointegrated and have a cointegrating vector of (1,-1). The column reports three numbers -- the "t-statistic" on $\hat{\rho}$ from equation (7), and the critical values at the 5 per cent and one per cent levels.

We see that we fail to reject the null for all but two series, v_{3t} and v_{5t}. The null that v_{5t} is not cointegrated with y_t with cointegrating vector [1,-1] is barely rejected at the five per cent level, while the same null for v_{3t} and y_t is easily rejected at the one per cent level.

The v_{5t} measure is an unusual one -- we would not expect the long-run real exchange rate to depend on monetary variables. However, note that v_{5t} also incorporates income variables. Perhaps their inclusion dominates the movement of v_{5t} over time. Or, perhaps the money supplies are endogenous and respond to some unmeasured real variables that more plausibly might drive the real exchange rate in the long run.

In any event, the relation between v_{3t} and y_t is clearly the strongest. The two variables seem to be capturing the same long-run movements. If in fact our measure of

18

the permanent component of the real exchange rate is accurate, this finding suggests that relative output measures may play a role in determining long-run real exchange rate movements.

6. Conclusions

Typically it is difficult to recover the permanent component of real exchange rates. The long-run movement of the real exchange rate is masked by the shorter run movement. To extract the long-run movement, we need a model of the short-run changes that is sufficiently rich -- particularly in capturing the different variance regimes. Here, we propose such a model. We overcome the difficulty in estimating the model by extending the techniques of Gibbs sampling to general state-space models.

We find:

1) The transitory component is highly persistent.

2) The regimes of variance of the transitory component of the real exchange rate are closely related to monetary phenomena, as Mussa (1986) has claimed.

3) Innovations in the permanent component have relatively low variance.

4) Relative per capita output levels may be important in understanding the behavior of the long-run real exchange rate.

In addition, we have shown that the size bias in unit roots test of real exchange rates that Engel (1996) discusses extends to the heteroskedastic model we consider.

The techniques we employ in this paper will be useful in further studies of real exchange rates. A next step that likely would be beneficial is to model jointly the behavior of real exchange rates and output levels. Since the real exchange rate and relative outputs appear to share a common trend, the data on each would be mutually reinforcing in detecting that trend.

References

Albert, James H., and Siddartha Chib, 1993, Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts, Journal of Business and Economic Statistics 11, 1-15.

Asea, Patrick K., and Enrique G. Mendoza, 1994, The Balassa-Samuelson model: A general equilibrium appraisal, Review of International Economics 2, 244-267.

Brock, Philip L., and Stephen J. Turnovsky, 1994, The dependent-economy model with both traded and non-traded capital goods, Review of International Economics 2, 306-325.

Carter, C.K. and P. Kohn, 1994, On Gibbs sampling for state space models, Biometrika, 81, 541-553.

Dornbusch, Rudiger, 1976, Expectations and exchange rate dynamics, Journal of Political Economy 84, 1161-1176.

Engel, Charles, 1995, Accounting for U.S. real exchange rate changes, National Bureau of Economics, working paper no. 5394.

Engel, Charles, 1996, Long-run PPP may not hold after all, National Bureau of Economics, working paper no.5646.

Engel, Charles and James D. Hamilton, 1990, Long swings in the dollar: Are they in the data and do markets know it?, American Economic Review 80, 689-713.

Friedman, Milton and Anna J. Schwartz, 1963, A Monetary History of the United States, 1867-1960 (Princeton).

Gelfand, Alan E., and Adrian F.M. Smith, 1990, Sampling-based approaches to calculating marginal densities, Journal of the American Statistical Association 85, 398-409.

Grilli, Vittorio and Graciela Kaminsky, 1991, Nominal exchange rate regimes and the real exchange rate: Evidence from the United States and Great Britain, 1885-1986, Journal of Monetary Economics 27, 191-212.

Hamilton, James D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica 57, 357-384.

Hansen, Bruce E., 1995, Rethinking the univariate approach to unit root testing: Using covariates to increase power, Department of Economics, Boston College.

Kim,Chang-Jin, 1993, Unobserved-component time-series models with Markov-switching heteroskedasticity: Changes in regime and the link between inflation rates and inflation uncertainty, Journal of Business and Economic Statistics 11, 341-349.

Kim,Chang-Jin, 1994, Dynamic linear models with Markov-switching, Journal of Econometrics 60, 1-22.

Kim,Chang-Jin, and Myung-Jig Kim, 1996, Transient fads and the crash of '87, Journal of Applied Econometrics 11, 41-58.

20

Kim, Chang-Jin; Charles R. Nelson; and Richard Startz, 1996, Testing for mean reversion in heteroskedastic data based on Gibbs-sampling augmented randomization, Journal of Empirical Finance, forthcoming.

Kremers, Jeroen J.M.; Neil R. Ericsson; and, Juan J. Dolado, 1992, The power of cointegration tests, Oxford Bulletin of Economics and Statistics 54, 325-348.

Mark, Nelson C., and Doo-Yull Choi, 1996, Real exchange-rate prediction over long horizons, Journal of International Economics, forthcoming.

Mussa, Michael, 1986, Nominal exchange rate regimes and the behavior of real exchange rates: Evidence and implications, in Karl Brunner and Allan Meltzer, eds., Carnegie-Rochester Series on Public Policy 25, 117-214.

Ng, Serena and Pierre Perron, 1995, Unit roots tests in ARMA models with data dependent methods for the truncation lag, Journal of the American Statistical Association 90, 268-281.

Obstfeld, Maurice, and Kenneth Rogoff, 1995, The intertemporal approach to the current account, in Gene M. Grossman, and Kenneth Rogoff, eds., Handbook of International Economics, (Amsterdam: North Holland).

Razin, Assaf, 1995, The dynamic optimizing approach to the current account: Theory and evidence, in Peter B. Kenen, ed., Understanding Interdependence: The Macroeconomics of the Open Economy (Princeton).

Rogoff, Kenneth, 1992, Traded goods consumption smoothing and the random walk behavior of the real exchange rate, Bank of Japan Monetary and Economic Studies 10, 1-29.

Rogoff, Kenneth, 1996, The purchasing power parity puzzle, Journal of Economic Literature 34, 647-668.

Stockman, Alan C., 1988, Real exchange-rate flexibility under pegged and floating exchange-rate systems, in Karl Brunner and Allan Meltzer, eds., Carnegie-Rochester Series on Public Policy 29: 259-294.

Stockman, Alan C. and Linda L. Tesar, 1995, Tastes and technology in a two-country model of the business cycle: Explaining international comovements, American Economic Review 85, 168-185.

Zivot, Eric, 1995, The power of single equation tests for cointegration when the cointegrating vector is known, Department of Economics, University of Washington.

Table 1

Parameter Estimates and 95% Confidence Bands

Parameter	Mean	Lower Band	Upper Band
p_{11}	0.9866	0.9774	0.9935
p_{12}	0.0037	0.0002	0.0101
p_{21}	0.0103	0.0005	0.0309
p_{22}	0.9735	0.9407	0.9934
p_{31}	0.1037	0.0408	0.1961
p_{32}	0.0531	0.0028	0.1549
ϕ_1	1.1308	1.0486	1.2115
ϕ_2	-0.143	-0.223	-0.064
$\phi_1 + \phi_2$	0.9876	0.9777	0.9917
σ_v^2	1.0014	0.7603	1.2911
σ_1^2	0.8167	0.6426	1.0057
σ_2^2	5.9347	4.4617	7.7379
σ_3^2	24.992	15.920	40.059

Parameters are from the model of equations (1)-(5). "Mean" is the mean of the estimates from the Gibbs samples. "Lower band" and "Upper band" refer to the 95 per cent confidence interval.

Table 2

Size Distribution of Augmented-Dickey Fuller Test

Nominal Size	0.01	0.025	0.05	0.10
Parameters Known	0.1079	0.1946	0.2951	0.4438
Parameter Uncertainty	0.0744	0.1272	0.1908	0.2791

The two lines give the true size of the ADF test, first when the parameters of the model (1)-(5) are treated as known; and, second, when a distribution of parameters is generated by the Gibbs sampler.

Table 3

Test of Cointegration of v_{it} and y_t

	$\hat{\beta}$	t-stat	5%	1%
v_{1t}	-0.026	-0.523	2.69	3.28
v_{2t}	0.825	1.403	2.75	3.35
v_{3t}	1.132	4.090	2.86	3.43
v_{4t}	0.140	1.742	2.75	3.36
v_{5t}	1.385	2.917	2.86	3.43
v_{6t}	-0.571	-0.513	2.86	3.43

"v_{it}" refers to measures of the long-run real exchange rate from section 5 of the paper. "$\hat{\beta}$" is the slope coefficient from a regression of v_{it} on y_t. "t-stat" refers to the Student statistic on ρ in regression (7). "5%" and "1%" are the 5 per cent and 1 per cent critical values, respectively.

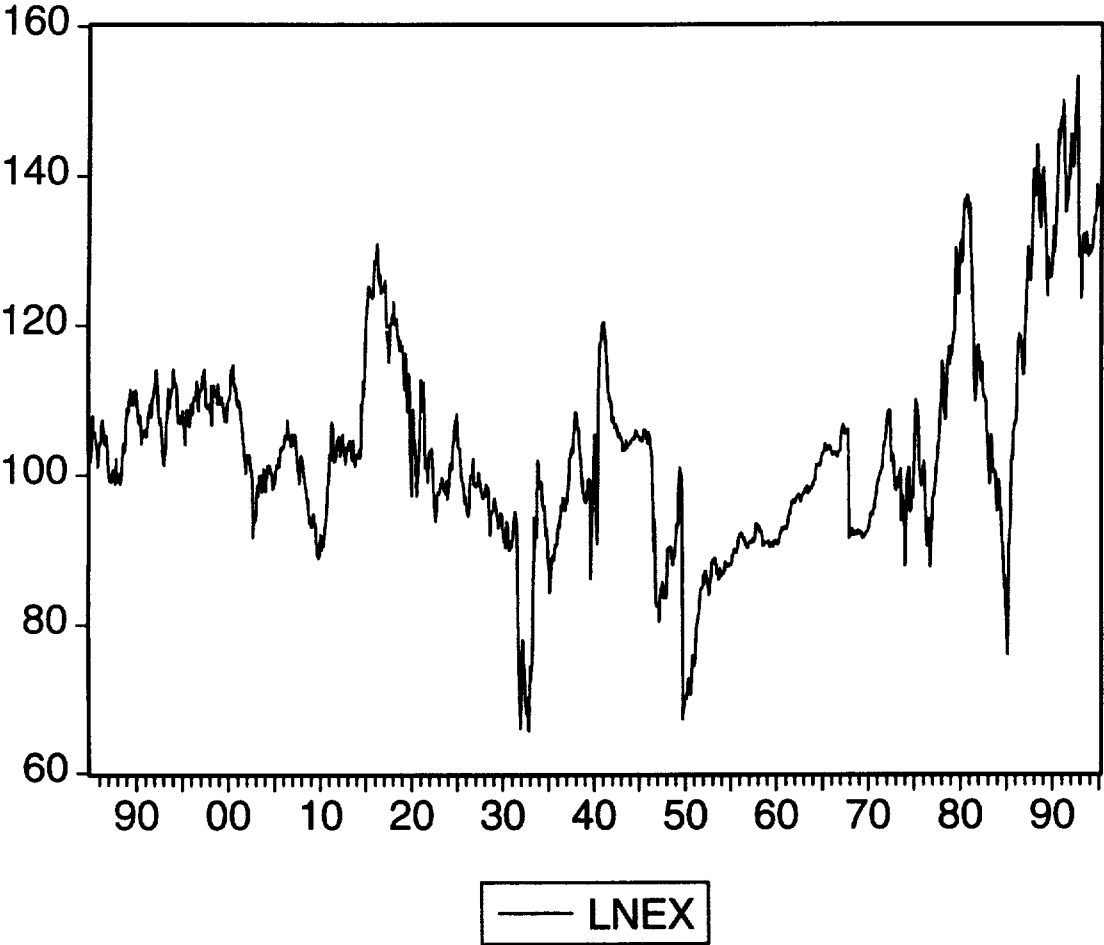Figure 1. Log of U.S./U.K Real Exchange Rates, 1885-1995.

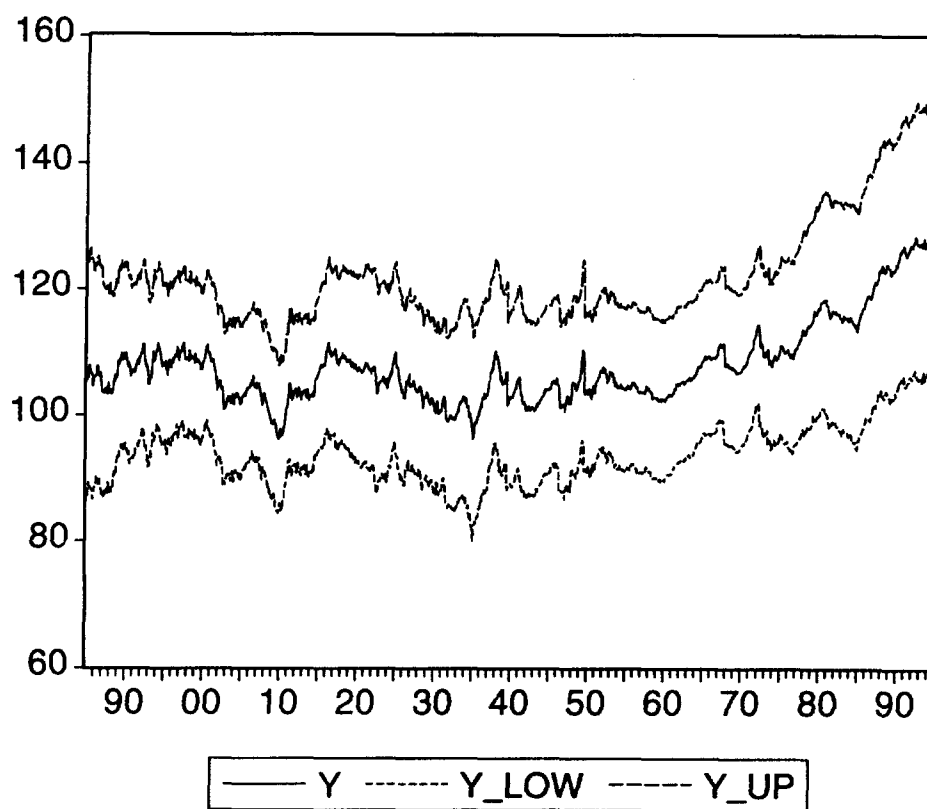Figure 2.1 Permanent Components (Yt) and their 95% Confidence Bands

Legend: Y -------- Y_LOW ------ Y_UP

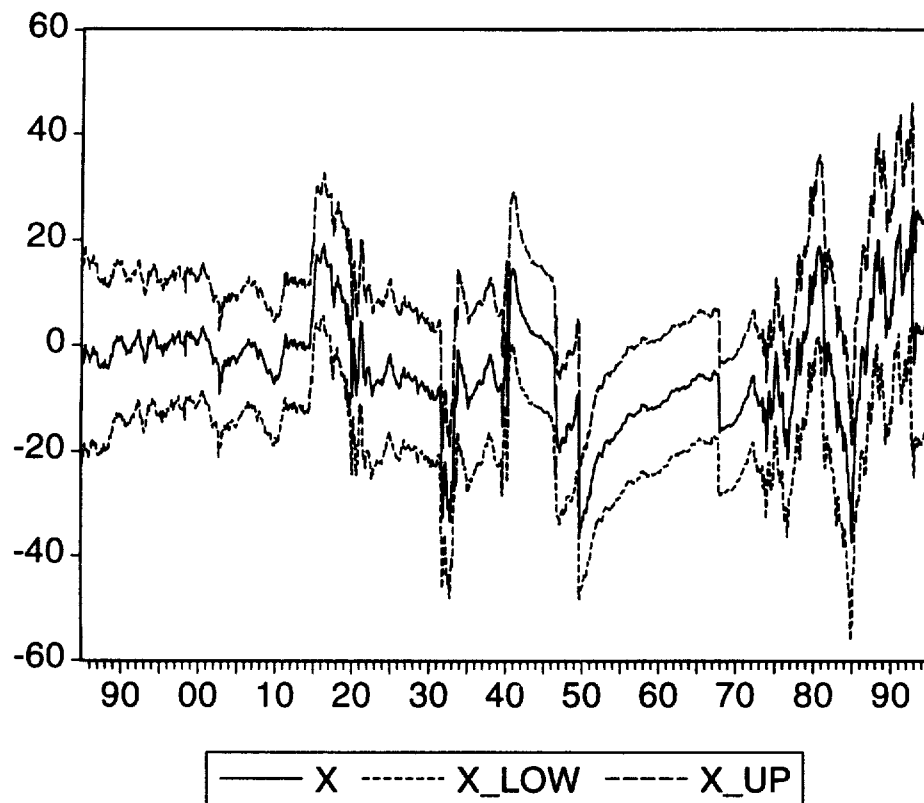Figure 2.2. Transitory Components (Xt) and their 95% Confidence Bands
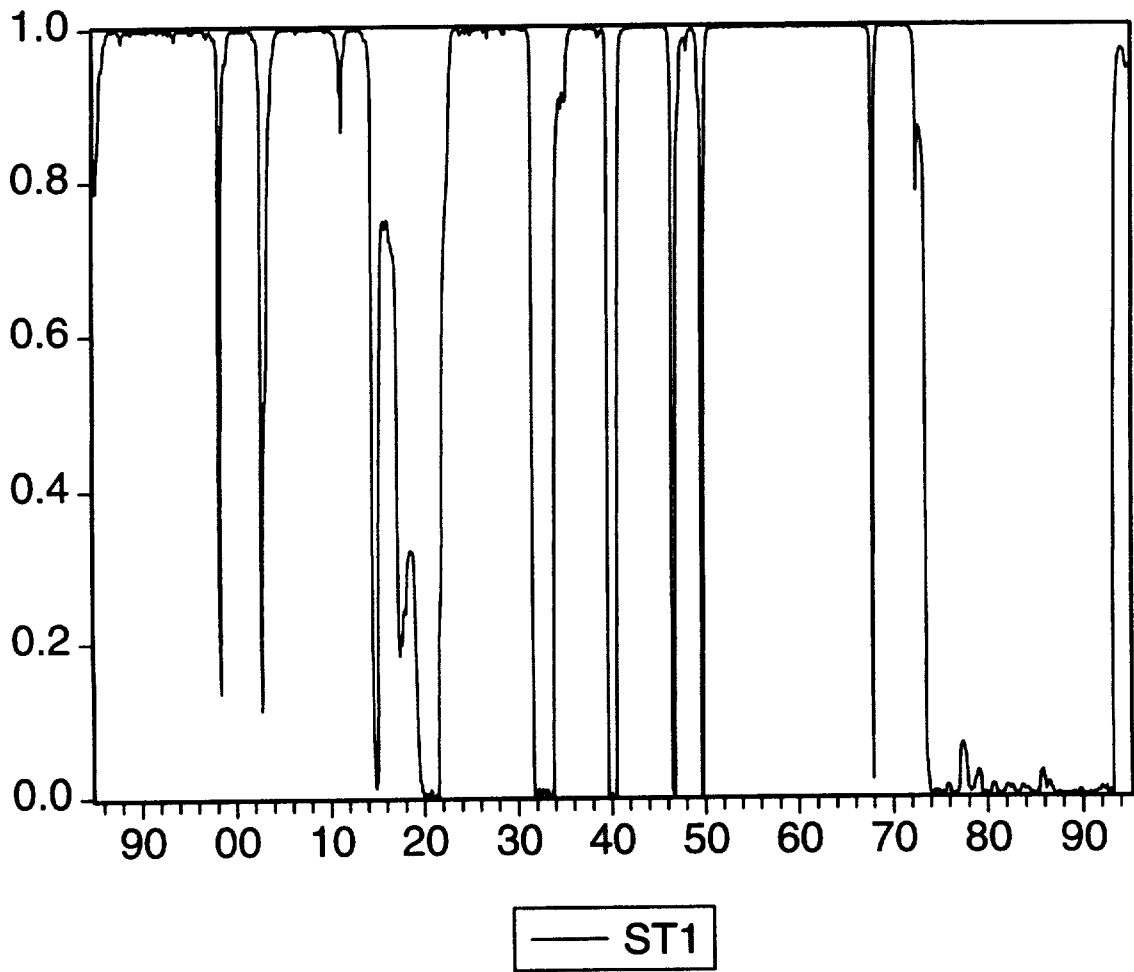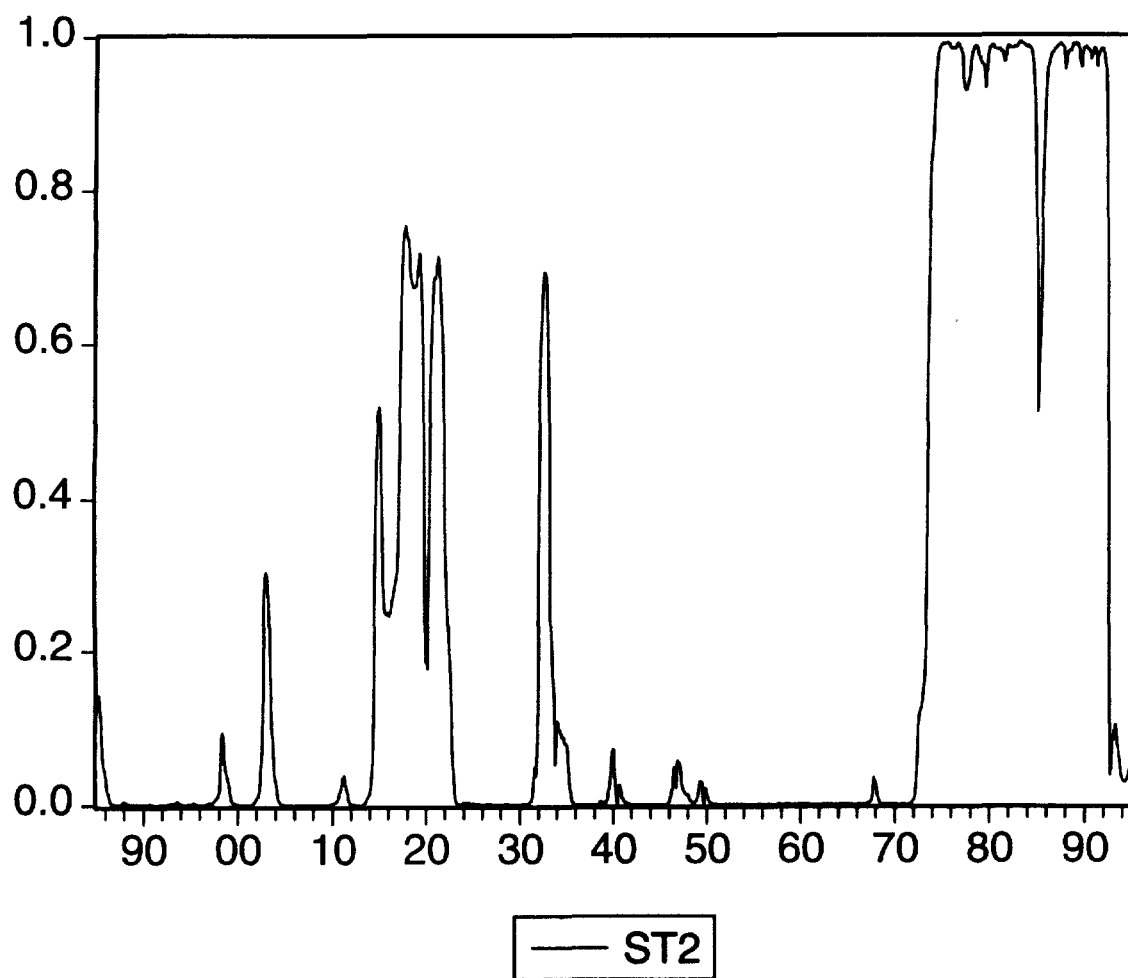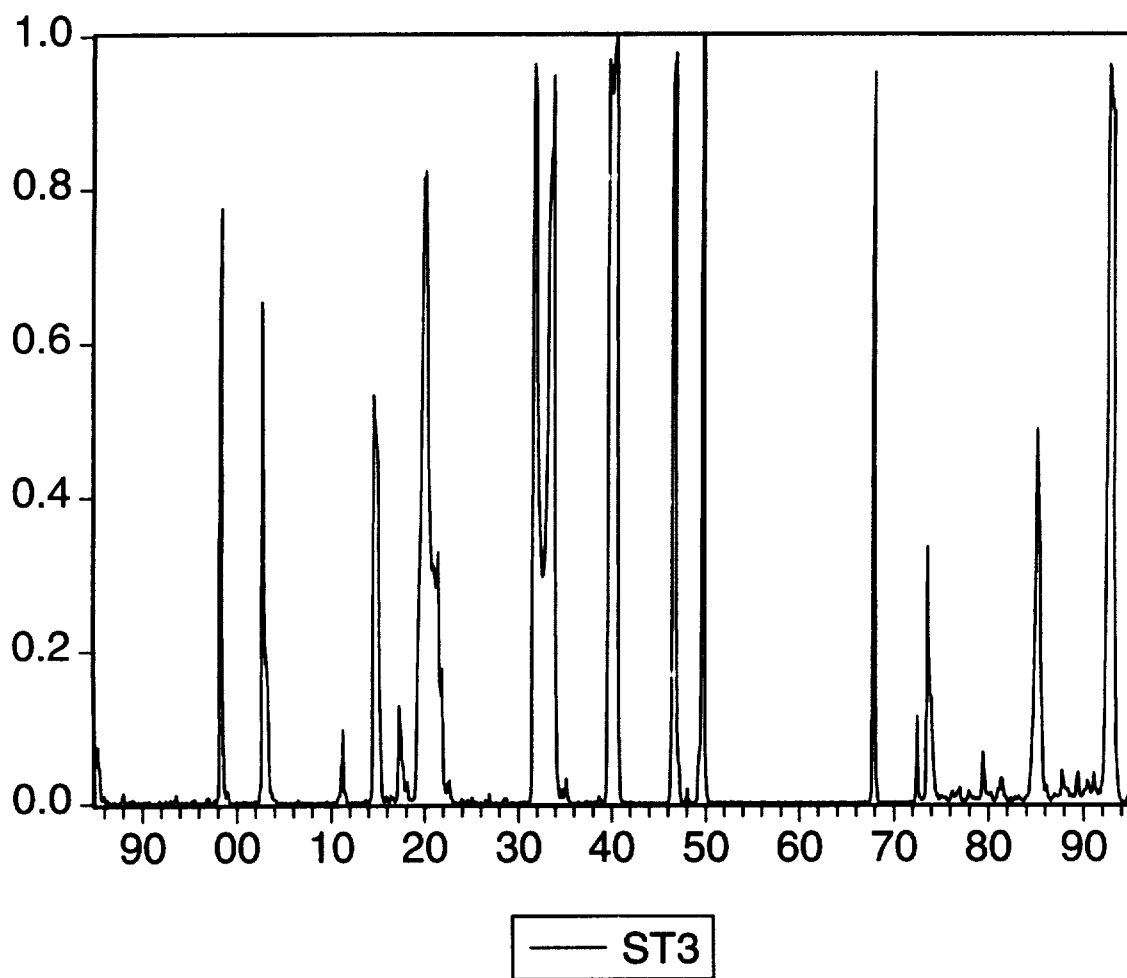
Figure3.1. Pr[St=1]

ST1

Figure 3.2. Pr[St=2]

Figure 3.3. Pr[St=3]

Equations (1)-(6) in the text describe the model to be estimated. Writing the model in state-space form, we have:

Measurement Equation :

(A1) $$q_t = Hc_t,$$

Transition Equation :

(A2) $$c_t = Fc_{t-1} + \varepsilon_t,$$

(A3) $$E(\varepsilon_t \varepsilon'_t) = R_t,$$

where

$$H = [1 \ 1 \ 0], \ c_t = \begin{bmatrix} y_t \\ x_t \\ x_{t-1} \end{bmatrix}, \ F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix}, \ \varepsilon_t = \begin{bmatrix} \nu_t \\ e_t \\ 0 \end{bmatrix}, \ \text{and} \ R_t = \begin{bmatrix} \sigma_\nu^2 & 0 & 0 \\ 0 & \sigma_{e,t}^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Inferences of the model

Denote $\tilde{Q}^t = \{q_1, q_2, ..., q_t\}$ as the vector of real exchange rates up to time t, $\tilde{C}^t = \{c_1, c_2, ..., c_t\}$ as the set of state vectors up to time t (here, the first and second rows of $\tilde{C}^t$ can be defined as $\tilde{Z}^t = \{z_1, z_2, ..., z_t\}$, where $z_t = [y_t, x_t]'$), $\tilde{S}^t = \{S_1, S_2, ..., S_t\}$ as the vector of Markov-switching variables up to time t, and finally, $\tilde{\theta} = \{\phi_1, \phi_2, \sigma_\nu^2, \sigma_{e,1}^2, \sigma_{e,2}^2, \sigma_{e,3}^2, p_{11}, p_{12}, p_{21}, p_{31}, p_{32}\}$ as the vector of all parameters of the model. Our purpose is to make inferences on $\tilde{C}^T$ (more specifically, $\tilde{Z}^T$, the first two rows of $\tilde{C}^T$), $\tilde{S}^T$, and on $\tilde{\theta}$ conditional on real exchange rates, $\tilde{Q}^T$. Maximum likelihood estimation of similar models and ways to make inferences on $\tilde{C}^T$ conditional on parameter estimates have been proposed by Kim (1994). (For applications of the maximum likelihood estimation of the unobserved components model with Markov-switching heteroskedasticity, see Kim (1993) and Kim and Kim (1996).) However, Kim's methods are based on approximations. In this

paper, we adopt the Bayesian Gibbs sampling approach proposed by Carter and Kohn (1994) in making inferences of the model based on the posterior density, $p(\widetilde{C}^T, \widetilde{S}^T, \widetilde{\theta} \mid \widetilde{Q}^T)$ of $\widetilde{C}^T$, $\widetilde{S}^T$, and $\widetilde{\theta}$.

As documented in Gelfand and Smith (1990) and Carter and Kohn (1994), Gibbs sampling in the present context generates $\widetilde{C}^T$, $\widetilde{S}^T$, and $\widetilde{\theta}$ from conditional densities $p(\widetilde{C}^T \mid \widetilde{S}^T, \widetilde{\theta}, \widetilde{Q}^T)$, $p(\widetilde{S}^T \mid \widetilde{C}^T, \widetilde{\theta}, \widetilde{Q}^T)$, and $p(\widetilde{\theta} \mid \widetilde{C}^T, \widetilde{S}^T, \widetilde{Q}^T)$ until eventually $\{\widetilde{C}^T, \widetilde{S}^T, \widetilde{\theta}\}$ is generated from the joint posterior density $p(\widetilde{C}^T, \widetilde{S}^T, \widetilde{\theta} \mid \widetilde{Q}^T)$. [Carter and Kohn, p. 542-543.] The following explains how each of $\widetilde{S}^T$, $\widetilde{\theta}$, and the first two rows of $\widetilde{C}^T$ can be generated from an appropriate conditional distribution, combining ideas in Carter and Kohn (1994), Albert and Chib (1993), and Kim, Nelson and Startz (1996).

A. Generating $\widetilde{Z}^T = \{\widetilde{Y}^T \ \widetilde{X}^T\}$ Conditional on $\widetilde{\theta}$ and $\widetilde{S}^T$, and on Data, $\widetilde{Q}^T$

Conditional on $\widetilde{S}^T$ and $\widetilde{\theta}$, the state-space model in (A1)-(A3) is linear, and we can adopt Carter and Kohn's (1994) multi-move Gibbs sampling to generate the first two rows of the state vectors $\widetilde{C}^T$. What follows is a detailed description of their algorithm in our context.

Assuming that F and R_t in (A1)-(A3) are known, the joint distribution of $\widetilde{C}^T$ = $\{c_1, c_2, ..., c_T\}$, given the data set $\widetilde{Q}^T$ and the prior distribution of c_0, is written as

(A4) $\qquad p(\widetilde{C}^T \mid \widetilde{Q}^T) = p(c_T \mid \widetilde{Q}^T) \prod_{t=1}^{T-1} p(c_t \mid \widetilde{Q}^t, c_{t+1}),$

where $\widetilde{Q}^t$ is information up to time t. Note that the second element of c_t is linked to the third element of c_{t-1} as an identity in the transition equation (A2). Therefore, the above equation suggests we can generate z_T, and then successively generate z_t conditional on z_{t+1} and $\widetilde{Q}^t$ for t = T-1, T-2,...,1. We can take advantage of the Gaussian Kalman filter to obtain $p(c_T \mid \widetilde{Q}^T)$ and

$p(c_t | \tilde{Q}^t, c_{t+1})$, as the state-space model is linear given F and R_t. Summarizing Carter and Kohn's (1994) algorithm in our context, we have the following steps:

Step 1:

Run the Kalman filter algorithm to calculate $c_{t|t} = E(c_t | \tilde{Q}^t)$ and $V_{t|t} = cov(c_t | \tilde{Q}^t)$ for $t = 1, 2, ..., T$ and save them. The last iteration of the Kalman filter provides us with $c_{T|T}$ and $V_{T|T}$. The first two elements of $c_{T|T}$ (denoted by $z_{T|T}$) and the first 2×2 block of $V_{T|T}$ (denoted by $V_{T|T}^*$) can be used to generate the first two elements of c_T ($z_T = [y_T \ x_T]'$) from a joint Normal distribution.

Step 2:

For $t = T-1, T-2, ..., 1$, given $c_{t|t}$ and $V_{t|t}$, if we treat $z_{t+1} = [y_{t+1} \ x_{t+1}]'$ generated from the previous iteration as an additional vector of observations to the system, the distribution $p(c_t | \tilde{Q}^t, z_{t+1})$ is easily derived by applying the updating equations of the Kalman filter. From equation (A2), since z_{t+1} is given by:

$$(A5) \qquad z_{t+1} = F^* c_t + \varepsilon^*_{t+1},$$

where F^* is the first two rows of F and ε^*_{t+1} is the first two elements of ε_{t+1}, updating equations are derived as:

$$(A6) \qquad c_{t|t,z_{t+1}} = c_{t|t} + V_{t|t} F^{*} \Omega_t^{-1} \eta_t,$$

$$(A7) \qquad V_{t|t,z_{t+1}} = V_{t|t} - V_{t|t} F^{*'} \Omega_t^{-1} F^* V_{t|t}.$$

Here, $\eta_t = z_{t+1} - F^* c_{t|t}$ is treated as a vector of forecast errors and is the difference between z_{t+1} and the forecast of the vector z_{t+1} conditional on information up to time t; and, $\Omega_t = F^* V_{t|t} F^{*'} + Cov(\varepsilon^*_{t+1})$, is the covariance of η_t. Then the first two elements of $z_{t|t,z_{t+1}}$, and the first 2×2 block of

$V_{t|t,z_{t+1}}$ can be used to generate z_t from a joint Normal distribution, for $t = T-1, T-2, \ldots, 1$.

B. Generating σ_ν^2 of the Random Walk Component, Conditional on $\widetilde{Y}^T$

Conditional on $\widetilde{Y}^T = [y_1 \ y_2 \ \ldots \ y_T]'$, the only parameter of the random walk, σ_ν^2, is independent of the data set, $\widetilde{Q}^T$, and is independent of the stationary component, x_t, by assumption. This allows us to focus on equation (2), by treating the generated $\widetilde{Y}^T$ as the data set. By choosing the inverse gamma distribution as the prior $(IG(\frac{\upsilon}{2},\frac{\delta}{2}))$, one can show that the conditional distribution from which σ_ν^2 is generated is given by:

$$[\sigma_\nu^2 | \widetilde{Y}^T] \sim IG(\frac{\upsilon+(T-1)}{2}, \frac{\delta + \Sigma_{t=2}^T(y_t-y_{t-1})^2}{2}).$$

We adopt $\upsilon = \delta = 0$ for our application.

C. Generating $\widetilde{S}^T$ and Parameters Associated with the Stationary Component, Conditional on $\widetilde{X}^T$

$\widetilde{X}^T = [x_1 \ x_2 \ \ldots \ x_T]'$, the Markov switching variable S_t and other parameters associated with the stationary component, x_t, are independent of the data set, $\widetilde{Q}^T$ and of the random walk component, y_t, by assumption. This allows us to focus only on equation (3), by treating generated $\widetilde{X}^T$ as a data set. The Bayesian Gibbs sampling approach to a two-state Markov-switching model has been suggested by Albert and Chib (1993). More recently, Gibbs sampling has been successfully implemented for a three-state Markov-switching model of stock returns by Kim, Nelson and Startz (1996). The following is based on Kim, Nelson and Startz's extension of Albert and Chib's algorithm.

C.1 Generating $\tilde{S}^t$, t = 1,2,...T, Conditional on $\tilde{X}^T$ and Parameters of the Stationary Component

Defining $\tilde{\theta}^*$ as a subset of $\tilde{\theta}$ with the exclusion of σ_ν^2, $\tilde{S}^T$ can be generated based on the following distribution, similar to that in (A4):

(A8) $$p(\tilde{S}^T | \tilde{X}^T, \tilde{\theta}^*) = p(S_T | \tilde{X}^T, \tilde{\theta}^*) \prod_{t=1}^{T-1} p(S_t | \tilde{X}^t, \tilde{\theta}^*, S_{t+1}).$$

In order to simulate $\tilde{S}^T$ from the above distribution, we first run Hamilton's (1989) basic filter for the model to get $p(S_t | \tilde{X}^t, \tilde{\theta}^*)$ and $p(S_t | \tilde{X}^{t-1}, \tilde{\theta}^*)$, for t = 1,2,...,T and save them. The last iteration of the filter provides us with $p(S_T | \tilde{X}^T, \tilde{\theta}^*)$ from which S_T is generated. Then, we can successively generate S_t from $p(S_t | \tilde{X}^t, \tilde{\theta}^*, S_{t+1})$, for t = T-1,T-2,...,1, using:

(A9) $$p(S_t | \tilde{X}^t, \tilde{\theta}^*, S_{t+1}) = \frac{p(S_{t+1} | S_t) p(S_t | \tilde{X}^t, \tilde{\theta}^*)}{p(S_{t+1} | \tilde{X}^t, \tilde{\theta}^*)}.$$

The uniform distribution can be used to generate the three-state Markov-switching variable, S_t, t = 1,2,...,T. Conditional on S_{t+1} = k, k = 1,2,3, define $p_j = p_{jk} \times p(S_t = j | \tilde{X}^t, \theta)$, j = 1,2,3, where p_{jk} is the transition probability defined in (3). We first generate a random number from the uniform distribution. If the generated number is greater than or equal to $\frac{p_1}{p_1+p_2+p_3}$, we set S_t = 1; if it is less than $\frac{p_1}{p_1+p_2+p_3}$, we generate another random number from the uniform distribution. And then, if the generated random number is greater than or equal to $\frac{p_2}{p_2+p_3}$, we set S_t = 2; if it is less than $\frac{p_2}{p_2+p_3}$, we set S_t = 3.

C.2 Generating σ_j^2, j = 1,2,3, Conditional on $\tilde{X}^T$, $\tilde{S}^T$, and on other Parameters of the Stationary Component

In order to impose the constraint that $\sigma_1^2 < \sigma_2^2 < \sigma_3^2$, we may redefine σ_2^2 and σ_3^2 in the following way:

(A10) $\sigma_2^2 = \sigma_1^2(1+h_2)$ and $\sigma_3^2 = \sigma_1^2(1+h_2)(1+h_3)$,

where $h_2 > 0$ and $h_3 > 0$. We first generate σ_1^2, then generate $1+h_2$ and $1+h_3$.

First, to generate σ_1^2, we transform equation (3) as follows:

(A11) $$x_{1t} = \frac{x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2}}{\sqrt{(1+S_{2t}h_2)(1+S_{3t}h_2)(1+S_{3t}h_3)}}.$$

By choosing the inverse gamma distribution as the prior $(IG(\frac{\upsilon_1}{2},\frac{\delta_1}{2}))$, one can show that the conditional distribution from which σ_1^2 is generated is given by:

(A12) $$[\sigma_1^2 | X^T, S^T, \tilde{\theta}^*_{j \neq \sigma_1^2}] \sim IG(\frac{\upsilon_1 + (T-2)}{2}, \frac{\delta_1 + \Sigma_{t=3}^{T} x_{1t}^2}{2}),$$

where $\tilde{\theta}^*_{j \neq \sigma_1^2}$ represents a vector of paramenters of the stationary component that excludes σ_1^2.

Second, to generate $\bar{h}_2 = 1 + h_2$, and thus σ_2^2, we transform equation (3) to get:

(A13) $$x_{2t} = \frac{x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2}}{\sqrt{\sigma_1^2(1+S_{3t}h_3)}}.$$

Here, we note that the likelihood function of h_2 depends on the values of x_{2t} for which $S_t = 2$ or 3. By derining $T_2 = \{t : S_t = 2$ or $3\}$ and chosing the inverse gamma distributions for the priors of $\bar{h}_2$ $(IG(\frac{\upsilon_2}{2},\frac{\delta_2}{2})I_{[\bar{h}_2>1]})$, one can show that the complete conditional is given by:

(A14) $$[\bar{h}_2 | X^T, S^T, \tilde{\theta}^*_{j \neq h_2}] \sim IG(\frac{\upsilon_2 + N_2}{2}, \frac{\delta_2 + \Sigma_{t=1}^{T_2} x_{2t}^2}{2})I_{[\bar{h}_2>1]},$$

where $\tilde{\theta}^*_{j \neq h_2}$ represents a vector of parameters of the model that excludes h_2; I is the indicator function on $[\bar{h}_2 > 1]$; N_2 are cardinalities of T_2 and the sum is over the elements of T_2.

Finally, to generate $\bar{h}_3 = 1 + h_3$, and thus, σ_3^2, we transform equation (3) to get:

(A15) $$x_{3t} = \frac{x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2}}{\sqrt{\sigma_1^2(1+S_{3t}h_2)}}.$$

Here, we note that the likelihood function of h_3 depends only on the values of x_{3t} for which $S_t = 3$. By defining $T_3 = \{t : S_t = 3\}$ and choosing the inverse gamma distributions for the priors of $\bar{h}_3$ ($IG(\frac{\upsilon_3}{2},\frac{\delta_3}{2})I_{[\bar{h}_3 > 1]}$), one can show that the complete conditional is given by

$$(A16) \qquad [\bar{h}_3 \mid \tilde{X}^T, \tilde{S}^T, \tilde{\theta}^{*}_{j \neq h_3}] \sim IG(\frac{\upsilon_3 + N_3}{2}, \frac{\delta_3 + \Sigma_{t=1}^{T_3} x_{3t}^2}{2})I_{[\bar{h}_3 > 1]},$$

where $\tilde{\theta}^{*}_{j \neq h_3}$ represents a vector of parameters of the model that excludes h_3; I is the indicator function on $[\bar{h}_3 > 1]$; N_3 are cardinalities of T_3, and the sum is over the elements of T_3.

The quantities υ_i, $i = 1,2,3$, represent the strength of the priors of σ_1^2, h_2 and h_3. For our application, we employ $\upsilon_i = 0$ and $\delta_i = 0$, for $i = 1,2,3$.

C.3 Generating $\tilde{\phi} = [\phi_1 \ \phi_2]'$, Conditional on $\tilde{X}^T$, $\tilde{S}^T$, and on other Parameters of the Stationary Component

Rewriting equation (3) to have homoskedastic errors, we have:

$$(A17) \qquad \frac{x_t}{\sigma_{e,t}} = \phi_1 \frac{x_{t-1}}{\sigma_{e,t}} + \phi_2 \frac{x_{t-2}}{\sigma_{e,t}} + e_t^{*}, \quad t = 3,4,...T,$$

where $e_t^{*} \sim$ iid $N(0,1)$. We adopt the multivariate Normal prior distribution for $\tilde{\phi}$, $\tilde{\phi} \sim N(\alpha, A^{-1})$, and denote $\tilde{X}$ to be the vector of left-hand-side variables and $\tilde{W}$ to be the matrix of right-hand-side variables. The posterior distribution from which $\tilde{\phi}$ is to be generated is given by:

$$(A18) \qquad \tilde{\phi} \sim N((A + \tilde{W}'\tilde{W})^{-1}(\alpha A + \tilde{W}'\tilde{X}), (A + \tilde{W}'\tilde{W})^{-1}).$$

We adopt $\alpha = [0 \ 0]'$ and $A^{-1} = 25$ for our application. In generating $\tilde{\phi}$ from the above posterior distribution, we adopt rejection sampling, so that the roots of $\phi(L) = 1 - \phi_1 L - \phi_2 L^2$ lie outside the unit circle.

C.4 Generating Transition Probabilities $(p_{11}, p_{12}, p_{21}, p_{31}, p_{32})$ Conditional on $\tilde{X}^T$, $\tilde{S}^T$, and on other Parameters of the Stationary Component

Conditional on $\tilde{S}^T$, the transition probabilities are independent of $\tilde{X}^T$ and other parameters, as in Albert and Chib (1993). For a two-state Markov-switching model, Albert and Chib derive the full conditional distributions of the transitions probabilities as a product of independent beta distributions. For a three-state Markov-switching model, such as in this paper, Kim, Nelson, and Startz (1996) adopt a slight modification of their approach.

Given $\tilde{S}^T$ and the initial state, let n_{ij}, $i,j = 1,2,3$, be the total number of transitions from state $S_{t-1} = i$ to $S_t = j$, $t = 1,2,...,T$. Define $\bar{p}_{ii} = \Pr(S_t \neq i | S_{t-1} = i)$ and $\bar{p}_{ij} = \Pr(S_t = j | S_{t-1} = i, S_t \neq i)$, $i = 1,2,3$, $j = 1,2$. Correspondingly, we have $p_{ij} = \bar{p}_{ij} \cdot \bar{p}_{ii}$ for $i \neq j$. Similarly, define $\bar{n}_{ii}$ to be the number of transitions from state $S_{t-1} = i$ to $S_t \neq i$, and $\bar{n}_{ij}$ to be the number of transitions from state $S_{t-1} = i$ to state $S_t = j$, conditional on $S_t \neq i$.

Then, as in Albert and Chib, by taking the beta family of distributions as conjugate priors, it can be shown that the posterior distributions of p_{ii} are given by

(A19) $\qquad [p_{ii} | \tilde{S}^T] \sim \text{beta}(u_{ii} + n_{ii}, \bar{u}_{ii} + \bar{n}_{ii}), \quad i = 1,2,3,$

where u_{ii} and $\bar{u}_{ii}$ are the hyperparameters of the prior. Once p_{ii}, $i = 1,2,3$, are generated from the above distribution, generation of the other parameters is straightforward. For example, given that p_{ii} is generated, p_{ij} can be calculated by $p_{ij} = \bar{p}_{ij} \cdot \bar{p}_{ii}$. $\bar{p}_{ij}$ can be generated from the following beta distribution:

(A20) $\qquad [\bar{p}_{ij} | \tilde{S}^T] \sim \text{beta}(u_{ij} + n_{ij}, u_{ik} + n_{ik}), \quad i \neq j \neq k,$

where u_{ij} and u_{ik} are the hyperparameters of the prior. The values of the hyperparameters that we employ are: $u_{ii} = 9$ and $\bar{u}_{ii} = 1$; and $u_{ij} = 1$ and $u_{ik} = 1$ for $i \neq j \neq k$.