1997

Statistical analysis of foreign exchange rates: application of cointegration model and regime-switching stochastic volatility model

Koji Kondo
Iowa State University

Statistical analysis of foreign exchange rates:

Application of cointegration model and regime-switching stochastic volatility model

by

Koji Kondo

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-majors: Economics; Statistics

Major Professors: Stefano Athanasoulis and F. Jay Breidt

Iowa State University

Ames, Iowa

1997

UMI Number: 9814658

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Koji Kondo

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Co-major Program

Signature was redacted for privacy.

For the Co-major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 GENERAL INTRODUCTION

As the world's economies move toward greater integration, foreign exchange rate determination has become an increasingly important area in international finance. Researchers have frequently used empirical research to monitor and predict exchange rate movements and many have sought to employ new statistical methods in their efforts to understand the complex movements of exchange rates. This dissertation applies two statistical models to foreign exchange rate data in two main parts. The first part, an application of the partial system model of cointegration developed by Johansen (1990), uses the concept of weak exogeneity to simplify the complex analysis. While a direct application of the cointegration approach with many variables is not easy to handle, the partial system model can reduce the number of the parameters to be estimated by identifying weakly exogenous variables. This method is illustrated utilizing a theoretical long-run model based on Dornbusch's sticky price model. In this part, the small country assumption is relaxed, so that both countries may be taken to be large. Furthermore, the model is also extended to include a third country.

The data set here consists of monthly exchange rates, countries' money supplies and GNPs for three countries; Germany, Japan and the United States. First, the full system cointegration model is estimated and the weakly exogenous variables are identified from the results of the full system model. Using the information from the weakly exogenous variables permits the number of the parameters to be reduced, thereby forming the partial system model. Estimation of the partial system model will provide information of long-run relations among the variables. Then, the next step is to interpret long-run relations among the parameters, an interpretation based on the modified Dornbusch's model. Because some of the relations may not be interpreted in an economically meaningful way, variance decomposition and impulse response analysis are conducted to investigate the short-run dynamics of the system.

In the second part, a regime-switching stochastic volatility (RSV) model is applied to daily exchange rate data to capture the possibly changing volatility of exchange rates over time. While more complicated to implement than other methods, the RSV model recommends itself as the most natural method to apply when compared to the ARCH and GARCH models.

Here, a Gibbs sampler technique is used to approximate the posterior distribution of all unknown model parameters. By imposing interest rate parity, the relationship between exchange rates and foreign and domestic interest rate differences is also simultaneously examined. The results indicate that the interest rate difference does not affect the level and the volatility of exchange rates, a finding which supports the random walk theory of exchange rates. On the other hand, two different regimes, a high-volatility regime and a low-volatility regime, are discovered and well modeled. The development of a forecasting model will be the subject for future studies.

PART I

COINTEGRATION ANALYSIS OF EXCHANGE RATES

2 INTRODUCTION

Exchange rate determination has been one of the important fields in international economics. Since the world economy moved to the floating exchange rate system early in 1970's, researchers have been especially concerned with how exchange rates are determined in the foreign exchange rate market and have presented many models exploring the questions of exchange rate determination. Among the most important models are the monetary model by Frenkel, Bilson and Mussa, the sticky price model by Dornbusch and the currency portfolio model by King, Putnam and Wilford. Many other variations are derived from these three main models.

Chapter 3 will review some structural models of exchange rate determination and discuss three traditional models; the monetary approach, the portfolio balance approach, and the currency substitution approach as well as some variations derived from these models. Chapter 4 presents a new theoretical model based on Dornbusch's sticky price model. This new model slightly modifies Dornbusch's model by adopting the large-country assumption, an assumption that permits all prices in the system to be endogenized. It also attempts to extend the model to the three-country case so that the third country effects can be analyzed.

Many economic variables contain a unit root or unit roots. With non-stationary variables, the traditional approach applies the first differenced variables, however, if there is a linear combination among the variables which is stationary, the traditional approach is no longer appropriate. In their seminal work, Nelson and Plosser (1982) point out that many economic variables contain unit roots that require special treatments in this case. Some special treatments are available because of recent developments in econometrics. Dickey and Fuller, and Phillips and Perron are among those who have developed unit root tests. In addition, Johansen's seminal paper (1988) developed a methodology to deal with the so-called cointegrated variables.

Chapter 5 examines statistical methodology. First, the vector autoregression (VAR) model discussed in Sims' seminal paper (1980) is reviewed, followed by identification issues and hypothesis testing. Then, the chapter explores the error correction model to deal with cointegrated relations among the variables.

Testing for the number of cointegrating relations among the variables will also be a topic in this chapter. To determine the numbers of the cointegrating relations, the trace and likelihood ratio test will be used.

One of the problems that the VAR-type analysis faces is that adding more variables to the system drastically increases the number of the parameters in the system. This will create some difficulties in estimating these parameters in terms of degrees of freedom. To reduce such difficulties, the partial system model will be applied to the data. While the full system model such as VAR treats all the variables in the system as endogenous, the partial system model treats some of the variables in the system as exogenous so that these variables can be modeled less carefully. To apply the partial system model, the concept of weak exogeneity is required. These issues also will be examined in Chapter 4.

The rest of the part reports empirical results based on the previously discussed theory and methodology. Chapter 6 reports the data set containing economic variables from Germany, Japan and the United States used for the empirical work. It presents the summary of the data as well as the results for the unit root tests.

In Chapter 7, both the error correction model and the partial system model will be estimated. It also reports the cointegrating relations among the variables. These cointegrating relations, which are considered to be economic long-run relations among the variables, will be examined and interpreted. After examining the two-country cases; the Germany-U.S. case and the Japan-U.S. case, the three-country case; i.e., the Germany-Japan-U.S. case is investigated. The long-run analysis is concerned with long-run equilibrium. Short-run dynamics and long-run effects will be issues of short-run analysis.

Chapter 8 reports the results for short-run analysis, as well as the variance decomposition and impulse response from the models estimated in Chapter 7. These analyses are conducted for the two-country cases and the three-country case. The results for variance decomposition analysis are reported for both the full system model and the partial system model. The results for impulse responses are reported only for the partial system model.

Chapter 9 presents conclusions and further research.

3 MONETARY APPROACH TO THE EXCHANGE RATE

Since the floating exchange rate system was adopted, researchers have focused on how the exchange rates are determined and how they behave. Among the many approaches to these problems, the monetary approach, the portfolio balance approach and the currency substitution approach are considered to be important. Many studies have been done using each approach. A general summaries of these three principal approaches are found in Dornbusch (1980a), Frankel (1983), Frenkel and Mussa (1985), MacDonald (1988), and Baillie and McMahon (1989). More detailed references for the monetary approach are Dornbusch (1976a), Dornbusch (1976b), Frenkel (1976), and Mussa (1984). For the portfolio balance approach, the readers are referred to McKinnon and Oats (1966), Branson (1968), Branson (1975), and McKinnon (1969). Finally, Kouri (1976), Kouri and de Macedo (1978), Calvo and Rodriguez (1977), and Frenkel and Rodriguez (1982) are good references for the currency substitution approach. This chapter will review two approaches to exchange rate determination; the monetary approach and the currency portfolio approach. Discussion of the monetary approach will include the flexible price model, the sticky price model, and the interest rate differential model. The following chapter, Chapter 4, will consider and extend the sticky price model as the theoretical model and focus of the analysis in this part.

3.1 The Monetary Approach

There are three principal version of the Monetary Approach to exchange rate determination; the flexible price monetary model by Frenkel (1976) and Bilson (1978a,b), the sticky price model of Dornbusch (1976a,b), and the real interest rate differential model due to Frankel (1979). These three models are similar in the sense that all the models adopt the so-called asset market view of exchange rate determination; Mussa (1984). This view considers the foreign exchange rate as an asset and prices the exchange rate like other financial assets. Frankel and Bilson's model is the basic model in the monetary approach, and Dornbusch's and Frankel's models modify Frenkel and Bilson's model by replacing some of the assumptions used by Frenkel and Bilson.

3.1.1 The Flexible Price Monetary Model

The following five assumptions are usually made in the monetary flexible approach: (a) goods prices are completely flexible, (b) there exists perfect substitutability between domestic and foreign assets, (c) capital is perfectly mobile, (d) the money supply and real income are exogenous variables and (e) domestic money is held by domestic residents only while foreign money is held by foreign residents only.

Since the exchange rate is considered as the relative price of one nation's money to another nation's money in the flexible price approach, it is determined where the supply of national monies equals the demand for these currencies. It emphasizes the importance of the stock aspect rather than flow aspect. This approach starts with the assumption of money market equilibrium. The real money demand function is written as:

$$\frac{M^d}{P} = L(Y, i) \qquad \text{where} \qquad \frac{\partial L}{\partial Y} > 0,\ \frac{\partial L}{\partial i} < 0 \tag{3.1}$$

where

M^d is the demand for money,

P is the domestic price level,

Y is the domestic income level,

i is the domestic short-term interest rate.

The above equation indicates that real money demand is a function of income and interest rates. Money demand is assumed to respond positively to domestic income and negatively to interest rates. The equation (3.1) often appears in the literature in logarithmic form:

$$m_t^d - p_t = k + \phi y_t - \lambda i_t \tag{3.2}$$

where

k = constant,

p_t = log of the domestic price level,

y_t = log of domestic income level,

i_t = the domestic short-term interest rate,

ϕ = the money semi-elasticity of the real income,

λ = the money elasticity of the interest rate.

The same equation is assumed to hold for a foreign country:

$$m_t^{*d} - p_t^* = k^* + \phi y_t^* - \lambda i_t^* \tag{3.3}$$

where the asterisks denote foreign variables. As in many theoretical and empirical works. the assumption is made here that both the money demand elasticity of the real income, ϕ, and the money demand semi-elasticity of the interest rate, λ, are the same for the domestic and foreign country. Equilibria in the money markets are described by:

$$m_t^d = m_t^s = m_t, \qquad m_t^{*d} = m_t^{*s} = m_t^* \tag{3.4}$$

Therefore, the following relationship is derived from (3.2), (3.3) and (3.4):

$$p_t - p_t^* = -(k - k^*) + (m_t - m_t^*) - \phi(y_t - y_t^*) + \lambda(i_t - i_t^*) \tag{3.5}$$

Another basic assumption in this approach is the purchasing power parity assumption, made from assumption (a) above:

$$e_t = p_t - p_t^* \tag{3.6}$$

The purchasing power parity condition links domestic and foreign money demand. e_t is defined here as the price of foreign currency in units of domestic currency. Substituting (3.6) into (3.5) gives:

$$e_t = -(k - k^*) + (m_t - m_t^*) - \phi(y_t - y_t^*) + \lambda(i_t - i_t^*) \tag{3.7}$$

The equation (3.7) is the simplest equation of exchange rate determination. According to this simplest of models, the exchange rate is determined by a linear combination of the differences of domestic and foreign fundamentals; that is, differences in money supplies, in incomes, and in interest rates. By considering assumptions (b) and (c), the covered interest parity condition can be introduced:

$$i_t - i_t^* = f_t - e_t \tag{3.8}$$

where f_t is a forward exchange rate. Then, (3.7) can be modified as:

$$e_t = -(k - k^*) + (m_t - m_t^*) - \phi(y_t - y_t^*) + \lambda(f_t - e_t) \tag{3.9}$$

This is Bilson's familiar exchange rate determination model. As Gardeazabal and Regulez (1992) point out, the equations (3.7) and (3.9) are equivalent under perfect capital mobility because the covered interest rate parity condition becomes the no-arbitrage condition. Furthermore, the assumption that the forward exchange rate is an unbiased efficient expectation of the future spot exchange rate, $f_t = \mathrm{E}_t e_{t+1}$, will introduce the uncovered interest parity condition:

$$i_t - i_t^* = \mathrm{E}_t e_{t+1} - e_t \tag{3.10}$$

By substituting (3.10) into (3.7):

$$e_t = -(k - k^*) + (m_t - m_t^*) - \phi(y_t - y_t^*) + \lambda(\mathrm{E}_t e_{t+1} - e_t) \tag{3.11}$$

Solving equation (3.11) above for the current exchange rate, e_t, then:

$$\begin{aligned} e_t &= \frac{1}{1+\lambda}[-(k - k^*) + (m_t - m_t^*) - \phi(y_t - y_t^*) + \lambda(\mathrm{E}_t e_{t+1})] \\ &= \frac{1}{1+\lambda} z_t + \frac{\lambda}{1+\lambda}\mathrm{E}_t e_{t+1} \end{aligned} \tag{3.12}$$

where $z_t = -(k - k^*) + (m_t - m_t^*) - \phi(y_t - y_t^*)$ are economic fundamentals. Assuming the expectations of e_{t+1} are formed rationally, then (3.12) can be solved recursively:

$$\begin{aligned} e_t &= \frac{1}{1+\lambda}\sum_{i=1}^{\infty}(\frac{\lambda}{1+\lambda})^i[-(k - k^*) + (m_{t+i} - m_{t+i}^*) - \phi(y_{t+i} - y_{t+i}^*)] \\ &= \frac{1}{1+\lambda}\sum_{i=1}^{\infty}(\frac{\lambda}{1+\lambda})^i z_{t+i} \end{aligned} \tag{3.13}$$

The result, (3.12), reveals that the current exchange rate, e_t, depends on the expected future levels of the foreign and domestic exogenous variables, k, k^*, m_{t+i}, m_{t+i}^*, y_{t+i} and y_{t+i}^*. Equation (3.12) also clarifies the relationship between the current exchange rate, e_t, and the expected future exchange rate, $\mathrm{E}_t e_{t+1}$. If the money semi-elasticity of the interest rate, λ, is large enough, then $\frac{\lambda}{1+\lambda}$ is close to 1, and e_t will be close to $\mathrm{E}_t e_{t+1}$, that is, e_t and $E_t e_{t+1}$ are closely correlated. Baillie and McMahon (1989) modify the model (3.7) by introducing an exchange rate adjustment mechanism. They assume that the exchange rate adjusts as follows:

$$e_t - e_{t-1} = \theta(\bar{e}_t - e_{t-1}) \tag{3.14}$$

where

$$\bar{e}_t = p_t - p_t^* \tag{3.15}$$

Then, (3.7) becomes:

$$e_t = -\theta(k - k^*) + \theta(m_t - m_t^*) - \theta\phi(y_t - y_t^*) + \theta\lambda(i_t - i_t^*) + (1-\theta)e_{t-1} \tag{3.16}$$

The equations (3.7) and (3.16) tell us that an increase in money supply is expected to lead to a depreciation of the exchange rate by the same proportion and that an increase in domestic real interest rates will cause a depreciation unlike that predicted by the standard Keynesian model. The monetary approach explains these results by stating that an increase in the domestic interest rate reduces the demand for money which creates an excess supply in the market. Another direction to extend this simple flexible price model (3.7) is to specify the stochastic processes governing the evolution of the exogenous variables. For instance, MacDonald (1988) assumes that the levels and rates of growth of the money supply, $M_t = m_t - m_t^*$, follow random walks:

$$\begin{cases} M_t = M_{t-1} + \eta_t + \varepsilon_t \\ \eta_t = \eta_{t-1} + \mu_t \end{cases} \tag{3.17}$$

where ε_t and μ_t are white noise disturbances; $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $\mu_t \sim N(0, \sigma_\mu^2)$. He shows that the important factor that determines the accuracy of exchange rate expectation is how well market participants in foreign exchange market can distinguish shocks to the level of the money supply process, ε_t, from shocks to the rate of growth of the money supply, μ_t. He explains this by distinguishing two cases; (a) the full information case where participants have all information on the stochastic processes of ε, η, μ, and M, and (b) partial information case where participants can't differentiate the sources of the unanticipated change in the money supply.

3.1.2 The Sticky Price Monetary Model

The above approach, what MacDonald calls the flexible-price monetary approach (FPMA), imposes some unrealistic assumptions. PPP is one of the crucial building-block assumptions. However many empirical researches have indicated that PPP holds under the hyper-inflationary situation or in the long run, but not in the short run. One way to reconcile the model with this fact is to assume that the goods market is slow to move back to the equilibrium due to the stickiness of goods prices, once the goods market deviates from the equilibrium, while the money market is quick to return to the equilibrium or is always in equilibrium. The sluggishness of goods prices still assumes that PPP holds in the long run because the goods market is also in equilibrium in the long run. The difference of adjustment speed in the two markets, goods market and money market, explains the failure for PPP to hold in the short run, and also the volatility of the exchange rate, that is, overshooting of the exchange rate. This sticky-price monetary approach was developed by Dornbusch (1976a). He changes some of the assumptions made for FPMA. Instead of assuming that goods prices are completely flexible (a) in the above, it is assumed that goods prices adjust to a new equilibrium with a lag and domestic and foreign goods are no longer perfect substitutes (a'). The different speeds of adjustment in the two markets allow a short-run change in the money supply to have real effects due to the terms of trade. The model is formulated as follows (Dornbusch (1976a)):

$$m - p = \phi y - \lambda i \tag{3.18}$$

$$i = i^{*} + \dot{e} \qquad \text{where} \qquad \dot{e} = \theta(\bar{e} - e) \tag{3.19}$$

$$d = u + \delta(e - p) + \gamma y - \sigma i \tag{3.20}$$

$$\begin{aligned} \dot{p} &= \Pi[d - y] \\ &= \Pi[u + \delta(e - p) + (\gamma - 1)y - \sigma i] \end{aligned} \tag{3.21}$$

The equation (3.18) is a conventional money demand equation. The equation (3.19) implies that capital is perfectly mobile. The equation (3.20) is a demand function which describes the dependency on e, p, y, and i. u is a shift parameter in this equation. The equation (3.21) is a representation of excess demand where Π is the speed of adjustment of excess demand. The following long-run relationships between the exchange rate and the price level exist, since $\dot{p} = 0$ in the long run:

$$\bar{e} = \bar{p} + \frac{1}{\delta}[\sigma i^{*} + (1 - \gamma)y - u] \tag{3.22}$$

$$e = \bar{e} - \frac{1}{\lambda\theta}[p - \bar{p}] \tag{3.23}$$

By using (3.22) and (3.23):

$$\begin{aligned} \dot{p} &= -\Pi(\frac{\delta + \sigma\theta}{\theta\lambda + \delta})[p - \bar{p}] \\ &= -\nu[p - \bar{p}] \end{aligned} \tag{3.24}$$

where $\nu \equiv \Pi(\frac{\delta+\sigma\theta}{\theta\lambda+\delta})$. Dornbusch solves the above system of differential equations and obtains the following solutions:

$$p_t = \bar{p} + (p_0 - \bar{p})e^{-\mu t} \tag{3.25}$$

$$\begin{aligned} e_t &= \bar{e} - \frac{1}{\lambda\theta}(p_0 - \bar{p})e^{-\mu t} \\ &= \bar{e} + (e_0 - \bar{e})e^{-\mu t} \end{aligned} \tag{3.26}$$

The equation (3.26) states that the exchange rate is above the long-run exchange rate if the starting value of p is below the long-run price level. He also derives the monetary expansion effect on the exchange rate:

$$\frac{\partial e}{\partial m} = 1 + \frac{1}{\lambda\theta} \tag{3.27}$$

where m is money supply. Apparently $\frac{\partial e}{\partial m} > 1$, since both λ and θ are positive. Since goods prices do not change instantaneously, exchange rates must swing quickly and must swing beyond target levels, which implies that the exchange rate overshoots its long-run level. A crucial point of this overshooting phenomenon is that the money market is in equilibrium constantly but goods market may not be in equilibrium in the short run. MacDonald (1988) and others point out some shortcomings of this model. First, the Dornbusch model assumes a one-time rise in the domestic interest rate leads to an infinite capital inflow. However, MacDonald points out that this is only applicable to the very short-run case. Secondly, the Dornbusch model allows domestic residents to hold domestic money only, not foreign assets. As King, Putnum, and Wilford (1986) point out, this assumption is not realistic. King et al. emphasize the importance of currency substitution while the Dornbusch model assumes the elasticity of substitution is zero. Thirdly, the country with expansionary monetary policy faces a current account surplus, which implies domestic residents are accumulating foreign assets. Hence, this state can not be an equilibrium. The portfolio balance model may be more appropriate to take this situation into account.

3.1.3 The Real Interest Differential Model

Frankel (1979) emphasizes a role for differences in secular rates of inflation in his model. He replaces the rational expectation of the future exchange rate with an observed proxy, the expected inflation differential. The real interest differential model of Frankel still assumes that goods prices are sticky and that PPP does not hold in the short run, as in the Dornbusch model. Frankel replaces (3.14) with the following equation:

$$\dot{e}_t = -\theta(e_t - \bar{e}_t) + \Pi_t^e - \Pi_t^{*e} \tag{3.28}$$

where Π_t^{*e} is the current rate of expected long-run inflation. Given uncovered interest parity, the long-run interest differential equals the expected long-run inflation differential:

$$\bar{i}_t - \bar{i}_t^* = \Pi_t^e - \Pi_t^{*e} \tag{3.29}$$

Therefore:

$$e_t - \bar{e}_t = -\frac{1}{\theta}[(\bar{i}_t - i_t) - (\bar{i}_t^* - i_t^*)] \tag{3.30}$$

$$\bar{e}_t = (\bar{m}_t - \bar{m}_t^*) - \phi(\bar{y}_t - \bar{y}_t^*) + \lambda(\Pi_t^e - \Pi_t^{e*}) \tag{3.31}$$

or

$$\bar{e}_t = (m_t - m_t^*) - \phi(y_t - y_t^*) + \frac{1}{\theta}(i_t - i_t^*) + (\frac{1}{\theta} + \lambda)(\Pi_t^e - \Pi_t^{e*}) \tag{3.32}$$

The reason this model is called real interest rate differential model is because the above equation can be written to have both nominal and real interest rate differentials as economic fundamentals. To estimate the model econometrically, many researchers use the following form:

$$\bar{e}_t = \alpha_1(m_t - m_t^*) + \alpha_2(y_t - y_t^*) + \alpha_3(i_t - i_t^*) + \alpha_4(\Pi_t^e - \Pi_t^{e*}) \tag{3.33}$$

In fact, this final equation (3.33) includes both the flexible-price monetary approach and sticky-price monetary approach as special cases. For instance, if $\alpha_3 = 0$ and $\alpha_4 > 0$, then model is flexible price monetary model. If $\alpha_3 < 0$ and $\alpha_4 = 0$, then model becomes the sticky-price monetary model.

3.2 The Currency Portfolio Approach

As previously mentioned, King, Putnum, and Wilford (1986) point out the importance of substitutability among different currencies. In the real world, banks and participants in foreign exchange markets tend to hold assets denominated in different currencies. These agents are considered to diversify assets in order to maximize their utility. The currency substitution model allows domestic residents to hold a basket of currencies depending on the risk and expected rates of return of the specific currencies. If the dollar is expected to depreciate, participants will substitute the dollars for other currencies, say, the German Mark. Since exchange controls were removed during 1970's, it has become much easier for market participants to hold multiple currencies.[1] Following King et al., the simple currency portfolio model will be reviewed in this section. Market participants have an incentive to hold various currencies. A money demand function can be written as follows:

$$\frac{M^d}{P} = \phi f(y, i, u) \tag{3.34}$$

where

M^d = domestic money demanded,

P = domestic price level,

ϕ = the proportion of money services provided by domestic money,

y = real income,

[1] MacDonald (1988) stresses the difference between the phenomenon of currency substitution and the capital mobility referred to by McKinnon (1982). However, the distinction between the currency substitution and capital mobility is subtle and difficult.

i = opportunity cost of holding money,

u = stochastic disturbance.

It is assumed that $0 < \phi < 1$. $1 - \phi$ is the proportion of money services provided by foreign money. Residents will allocate their holdings of currencies depending on the degree of substitutability among currencies. The question is, what factors determine ϕ? King et al. answer that the integration of world markets for goods and financial assets, I, will determine the degree of substitutability. Although currencies are ultimately imperfect substitutes because domestic currency dominates in domestic transaction, the integration of world markets for goods and services will increase substitutability among various currencies. King et al. formulate the elasticity of currency substitutability in the following fashion:

$$\sigma = k(I) \tag{3.35}$$

where

σ = elasticity of currency substitution,

I = the intensity of world market integration.

It is assumed, from the above discussion, that $\frac{d\sigma}{dI} > 0$. Although the intensity of integration is assumed to be constant for simplicity, they point out that the intensity of integration depends on several factors such as trade barriers, T, capital controls, C, transportation cost, θ, and information available concerning goods and financial markets, λ:

$$I = f(T, C; \theta, \lambda) \tag{3.36}$$

where

$$\frac{\partial I}{\partial T} < 0, \quad \frac{\partial I}{\partial C} < 0, \quad \frac{\partial I}{\partial \theta} < 0, \quad \frac{\partial I}{\partial \lambda} > 0. \tag{3.37}$$

Trade barriers, T, will hinder the intensity of the integration of world markets, while the availability of information concerning goods and asset markets, λ, will enhance the intensity of the integration.

Given the intensity of integration, I, the proportion of monetary services provided by domestic money, ϕ, is mainly determined by two factors; the expected exchange rate relative to the current spot exchange rate, e^e, and uncertainty associated with exchange rate expectations, V. The expected exchange rate relative to the current spot exchange rate, e^e, will affect the behavior of currency holders. If domestic currency holders expect the currency to depreciate, they may shift their portfolio from domestic to foreign currency. An increase in uncertainty associated with the exchange rate expectations, V, will discourage the domestic currency holders from holding domestic currency. The proportion of monetary services provided by domestic money, ϕ, is formulated as:

$$\phi = k(e^e, V|I) \tag{3.38}$$

The next question to ask is: how the expectation of exchange rate, e^e, and uncertainty associated with exchange rate expectations, V, are determined? The simplest way of formulating these two factors is:

$$e^e = l(m^e|m_w^e, I) \qquad \frac{\partial e^e}{\partial m^e} > 0 \tag{3.39}$$

where

m^e = expected domestic money supply,

m_w^e = expected foreign monetary expansion.

The expected foreign monetary expansion is considered to be given by the equation:

$$V = v[\text{var}(m^e)|m_w^e, I] \qquad \frac{\partial V}{\partial \text{var}(m^e)} > 0 \tag{3.40}$$

where

$\text{var}(m^e)$ = variance associated with expected domestic monetary policy.

An increase in the variance of expected domestic monetary policy raises uncertainty associated with exchange rate expectations. Substituting (3.39) and (3.40) into (3.38) provides:

$$\phi = h(m^e, \text{var}(m^e)|m^{*e}, I) \qquad \frac{\partial \phi}{\partial m^e} < 0, \quad \frac{\partial \phi}{\partial \text{var}(m^e)} < 0 \tag{3.41}$$

An increase in expected money supply leads to a depreciation of the currency and an increase in the proportion of foreign currency. A larger variance of monetary policy raises the holding cost of domestic currency and increases the proportion of foreign currency. Now, a generalized model of exchange rate determination can be constructed. First, the following specific money demand function is given:

$$\frac{M^d}{P} = \phi y^\alpha e^{\gamma i} e^u \tag{3.42}$$

King et al. use a growth form instead of a logarithmic form:

$$g(M^d) - g(P) = g(\phi) + \alpha g(y) + \gamma d(i) + u \tag{3.43}$$

where

$g(x) = \frac{dx/dt}{x}$.

α = real income elasticity of demand for money, $\alpha > 0$,

γ = a semi-log parameter of demand for money, $\gamma < 0$.

Money supply, M^s, is determined by the money authority and it satisfies:

$$g(M^s) = g(M) \tag{3.44}$$

Money equilibrium gives:

$$g(M^s) = g(M^d) = g(M) \tag{3.45}$$

Assuming highly integrated goods and asset markets, PPP and interest parity conditions are given in growth terms:

$$g(P) = g(P^*) + g(e) \tag{3.46}$$

$$i = i^* + g(e^e) \tag{3.47}$$

where

e = a spot exchange rate,

e^e = an expected exchange rate.

By solving this system of equations, an exchange rate in growth terms is obtained:

$$g(e) = g(P^*) - \alpha g(y) - \gamma d(i^*) - g(\phi) - \gamma dg(e^e) + g(M) + u \tag{3.48}$$

This equation includes the proportion of money, ϕ, directly. By using (3.39) and (3.40):

$$\begin{aligned} dg[e^e|M^{*e}, I] &= mdg(M^e) \\ m &= \frac{\partial e^e}{\partial M^e} > 0 \end{aligned} \tag{3.49}$$

$$\begin{aligned} g(\phi|I) &= k_1 g(M^e) + k_2 g[\text{var}(M^e)] \\ \phi &= h(m^e, \text{var}(m^e)|m^e_w, I) \end{aligned} \tag{3.50}$$

The final reduced-form becomes:

$$\begin{aligned} g[e|g(M^{*e}), I] \;=\; & -g(P^*) - \alpha g(y) - \gamma d(i^*) - k_1 g(M^e) \\ & - \gamma mdg(M^e) - k_2[\text{var}g(M^e)] + g(M) + u \\ & k_1 < 0, \quad K_2 < 0, \quad \gamma m < 0 \end{aligned} \tag{3.51}$$

This yields a general form of the monetary approach to exchange rate determination. The final equation states that a decrease in world price, P^*, an increase in world interest rate, i^*, and a decrease in domestic income, y, will result in the exchange depreciation.

4 THE DORNBUSCH STICKY PRICE MODEL: LARGE-COUNTRY CASE

As was discussed in the previous chapter, the Dornbusch's sticky price model explains the overshooting phenomena by introducing differences in the adjustment speed of the money market and goods market. Since the speed of price adjustment in the goods market is assumed to be slower than the speed of price adjustment in the money market, the exchange rate overshoots its long-run target to compensate for slow adjustment in the goods market. This chapter will consider a simple variation of the Dornbusch sticky price model. Further, it will introduce a new assumption to the model, that is, that both domestic and foreign countries are large countries. This enables prices in both countries in the model to be endogenized. The first section will consider Dornbusch's two-country sticky price model, then the second section will extend the model to the three-country case where all three countries are considered to be large.

4.1 Two-Country Case

Here, Dornbusch's small-country assumption in the two-country case is replaced with the large-country assumption. Consider two large countries, such as Germany and the United States, since these countries are large, prices are no longer assumed to be given in either country. The model will attempt to endogenize prices in both countries. Following Dornbusch (1976a), four markets will be introduced; a domestic and a foreign money market and a domestic and a foreign goods market. It is assumed that two countries influence each other only through the goods markets. The money markets are introduced below.

4.1.1 The Money Markets

In the money markets, domestic and foreign interest rates will be determined in equilibrium. As in Dornbusch (1976a), the conventional money demand functions are:

$$m^d - p = \alpha y - \beta i \tag{4.1}$$

$$m^{*d} - p^* = \alpha^* y^* - \beta^* i^* \tag{4.2}$$

where

i = domestic nominal interest rate,

m^d = log of the domestic nominal quantity of money,

p = log of the domestic price level,

y = log of domestic real income.

Note that the asterisks indicate foreign variables in the rest of the chapter.

Domestic real money demand depends only on domestic variables, y, p, and i. Foreign real money demand also depends only on foreign variables. Thus, money market equilibrium in both domestic and foreign money market creates the following equations:

$$m^d = m \tag{4.3}$$

$$m^{*d} = m^* \tag{4.4}$$

where m and m^* are money supply in the money market in each country.

4.1.2 The Goods Markets

One assumption made in the goods market is that domestic demand depends on the relative price of domestic goods, real income, interest rate, and shift variables in the goods market. Since both countries are assumed to be large countries, the relative price of domestic goods is $e - p + p^*$ where e is log of exchange rate which is the dollar price of foreign currency. In Dornbusch's seminal paper, he normalizes, without loss of generality, the log of the price of foreign goods ($p^* = 0$) using his small-country assumption. The demand function for domestic and foreign goods is assumed to be:

$$d = -\delta(e - p + p^*) + \phi y - \lambda i + \mu \tag{4.5}$$

$$d^* = \delta^*(e - p + p^*) + \phi^* y^* - \lambda^* i^* + \mu^* \tag{4.6}$$

where

$d =$ log of domestic demand,

$e =$ log of exchange rate,

$\mu =$ shift parameter.

The price change is proportional to the excess demand for goods:

$$\dot{p} = \Pi[d - y] = \Pi[-\delta(e - p + p^*) + (\phi - 1)y - \lambda i + \mu] \tag{4.7}$$

$$\dot{p^*} = \Pi^*[d^* - y^*] = \Pi^*[\delta^*(e - p + p^*) + (\phi^* - 1)y^* - \lambda^* i^* + \mu^*] \tag{4.8}$$

Finally, uncovered interest rate parity is introduced:

$$\dot{e} = i - i^* \tag{4.9}$$

From (4.1) and (4.7)[1]:

$$\dot{p} = \Pi[-\delta e - (-\delta + \frac{\lambda}{\beta})p - \delta p^* + (\phi - 1 - \frac{\lambda\alpha}{\beta})y + \frac{\lambda}{\beta}m + \mu] \tag{4.10}$$

(4.2) and (4.8) provide:

$$\dot{p^*} = \Pi^*[\delta^* e - \delta^* p + (\delta^* - \frac{\lambda^*}{\beta^*})p^* + (\phi^* - 1 - \frac{\lambda^*\alpha^*}{\beta^*})y^* + \frac{\lambda^*}{\beta^*}m^* + \mu^*] \tag{4.11}$$

Combining (4.1), (4.2) and (4.9) yields:

$$\dot{e} = \frac{1}{\beta}p - \frac{1}{\beta^*}p^* - \frac{1}{\beta}m + \frac{1}{\beta^*}m^* + \frac{\alpha}{\beta}y - \frac{\alpha^*}{\beta^*}y^* \tag{4.12}$$

In matrix form, the three equations above are written as follows:

$$\begin{bmatrix} \dot{e} \\ \dot{p} \\ \dot{p^*} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{\beta} & -\frac{1}{\beta^*} \\ -\Pi\delta & -\Pi(-\delta + \frac{\lambda}{\beta}) & \Pi\delta \\ \Pi^*\delta^* & -\Pi^*\delta^* & \Pi^*(\delta^* - \frac{\lambda^*}{\beta^*}) \end{bmatrix} \begin{bmatrix} e \\ p \\ p^* \end{bmatrix} +$$

$$\begin{bmatrix} -\frac{1}{\beta} & \frac{\alpha}{\beta} & 0 & \frac{1}{\beta^*} & -\frac{\alpha^*}{\beta^*} & 0 \\ \frac{\Pi\lambda}{\beta} & \Pi(\phi - 1 - \frac{\lambda\alpha}{\beta}) & \Pi & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Pi^*\lambda^*}{\beta^*} & \Pi^*(\phi^* - 1 - \frac{\lambda^*\alpha^*}{\beta^*}) & \Pi^* \end{bmatrix} \begin{bmatrix} m \\ y \\ \mu \\ m^* \\ y^* \\ \mu^* \end{bmatrix} \tag{4.13}$$

[1] The equation (4.1) is solved for i and substituted in (4.7).

The change in exchange rate, $\dot{e}$, will be directly affected by both domestic and foreign variables; m, y, m^* and y^*. However, the change in domestic price, $\dot{p}$, will be directly affected by domestic variables, m and y, and indirectly by foreign variables, m^* and y^*, through foreign prices. This is because two countries are not directly linked in the money markets. The paths of the exchange rate, domestic price, and foreign price will be obtained by solving the above system of differential equations simultaneously. Later, the above model is applied to the data to investigate long-run relations and dynamics among the variables. For this purpose, theoretical long-run relations must be derived from the above model. Since, in the long run, all economic variables are in the state of equilibrium, the left-hand side of (4.10), (4.11) and (4.12) are set to zero, i.e., $\dot{e} = \dot{p} = \dot{p}^* = 0$. Then, the following 3 long-run relations are obtained:

$$\frac{1}{\beta}p - \frac{1}{\beta^*}p^* - \frac{1}{\beta}m + \frac{1}{\beta^*}m^* + \frac{\alpha}{\beta}y - \frac{\alpha^*}{\beta^*}y^* = 0 \tag{4.14}$$

$$-\delta e - (-\delta + \frac{\lambda}{\beta})p - \delta p^* + (\phi - 1 - \frac{\lambda\alpha}{\beta})y + \frac{\lambda}{\beta}m + \mu = 0 \tag{4.15}$$

$$\delta^* e - \delta^* p + (\delta^* - \frac{\lambda^*}{\beta^*})p^* + (\phi^* - 1 - \frac{\lambda^*\alpha^*}{\beta^*})y^* + \frac{\lambda^*}{\beta^*}m^* + \mu^* = 0 \tag{4.16}$$

(4.14) is a long-run equilibrium in the money market. (4.15) and (4.16) are equilibria in the two goods markets. If the following 3 variables, real exchange rate, real money supply and real GNP, are defined as:

$$E \equiv e - p + p^*,$$
$$M \equiv m - p,$$
$$Y \equiv y.$$

then (4.14), (4.15) and (4.16) will be rewritten as follows[2]:

$$-\frac{1}{\beta}M + \frac{1}{\beta^*}M^* + \frac{\alpha}{\beta}Y - \frac{\alpha^*}{\beta^*}Y^* = 0 \tag{4.17}$$

$$-\delta E + \frac{\lambda}{\beta}M + (\phi - 1 - \frac{\lambda\alpha}{\beta})Y = 0 \tag{4.18}$$

$$\delta^* E + \frac{\lambda^*}{\beta^*}M^* + (\phi^* - 1 - \frac{\lambda^*\alpha^*}{\beta^*})Y^* = 0 \tag{4.19}$$

(4.17), (4.18) and (4.19) still have the same interpretations as (4.14), (4.15) and (4.16). These are long-run relations that will be examined using the data set.

[2]Shift parameters are omitted from the equations for simplicity.

4.2 Three-Country Case

In this section, the two country model is extended to the three-country case by adding a third country where all three countries are assumed to be large countries. In the example, the home country is the United States and the two foreign countries are Germany and Japan. The money and goods market will be introduced for each country. It is assumed that these three countries interact only in the goods markets.

4.2.1 The Money Markets

The money demand depends only on domestic variables in each country. It will not depend on any foreign variables:

$$m^d - p = \alpha y - \beta i \tag{4.20}$$

$$m^{*d} - p^* = \alpha^* y^* - \beta^* i^* \tag{4.21}$$

$$m^{**d} - p^{**} = \alpha^{**} y^{**} - \beta^{**} i^{**} \tag{4.22}$$

All the variables given here are defined as in the previous section. The single asterisk indicates the variables of the first foreign country (here, Germany) and the double asterisks indicate those of the second foreign country (Japan). Again, in the equilibrium, money demand and money supply are equal in the money market of all three countries:

$$m^d = m \tag{4.23}$$

$$m^{*d} = m^* \tag{4.24}$$

$$m^{**d} = m^{**} \tag{4.25}$$

4.2.2 The Goods Markets

The demand in each country depends on real income, interest rates, relative prices of domestic goods and shift parameters:

$$d = -\delta(e_1 - p + p^*) - \sigma(e_2 - p + p^{**}) + \phi y - \lambda i + \mu \tag{4.26}$$

$$d^* = \delta^*(e_1 - p + p^*) + \sigma^*(e_1 - e_2 + p^* - p^{**}) + \phi^* y^* - \lambda^* i^* + \mu^* \tag{4.27}$$

$$d^{**} = \delta^{**}(e_2 - p + p^{**}) - \sigma^{**}(e_1 - e_2 + p^* - p^{**}) + \phi^{**} y^{**} - \lambda^{**} i^{**} + \mu^{**} \tag{4.28}$$

where the parameters δ, σ, ϕ, λ, ϕ^*, λ^*, σ^{**}, ϕ^{**}, λ^{**}, σ^* and δ^{**} are assumed to be positive. Here again, all variables are defined as in the previous section. Since there are three countries in this model, demand for domestic goods has three sources; demand for domestic goods in the domestic market which depends on domestic real income and interest rate, demand in the first foreign country, and demand in the second country. The term $e_1 - p + p^*$ captures the relative price of the domestic goods to the first country's goods and explains the first country's demand for the domestic goods. The term $e_2 - p + p^{**}$ captures the relative price of the domestic goods to the second foreign country's goods. The price change is proportional to the excess demand in each country:

$$\dot{p} = \Pi[d - y] = \Pi[-\delta(e_1 - p + p^*) - \sigma(e_2 - p + p^{**}) + (\phi - 1)y - \lambda i + \mu] \tag{4.29}$$

$$\dot{p}^* = \Pi^*[d^* - y^*] = \Pi^*[\delta^*(e_1 - p + p^*) + \sigma^*(e_1 - e_2 + p^* - p^{**}) + (\phi^* - 1)y^* - \lambda^* i^* + \mu^*] \tag{4.30}$$

$$\dot{p}^{**} = \Pi^{**}[d^{**} - y^{**}] = \Pi^{**}[\delta^{**}(e_2 - p + p^{**}) - \sigma^{**}(e_1 - e_2 + p^* - p^{**}) + (\phi^{**} - 1)y^{**} - \lambda^{**} i^{**} + \mu^{**}] \tag{4.31}$$

Combining (4.20) and (4.29) provides:

$$\dot{p} = \Pi[-\delta e_1 - \sigma e_2 + (\delta + \sigma - \frac{\lambda}{\beta})p - \delta p^* - \sigma p^{**} + (\phi - 1 - \frac{\lambda\alpha}{\beta})y + \frac{\lambda}{\beta}m + \mu] \tag{4.32}$$

From (4.21) and (4.30):

$$\dot{p}^* = \Pi^*[(\delta^* + \sigma^*)e_1 - \sigma^* e_2 - \delta^* p + (\delta^* + \sigma^* - \frac{\lambda^*}{\beta^*})p^* - \sigma^* p^{**} + (\phi^* - 1 - \frac{\lambda^*\alpha^*}{\beta^*})y^* + \frac{\lambda^*}{\beta^*}m^* + \mu^*] \tag{4.33}$$

(4.22) and (4.31) gives:

$$\dot{p}^{**} = \Pi^{**}[-\sigma^{**} e_1 + (\delta^{**} + \sigma^{**})e_2 - \delta^{**} p - \sigma^{**} p^* + (\delta^{**} + \sigma^{**} - \frac{\lambda^{**}}{\beta^{**}})p^{**} + (\phi^{**} - 1 - \frac{\lambda^{**}\alpha^{**}}{\beta^{**}})y^{**} + \frac{\lambda^{**}}{\beta^{**}}m^{**} + \mu^{**}] \tag{4.34}$$

Assuming that uncovered interest parity holds for the two exchange rates:

$$\dot{e_1} = i - i^{*} \tag{4.35}$$

$$\dot{e_2} = i - i^{**} \tag{4.36}$$

Then, the following equations are obtained:

$$\dot{e_1} = -\frac{1}{\beta}[m - p - \alpha y] + \frac{1}{\beta^{*}}[m^{*} - p^{*} - \alpha^{*} y] \tag{4.37}$$

$$\dot{e_2} = -\frac{1}{\beta}[m - p - \alpha y] + \frac{1}{\beta^{**}}[m^{**} - p^{**} - \alpha^{**} y^{**}] \tag{4.38}$$

Putting the above five equations in a matrix form:

$$\dot{X} = \Phi X + \Theta Y \tag{4.39}$$

where

$$\dot{X}' = \begin{bmatrix} \dot{e_1} & \dot{e_2} & \dot{p} & \dot{p}^{*} & \dot{p}^{**} \end{bmatrix},$$

$$X' = \begin{bmatrix} e_1 & e_2 & p & p^{*} & p^{**} \end{bmatrix},$$

$$Y' = \begin{bmatrix} m & y & \mu & m^{*} & y^{*} & \mu^{*} & m^{**} & y^{**} & \mu^{**} \end{bmatrix},$$

$$\Phi = \begin{bmatrix} 0 & 0 & \frac{1}{\beta} & -\frac{1}{\beta^{*}} & 0 \\ 0 & 0 & \frac{1}{\beta} & 0 & -\frac{1}{\beta^{**}} \\ -\Pi\delta & -\Pi\sigma & \Pi(\delta + \sigma - \frac{\lambda}{\alpha}) & -\Pi\delta & -\Pi\sigma \\ \Pi^{*}(\delta^{*} + \sigma^{*}) & -\Pi^{*}\sigma^{*} & -\Pi^{*}\delta^{*} & \Pi^{*}(\delta^{*} + \sigma^{*} - \frac{\lambda^{*}}{\beta^{*}}) & -\Pi^{*}\sigma^{*} \\ -\Pi^{**}\sigma^{**} & \Pi^{**}(\delta^{**} + \sigma^{**}) & -\Pi^{**}\delta^{**} & -\Pi^{**}\sigma^{**} & \Pi^{**}(\delta^{**} + \sigma^{**} - \frac{\lambda^{**}}{\beta^{**}}) \end{bmatrix},$$

$$\Theta = \begin{bmatrix} -\frac{1}{\beta} & \frac{\alpha}{\beta} & 0 & \frac{1}{\beta^{*}} & -\frac{\alpha^{*}}{\beta^{*}} & 0 & 0 & 0 & 0 \\ -\frac{1}{\beta} & \frac{\alpha}{\beta} & 0 & 0 & 0 & 0 & \frac{1}{\beta^{**}} & -\frac{\alpha^{**}}{\beta^{**}} & 0 \\ \frac{\Pi\lambda}{\beta} & \Pi(\phi - 1 - \frac{\lambda\alpha}{\beta}) & \Pi & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Pi^{*}\lambda^{*}}{\beta^{*}} & \Pi^{*}(\phi^{*} - 1 - \frac{\lambda^{*}\alpha^{*}}{\beta^{*}}) & \Pi^{*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\Pi^{**}\lambda^{**}}{\beta^{**}} & \Pi^{**}(\phi^{**} - 1 - \frac{\lambda^{**}\alpha^{**}}{\beta^{**}}) & \Pi^{**} \end{bmatrix}.$$

In matrix Θ, it can be seen that foreign variables do not affect price change in domestic price directly. However, foreign variables do affect them indirectly through foreign price levels, as seen in the previous section. The paths of all five variables will be found by solving this system of the 5 equations simultaneously. Here, it is also possible to derive long-run relations among the variables. Setting the left-hand of (4.32), (4.33), (4.34), (4.37), and (4.38) equal to zero obtains the following 5 long-run relations:

$$-\frac{1}{\beta}[m - p - \alpha y] + \frac{1}{\beta^{*}}[m^{*} - p^{*} - \alpha^{*} y] = 0 \tag{4.40}$$

$$-\frac{1}{\beta}[m-p-\alpha y]+\frac{1}{\beta^{**}}[m^{**}-p^{**}-\alpha^{**}y^{**}]=0 \tag{4.41}$$

$$-\delta e_1-\sigma e_2+(\delta+\sigma-\frac{\lambda}{\beta})p-\delta p^{*}-\sigma p^{**}+\frac{\lambda}{\beta}m+(\phi-1-\frac{\lambda\alpha}{\beta})y+\mu=0 \tag{4.42}$$

$$\begin{aligned}(\delta^{*}+\sigma^{*})e_1-\sigma^{*}e_2-\delta^{*}p+(\delta^{*}+\sigma^{*}-\frac{\lambda^{*}}{\beta^{*}})p^{*}-\sigma^{*}p^{**}\\+\frac{\lambda^{*}}{\beta^{*}}m^{*}+(\phi^{*}-1-\frac{\lambda^{*}\alpha^{*}}{\beta^{*}})y^{*}+\mu^{*}=0\end{aligned} \tag{4.43}$$

$$\begin{aligned}-\sigma^{**}e_1+(\delta^{**}+\sigma^{**})e_2-\delta^{**}p-\sigma^{**}p^{*}+(\delta^{**}+\sigma^{**}-\frac{\lambda^{**}}{\beta^{**}})p^{**}\\+\frac{\lambda^{**}}{\beta^{**}}m^{**}+(\phi^{**}-1-\frac{\lambda^{**}\alpha^{**}}{\beta^{**}})y^{**}+\mu^{**}=0\end{aligned} \tag{4.44}$$

(4.40) and (4.41) are long-run equilibria in the two money markets. (4.42), (4.43) and (4.44) are goods market equilibria. Rewriting the above system of five equations in terms of real variables[3]; E_1, E_2, M, M^{*}, M^{**}, Y, Y^{*}, and Y^{**} [4] yields:

$$-\frac{1}{\beta}M+\frac{1}{\beta^{*}}M^{*}+\frac{\alpha}{\beta}Y-\frac{\alpha^{*}}{\beta^{*}}Y^{*}=0 \tag{4.45}$$

$$-\frac{1}{\beta}M+\frac{1}{\beta^{**}}M^{**}+\frac{\alpha}{\beta}Y-\frac{\alpha^{**}}{\beta^{**}}Y^{**}=0 \tag{4.46}$$

$$-\delta E_1-\sigma E_2+\frac{\lambda}{\beta}M+(\phi-1-\frac{\lambda\alpha}{\beta})Y=0 \tag{4.47}$$

$$(\delta^{*}+\sigma^{*})E_1-\sigma^{*}E_2+\frac{\lambda^{*}}{\beta^{*}}M^{*}+(\phi^{*}-1-\frac{\lambda^{*}\alpha^{*}}{\beta^{*}})Y^{*}=0 \tag{4.48}$$

$$-\sigma^{**}E_1+(\delta^{**}+\sigma^{**})E_2+\frac{\lambda^{**}}{\beta^{**}}M^{**}+(\phi^{**}-1-\frac{\lambda^{**}\alpha^{**}}{\beta^{**}})Y^{**}=0 \tag{4.49}$$

Later, this part will examine the relations among these real variables and will use the error correction model to investigate long-run relations and short-run dynamics among these real variables. The purpose then will be to apply the cointegration techniques to the data set.

[3] Again, shift variables are omitted from the equations.
[4] Each variable is defined as previously.
$E_1 \equiv e-p+p^{*}$, $E_2 \equiv e-p+p^{**}$, $M \equiv m-p$,
$M^{*} \equiv m^{*}-p^{*}$, $M^{**} \equiv m^{**}-p^{**}$, $Y \equiv y$,
$Y^{*} \equiv y^{*}$, $Y^{**} \equiv y^{**}$.

5 ERROR CORRECTION MODEL (ECM)

Since Sims' influential work (1980), many researchers have analyzed the dynamics of economic systems by using a vector autoregression (VAR) model. However, when some of the variables in the system are integrated of order d, i.e., the system contains d unit roots, the VAR model in level is no longer appropriate. To deal with the variables integrated of order d, the error correction model (ECM) is introduced. In the ECM, both terms in the level and in the difference are included. The first section discusses VAR model, and some issues associated with the model, while the error correction model and cointegration analysis will be discussed in the second section. In econometrics, some variables of prime interest are analyzed by means of the information of other variables. Usually, the former is called an endogenous variable while the latter is called an exogenous variable. In other words, endogenous variables are modeled conditioned on exogenous variables. It would be easier to use a conditional model and leave the exogenous variables unspecified or at least model them less carefully. Some researchers have combined the concept of weak exogeneity and the error correction model and thus are able to analyze the cointegration in the reduced dimensions. Weak exogeneity and the partial system model are the topics of the last section.

5.1 Vector Autoregression Model

VAR is a popular technique to analyze the dynamics of economic systems. Good references on the VAR model are Sims (1980), Hamilton (1994) and Watson (1994). In this section, some of the properties of a VAR are discussed. Suppose that y_t is an $(n \times 1)$ vector and ε_t is an $(m \times 1)$ vector. Consider the following model:

$$y_t = C(L)\varepsilon_t \tag{5.1}$$

The model (5.1) is called the structural moving average model. The vector y_t is understood to contain endogenous economic variables and ε_t are exogenous shocks to the economy. ε_t is not directly observed, however, it can be observed indirectly through its effects on y_t. $C(L)$ is assumed to be a lag polynomial matrix:

$$C(L) = c_0 + c_1 L + c_2 L^2 + \cdots = \sum_{k=0}^{\infty} c_k L^k \tag{5.2}$$

where c_k is an $(n \times m)$ matrix and L is a lag operator. A typical element of c_k is denoted by $c_{ij,k}$ obtained from (5.1) and (5.2) as follows:

$$c_{ij,k} = \frac{\partial y_{i,t}}{\partial \varepsilon_{j,t-k}} = \frac{\partial y_{i,t+k}}{\partial \varepsilon_{j,t}} \tag{5.3}$$

$y_{i,t}$ is the i th element of y_t, and $\varepsilon_{j,t}$ is the j th element of ε_t. This $c_{ij,k}$ is called the impulse response function of $\varepsilon_{j,t}$ for $y_{i,t}$ if viewed as a function of k. $c_{ij,k}$ tells how much the i th element of y_{t+k} will be affected by a change in the j th element of ε_t. Now, assuming that ε_t is independently identically distributed, then $\varepsilon_t \sim iid(0, \Omega)$. This implies that serial correlation among the exogenous variables will be captured by $C(L)$. Inverting $C(L)$ in (5.1) gives the structural VAR representation:

$$A(L) y_t = \varepsilon_t \tag{5.4}$$

where $A(L) = A_0 - \sum_{k=1}^{\infty} A_k L^k$. In other words, exogenous variables ε_t can be written as a function of current and lagged endogenous variables y_t. In most of the cases, especially for empirical purposes, a finite order polynomial is used. Note that the invertibility of $C(L)$ is not necessarily the case. For instance, if $n \leq m$, $C(L)$ is not invertible. By assuming $n = m$, $C(L)$ is a square matrix and as long as all the roots of $|C(z)| = 0$ are outside the unit circle, $C(L)$ is invertible.[1] Assuming that the lag polynomial of $A(L)$ is finite and of order p, then (5.4) will be written as:

$$A_0 y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t \tag{5.5}$$

Watson points out that (5.5) is different from the standard simultaneous equation setting because no observable exogenous variables are included in the equation. However, standard techniques can be applied to the equations for estimation purposes by treating exogenous and predetermined variables equally. Dividing both sides of (5.5) by A_0 gives:

$$y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_p y_{t-p} + e_t \tag{5.6}$$

where $\Phi_i = A_0^{-1} A_i$ and $e_t = A_0^{-1} \varepsilon_t$. It is assumed that $e_t \sim iid(0, \Sigma)$ where $\Sigma = (A_0^{-1}) \Omega (A_0^{-1})'$. The number of parameters to estimate is $(n \times n) \times p + n \times (n+1)/2$. On the other hand, in the structural model (5.5) there are $n^2 \times (p+1) + n \times (n+1)/2$ parameters to estimate. Hence, at least n^2 restrictions must be imposed for identification.[2]

[1] Watson (1994) discusses the problems when the roots of $|C(z)| = 0$ are inside and on the unit circle.

[2] For identification issue, see Johnston (1983).

5.2 Identification Issues

Typically, researchers impose the restriction that the diagonal elements of A_0 are equal to 1 and the rest of the $(n-1)$ restrictions are based on economic models. There are mainly two ways of imposing restrictions. One is to impose restrictions on the coefficients, for example, if economic theory predicts that some variables should not be included in the model, the coefficients of these variables can be set equal to zero. Another way, a point made by Sims (1980), is to impose restrictions on the covariance matrix of the structural shocks Ω, the matrix of contemporaneous coefficients A_0 and the matrix of long-run multipliers $A(1)^{-1}$. If the structural shocks are considered to be uncorrelated, the restriction on diagonal Ω can be imposed. This requires $n \times (n-1)$ restrictions and $n \times (n-2)/2$ additional necessary conditions are needed. The additional restrictions may come from the matrix A_0. Watson gives us some examples in the bivariate case in his paper. For instance, if one exogenous shock, say ε_2, does not affect an endogenous variable, y_1, in a bivariate setting, a lower triangular structure can be imposed on the matrix A_0.[3] This will give $n(n-1)/2$ more restrictions. Other non-triangular type of restrictions are used by many researchers[4] while researchers, such as Blanchard and Quah (1989) and King et al. (1991), prefer alternative restrictions on the matrix $A(1)^{-1}$. In any case, finding $n(n-1)/2$ extra restrictions on the long-run multiplier enables the system to be identified

5.3 Estimation

There are several techniques to estimate the parameters of the structural VAR; for example, generalized method of moments (GMM) or the maximum likelihood (ML) method. The simplest GMM estimator is the indirect least squares method. The GMM technique uses the following relations with the OLS estimators of the reduced form:

$$A_0^{-1}A_i = \phi_i \tag{5.7}$$

$$(A_0)\Omega(A_0)' = \Sigma \tag{5.8}$$

As long as it can be assumed that the model is exactly identified, OLS can be applied to the reduced form to obtain $\hat{\phi}_i$, $\hat{\Sigma}$. Given $\hat{A}_0$, then, $\hat{A}_i = \hat{A}_0\hat{\phi}_i$, and $\hat{\Omega} = (\hat{A}_0^{-1})\hat{\Sigma}(\hat{A}_0^{-1})'$. Readers are referred to Hamilton (1994) for the details. The maximum likelihood method is more frequently used for VAR estimation. Consider the following model:

$$y_t = c + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \varepsilon_t \tag{5.9}$$

[3] See Watson (1994) and Sims (1980) for the detail.

[4] See Watson (1986), Sims (1986) and Blanchard and Watson (1986).

where $\varepsilon_t \sim iidN(0, \Sigma)$. Note that ε_t is assumed to be normally distributed. When the conditional likelihood function is introduced, then:

$$f_{Y_T, Y_{T-1}, \cdots, Y_1 | Y_0, Y_{-1}, \cdots, Y_{-p+1}}(y_T, y_{T-1}, \cdots, y_1 | y_0, y_{-1}, \cdots, y_{-p+1} : \theta) \tag{5.10}$$

The first p observations are conditioned for this function while the last T observations serve as a basis for estimation. The introduction of the normality assumption gives:

$$\begin{aligned} y_t | y_{t-1}, \cdots, y_{-p+1} &\sim N(c + \Phi_1 y_{t-p} + \cdots + \Phi_p y_{t-p}, \Sigma) \\ &\sim N(\Pi' x_t, \Sigma) \end{aligned} \tag{5.11}$$

where $\Pi' \equiv [c\ \Phi_1\ \Phi_2 \cdots \Phi_p]$ and $x_t' \equiv [1\ y_{t-1}' \cdots y_{t-p}']$. Hence, the following is obtained:

$$\begin{aligned} & f_{Y_T | Y_{T-1}, \cdots, Y_{-p+1}}(y_T | y_{T-1}, \cdots, y_{-p+1} : \theta) \\ &= (2\Pi)^{-n/2} |\Sigma^{-1}|^{1/2} \exp[-\frac{1}{2}(y_t - \Pi' x_t)' \Sigma^{-1} (y_t - \Pi' x_t)] \end{aligned} \tag{5.12}$$

The joint density of observations from 1 to t, conditioned on $y_0, \cdots, y_{-p+1}$, is:

$$\begin{aligned} & f_{Y_1, Y_2, \cdots, Y_t | y_0, \cdots, y_{-p+1}}(y_T, \cdots, y_1 | y_0, y_{-1}, \cdots, y_{-p+1} : \theta) \\ &= \Pi_{t=1}^{T} f_{Y_t | Y_{t-1}, \cdots, Y_{-p+1}}(y_t | y_{t-1}, \cdots, y_{-p+1} : \theta) \end{aligned} \tag{5.13}$$

The log likelihood function will be:

$$\begin{aligned} & \sum_{t=1}^{T} \log f_{Y_t | Y_{t-1}, \cdots, Y_{-p+1}}(y_t | y_{t-1}, \cdots, y_{-p+1} : \theta) \\ &= -\frac{Tn}{2} \log(2\Pi) + \frac{T}{2} \log|\Sigma^{-1}| - \frac{1}{2} \sum_{t=1}^{T} [(y_t - \Pi' x_t)' \Sigma^{-1} (y_t - \Pi' x_t)] \end{aligned} \tag{5.14}$$

To find the ML estimators of Π and Σ, a derivative with respect to Π is taken and then the derivative is set equal to zero:

$$\hat{\Pi}' = [\sum_{t=1}^{T} y_t x_t']' [\sum_{t=1}^{T} x_t x_t']^{-1} \tag{5.15}$$

The i th row of $\hat{\Pi}'$ is:

$$\hat{\Pi}_i' = [\sum_{t=1}^{T} y_{it} x_t']' [\sum_{t=1}^{T} x_t x_t']^{-1} \tag{5.16}$$

which implies that this is an OLS estimator by regressing y_{it} on x_t. Hence, the parameters in the VAR model can be estimated by applying OLS to each equation, that is, by regressing each y_{it} on a constant and p lags of all the variables in the system. To find the ML estimator of Σ, first, the likelihood function at $\hat{\Pi}$ is evaluated:

$$\sum_{t=1}^{T} \log f_{Y_t|Y_{t-1},\cdots,Y_{-p+1}}(y_t|y_{t-1},\cdots,y_{-p+1}:\theta)$$
$$= -\frac{Tn}{2}\log(2\Pi) + \frac{T}{2}\log|\Sigma^{-1}| - \frac{1}{2}\sum_{t=1}^{T}[(y_t - \Pi' x_t)'\Sigma^{-1}(y_t - \Pi' x_t)] \tag{5.17}$$

$$L(\hat{\Pi}) = -\frac{Tn}{2}\log(2\Pi) + \frac{T}{2}\log|\Sigma^{-1}| - \frac{1}{2}\sum_{t=1}^{T}\hat{\varepsilon}_t'\Sigma^{-1}\hat{\varepsilon}_t \tag{5.18}$$

Taking a derivative with respect to Σ and setting the derivative equal to zero, yields:

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t' \tag{5.19}$$

The i th row and i th column of $\hat{\Sigma}$, σ_{ii}, is the ML estimator of the variance for the i th equation. The i th row and j th column of $\hat{\Sigma}$, σ_{ij} is the ML estimator of the covariance between the equation i and j.

5.4 Hypothesis Testing

The matrix $\hat{\Sigma}$ can be used to conduct a simple likelihood ratio test. Suppose that the number of lags for the variables to be included in the model must be determined. The null hypothesis is that the number of lags to be included is P_0 and the alternative hypothesis is the number of lags is P_1, where $P_0 < P_1$. Two sets of n OLS regressions can be performed, one of which has a constant and P_0 lags of the variables and the other, a constant and P_1 lags of the variables. These sets of OLS regressions yield the equations; $\hat{\Sigma}_0 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t(P_0)\hat{\varepsilon}_t(P_0)'$, the variance-covariance matrix from the first set of n OLS regressions, and $\hat{\Sigma}_1 = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_t(P_1)\hat{\varepsilon}_t(P_1)'$, the variance-covariance matrix from the second set of n OLS regressions. Likelihood ratio statistic for this test is computed by:

$$2(\ell_1^* - \ell_0^*) = T[\log|\hat{\Sigma}_0| - \log|\hat{\Sigma}_1|] \tag{5.20}$$

where ℓ_0^* is a likelihood function evaluated at $\Sigma = \hat{\Sigma}_0$ and ℓ_1^* is a likelihood function evaluated at $\Sigma = \hat{\Sigma}_1$. It can be proved that this statistic asymptotically has χ^2 with $n^2(P_1 - P_0)$ degrees of freedom. To take account of small-sample bias, Sims (1980) recommends:

$$(T - K)[\log|\hat{\Sigma}_0| - \log|\hat{\Sigma}_1|] \tag{5.21}$$

where $K = 1 + nP_1$ instead of (5.20). The ML estimators, $\hat{\Pi}$ and $\hat{\Sigma}$, are consistent estimators even if ε_t is not normally distributed. Therefore, if ε_t is independently and identically distributed with mean

0 and variance Σ and the fourth moment of ε_t is finite, then:

$$\varepsilon_t \sim iid(0, \Sigma) \tag{5.22}$$

$$E(\varepsilon_{it}\varepsilon_{jt}\varepsilon_{kt}\varepsilon_{lt}) < \infty \qquad \forall i, j, k, l \tag{5.23}$$

and if the roots of $|I_n - \Phi_1 z - \cdots - \Phi_p z^p| = 0$ are outside the unit circle, then the following results hold:

$$\frac{1}{T}\sum_{t=1}^{T} x_t x_t' \xrightarrow{P} Q = E(x_t x_t') \tag{5.24}$$

$$\hat{\pi}_T \xrightarrow{P} \pi \tag{5.25}$$

where $\hat{\pi}_T = vec(\hat{\Pi}_t)$

$$\hat{\Sigma}_T \xrightarrow{P} \Sigma \tag{5.26}$$

$$\begin{pmatrix} \sqrt{T}(\hat{\Pi}_T - \Pi) \\ \sqrt{T}(vech(\hat{\Sigma}_T) - vech(\Sigma)) \end{pmatrix} \xrightarrow{L} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} \Sigma \otimes Q^{-1} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \right) \tag{5.27}$$

where $vech$ is a transformation operator that transform an $(n \times n)$ matrix into an $(n(n+1)/2 \times 1)$ vector by stacking these elements on or below the principal diagonal. The element of Σ_{22} is given by $\sigma_{il}\sigma_{jm} + \sigma_{im}\sigma_{jl}$ for all $i, j, l, m = 1, \cdots, n$. The above reveals that the usual OLS t and F test can be applied asymptotically to the coefficients in each equation in VAR system. For instance, to test some restrictions on the coefficients, say, $R\Pi = r$, then:

$$\sqrt{T}(R\hat{\Pi}_T - r) \xrightarrow{L} N(0, R(\Sigma \otimes Q^{-1})R') \tag{5.28}$$

which implies:

$$\sqrt{T}(R\hat{\Pi}_T - r) \xrightarrow{P} N(0, R(\hat{\Sigma}_T \otimes \hat{Q}^{-1}{}_T)R') \tag{5.29}$$

where $\hat{\Sigma}_T = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t\hat{\varepsilon}_t'$, and $Q_T = \frac{1}{T}\sum_{t=1}^{T} x_t x_t'$. Consider the following statistic:

$$(R\hat{\Pi}_T - r)'\{R[\hat{\Sigma}_T \otimes (\sum_{t=1}^{T} x_t x_t')^{-1}]R'\}(R\hat{\Pi}_T - r) \tag{5.30}$$

This statistic asymptotically follows a χ^2 distribution with m degrees of distribution where m is the number of restrictions. So far, only the unrestricted VAR model has been discussed. In other words, all the equations in the VAR have the same regressors, that is, a constant and lags of all the variables. If restrictions are imposed on the coefficients, the coefficient estimation changes slightly. If the restriction that some of the variables do not have explanatory power in predicting other variables is imposed, then exactly the same variables will not be found in all equations. If y_t is divided into two groups; y_{1t} which is $(n_1 \times 1)$ vector and y_{2t} which is $(n_2 \times 1)$ vector where $n_1 + n_2 = n$, corresponding lags, x_{1t} and x_{2t}, are also defined that is, $x_{1t} \equiv [y'_{1t-1}\ y'_{1t-2} \cdots y'_{1t-p}]'$ and $x_{1t} \equiv [y'_{2t-1}\ y'_{2t-2} \cdots y'_{2t-p}]'$, then (5.9) can be written as follows:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} A'_1 & A'_2 \\ B'_1 & B'_2 \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{5.31}$$

where c_1 and c_2 are $(n_1 \times 1)$ and $(n_2 \times 1)$ vector of constants respectively. A_1, A_2, A_3, and A_4 are matrices of coefficients. If the lagged variables of y_2 help to predict y_{1t}, the restrictions that $A_2 = 0$ can be imposed. If this restriction is true, then y_1 is called block-exogenous with respect to y_2. By grouping y_{1t} and y_{2t}, the log likelihood function is written as follows:

$$\ell(\theta) = \sum_{t=1}^{T} l_{1t} + \sum_{t=1}^{T} l_{2t} \tag{5.32}$$

$$\begin{aligned} l_{1t} &= -\frac{n_1}{2}\log(2\Pi) - \frac{1}{2}\log|\Sigma_{11}| - \\ &\quad \frac{1}{2}(y_{1t} - c1 - A'_1 x_{1t} - A'_2 x_{2t})'\hat{\Sigma}^{-1}(y_{1t} - c1 - A'_1 x_{1t} - A'_2 x_{2t}) \end{aligned} \tag{5.33}$$

$$\begin{aligned} l_{2t} &= -\frac{n_2}{2}\log(2\Pi) - \frac{1}{2}\log|H| - \frac{1}{2}(y_{2t} - d - D'_0 y_{1t} - D'_1 x_{1t} - D'_2 x_{2t})' \\ &\quad H'(y_{2t} - d - D'_0 y_{1t} - D'_1 x_{1t} - D'_2 x_{2t}) \end{aligned} \tag{5.34}$$

where $H = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$, $d = c_2 - \Sigma_{21}\Sigma_{11}^{-1}c_1$, $D'_0 = \Sigma_{21}\Sigma_{11}^{-1}$, $D_1 = B'_1 - \Sigma_{21}\Sigma_{11}^{-1}A_1$, and $D_2 = B'_2 - \Sigma_{21}\Sigma_{11}^{-1}A'_2$. Now, the log likelihood function (5.32) is maximized with respect to c_1, A_1, A_2, d, D_0, D_1 D_2, Σ_{11} and H and transformed back to c_1, c_2, A_1, A_2, B_1, B_2, Σ_{11} Σ_{12}, and Σ_{22}. Note that (c_1, A_1, A_2, and Σ_{11}) and (d, D_0, D_1 D_2, and H) appear in l_{1t} and l_{2t} only respectively. Therefore, the OLS regression of y_{1t} on a constant, x_{1t} and x_{2t}, can be used to obtain ML estimators of c_1, A_1, A_2, and Σ_{11}. $\hat{\Sigma}_{11}$ is a sample variance-covariance matrix of residuals from these regressions. To obtain d, D_0, D_1 D_2, and H, y_{2t} can be regressed on a constant, y_{1t}, x_{1t} and x_{2t}. It is important to note that the residuals from the second set of regressions, $\hat{\nu}_{2t} \equiv y_{2t} - \hat{d} - \hat{D}'_0 y_{1t} - \hat{D}'_1 x_{1t} - \hat{D}'_2 x_{2t}$ are uncorrelated with the residuals from the first set of regressions, $\hat{\varepsilon}_{1t} \equiv y_{1t} - \hat{c}_1 - \hat{A}'_1 x_{1t} - \hat{A}'_2 x_{2t}$. Again, consider the case of $A_2 = 0$, block-exogeneity of y_{1t}. If $A_2 = 0$, then l_{1t} is:

$$l_{1t} = -\frac{n_1}{2}\log(2\Pi) - \frac{1}{2}\log|\Sigma_{11}| - \frac{1}{2}(y_{1t} - c1 - A'_1 x_{1t})'\hat{\Sigma}^{-1}(y_{1t} - c1 - A'_1 x_{1t}) \tag{5.35}$$

Therefore, ML estimators, $\hat{c}_1$ and $\hat{A}_1$ are obtained by using OLS regression of y_{1t} on a constant and x_{1t}, its own lagged terms. $\hat{d}$, $\hat{D}_0'$, $\hat{D}_1$, $\hat{D}_2$ and $\hat{H}$ are obtained by regressing y_{2t} on a constant, y_{1t}, x_{1t} and x_{2t}. Now, notice that $\hat{B}_2' = \hat{D}_2'$, $\hat{B}_1' = \hat{D}_1' + \hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{A}_1'$, $\hat{c}_2 = \hat{d} + \hat{\Sigma}_{21}\hat{\Sigma}_{11}^{-1}\hat{c}_1$. The likelihood ratio test is used to test the null hypothesis that $A_2 = 0$, again, using the statistic:

$$2[\ell(\hat{\theta}) - \ell(\hat{\theta}_0)] = T[\log\hat{\Sigma}_{11} - \log\hat{\Sigma}_{11,0}] \tag{5.36}$$

This will asymptotically follow a χ^2 distribution with $n_1 n_2 P$ degrees of freedom. Another way of testing a dependency between y_{1t} and y_{2t} is Geweke's measure of linear dependency. For the details, see Geweke (1982) and Hamilton (1994). If the restrictions on the coefficients can not be described as block-recursive form, then the SUR method can be applied to the VAR.

5.5 Error Correction Model (ECM)

As the next chapter reveals, the variables in the data set have a unit root. That is, the variables in the data set are integrated of order 1. When a series of variables has a unit root[5], VAR representation in level (5.5) is no longer appropriate. A variable y_t is called integrated of order d, written as $I(d)$, $d = 1, 2, \cdots$, if $\Delta^d y_t$ is $I(0)$. Δ^d is the d-th difference and $I(0)$ variable is stationary. When the $(n \times 1)$ vector y_t has a series containing a unit root, y_t is called cointegrated if some linear combination of the individual elements of y_t, $\beta' y_t$, is stationary. In other words, if $y_t \sim I(1)$ and there exists some vector β' such that $\beta' y_t \sim I(0)$, then y_t is called cointegrated.[6] β is called the cointegrating vector. The cointegrating rank is the number of linearly independent cointegrating relations and the space spanned by the cointegrating relations is called the cointegrating space.[7] When y_t is $I(1)$ or contains some non-stationary series, the traditional methodology is to take the first difference, Δy_t. However, developments in the non-stationary time series area have shown that it is not correct to fit a vector autoregression to the differenced data if y_t is cointegrated. When y_t is cointegrated, a VAR in level can be still used with some modification while the VAR presentation in level (5.5) is not appropriate. It is the error correction model that will be used for the variables which are cointegrated. Here is a brief review of the error correction model that will apply to the data set in later chapters. Error correction representation is derived from the fact that VAR representation (5.5) can be written as:

$$y_t = \eta_1 \Delta y_{t-1} + \eta_2 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + c + \rho y_{t-1} + \varepsilon_t \tag{5.37}$$

[5] The series could have more than one unit root. However the data indicate that none of the series contains more than one unit root.

[6] For y_t to be cointegrated, it is not required that all components of y_t are $I(1)$. Some components can be $I(0)$. Only $I(1)$ variables are considered since the data set contains only $I(1)$ variables.

[7] More formally, let y_t be $I(d)$. If $\beta \neq 0$ is found such that $\beta' y_t$ is $I(d-b)$, then y_t is called cointegrated $CI(d-b)$. β is called the cointegrating vector. In this case, $b = d = 1$.

where

$$y_t = \text{an } (n \times 1) \text{ vector},$$

$$\rho = \Phi_1 + \cdots + \Phi_p,$$

$$\eta_s = -[\Phi_{s+1} + \Phi_{s+2} \cdots + \Phi_p] \qquad \text{for } s = 1, 2, \cdots, p-1.$$

Subtracting y_{t-1} from both sides of (5.37), yields:

$$\Delta y_t = \eta_1 \Delta y_{t-1} + \eta_2 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + c + \eta_0 y_{t-1} + \varepsilon_t \tag{5.38}$$

where

$$\eta_0 \equiv \rho - I \equiv -[\Phi_1 + \cdots + \Phi_p - I] \equiv -\Pi(1),$$

$$\Pi(z) \equiv I - \Phi_1 z - \cdots - \Phi_p z^p.$$

If y_t has h cointegrating relations, then:

$$\Pi(1) = \alpha\beta' \tag{5.39}$$

where β' is the $(h \times n)$ matrix and α is the $(n \times h)$ matrix and each row of β', b_i', is called a cointegrating vector. Therefore, $z_t \equiv \beta' y_t$ is a stationary $(n \times 1)$ vector. Hence, (5.38) can be written as:

$$\Delta y_t = \eta_1 \Delta y_{t-1} + \eta_2 \Delta y_{t-1} + \cdots + \eta_{p-1} \Delta y_{t-p+1} + c - \alpha\beta' y_{t-1} + \varepsilon_t \tag{5.40}$$

The equations (5.38) and (5.40) are called an error correction representation. Note that all terms in (5.38) and (5.40) are stationary because all the first differenced terms and $\beta' y_{t-1}$ are stationary. If y_t is not cointegrated, then $\Pi(1) = 0$ and (5.38) becomes VAR representation in difference. When the error correction model is fitted to the data, the first thing that should be done is to determine the number of cointegrating relations among the variables, i.e., determine the rank of Π. Once the rank of Π is found, the long-run coefficient matrix, β and adjustment matrix, α can be identified. As noted, the matrix β is interpreted as the long-run relation that holds among the variables and the matrix α is interpreted as the speed of adjustment back to the long-run equilibrium once the variables deviate away from the long-run relation. Note that matrices α and β are not uniquely determined. To determine the rank of Π, there are two ways of testing the number of cointegrating relations; the trace test and the likelihood ratio test. First, consider the following hypotheses:

H_0: Exactly h cointegrating relations among the variables exist,

H_A: There are n cointegrating relations where n is the number of elements of y_t.

Before writing out the maximum likelihood function, the following two auxiliary regressions must be considered:

$$\Delta y_t = \hat{c}_0 + \hat{\Pi}_1 \Delta y_{t-1} + \cdots + \hat{\Pi}_{p-1} \Delta y_{t-p+1} + \hat{u}_{0t} \tag{5.41}$$

$\hat{u}_{0t}$ is an $(n \times 1)$ vector of OLS residual from the above regression (5.41). The other regression is:

$$y_{t-1} = \hat{c}_1 + \hat{\chi}_1 \Delta y_{t-1} + \cdots + \hat{\chi}_{p-1} \Delta y_{t-p+1} + \hat{u}_{1t} \tag{5.42}$$

$\hat{v}_t$ is an $(n \times 1)$ vector of OLS residual from the above regression (5.42). We define $\hat{S}_{ij} \equiv \frac{1}{T} \sum_{t=1}^{T} \hat{u}_{it} \hat{u}'_{jt}$ where $i, j = 0, 1$. The ith eigenvalue $\hat{\lambda}_i$ is obtained from the following equation, constructed by using the two residuals from the above auxiliary regressions:

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0 \tag{5.43}$$

Then, the log maximum likelihood function under H_0 is written as:

$$\ell_0^* = -(\frac{Tn}{2})\log(2\Pi) - (\frac{Tn}{2}) - (\frac{T}{2})\log|\hat{S}_{00}| - (\frac{T}{2}) \sum_{i=1}^{h} \log(1 - \hat{\lambda}_i) \tag{5.44}$$

Under H_A, the log maximum likelihood function is:

$$\ell_A^* = -(\frac{Tn}{2})\log(2\Pi) - (\frac{Tn}{2}) - (\frac{T}{2})\log|\hat{S}_{00}| - (\frac{T}{2}) \sum_{i=1}^{n} \log(1 - \hat{\lambda}_i) \tag{5.45}$$

Hence, the likelihood ratio test of H_0 against H_A is computed by:

$$2(\ell_0^* - \ell_A^*) = -(\frac{T}{2}) \sum_{i=h+1}^{n} \log(1 - \hat{\lambda}_i) \tag{5.46}$$

This is called the trace test statistic. Usually the trace test is used to determine the maximum number of the cointegrating relations among the variables. Another test, so-called likelihood ratio test, uses the following hypotheses:

H_0: h cointegrating relations exist among the variables,

H_A: $h + 1$ cointegrating relations exist.

For this hypothesis, the log likelihood ratio can be written as follows:

$$2(\ell_0^* - \ell_A^*) = -T\log(1 - \hat{\lambda}_{h+1}) \tag{5.47}$$

When these statistics are smaller than the critical values, the hypotheses will be accepted. If the statistics are larger than the critical values, the hypotheses will be rejected. Later chapters will demonstrate how to perform these two tests. It should be noted that these two statistics do not follow a standard distribution. Note, also, that these statistics are very sensitive with the estimated model, i.e., inclusion of a constant term or inclusion of a time trend. The tables for the critical values are available in, for instance, Table B.10 and B.11 in Hamilton (1994). Suppose the rank of Π is determined to be h, that is, h cointegrating relations among the variables. Once the rank of Π is determined, the two matrices α $(n \times h)$ and β $(n \times h)$ such that $\alpha\beta' = \Pi$ can be found. It is obvious that α and β are not uniquely determined. If restrictions are imposed on β, typically normalization of one of the elements in β, then, β and α corresponding to such a β can be obtained. In economics, β is interpreted as the long-run relation to hold among the variables and $\beta' y_t$ is considered to be the deviation from the long-run relations. On the other hand, α is the speed of the adjustment back to the long-run relation once the variables deviate from the long-run relation. The model can be written out as in (5.40). When the system includes many variables, it becomes difficult to model all the variables in the full systems, especially if the number of the parameters being estimated increases rapidly. One way to avoid this difficulty is to introduce the partial system model, where some of the variables, called endogenous variables, are modeled conditioned on the other variables, called exogenous variables. In the partial system, the latter is considered as strongly or, at least, weakly exogenous for the parameters of interest. The advantage of this method is that the dimensions in the system may be reduced without causing any loss of information. Of course, there is always a risk of imposing the wrong exogeneity assumptions in setting up the partial systems.

5.5.1 Weak Exogeneity

The concept of exogeneity is developed in detail in Richard (1980), Engle, Hendry and Richard (1983) and Hendry (1995). Hendry describes the exogeneity issues, in comparison with the causality issues, as follows:

> *Causality issues arise when marginalizing with respect to variables and their lags. Exogeneity issues arise when seeking to analyse a subset of the variables given the behaviour of the remaining variables.*

Exogeneity issues arise when an attempt is made to model some variables, given the information of the other variables. There are three different concepts of exogeneity; weak exogeneity, strong exogeneity and super exogeneity. To construct the partial system model, only the concept of weak exogeneity is required, so it is the only one reviewed here. Consider the joint density at time t for $y_t = (x_t, z_t)'$ conditional on $Y_{t-1} = (Y_0, y_1, \cdots, y_{t-1})$:

$$D_Y(y_t|Y_{t-1}, \theta) \equiv D_Y(x_t, z_t|Y_{t-1}, \theta) \tag{5.48}$$

where $\theta = (\theta_1, \cdots, \theta_n)'$. θ is n parameters in the joint density. Suppose that a one-to-one transformation f from the original n parameters θ to any new subset of n parameters $\lambda \in \Lambda$ exists:

$$\lambda = f(\theta) \tag{5.49}$$

where $\theta \in \Theta$ and $\lambda \in \Lambda$. Let $\lambda = (\lambda_1, \lambda_2)$ be partitioned, such that λ_i $(n_i \times 1)$, where $n_1 + n_2 = n$, corresponds to the factorization of the joint density into a conditional density and a marginal density:

$$D_Y(x_t, z_t | Y_{t-1}, \theta) = D_{x|z}(x_t | z_t, Y_{t-1}, \lambda_1) D_z(z_t | Y_{t-1}, \lambda_2) \tag{5.50}$$

Note that the number of the parameters in the factorization equals the number of the original parameters. The factorization can always be achieved if λ_1 and λ_2 are defined to support it. Suppose that the joint density under analysis involves a subset, ψ $(k \times 1)$, of the parameters λ where $k \leq n_1$ parameters of interest. For z_t to be weakly exogenous, the parameters of interest ψ must be a function of λ_1 only:

$$\psi = g(\lambda_1) \tag{5.51}$$

λ_2 can not provide any information on the parameters of interest ψ. It also requires that λ_1 does not depend on λ_2 so that λ_2 can not be even indirectly informative to learn about ψ:

$$(\lambda_1, \lambda_2) \in (\Lambda_1 \times \Lambda_2) \tag{5.52}$$

That is, $(\lambda_1,\ \lambda_2)$ are variation free. Hence, the parameters of interest ψ might be learned from the *conditional* density but not from the marginal density. When the above two requirements are met, z_t is called weakly exogenous with respect to the parameters of interest. It is noted, as Urbain (1988) pointed out, that the above definition of weak exogeneity does not exclude relation between lagged x_t and z_t. Now to reconsider the equation (5.40). Johansen (1988, 1991a, 1991b) and Johansen and Juselius (1992) developed maximum likelihood method in the full system model. Following the maximum likelihood framework, Johansen (1992) and Urbain (1993) developed a test for weak exogeneity. It turned out that testing restrictions on the matrix α provided a test for weak exogeneity if the parameters of interest are only the long-run parameters. Testing exclusion of the row of the matrix α indicates the weak exogeneity of the corresponding variables. For instance, if the l-th row of the matrix α is 0, then it will be concluded that the corresponding l-th variable in y_t can be treated as a weakly exogenous variable. Urbain also noted that even if there is interest in the short-run parameters, the above procedure may be sufficient for the rejection of weak exogeneity. In his paper, Urbain also discusses a test for weak exogeneity when our parameters of interest are both long-run and short-run parameters. In this part, the parameters of interest are the long-run parameters only, so restrictions are simply imposed on the matrix α.

5.5.2 Partial System Model

After identifying weakly exogenous variables, the full system model (5.40) is reformulated into the partial system model. Suppose that y_t $(n \times 1)$ is partitioned into x_t $(n_x \times 1)$ and z_t $(n_z \times 1)$, where

$n_x + n_z = n$. x_t denotes the endogenous variables and z_t the weakly exogenous variables. The model (5.40) is rewritten as:

$$\Delta y_t = \begin{pmatrix} c^x \\ c^z \end{pmatrix} + \sum_{i=1}^{p-1} \begin{pmatrix} \eta_i^x \\ \eta_i^z \end{pmatrix} \Delta y_{t-i} + \begin{pmatrix} \alpha^x \beta' \\ \alpha^z \beta' \end{pmatrix} y_{t-1} + \begin{pmatrix} \varepsilon_t^x \\ \varepsilon_t^z \end{pmatrix} \tag{5.53}$$

When z_t is weakly exogenous, $\alpha^z = 0$ and the equation (5.53) can be written as:

$$\Delta y_t = \begin{pmatrix} c^x \\ c^z \end{pmatrix} + \sum_{i=1}^{p-1} \begin{pmatrix} \eta_i^x \\ \eta_i^z \end{pmatrix} \Delta y_{t-i} + \begin{pmatrix} \alpha^x \beta' \\ 0 \end{pmatrix} y_{t-1} + \begin{pmatrix} \varepsilon_t^x \\ \varepsilon_t^z \end{pmatrix} \tag{5.54}$$

The Gaussian error terms, ε^x and ε^z have marginal variances Σ_{xx}, Σ_{zz} and covariance Σ_{xz}. The partial model is then given by the model for Δx_t conditional on Δz_t and the past:

$$\Delta x_t = c^{x \cdot z} + \alpha^x \beta' y_{t-1} + \sum_{i=1}^{p-1} \eta_i^{x \cdot z} \Delta y_{t-i} + \omega \Delta z_t + \varepsilon_t^{x \cdot z} \tag{5.55}$$

and the marginal model is given by:

$$\Delta z_t = c^z + \sum_{i=1}^{p-1} \eta_i^z \Delta y_{t-i} + \varepsilon_t^z \tag{5.56}$$

$\varepsilon_t^{x \cdot z}$ and ε_t^z are independently mean zero and Gaussian-distributed with variances $\Sigma_{xx \cdot z} = \Sigma_{xx} - \Sigma_{xz}\Sigma_{zz}^{-1}\Sigma_{zx}$ and Σ_{zz}. $\omega = \Sigma_{xz}\Sigma_{zz}^{-1}$, $\eta_i^{x \cdot z} = \eta_i^x \omega \eta_i^z$ and $c^{x \cdot z} = c^x - \omega c^z$. Johansen (1995) shows that the maximum likelihood estimators of β and α^x can be calculated from the conditional model. It is not necessary to find the rank of β in the partial system model. As Johansen points out, in general, it is advisable to determine the rank in the full system since the asymptotic analysis becomes much simpler. Many researchers use the results for the rank obtained from the full system model (Johansen (1992), Urbain (1993)). Harboe, Johansen and Hansen (1995) developed the test for the rank in the partial system model. As the testing distribution is very complicated, depending on the nuisance parameters, the test for the rank of β in the partial system model is not discussed here. Readers are referred to their paper. As reported in the next chapter, both results for testing the rank in the full system and in the partial system are the same with the data used here.

Please Note

Page(s) not included with original material and unavailable from author or university. Filmed as received.

38

newly constructed data. The mean, standard deviations, skewness[3], and excess kurtosis[4] for each time series are found in the table. Positive (negative) skewness indicates that the distribution is skewed to the left (right). If skewness is zero, the distribution is symmetric about its mean. The distribution with excess kurtosis greater than 0 has more mass in the tails than a Gaussian distribution with the same variance.

Table 6.1 Data Summary

Variables[a]	Mean	Standard Error	Skewness	Excess Kurtosis
EG	-0.56	0.17	-0.98	0.44*
EJ	-5.09	0.21	0.14*	-1.37
MG	21.95	0.27	0.63	-0.81*
MJ	27.57	0.17	0.44*	-1.30
MUS	22.76	0.16	0.58	-0.83*
GG	23.73	0.15	0.37*	-1.08*
GJ	28.97	0.09	0.33*	-1.47
GUS	24.59	0.13	-0.06*	-1.28

[a] All the variables are in a logarithm. G for Germany, J for Japan and US for the United States.

The asterisk in the table indicates that the statistic is not significantly different from zero. For example, the skewness of Japanese exchange rate is not significantly different from zero (0.14*). It means the distribution of Japanese exchange rate is considered to be symmetric. Thus, most of the series show evidence that the distributions are symmetric and that half of the variables have normal tails.

6.2 Unit Root Tests

Many empirical researchers have found that some macroeconomic variables are integrated of order one or more. Some economic time series have one or more unit roots.[5] When a time series has one or more unit roots, characteristics of the series may be different from those of the stationary series. Exogenous shocks to the non-stationary variable tend to last longer than an exogenous shock to the stationary variables. This fact, with the presence of unit root, makes traditional estimation

[3] The skewness is calculated by:

$$s_k = \frac{T^2}{(T-1)(T-2)} \frac{m_3}{s^3}$$

where $m_k = \frac{1}{T}\sum_{t=1}^{T}(y_t - \bar{y})^k, k = 1, 2, \cdots$, and $\bar{y} = \sum_{t=1}^{T} y_t$.

[4] The kurtosis is estimated by:

$$k_u = \frac{T^2}{(T-1)(T-2)(T-3)} \frac{(T+1)m_4 - 3(T-1)m_2^2}{s^4}.$$

[5] See Nelson and Plosser (1982) for instance.

methods inappropriate. That is, when there is more than one non-stationary variable in the system, the conventional VAR methodology will no longer be appropriate. Hence, searching for unit root(s) is an important step before deciding the estimation methods. In this section, unit root tests will be briefly reviewed and the three main unit root tests will be discussed, those that will later be applied to the data set to examine the existence of a unit root or unit roots in the time series.

There are three main tests for unit roots: the Dickey-Fuller test (the DF test), the Augmented Dickey-Fuller test (the ADF test) and the Phillips and Perron test (the PP test). The Dickey-Fuller t-test is the simplest test among these tests. Here, in the empirical work, the Augmented Dickey-Fuller test will be applied. In performing unit root tests, it is important to keep in mind what the true model and the estimated model are. Suppose that the data are generated by a random walk model. The true model is assumed to be a random walk model:

$$y_t = y_{t-1} + \varepsilon_t \tag{6.1}$$

An AR(1) model without an intercept term, however, is considered as the estimated model:

$$y_t = \rho y_{t-1} + \varepsilon_t \tag{6.2}$$

where ε_t is i.i.d. with mean zero and variance σ^2. The ρ in (6.2) is estimated by using an OLS estimation. Then, an OLS estimate, $\hat{\rho}_T$, is calculated by:

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_t^2} \tag{6.3}$$

The t-statistic is constructed by using this OLS estimate, $\hat{\rho}_T$, as follows:

$$t_T = \frac{(\hat{\rho}_T - 1)}{\hat{\sigma}_{\hat{\rho}_T}} = \frac{(\hat{\rho}_T - 1)}{\{s_T^2 \div \sum_{t=1}^{T} y_t^2\}^{1/2}} \tag{6.4}$$

where $\hat{\sigma}_{\hat{\rho}_T}$ is the usual OLS standard error for the estimated coefficient $\hat{\rho}_T$ and s_t^2 is the OLS estimate of the residual variance. Although the t-statistic t_T in (6.4) is constructed in the normal way, t_T has the following limiting distribution:

$$t_T \longrightarrow \frac{(1/2)\sigma^2\{[W(1)]^2 - 1\}}{\{\sigma^2 \int_0^1 [W(r)]^2\, dr\}^{1/2}\{\sigma^2\}^{1/2}} = \frac{(1/2)\{[W(1)]^2 - 1\}}{\{\int_0^1 [W(r)]^2\, dr\}^{1/2}} \tag{6.5}$$

where $W(\cdot)$ is a Wiener process. For the derivations, see Dickey and Fuller (1981) and Hamilton (1994). In other words, t_T no longer follows the ordinary t-distribution. Dickey and Fuller have constructed tables of the critical values by running a Monte Carlo simulation.

In sum, when a model without an intercept term (6.2) is fitted, the t-statistics still can be obtained in an ordinary way but different tables should be used to find the critical values; for instance, table B.6, case 1 in Hamilton (1994).

When a model with an intercept term is fitted, the basic procedure still follows the same steps. The assumption that the true model is a random walk model still holds:

$$y_t = y_{t-1} + \varepsilon_t \tag{6.6}$$

An AR(1) model with an intercept term is used as the estimated model:

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \tag{6.7}$$

The t-statistic constructed as in (6.4) is distributed in the limit as follows[6]:

$$t_T \longrightarrow \frac{(1/2)\{[W(1)]^2 - 1\} - W(1) \cdot \int_0^1 W(r)\,dr}{\{\int_0^1 [W(r)]^2\,dr - [\int_0^1 W(r)\,dr]^2\}^{1/2}} \tag{6.8}$$

An example of the critical values for this case are tabulated in table B.6, case 2 in Hamilton.

The above discussion does not take account of serial correlation in errors or it is assumed that there was no correlation in errors (ε_t is i.i.d.). When there is a possibility of serial correlation in errors, other methods are required. The Phillips-Perron unit root test[7] (the PP test) controls serial correlation by introducing correction terms into the t-statistic. The PP test adds some correction terms to the t-statistics, using the same simple AR(1) models (6.2) and (6.7) as in the DF test.

Assuming that the data are generated by a random walk (6.1), suppose that an AR(1) model with an intercept term (6.7) is fitted:

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t \tag{6.9}$$

Now, ε_t is assumed to be serially correlated and possibly heteroscedastic. If ρ equals 1, the convergence rate, T, ensures that the OLS estimate, $\hat{\rho}_T$, converges in probability to 1, even if ε_t is serially correlated. The t-statistic will be:

$$\begin{aligned} t &= \frac{(\hat{\rho}_T - 1)}{\sigma_{\hat{\rho}_T}} = \frac{T(\hat{\rho}_T - 1)}{\{T^2 \sigma^2_{\hat{\rho}_T}\}^{\frac{1}{2}}} \\ &\longrightarrow \{\frac{(1/2)\{[W(1)]^2 - 1\} - W(1) \cdot \int_0^1 W(r)\,dr}{\int_0^1 [W(r)]^2\,dr - [\int_0^1 W(r)\,dr]^2} + \frac{1}{2} T^2 \frac{\hat{\sigma}^2_{\hat{\rho}_T}}{s_T^2}(\lambda^2 - \gamma_0)\} \\ &\quad \div \{T^2 \hat{\sigma}^2_{\hat{\rho}_T}\}^{\frac{1}{2}} \end{aligned} \tag{6.10}$$

The first term in the first parenthesis is the limiting distribution of $T(\hat{\rho}_T - 1)$, if ε_t is i.i.d. The second term in the parenthesis is the estimate of the correction for serial correlation. It can be shown:

$$T^2 \hat{\sigma}^2_{\hat{\rho}_T} \longrightarrow (\frac{s_T^2}{\lambda^2}) \frac{1}{\int_0^1 [W(r)]^2\,dr - [\int_0^1 W(r)\,dr]^2} \tag{6.11}$$

[6]See Hamilton (1994) for the detailed derivations.

[7]See Phillips (1987) and Phillips and Perron (1988).

and

$$s_T^2 = \frac{1}{T-2}\sum_{t=1}^{T}(y_t - \hat{\alpha}_T - \hat{\rho}_T y_{t-1})^2 \longrightarrow E(\varepsilon_t^2) = \gamma_0 \tag{6.12}$$

Now to construct the modified statistic:

$$(\frac{\gamma_0}{\lambda^2})^{\frac{1}{2}} t_T - \{\frac{1}{2}(\lambda^2 - \gamma_0)/\lambda\} \times \{T\hat{\sigma}_{\hat{\rho}_T} \div s_T\} \tag{6.13}$$

This modified statistic will have the same limiting distribution as (6.8) and the same table can be used for the critical values.

When an AR(1) model is fitted without an intercept term (6.2), provided that the true model is a random walk (6.1), the statistic is obtained by including some correction terms. In other words, the t-value is corrected by using correction terms and consulting a different table of the critical values; see table B6 case 1 in Hamilton, for an example.

The Augmented Dickey-Fuller test has the same purpose as the PP unit root test. It also takes into account a possible serial correlation in errors by including higher-order autoregressive terms.

Suppose that the data are generated by the following AR(p) model:

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)y_t = \varepsilon_t \tag{6.14}$$

where ε_t is i.i.d. with mean zero, variance σ^2, and a finite fourth moment. The equation (6.14) can be also written as:

$$y_t = \rho y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1}\Delta y_{t-p+1} + \varepsilon_t \tag{6.15}$$

where

$$\rho \equiv \phi_1 + \phi_2 + \cdots + \phi_p,$$

and

$$\zeta_s \equiv -[\zeta_{s+1} + \zeta_{s+2} + \cdots + \zeta_p] \qquad \text{for} \qquad s = 1, 2, \cdots, p-1.$$

The advantage of using (6.15) over (6.14) is that only one of the regressors, y_t, is integrated of order one, I(1), while the others, $\Delta y_{t-1}, \cdots, \Delta y_{t-p+1}$, are stationary in (6.15).

Suppose that the process contains a single unit root. Then, the model is estimated using (6.15). Under the null hypothesis that $\alpha = 0$ and $\rho = 1$ in (6.7), the coefficients of Δy_{t-i} for $i = 1, 2, \cdots, p-1$

satisfy:

$$\sqrt{T}\begin{bmatrix} \hat{\zeta}_{1T}-\zeta_1 \\ \hat{\zeta}_{2T}-\zeta_2 \\ \vdots \\ \hat{\zeta}_{p-1T}-\zeta_{p-1} \end{bmatrix} \longrightarrow N\left[\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \sigma^2\begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \vdots & & & \\ \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{pmatrix}\right] \tag{6.16}$$

where $\gamma_j = E[(\Delta y_t)(\Delta y_{t-j})]$. Then, the null hypotheses on the coefficients $(\zeta_1, \zeta_2, \cdots, \zeta_{p-1})$ can be tested by using the standard t and F-statistic asymptotically.

If the null hypothesis is $\rho = 1$, then the limiting distribution of the t-statistic for this null hypothesis is computed as follows:

$$\rho \longrightarrow \frac{(1/2)\{[W(1)]^2 - 1\} - W(1)\cdot\int_0^1 W(r)\,dr}{\{\int_0^1 [W(r)]^2\,dr - [\int_0^1 W(r)\,dr]^2\}^{\frac{1}{2}}} \tag{6.17}$$

Note that this is the same limiting distribution as (6.8) and that the same table will be used to find the critical values, as in the previous case. Since the lagged values of Δy take into account the possible serial correlation in errors, no correction on the t-statistic will be necessary.

It is also interesting to test the joint null hypothesis that $\alpha = 0$ and $\rho = 1$. The F-statistic for this hypothesis can be constructed as:

$$F_T = (b_T - \beta)' R' \{s_T^2 R(\sum x_t x_t') R'\}^{-1} R(b_T - \beta)/2 \tag{6.18}$$

where

$$\begin{aligned} x_t &= [\Delta y_{t-1}, \cdots, \Delta y_{t-p+1}, 1, y_{t-1}]', \\ \beta &= [\zeta_1, \zeta_2, \cdots, \zeta_{p-1}, \alpha, \rho]', \\ R &= [0, I_2], \\ \gamma_T &= \begin{bmatrix} T^{\frac{1}{2}} & 0 \\ 0 & T \end{bmatrix}. \end{aligned}$$

Then the F_T statistic will be compared with the critical values on the table B7 case 2 in Hamilton. For other cases such as the estimated regression with trend and the regression without an intercept term, the t-statistic and F-statistic are formed in the similar fashion and the corresponding tables of the critical values will be applied.

There is another test called ρ test based on (6.10). This test uses the following ρ statistic:

$$\rho = \frac{T\cdot(\hat{\rho}_T - 1)}{(1 - \hat{\delta}_1 - \hat{\delta}_2 - \cdots - \hat{\delta}_{p-1})} \tag{6.19}$$

The ρ-statistic is, in the limit, distributed as:

$$\rho \longrightarrow \frac{(1/2)\{[W(1)]^2 - 1\} - W(1)\cdot\int_0^1 W(r)\,dr}{\int_0^1 [W(r)]^2\,dr - [\int_0^1 W(r)\,dr]^2} \tag{6.20}$$

For this case, the table 5.B case 1 or case 2 will be used depending on whether the estimated model has an intercept.

6.2.1 Practical Procedures

This section discusses the practical procedures that will be followed in applying the Augmented Dickey-Fuller (AFD) unit root test to the data set. Hossain (1995) summarizes sequential procedure in performing the unit root tests as follows. Beginning with the least restrictive model with an intercept term and a time trend:

$$\Delta y_t = \alpha + \beta t + \rho y_{t-1} + \sum_{i=2}^{p} \gamma_i \Delta y_{t+1-i} + \varepsilon_t \tag{6.21}$$

The OLS estimation is used to estimate the model (6.21) and the t_1-statistic is constricted to test the null hypothesis that the time series includes a unit root. The null and alternative hypothesis are written as:

$$\begin{aligned} H_0 &: \rho = 0 \\ H_A &: \rho < 0 \end{aligned} \tag{6.22}$$

Hossain points out that, since unit root tests usually have lower power to reject the null hypothesis, it could be concluded that the series does not contain a unit root if the null hypothesis is rejected. If the null hypothesis is not rejected, then the significance of a time trend in the presence of a unit root must be tested. The null hypothesis is expressed as:

$$\begin{aligned} H_0 &: \rho = \beta = 0 \\ H_A &: \text{not} \quad H_0 \end{aligned} \tag{6.23}$$

The ϕ_1-statistic is used to test the above null hypothesis. If the null hypothesis is not rejected, a regression without a time trend is estimated:

$$\Delta y_t = \alpha + \rho y_{t-1} + \sum_{i=2}^{p} \gamma_i \Delta y_{t+1-i} + \varepsilon_t \tag{6.24}$$

If the null hypothesis is rejected, then the t_1-statistic is compared with the normal by estimating (6.24) with OLS and obtaining the t_2-statistic. The null hypothesis is the same as the null hypothesis in (6.22). If the null hypothesis in not rejected, it must be determined whether or not a constant term is significantly different from zero with a unit root in the variable:

$$\begin{aligned} H_0 &: \rho = \alpha = 0 \\ H_A &: \text{not} \quad H_0 \end{aligned} \tag{6.25}$$

The ϕ_2-statistic is used for this significance test. Finally, if this null hypothesis is not rejected, the regression without a constant term should be reestimated:

$$\Delta y_t = \rho y_{t-1} + \sum_{i=2}^{p} \gamma_i \Delta y_{t+1-i} + \varepsilon_t \tag{6.26}$$

To test the presence of a unit root the t_3-statistic is used.

6.3 Some Empirical Results

This section will examine the empirical results of the unit root tests following the procedure that was outlined above. The unit root tests used a total of 8 variables; 3 variables for Germany and Japan and 2 variables for the U.S. and each as tested for the existence of a unit root in the series.

6.3.1 Germany

The results for the unit root tests on German data are given in Table 6.2. Considering all variables creates 82 observations. First the model is estimated (6.21).[8] The third column of the table gives the t_1-statistic for the coefficient ρ from the regression. The fourth column is the ϕ_1-statistic for the null hypothesis that the series contains a unit root but no time trend (6.23). Then, the second regression (6.24) is used and the t_2 and ϕ_2-statistic are obtained. The ϕ_2-statistic is used to test the null hypothesis that the series includes a unit root but no constant (6.25). Finally, the seventh column provides the t_3-statistic from the regression (6.26).

Table 6.2 Unit Root Test: Germany

Variables	No.of obs.	t_1	ϕ_1	t_2	ϕ_2	t_3
EG	82	-1.90	1.88	-1.94	1.89	-0.61
MG	82	-1.31	1.76	0.70	2.86	2.30
GG	82	-1.71	1.46	0.42	2.49	2.20
Critical value		-3.47	6.58	-2.91	4.76	-1.95

6.3.1.1 Exchange Rate

The plots of German real exchange rate are shown in Figure 6.1. From the upper plot it is observed that German real exchange rate gradually decreased during the first half of the 1980s and reached its bottom around 1985. Since 1985[9], the exchange rate has been increasing. In 1990s, the exchange rate

[8] Lag 4 is chosen as the result of lag length test.

[9] In 1985 the G-7 countries agreed at the Plaza meeting that the U.S. dollar was overvalued and that they should intervene in the market to help the dollar depreciate.

shows continuous ups and downs. In the second plot, the first difference of the exchange rate is given. The first difference series indicates that the series is stationary. The exchange rate is volatile over the sample period. No strong evidence is seen that the exchange rate has become more volatile than it used to be, that is, the exchange rate has been always volatile over the sample period. The t_1-statistic from the regression (6.21) is -1.90, which is smaller than the 5% critical value -3.47[10] in absolute value.[11] The null hypothesis that the series contains a unit root in (6.21) is not rejected. The null hypothesis that the series includes no time trend in the presence of a unit root is also tested. This result, the ϕ_1-statistic, is shown in the fourth column of the table. The ϕ_1 statistic 1.88 is less than the 5% critical value 6.58, which leads to the conclusion not to reject the null hypothesis. The exchange rate does not contain a time trend in the presence of a unit root.

Next, a regression with the time trend (6.24) is run. The t_2-statistic for ρ obtained from the model(6.24) is -1.94. At the 5% significant level, the t_2-statistic is smaller than the critical value -2.91[12] in absolute value. The null hypothesis that the series contains a unit root in the specification of (6.24) will not be rejected. Furthermore, it is necessary to determine whether or not a constant term should be included with a unit root. The null hypothesis is expressed in (6.25). The ϕ_2-statistic is 1.89, which is smaller than 4.76. The conclusion is that the null hypothesis of no constant term should not be rejected. Further testing indicates that the German exchange rate follows a specification of simple random walk (6.1).[13]

6.3.1.2 Money Supply

German real money supply, plotted in Figure 6.2, exhibits its decrease from late 1970s to the mid-1980s and shows a gradual increase during the mid-1980s until recently. Around 1990, it recorded a big drop, possibly explained by political changes in the country.[14] From the regression (6.21), the t_1-statistic (-1.31) for ρ is obtained. The result does not indicate the rejection of the null hypothesis that the series includes a unit root in the specification of (6.21). The ϕ_1-statistic gives some evidence to support the contention that there is no time trend with the presence of a unit root since $\phi_1 = 1.76 < 6.58$. Now to examine the regression (6.24). The regression (6.24) gives the two statistics t_2 and ϕ_2 that imply the null hypothesis that German real money supply contains a unit root without the presence of constant

[10]The sample size is $T = 82$. In Hamilton (1994), when $T = 50$, the 5% critical value is -3.50 and when $T = 100$, it is -3.45. Here, extrapolate to obtain -3.47. All other critical values are obtained in the same way.

[11]We compare the test statistics and the critical values in absolute value.

[12]The critical value is found in B.6 case2 in Hamilton.

[13]The null hypothesis that all lag terms in the first difference are not significant is accepted though not shown in the table.

[14]Of course, German reunification is an important factor for this.

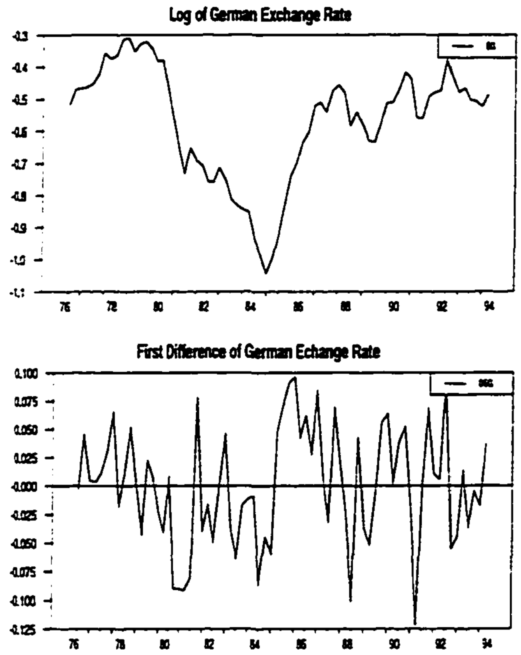

Figure 6.1 German Real Exchange Rate

term since both statistics are smaller than the critical values. Finally, the t_3-statistic from the model (6.26) leads to the rejection of the null hypothesis that German real money supply contains a unit root with the specification of (6.26).

6.3.1.3 GNP

The plots of German GNP are found in Figure 6.3. From the plot in level, it is noted that German GNP decreases during and after the oil crisis (1979-1982). The same information is found in the plot in first difference; otherwise German GNP increased over the sample period. The plot in difference shows that the series is stationary. From the t_1-statistic, there is some evidence that German GNP includes a unit root ($|-1.71| < |-3.47|$). The ϕ_1-statistic also indicates that the null hypothesis of no time trend (6.23) is not rejected since $\phi_1 = 1.46 < 6.58$. The t_2-statistic shows that the hypothesis of a unit root in the specification of (6.24) is accepted ($t_2 = 0.42 < |-2.91|$). Similarly, the ϕ_2-statistic indicates that the hypothesis of no constant in the presence of a unit root should also be accepted. However, from the t_3-statistic, the existence of a unit root is rejected in the specification of (6.26).

In sum, all three German variables have shown the evidence that the variables may contain a unit root.

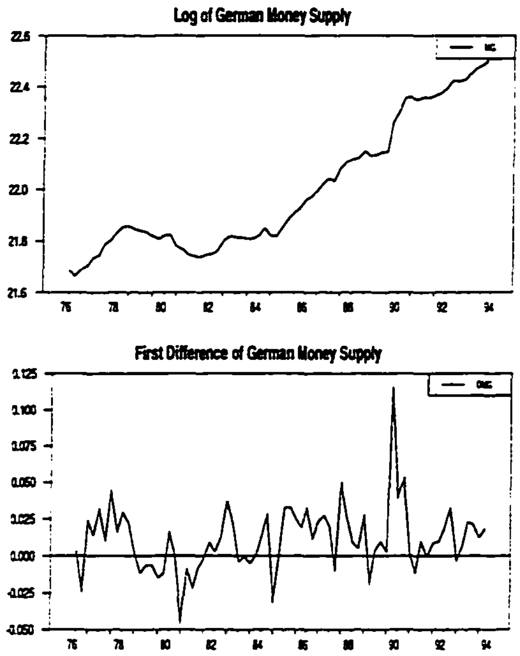

Figure 6.2 German Real Money Supply

6.3.2 Japan

The results for the unit root tests on Japanese variables are given in Table 6.3, again using the same 3 variables; real exchange rate, real money supply and real GNP. The number of observations is also 82 as German variables. Table 6.3 should be read in the same way as Table 6.2.

Table 6.3 Unit Root Test: Japan

Variables	No. of obs.	t_1	ϕ_1	t_2	ϕ_2	t_3
EJ	82	-2.27	2.65	-1.42	1.38	-0.90
MJ	82	-2.34	3.00	-0.06	1.75	1.88
GJ	82	-2.08	4.02	0.59	1.59	1.69
Critical value		-3.47	6.58	-2.91	4.76	-1.95

6.3.2.1 Exchange Rate

The plots of Japanese real exchange rate are given in Figure 6.4. From Figure 6.4, the same tendency in the exchange rate movement as in Figure 6.1 can be observed. The real exchange rate gradually decreased during the first half of the 1980s, while since 1985, the exchange rate has been increasing except for the period of 1988-1990. The plot of the first difference also indicates the stationarity of the

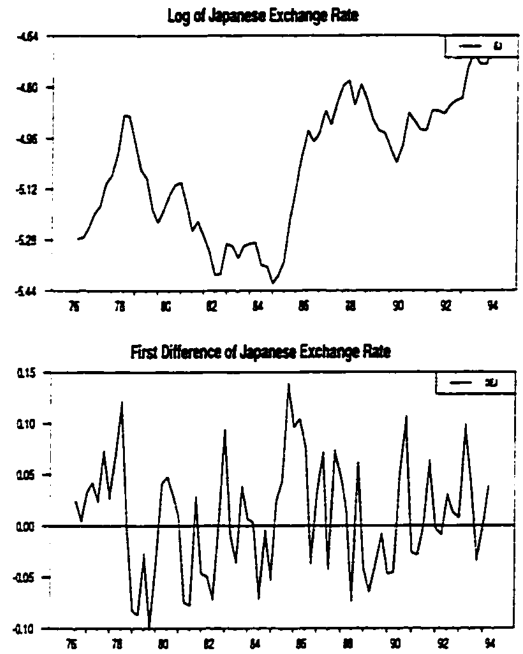

Figure 6.4 Japanese Real Exchange Rate

changes around 1980 and the period of 1989-1991. A relatively large drop in money supply is noted during the oil crisis.

The regression (6.21) gives the t_1-statistic (-2.34) for ρ, thus, the null hypothesis that the Japanese money supply includes a unit root cannot be rejected. The t_2-statistic permits the conclusion that there is no time trend with the presence of a unit root to be made, since $\phi_1 = 3.00 < 6.58$. So, the model without the time trend (6.24) will be used. The t_2 and ϕ_2-statistics support the hypothesis that the money supply contains a unit root without constant term ($\phi_2 = 1.75 < 4.76$), in fact, it can be concluded that the Japanese money supply contains a unit root with the specification of no time trend and no constant (6.26).

6.3.2.3 GNP

Japanese GNP is plotted in both level and difference in Figure 6.6. The plot shows a constant increase throughout the sample years. During and after the oil crisis and after 1992, Japan experienced recessions, when, it is also noted, that the plot of GNP is stationary. From the t_1 and ϕ_1-statistic, it is not possible to reject either hypothesis, (6.22) or (6.23) and the time trend is not included in the model. The t_2-statistic from the regression (6.24) is 0.95, which is smaller than the 5% critical value -2.91 in absolute value and thus precludes the rejection of the hypothesis that the series contains a unit root

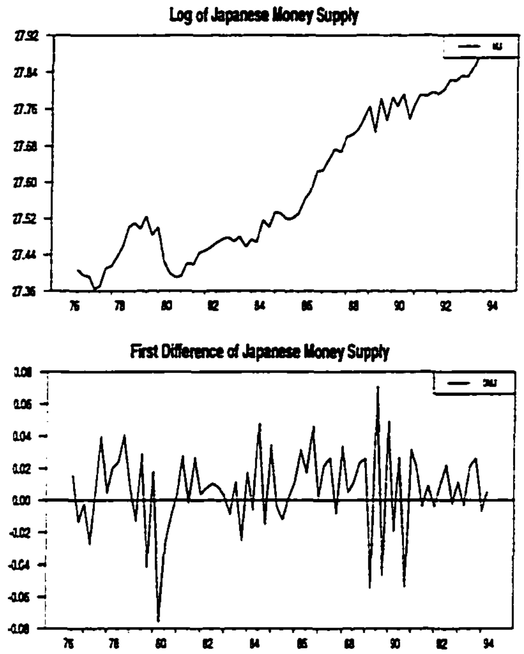

Figure 6.5 Japanese Real Money Supply

under the specification (6.24). The ϕ_2-statistic (1.59) is also smaller than its critical value (4.76). The null hypothesis that no constant term is needed with the presence of a unit root is accepted. Further examination implies that Japanese GNP follows a random walk.

In sum, all three Japanese variables have shown evidence that they contain a unit root.

6.3.3 The United States

Finally, the U.S. variables are examined, as in the previous cases, all U.S. variables include 82 observations. Here, however, only two variables in the U.S. data set; real money supply and GNP are tested. The results of the unit root tests are provided in Table 6.4.

Table 6.4 Unit Root Test: U.S.

Variables	No.of obs.	t_1	ϕ_1	t_2	ϕ_2	t_3
MUS	82	2.10	3.05	0.02	0.85	1.31
GUS	82	-2.41	2.92	-0.59	2.13	1.98
Critical value		-3.47	6.58	-2.91	4.76	-1.95

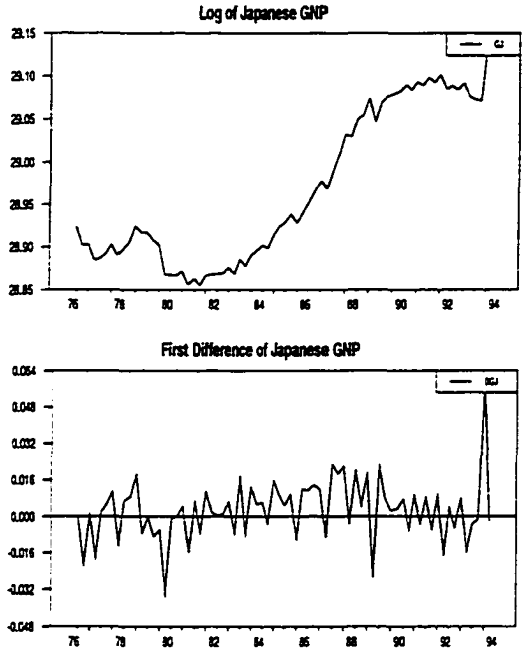

Figure 6.6 Japanese Real GNP

6.3.3.1 Money Supply

The U.S. real money supply is plotted in Figure 6.7. It is observed that the U.S. real money supply has been increasing gradually over the years. The United States experienced a decrease in money supply from 1978 to 1982 and again around 1988. The plot of the first difference indicates the stationarity of the series and that the variability of money supply increased in the middle of 1980s.

The result exhibits some evidence of a unit root since the t_1-statistic (2.10) is smaller than the absolute value of the critical value (-3.47), so the null hypothesis that the series contains a unit root with no time trend ($\phi_1 = 3.05 < 6.58$) can not be rejected. The two statistics, t_2 and ϕ_2, suggest that the series contain no constant with a unit root. In fact, the t_3-statistics indicates that money supply also follows a random walk.

6.3.3.2 GNP

The plots of GNP are given in Figure 6.8, which show that GNP has been increasing since the mid-1970s, except for the oil crisis, the beginning of the Reagan administration and then again around 1991. The second plot exhibits that GNP is a stationary series. It is also noted from the second plot that there was a big drop in GNP during the oil crisis. Because the t_1-statistic (-1.79) is smaller than

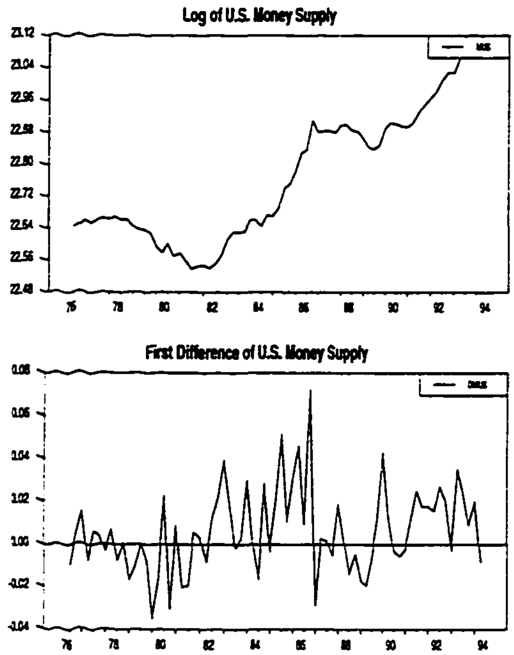

Figure 6.7 U.S. Real Money Supply

the critical value (-3.47), the null hypothesis that the GNP series includes a unit root is not rejected. Similarly, the null hypothesis that the series follows the specification with a unit root and no time tend (6.24) is accepted, since the $\phi_1 = 2.92 < 6.58$. The t_2 and ϕ_2-statistic imply that GNP does not need a constant with the presence of a unit root. However, the unit root test under the specification (6.26) is rejected since the t_3-statistic is larger than its critical value. The comparison of the t_2-statistic with the normal value confirms the unit root under the specification (6.24).

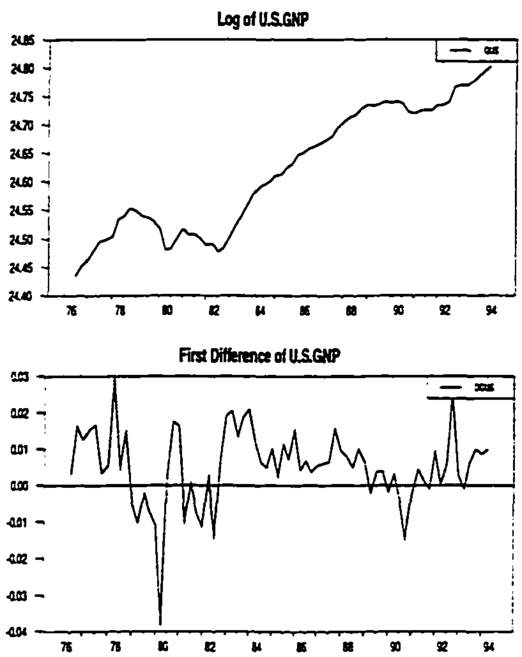

Figure 6.8 U.S. Real GNP

7 EMPIRICAL RESULTS: COINTEGRATION ANALYSIS

In this chapter, the focus will be on long-run relations among the variables created by the empirical results from cointegration analysis, weakly exogeneity and hypothesis testing on long-run relations. In the next chapter, short-run dynamics among the variables will be reported.

The analysis proceeds in the following way. First, the error correction model (5.37), discussed in Chapter 5, will be fitted to the data set. One characteristic of this model is that there is no differentiation between exogenous variables and endogenous variables, all the variables are treated equally at this stage. In other words, the full system model is estimated. Suppose that y_t contains n variables, i.e., y_t is a $(n \times 1)$ vector. The model is written as:

$$\Delta y_t = \eta_1 \Delta y_{t-1} + \eta_2 \Delta y_{t-1} + \cdots + \eta_{k-1} \Delta y_{t-k+1} + c - \Pi y_{t-1} + \varepsilon_t \tag{7.1}$$

After estimating the model (7.1), the trace and likelihood ratio tests will be used to determine the rank of Π, r. As many research papers have found, these tests are very sensitive and the determination of the rank of Π is a very difficult task. When the two tests give two different results, there is no straightforward way to draw conclusions.[1] The trace statistic is calculated by $-T\sum_{i=r+1}^{p} \ln(1-\hat{\lambda}_i)$ and the likelihood ratio statistic is computed by $-T\ln(1-\hat{\lambda}_{r+1})$ where $\hat{\lambda}_i$'s are obtained from the following equations:

$$|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0 \tag{7.2}$$

where $S_{ij} = T^{-1}\sum_{t=1}^{T} R_{it} R_{jt}'$ and $i, j = 0, 1$. R_{0t} and R_{1t} are the residuals obtained by regressing Δy_t and y_{t-1} on the lagged differences $\Delta y_{t-1}, \cdots, \Delta y_{t-k+1}$ and c.

Once the rank of Π is determined, the two matrices α and β can be found such that $\alpha\beta' \equiv \Pi$ where α and β are $(n \times r)$ matrices. Obviously, α and β are not uniquely determined since there always exists a nonsingular matrix Γ such that $\Gamma\Gamma' = I$. Hence, $\alpha\beta' = (\alpha\Gamma)(\beta\Gamma)'$. It is necessary to normalize β by setting one of the elements to one. Then, it is possible to rewrite (7.1) by using α and β as follows:

$$\Delta y_t = \eta_1 \Delta y_{t-1} + \eta_2 \Delta y_{t-1} + \cdots + \eta_{k-1} \Delta y_{t-k+1} + c - \alpha\beta' y_{t-1} + \varepsilon_t \tag{7.3}$$

[1] Many researchers use the trace test to determine the maximum rank of Π. However, this is not always the case. Refer to some examples in Johansen (1995).

After estimating the model (7.3) by imposing a restriction on the rank of Π, the next step is to test for the existence of weak exogenous variables in the system. As in the previous chapter, the weak exogeneity test is performed by imposing restrictions on the matrix, α, since the parameters of interest are long-run parameters only. Then, it is necessary to check whether the error correction term or deviation from long-run relations, $\beta' y_{t-1}$, should be included in each equation in the system. In other words, if the entire row of α is 0, then the error correction term, $\beta' y_{t-1}$, should not be included in the equation corresponding to the row of α. Therefore, it should be concluded that the corresponding variable can be treated as weakly exogenous.

Having determined the weakly exogenous variables, the full system model (7.1) is now reformulated to the partial system model since removing the exogenous variables from the full system model does not cause any loss of information and it reduces dimensions in the system. The only information that is needed about weakly exogenous variables is the marginal information on how the variables are generated. The new system is:

$$\begin{aligned} \Delta x_t &= \omega \Delta z_t - \alpha_1 \beta' y_{t-1} + \sum_{i=1}^{k-1} \eta_{1i} \Delta y_{t-i} + c_1 + \varepsilon_{1t} \\ \Delta z_t &= \sum_{i=1}^{k-1} \eta_{2i} \Delta y_{t-i} + c_2 + \varepsilon_{2t} \end{aligned} \tag{7.4}$$

where $y_t = (x_t, z_t)$ and x_t is a $(n_x \times 1)$ vector of endogenous variables and z_t is a $(n_z \times 1)$ vector of weakly exogenous variables, $n_x + n_z = n$.

Now, to explore long-run relations among the variables in this partial system model, beginning with the rank test in this partial system framework once again, since the distribution of the test statistic has been changed by reformulating the model. Here, new sets of critical values, given in Harboe et al. (1995) will be used with the hope that this process will reduce the number of cointegrations if many cointegrations are found in the full system model.

Hypothesis testing to interpret the long-run relation matrix, β, will help to determine whether the theoretical long-run relationships that were derived in Chapter 4 will be supported by the data; that is the 3 relations in the two-country case (4.17) - (4.19) and the 5 relations in the three-country case (4.45) - (4.49). Recall that, in performing hypothesis testing, the important thing is that one could test on the cointegrating space but not on the cointegrating vectors (Johansen 1988, 1991a).

The final step of the analysis is to analyze short-run dynamics among the variables using the partial system model, which will be discussed in the following chapter.

This chapter analyzes the two-country case; Germany-U.S., and Japan-U.S. case, first, followed by the three-country case: Germany-Japan-U.S. case. In the two-country cases, 5 variables will be used; real exchange rate, home country's real money supply, foreign country's real money supply, home country's real GNP and foreign country's real GNP. For the three country case, 8 variables will be used; 2 real exchange rates, 3 real money supplies (home and two foreign countries) and 3 real GNPs. In all cases, the United States is always the home country.

Analytical procedures to apply to the data are essentially the same for all the cases. As described in the above, first, the number of the cointegrating relations among the variables which are considered to be long-run relationships is determined. Then, the existence of weakly exogenous variables is investigated and, if any exist, the partial system model is reformulated. Based on the partial system model, the cointegrating relations will be interpreted. It is hoped that long-run relations suggested by the data set will be explained by the theoretical model, however, it is a very difficult and sensitive task to determine and interpret long-run relations.

7.1 Other Empirical Researches

Before reporting empirical results, here is a brief review of some of the alternatives empirical research in the field.

Johansen's maximum likelihood method in cointegration framework has become more and more popular since his seminal work (1988). Johansen, and other researchers, illustrate how to use maximum likelihood methodology to estimate the rank of Π and the parameters in α and β using empirical data sets. The readers are referred to Johansen (1988, 1991a,b, 1992, 1995), Johansen and Juselius (1990, 1992), Hansen and Juselius (1995), Hendry (1995), Hatanaka (1996) and Benerjee et al. (1993). For instance, in his book (1995), Johansen uses the Australian and U.S. data to test the PPP and UIP. His data set consists of the quarterly data of log consumer indexes (P^{Au} and $P^{U.S.}$), the exchange rate ($exch$), five-year treasury bond rate in both countries (i^{Au} and $i^{U.S.}$) from 1972:1 to 1991:1. He illustrates the procedure for finding cointegrating relations and formulating simple economic hypotheses in terms of the parameters. First, he fits the data to the model (7.1) with lag of 2. Cointegrating analysis finds two cointegrating relations among the variables.[2] He tests the hypothesis that the interest rate differential is stationary and finds that the likelihood ratio test is significant in χ^2 distribution. He also tests the hypothesis that one equation contains the interest rate differential and the other contains the real exchange rate. The result of this test is not significant.

Later cointegration analysis was combined with the concept of exogeneity. The concept of exogeneity is discussed in detail in Engle et al. (1983) and Hendry (1995). There are several concepts of exogeneity;

[2] In fact, he finds that $P^{U.S.}$ and $i^{U.S.}$ can be treated as weakly exogenous variables.

weak exogeneity, strong exogeneity and super exogeneity. Here, the concept of weak exogeneity is particularly interesting. By introducing weak exogeneity into the model, it is possible to formulate the partial system model, to make inferences on the cointegrating rank in the partial system and estimate β and, finally, to test hypotheses on β. The issue of the partial system is discussed in Urbain (1992, 1993), Johansen (1992) and Hendry (1995). Harboe et al. (1995) demonstrate how difficult it is to determine the cointegrating rank without modeling full system even with the assumption of weak exogeneity. Urbain (1993) applies the partial system model to model Belgium aggregate imports. His data set consists of quarterly time series of import price (pm), domestic price (pd), import volume (m) and real income(y) from 1964:2 to 1990:1. He applies Johansen's procedure to the data, allowing the lag length to vary from 3 to 7. After examining the residuals in each case, he chooses 5 lags. He finds one cointegrating relation among the variables as the result of cointegration analysis. His focus at this stage is to test for the existence of weakly exogenous variables. Since his parameters of interest are long-run parameters only, he performs hypothesis testing on the matrix α in (7.3).[3] Then, he treats import price (pm), domestic price (pd) and real income(y) as weakly exogenous variables. He sets up the partial model, taking into account these weakly exogenous variables.

There are also many papers that attempted to analyze the behavior of the exchange rate using the idea of cointegration analysis. Baillie and McMahon (1989) cite some earlier work. The idea of cointegration was exploited in many papers on PPP. These papers use exchange rates and domestic and foreign prices that are considered to be I(1) process. They apply OLS and Dickey-Fuller methodology to find a single cointegrating relation among the variables (see Baillie and Selover (1987) and Taylor and McMahon (1988)). There are not many papers dealing with exchange rate determination in multiple cointegration framework; but among those who have examined this topic are Dibooglu (1993) and Dibooglu and Enders (1994). Dibooglu (1993) and Dibooglu and Enders (1994) analyze exchange rate determination by using multiple cointegration analysis by applying Johansen's maximum likelihood procedure, variance decomposition and impulse response to the empirical data. In their research, they investigate the two-country cases; the France-U.S. and Italy-U.S. case. The data set consists of money supply differential ($m_t - m_t^*$), price differential ($p_t - p_t^*$), GNP differential ($y_t - y_t^*$), interest rate differential ($r_t - r_t^*$) relative productivity differential ($pr_t - pr_t^*$) and exchange rate (s_t) from 1971:3 to 1990:4. Their theoretical model is based on Dornbusch's dependent economy model. Dornbusch's

[3] Urbain also investigates the case where the parameters of interest are not only long-run parameters but also short-run parameters. He discusses the testing procedure will be more complicated in this case than just performing hypothesis testings on α.

dependent economy model is expressed by the following two equations:

$$m_t - m_t^* = k - k^* + (p_t - p_t^*) + \eta(y_t - y_t^*) - \lambda(r_t - r_t^*) \tag{7.5}$$

$$s_t = (p_t - p_t^*) - (1-\theta)(\rho_t - \rho_t^*) \tag{7.6}$$

where $\rho_t = p_t^N - p_t^T$, expressing the relative price of non-traded goods to traded goods. The readers are referred to the derivations in Chapter 4 in Dibooglu (1993). First, they performed unit root tests on individual variables and confirmed that all the variables have one unit root.

Next, they applied the full system model (7.1) to the data and found two cointegrating relations among the variables for the France-U.S. case and three cointegrating relations for the Italy-U.S. case.[4] Following the Dibooglu-derived version of Dornbusch's dependent economy model, they interpreted these cointegrating vectors as the money market equilibrium and the modified PPP. Although they did not reject the money market equilibrium and the modified PPP when they imposed them individually on each cointegrating vector, they rejected both restrictions imposed on both vectors simultaneously.

Then, they applied Choleski variance decomposition and impulse response function technique to the full system model in order to analyze short-run dynamics of the model. They applied the above techniques to the restricted model, imposing some structures on the long-run parameters and the unrestricted model, which does not impose any restrictions on the long-run parameters other than the rank restriction. The comparison of the two models reveals that there are some changes in the results. The changes indicate that the restricted model explains better than the unrestricted model.

This part was inspired by their work. However, there are some differences between their work and this part. Here, the theoretical model is based on the modified Dornbusch sticky price model. It introduces the assumption that the two countries are large countries, which makes it possible to endogenize the two prices. It also permits the model to be extended to the three-country case, maintaining the large country assumption. While the model is estimated in the partial system framework, it uses the full system model that Dibooglu et al. applied to test for the existence of weakly exogenous variables. The partial system model is used to estimate the parameters and later to perform variance decomposition and impulse response analysis. The following three sections will present the empirical results.

[4] Dibooglu (1993) adds long-run interest rate differential $(i_t - i_t^*)$ to the model for the Italy-U.S. case. Hence, the model contains 7 variables for the Italy-U.S. case while the model includes only 6 variables for the France-U.S. case.

7.2 Cointegration Analysis: Two-Country Case

7.2.1 Germany-U.S.

This section will discuss results from the German and U.S. data. First, the full system model (7.1) is estimated, applying a lag of 2, i.e., $k = 2$, to keep the number of the estimated parameters small. There is no interest in estimates of the parameters at this stage, later, however, the residuals are checked to see if the number of lags in the model is appropriate. Table 7.1 displays the univariate diagnostic

Table 7.1 The Univarariate Diagnostic Statistics: Germany-U.S.

Equations	Mean	Std.Dev	Skewness	Kurtosis	ARCH(2)	Normality	R^2
EG	0.000000	0.045449	-0.238534	2.871676	0.338	0.863	0.202
MG	0.000000	0.017119	0.991425	6.739981	0.833	18.911	0.375
MUS	0.000000	0.015583	0.633006	4.531138	0.437	8.150	0.338
GG	0.000000	0.011646	-0.384460	2.931215	3.164	2.274	0.297
GUS	0.000000	0.008793	-0.128247	3.698648	2.434	3.832	0.330

statistics of the estimated residuals from the 5 equations; EG (German exchange rate) equation, MG (German money supply) equation, MS (U.S. money supply) equation, GG (German GNP) equation and GUS (U.S. GNP) equation. It presents the mean, standard deviation, skewness, and kurtosis of these 5 residuals, where the means of the residuals from all 5 equations are observed to be essentially zero. Most estimates of skewness are close to zero except for the residual from the MG equation. Kurtoses of the residuals from the MG and the MUS are not close to 3, indicating that the distributions of these residuals may have fatter tails than the normal distribution. In the sixth column, ARCH(2), the test statistic for ARCH effects in the residuals, is shown. It follows that χ^2 with 2 degrees of freedom.[5] None of the residuals from the equations are seen to have ARCH effect. No residuals indicate evidences of ARCH effects.[6] The individual normality test is presented in the seventh column. The test statistic follows χ^2 with 2 degrees of freedom (Shenton and Bowman (1977)) and the residuals from the MG and MUS equation show some indication of violation of the normality assumption (18.91 and 8.15).

Table 7.2 introduces the multivariate statistics of the residuals from all the equations. Here, the residual autocorrelations are checked to see if the description of the data is consistent with the assumption of white noise errors. The methods applied here are based on the Gaussian likelihood but the

[5] The ARCH(q) statistic is computed by $(T-k) \times R^2$, where R^2 is from the auxiliary regression:

$$\hat{\varepsilon}_{it}^2 = \gamma_0 + \sum_{j=1}^{q} \gamma_j \hat{\varepsilon}_{it-j}^2 + \eta_{ij}$$

In this case $q = 2$. In general, ARCH(q) statistic follows χ^2 with q degrees of freedom. See Engle(1982) and Enders(1994).

[6] In this case,

H_0: ARCH effect exists. v.s. H_A: No ARCH effect exists.

The critical values are $\chi^2_{2,.1} = 4.61$ and $\chi^2_{2,.05} = 5.99$.

Table 7.2 The Multivariate Diagnostic Statistics: Germany-U.S.

LB(20)	LM(1)	LM(4)	Normality
531.530	13.078	27.416	38.483
0.01	0.98	0.34	0.00

asymptotic properties of the methods only depend on the i.i.d. assumption of the error, so that the violation of the normality assumption is not so serious for the conclusions. The autocorrelation and ARCH effects are of greater concern.

The second row in the table provides the test statistics and the third row presents the corresponding p-values. LB(20) is the Ljung-Box test for residuals to check if the residuals are autocorrelated. This statistic is considered to approximately follow the χ^2 distribution. The LM tests for the first and fourth order autocorrelation are calculated using an auxiliary regression proposed by Godfrey (1988). The fourth column, multivariate normality test, is the sum of 5 univariate tests, based on system residuals.[7]

While the Ljung-Box test indicates that the residuals are autocorrelated ($p-value$=0.01), the LM tests show some evidence that they are not autocorrelated at the first and fourth lag ($p-value$=0.98 and $p-value$=0.34). The normality test rejects the null hypothesis that all residuals are multivariately normally distributed, mainly because the residuals from the MG equation shows a big deviation from the normality. However, this violation is not so serious for the following analysis.

The hypothesis $k = 2$ is also tested in the model with $k = 3$ lags and yields a likelihood ratio test of $LR = (T - kp)\log(|\hat{\Sigma}_2|/|\hat{\Sigma}_3|) = 23.47$.[8] This is asymptotically distributed as χ^2 with 25 degrees of freedom and gives no hint of misspecification.

Next, a cointegration analysis is performed on German and U.S. variables in the full system model. Table 7.3 presents the results of testing the number of cointegrating relations in the full system model. The first column gives eigenvalues obtained from the equation (7.2) and these eigenvalues are arranged in a descending order. The second and third column are the likelihood ratio statistic and the trace statistic. The 90% quantiles corresponding to each statistic are found in the sixth and seventh column. The hypothesis testing is advanced by comparing λ_{max} and $\lambda_{max}(90)$ and λ_{trace} and $\lambda_{trace}(90)$. The λ_{max} statistic is used for the null and alternative hypothesis:

[7] The system residuals are defined as:

$$\hat{u}_t = V\Lambda^{-1}V'\mathrm{diag}(\hat{\sigma}_i^{-1/2})(\hat{\varepsilon}_t - \bar{\varepsilon})$$

where Λ is a diagonal matrix of eigenvalues of the correlation matrix of the residuals and V are the eigenvalues. See more details in the CATS manual (1995).
The test statistic is approximately χ^2-distributed with 10 degrees of freedom.

[8] In general, the likelihood ratio statistic is calculated as $LR = (T - kp - m)\log(|\hat{\Sigma}_2|/|\hat{\Sigma}_3|)$ if the model includes seasonal dummies, where m is the number of seasonal dummies.

H_0: $r = h$ cointegrating relations exist,

H_A: $r = h + 1$ cointegrating relations exist.

If λ_{max} is larger than $\lambda_{max}(90)$, then we reject the null hypothesis. The following null and alternative hypothesis are tested by the trace statistic:

H_0: At most $r = h$ cointegrating relations exist,

H_A: More than $r = h$ cointegrating relations exist.

Many researchers use the λ_{trace} test to determine the maximum number of the cointegrating relations. They perform the λ_{trace} test and determine the upper bound for the number of the cointegrating relations and use the λ_{max} test to confirm or determine the number of cointegrating relations. On the other hand, some other researchers use the above rank tests just for their guideline. They also use some other information such as plots of $\beta' y_t$. It is, in fact, very difficult to determine the rank if the two tests show different results. The rank should be carefully determined in reference to other information as well.[9]

Table 7.3 The Results of Testing Cointegrating Relations: Germany and U.S.

Eigenvalues	λ_{max}	λ_{trace}	$H_0: r = h$	$n - h$	$\lambda_{max}(90)$	$\lambda_{trace}(90)$
0.3080	29.45	71.89	0	5	20.90	64.74
0.2164	19.51	42.44	1	4	17.14	43.84
0.1574	13.70	22.93	2	3	13.39	26.70
0.1020	8.61	9.23	3	2	10.60	13.31
0.0077	0.62	0.62	4	1	2.71	2.71

Here, the following null and alternative hypothesis serves as a start:

H_0: $r = 0$ v.s. H_A: $r > 0$.

To test this hypothesis, the trace statistics are used and The λ_{trace} statistic corresponding to the null hypothesis is 71.89 which is larger than $\lambda_{trace}(90) = 64.74$. This result implies that the null hypothesis should be rejected because there is no cointegrating relation among the 5 variables. So, it is must be concluded that there exists at least one cointegrating relation among the variables. The next formulated hypothesis is:

H_0: $r = 1$ v.s. H_A: $r > 1$.

For this null hypothesis $\lambda_{trace} = 42.44$ is smaller than $\lambda_{trace}(90) = 43.84$, which leads to the acceptance of the null hypothesis. In the λ_{trace} test, there is evidence that there is at most one cointegrating relation among those 5 variables. Now, to perform the λ_{max} test to conduct the following test:

[9] CATS will provide some useful information such as plot of the error correction term, $\beta_i y_{t-1}$. For instance we can check if the i-th error correction term is stable.

$$H_0\colon r = 1 \text{ v.s. } H_A\colon r = 2.$$

The λ_{max} test indicates that the null hypothesis should be rejected against the alternative hypothesis since $\lambda_{max} = 19.51 > \lambda_{max}(90) = 17.14$. Actually, the λ_{max} test leads to the conclusion that there exist 3 cointegrating relations ($\lambda_{max} = 8.61 < \lambda_{max}(90) = 10.60$) and so, it is necessary to choose one cointegrating relation among the variables.[10] After implementing the restriction of one cointegrating relation on Π in (7.1), the error-correction model (7.1) is reestimated. That is, the restriction that the rank of Π in (7.1) is one is imposed and α and β in (7.3) are estimated. The estimated adjustment parameters, $\hat{\alpha}$ and the estimated long-run parameters, $\hat{\beta}$, are shown in Table 7.4. Note that no other restrictions than the number of cointegrating relations have been imposed on the matrix β. Since the number of cointegrating relations is one, α and β are (5×1) column vectors, these estimated column vectors will be called $\hat{\alpha}_1$ and $\hat{\beta}_1$. In general, β is interpreted as a long-run relation among the variables and α is interpreted as the speed of the adjustment toward long-run relations. However, this section will not attempt to examine the long-run relation β, since the focus is on the partial system model, not on the full system model.

Table 7.4 The Estimates of the Adjustment and Long-Run Parameters: $\hat{\alpha}$ and $\hat{\beta}$

Variables	$\hat{\alpha}_1$	t-values for $\hat{\alpha}_1$	$\hat{\beta}_1$
EG	-0.002	-0.071	1.000
MG	0.019	1.482	-1.954
MUS	-0.058	-5.045	0.975
GG	-0.010	-1.101	0.581
GUS	-0.004	-0.613	1.594

Moving on to the partial system model, the existence of weakly exogenous variables is the first issue to be examined. If the parameters of interest are long-run parameters only, the existence of weakly exogenous variables can be tested by imposing restrictions on α. If the i-th row of α is 0, then the i-th equation of the system does not contain the error correction term, βy_{t-1}. The i-th variables can be treated as weakly exogenous. The t-values in the third column of Table 7.4 will give some idea of which variables may be weakly exogenous. German exchange rate (EG), German money supply (MSG), German GNP (GG) and U.S. GNP (GUS) could all be weakly exogenous. Formally, this test will use χ^2-statistics. First, it is necessary to test to see if each row of α is individually 0, that is, to see if individual variables are weakly exogenous. The results of this test show that only the hypothesis that

[10] Researchers often encounter the cases where the two rank tests give two different conclusions. It will be a good idea to investigate several cases and check if the results will change drastically.

the third row of α is 0 is rejected ($\chi^2(1) = 9.92$ and $p-value = 0.00$) as expected. The third variable, U.S. money supply, can not be treated as weakly exogenous; the other 4 variables listed in the above can be individually weakly exogenous. The other 4 variables are tested to see if they can be simultaneously weakly exogenous, the hypothesis tested here is whether or not the 4 rows of α are simultaneously 0. The results that $\chi^2 = 2.45$ and $p-value = 0.65$ imply that the hypothesis can not be rejected. Hence, the other 4 variables, EG, MSG, GG and GUS will be simultaneously treated as weakly exogenous.

Now that the 4 weakly exogenous variables are identified, the full system model is reformulated into the partial system model (7.4). Since there are one endogenous variable and 4 weakly exogenous variables, x_t in (7.4) consists of only one variable and the z_t contains 4 variables. That is, z_t is a (4×1) vector. The rank test is performed in the partial system model, not in the full system model.

Table 7.5 The Results of Testing Cointegrating Relations in the Partial System: Germany and U.S.

Eigenvalues	$Trace$	$H_0 : r = h$	n_z	$n_y - r$	$Trace(90)$
0.2865	27.00	0	4	1	18.1

Table 7.5 presents the result of the rank test in the partial system model. n_z in the fourth column is the number of weakly exogenous variables, here, 4. n_y in the fifth column is the number of endogenous variables, which is 1. The critical value $Trace(90)$ is taken from Harboe et al. (1995). Since $Trace = 27.00 > 18.1 = Trace(90)$, the hypothesis that there is no cointegration is rejected, that is, the existence of one cointegration is accepted since r can not be larger than $n_y = 1$.

Table 7.6 shows the estimates of the long-run relation in the partial system. This long-run relation is only included in the MUS equation.

The order of the variables in the first row of the table has changed to emphasize the fact that only MUS is endogenous and the other 4 variables are being treated as weakly exogenous. The column vector

Table 7.6 The Estimates of Long-Run Parameters in the Partial System: $\hat{\beta}$

Variables	$\hat{\beta}_1$
MUS	0.472
EG	1.000
MG	-0.988
GG	-1.472
GUS	2.489

$\hat{\beta}_1$ indicates that the following relation exists among the 5 variables:

$$0.472MUS + EG - 0.988MG - 1.472GG + 2.489GUS = 0 \tag{7.7}$$

The coefficient of German exchange rate in the above β is normalized. From the above relation, it can be seen that the German real exchange rate is negatively related to the U.S. money supply and, also, that both the German money supply and GNP have positive impacts on exchange rates while the U.S. GNP are positively related to the exchange rate.

Now, to examine the residuals from the partial system model. Attention goes to the i.i.d. assumption, i.e., autocorrelation of the residuals. No indication of autocorrelation is found.[11]

The next task is to interpret the estimated long-run relation in β. The theoretical model predicts the 3 long-run relations among the variables, as was the case in the previous chapter. For convenience, here are these 3 long-run relations again:

$$M - \frac{\beta}{\beta^*}M^* - \alpha Y + \frac{\beta\alpha^*}{\beta^*}Y^* = 0 \tag{7.8}$$

$$E - \frac{\lambda}{\delta\beta}M - (\frac{\phi}{\delta} - \frac{1}{\delta} - \frac{\lambda\alpha}{\delta\beta})Y = 0 \tag{7.9}$$

$$E + \frac{\lambda^*}{\delta^*\beta^*}M^* + (\frac{\phi^*}{\delta^*} - \frac{1}{\delta^*} - \frac{\lambda^*\alpha^*}{\delta^*\beta^*})Y^* = 0 \tag{7.10}$$

Note that, in each relation, the coefficient of the first variable is normalized. All the parameters are assumed to be positive and no other assumptions are made. For example, in the first relation (7.8), the coefficient of M^*, $\frac{\beta}{\beta^*}$, is positive, while it is unknown whether it is greater or lesser than one, depending on the magnitude of β and β^*. Table 7.7 shows the possible signs of the parameters in the relations. The first relation (7.8), the money market relation, does not include exchange rate and describes the relation among money supplies and GNPs. The second relation (7.9) excludes the foreign money supply and GNP and the third relation (7.10) rules out the domestic variables. To more thoroughly examine the empirical long-run relation, restrictions are imposed on the long-run parameters, β, in the model (7.4).

To implement restrictions on β, a restriction matrix R_k is constructed, where all three relations in the above are described by linear restrictions. Using the restriction matrix R_k, the null and alternative hypothesis can be written as:

$$H_0\colon R_k\beta_1 = 0 \qquad k = 1, 2, 3,$$
$$H_A\colon R_k\beta_1 \neq 0 \qquad k = 1, 2, 3.$$

where k implies each theoretical relation.

[11] The results are not shown here. They are similar results to Table 7.1 and Table 7.2.

First, the first relation (7.8) is examined to see if it explains the estimated long-run parameter β_1. Although a predicted pattern in signs is shown in Table 7.5, nonetheless, the restriction of exclusion of the exchange rate is imposed. The restriction matrix R_1 (1×5) is as follows:

$$R_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{7.11}$$

The test statistic follows χ^2 distribution. The results are $\chi^2 = 15.51$ and $p - value = 0.00$ the first relation is rejected. The reestimated long-run parameters with exclusion of exchange rate are shown in Table 7.8.

Similarly, the second (7.9) and third long-run relations (7.10) could be imposed on β_1. Since the second relation (7.9) excludes foreign variables, this relation requires two restrictions; exclusion of foreign money supply and GNP. The matrix R_2 (2×5) will be written as:

$$R_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{7.12}$$

Here, $\chi^2 = 12.98$ and $p - value = 0.00$ and again, the null hypothesis that β_1 satisfies the second long-run relation is rejected. Table 7.9 estimate the long-run parameters with the second restriction. The coefficient of U.S. money supply is negative and this is a correct sign. The sign of U.S. GNP is negative and this was not predicted by the model.

The third relation (7.10) requires only exchange rate, foreign money supply and GNP. The R_3 matrix is as follows:

$$R_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.13}$$

Again, the third relation requires two exclusion restrictions and R_3 is a (2×5) matrix. The third long-run relation (7.10) is rejected because $\chi^2 = 8.41$ and $p - value = 0.01$. Table 7.10 shows the estimated parameters with the third relation. The sign of the coefficient of German money supply is not correct.

Table 7.7 The Possible Signs of Coefficients

Equation	E	M^*	M	Y^*	Y
(7.8)		-	+	+	-
(7.9)	+		-	?	
(7.10)	+	+			?

Table 7.8 The Estimates of the Long-Run Parameters: Exclusion of Exchange Rate

Variables	$\hat{\beta}_1$
MUS	1.000
EG	
MG	0.837
GG	-4.131
GUS	1.146

Table 7.9 The Estimates of the Long-Run Parameters: Exclusion of Foreign Variables

Variables	$\hat{\beta}_1$
MUS	-0.313
EG	1.000
MG	
GG	
GUS	-0.641

None of the above three relations explains the long-run relation by itself. However, it may be possible to interpret that the cointegrating space supports some linear combination of these 3 relations. If this is indeed the case, it can be concluded that the long-run relation is not explained by a single relation listed in the above.

7.2.2 Japan-U.S.

The next two-country case is Japan and U.S.. Here, again, the 5 variables are; the Japanese exchange rate, two money supplies and two GNPs. First, the data is used to estimate the full system model (7.1), again, applying a lag of 2 to keep the model simple. The residuals from the full system model are checked to see if the i.i.d assumption is retained. In Table 7.11, the univariate diagnostic statistics of the residuals from the system is found.

Most estimates of the skewness are close to 0. The kurtoses of the MUS, GJ and GUS equation are slightly away from 3. No residuals show ARCH effects. The residuals from the last three equations also individually violate normality assumption, however, it is not too serious for the analysis of this work.

Table 7.12 presents the multivariate diagnostic statistics. No evidence of autocorrelations among the residuals from the LM test is observed but, since some residuals individually violate the normality assumption, the multivariate normality assumption is not satisfied.

Table 7.10 The Estimates of the Long-Run Parameters: Exclusion of Domestic Variables

Variables	β_1
MUS	
EG	1.000
MG	-0.645
GG	0.193
GUS	

Table 7.11 The Univarariate Diagnostic Statistics: Japan-U.S.

Equations	Mean	Std.Dev	Skewness	Kurtosis	ARCH(2)	Normality	R^2
EJ	0.000000	0.043895	0.269863	3.193700	0.846	1.519	0.305
MJ	0.000000	0.019166	-0.537635	2.982460	3.741	4.951	0.395
MUS	0.000000	0.015927	0.653248	4.694602	0.323	9.020	0.309
GJ	0.000000	0.011456	0.125240	4.407140	1.681	9.606	0.408
GUS	0.000000	0.008522	-0.550216	4.759572	0.052	10.085	0.371

When the hypothesis $k = 2$ in the model with $k = 3$ lags is tested, the likelihood ratio test yields $LR = (T - kp)\log(|\hat{\Sigma}_2|/|\hat{\Sigma}_3|) = 23.81$. This is asymptotically distributed as χ^2 with 25 degrees of freedoms and gives no hint of misspecification.

Table 7.13 is the results of cointegration analysis of the full system model (7.1). This will be read in the same way as Table 7.3.

From Table 7.13, it can be concluded that there are two cointegrations among these 5 variables. The λ_{trace} test demonstrates that 2 cointegrating relations exist since $\lambda_{trace} = 26.66 < 26.70 = \lambda_{trace}(90)$. Although the λ_{max} test rejects 2 cointegrating relations against 3 cointegrating relations, two cointegrating relations is concluded.

Now, implementing the restriction that the rank of Π is 2 on the full system model (7.1) and reestimating the model (7.2) to obtain α and β yields the results shown in Table 7.14.

Since 2 cointegrating relations have been found, α and β are (5×2) matrices. The first three columns are associated with the first column vector of α and β and the second three columns are the second column of α and β. At this stage we are not interested in the estimates of β but in identifying which variables can be treated as weakly exogenous. Again, the parameters of interest are long-run ones only so that we can identify weakly exogenous variables by testing α. By looking at the t-values for α (the third and sixth column of the table) it is suspected that EJ and GUS can be treated as

Table 7.12 The Multivariate Diagnostic Statistics: Japan-U.S.

LB(18)	LM(1)	LM(4)	Normality
513.789	25.269	37.274	39.200
0.02	0.45	0.05	0.00

Table 7.13 The Results of Testing Cointegrating Relations: Japan-U.S.

Eigenvalues	λ_{max}	λ_{trace}	$H_0 : r = h$	$n - r$	$\lambda_{max}(90)$	$\lambda_{trace}(90)$
0.4134	42.68	105.94	0	5	20.90	64.74
0.3671	36.60	63.26	1	4	17.14	43.84
0.1987	17.73	26.66	2	3	13.39	26.70
0.1056	8.93	8.94	3	2	10.60	13.31
0.0001	0.01	0.01	4	1	2.71	2.71

weakly exogenous since both t-values are small. In fact, when checking weak exogeneity one by one, it is obtained evidence that EJ and GUS are weakly exogenous ($\chi^2 = 2.75$, $p - value = 0.25$ for EJ and $\chi^2 = 2.48$, $p - value = 0.29$ for GUS). When simultaneously testing weak exogeneity of these 2 variables, the results are obtained that $\chi^2 = 4.66$, $p - value = 0.32$. Hence both Japanese exchange rate and U.S. GNP will be treated as weakly exogenous variables.

Now, the model is reformulated into the partial system model, taking account of the existence of the two weakly exogenous variables. Since, there are three endogenous and two weakly exogenous variables in the system, the first equation in (7.3) contains 3 equations and the second equation consists of 2 equations.

Performing the rank test in the partial system model gives the results below in Table 7.15. Table 7.15 presents the result of the rank test in the partial system model. Since $Trace = 38.49 > 28.0 = Trace(90)$, the hypothesis that there exists one cointegration is rejected. However $Trace = 4.70 < 13.2 = Trace(90)$ implies that two cointegrations will not be rejected.

Table 7.16 presents the estimates of the long-run relations in the partial system. Since there are two cointegrating relations, β is a (5×2) matrix. Note also that the order of the variables have been changed because only the first 3 variables are treated as endogenous variables. The estimates are normalized by the coefficient of Japanese exchange rate.

To implement restrictions on long-run relations, the three restrictions that were used in the previous section are imposed on the vectors simultaneously. In other words, the three restrictions are tested to see if they will be supported by the cointegration vectors.

Table 7.14 The Estimates of the Adjustment and Long-Run Parameters: $\hat{\alpha}$ and $\hat{\beta}$

Variables	$\hat{\alpha}_1$	t-values for $\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	t-values for $\hat{\alpha}_2$	$\hat{\beta}_2$
EJ	-0.001	-0.446	1.000	-0.096	-2.079	1.000
MJ	0.002	2.753	-7.346	0.069	3.655	-2.514
MUS	0.001	2.012	1.621	-0.068	-4.372	-0.396
GJ	0.002	6.021	-37.009	0.008	0.725	1.381
GUS	-0.000	-1.587	50.273	-0.007	-0.795	1.485

Table 7.15 The Results of Testing Cointegrating Relations in the Partial System: Japan and U.S.

Eigenvalues	$Trace$	$H_0 : r = h$	n_z	$n_y - r$	$Trace(90)$
0.3997	79.32	0	2	3	46.0
0.3445	38.49	1	2	2	28.0
0.0570	4.70	2	2	1	13.2

The first relation (7.8), the money market relation, is imposed on both vectors of β, using the same restriction matrix R_1 in (7.11). The test results are $\chi^2 = 16.78$ and $p - value = 0.00$. Hence, the first relation is not supported by the cointegration vectors β. To test the second relation (7.9), the restriction matrix R_2 in (7.12) is imposed on the vectors of β and shows that $\chi^2 = 26.94$ and $p - value = 0.00$, which indicates that the second relation, the exclusion of foreign variables, is not supported by the cointegration vectors. However, the hypothesis that the vector supports the third relation (7.10) is rejected, because of the exclusion of domestic variables at 5% significance level but not at 1% significance level. The results obtained are $\chi^2 = 11.61$ and $p - value = 0.02$.

Table 7.17 presents the estimates of the long-run parameters with the relation (7.10) implemented. The signs of MJ in both vectors are not consistent with the predicted sign in Table 7.7.

Next, the first relation, (7.8), and the third relation, (7.10), are simultaneously implemented on the two vectors. When the first relation (7.8) is implemented on the first vector β_1 and the third relation (7.10) is implemented on the second vector β_2 the results $\chi^2 = 0.02$ and $p - value = 0.90$ are obtained. These two relations are accepted by the long-run relations β. However, as in Table 7.18, the signs of the coefficients are not as predicted. The signs of GJ and GUS are incorrect in β_1 and the sign of MJ is not correct in β_2.

Table 7.16 The Estimates of Long-Run Parameters in the Partial System: $\hat{\beta}$

Variables	$\hat{\beta}_1$	$\hat{\beta}_2$
MJ	-1.212	-3.433
MUS	-2.800	0.423
GJ	14.178	0.829
EJ	1.000	1.000
GUS	-10.807	2.155

Table 7.17 The Estimates of the Long-Run Parameters: Japan-U.S. imposing the relation (7.10)

Variables	$\hat{\beta}_1$	$\hat{\beta}_2$
MJ	-4.248	-0.462
MUS		
GJ	4.623	2.926
EJ	1.000	1.000
GUS		

When the relations are reversed, i.e., the second relation on β_1 and the first relation on β_2, the hypothesis is again accepted although the sign patterns are not correct.

Other combinations of the relations were attempted, however, in all cases, the hypotheses were not accepted. There was strong evidence of the existence of some relations among the variables and this was not what was expected.

Table 7.18 The Estimates of Long-Run Parameters: Japan-U.S. imposing the relations (7.8) and (7.10)

Variables	$\hat{\beta}_1$	$\hat{\beta}_2$
MJ	1.000	-3.100
MUS	-1.367	
GJ	6.125	2.926
EJ		1.000
GUS	-6.115	

7.3 Cointegration Analysis: Three-Country Case

This section of the chapter reports results for the three-country case, the Germany-Japan-U.S. case. Instead of 5 variables, there are 8 variables included here: two real exchange rates, three money supplies and three GNPs and the data is applied to the full system model (7.1), allowing the same number of lags as in the previous two-country cases, i.e., 2 lags. The only difference from the previous cases is the number of the variables contained in y_t. This increase in the number of variables in y_t brings about some problems. This larger number of variables potentially increases the rank of Π, the number of cointegrating relations. As already seen, the larger the number of existing cointegrating relations becomes, the more difficult it is to interpret the relations. This is actually what is seen in this section. The methodology in this section is the same one that has been applied previously. First, the full system model (7.1) is estimated and checked for residuals, especially for the i.i.d. assumption. Then, the number of the cointegrating relations among the 8 variables is determined and α and β estimate.

Then, the weakly exogenous variables are identified to reduce the dimensionality of the system, which leads to the partial system model (7.4). The focus here is on the long-run relations in the partial system model and the attempt to interpret them by implementing some restrictions. The following chapter will investigate short-run dynamics among the variables, based on the partial system model, for the three-country case as well as the two-country cases.

7.3.1 Germany-Japan-U.S.

This section presents the data from Germany, Japan and U.S. where Germany will be treated as the first foreign country (one asterisk), Japan as the second foreign country (two asterisks) and U.S. as a home country (no asterisk).[12] The variables being used in this section are German exchange rate, Japanese exchange rate and Germany, Japanese and U.S. money supply and GNP.

Table 7.19 displays the univariate diagnostic statistics of the estimated residual from each of the 8 equations after fitting the full system model (7.1), these include the mean, standard deviation, skewness, and kurtosis of those residuals. The means of the residuals from all 8 equations are essentially zero while most of skewnesses are close to zero. Kurtoses of the residuals from most of equations are close to 3 except for the MG and MUS equation, indicating that the distributions of most of the residuals have normal tails. None of the residuals from the system has ARCH effects, which is what was hoped. The individual normality test is presented in the seventh column. The test statistic follows χ^2 with 2 degrees of freedom as previously. The residuals from the MG and MUS equation show some indication of

[12] Recall that asterisks denoted two foreign countries in the theoretical model in Chapter 4.

Table 7.19 The Univarariate Diagnostic Statistics: Germany-Japan-U.S.

Equation	Mean	Std.Dev	Skewness	Kurtosis	ARCH(2)	Normality	R^2
EG	0.000000	0.044648	-0.178278	3.212537	0.589	1.281	0.230
EJ	0.000000	0.040629	0.137878	3.425727	1.731	2.185	0.404
MG	0.000000	0.015784	0.346324	4.664655	0.258	10.892	0.469
MJ	0.000000	0.016543	-0.567240	3.223383	0.223	4.575	0.549
MUS	0.000000	0.015179	0.578441	4.315739	0.211	7.070	0.372
GG	0.000000	0.010839	-0.527354	3.315739	2.368	3.925	0.391
GJ	0.000000	0.010365	-0.109107	3.850945	4.157	4.929	0.515
GUS	0.000000	0.008165	-0.113216	3.604877	0.170	3.223	0.423

Table 7.20 The Multivariate Diagnostic Statistics

LB(20)	LM(1)	LM(4)	Normality
1341.708	60.644	80.000	49.787
0.01	0.60	0.09	0.00

violation of the normality assumption; the statistic for the residual from the MG equation is particularly large (10.89).

Table 7.20 presents the multivariate statistic of residuals from all the equations. The first row provides the test statistics and the second row presents the corresponding p-values. LB(20) is the Ljung-Box test for residuals to check if the residuals are autocorrelated and this statistic is considered to approximately follow the χ^2 distribution. The fourth column, multivariate normality test, is the sum of 8 univariate tests, based on system residuals.[13] While the Ljung-Box test indicates that the residuals are autocorrelated ($p-value = 0.01$), the LM tests show no evidence that they are autocorrelated at the first and fourth lag. The normality test rejects the null hypothesis that all residuals are multivariately normally distributed. This is mainly because the residuals from the MG equation show a deviation from normality. Again, the violation of the normality assumption is not so serious for the rest of the analysis since it relies on the asymptotic i.i.d. assumption.

Once more, it is necessary to test the hypothesis $k = 2$ in the model with $k = 3$ lags and to find likelihood ratio test $LR = (T - kp)\log(|\hat{\Sigma}_2|/|\hat{\Sigma}_3|) = 69.33$. This is asymptotically distributed as χ^2 with 64 degrees of freedoms (83.66) and betrays no hint of misspecification.

Table 7.21 presents the results of testing the number of cointegrating relations among these 8 variables in the full system model (7.1) using the same explanation as in the previous section. First, when the third column λ_{trace} and the seventh column $\lambda_{trace}(90)$ are compared the λ_{trace} test indicates that 3

[13] The test statistic is approximately χ^2-distributed with 16 degrees of freedom.

Table 7.21 The Results of Testing Cointegrating Relations: Germany-Japan-U.S.

Eigenvalues	λ_{max}	λ_{trace}	$H_0: r = h$	$p - r$	$\lambda_{max}(90)$	$\lambda_{trace}(90)$
0.5966	72.62	211.86	0	8	32.26	149.99
0.4215	43.79	139.24	1	7	28.36	117.73
0.3344	32.57	95.45	2	6	24.63	89.37
0.2319	21.11	62.88	3	5	20.90	64.74
0.1880	16.66	41.77	4	4	17.14	43.84
0.1597	13.92	25.12	5	3	13.39	26.70
0.1229	10.49	11.20	6	2	10.60	13.31
0.0087	0.70	0.70	7	1	2.71	2.71

cointegrating relations against 4 cointegrating relations should not be rejected because $\lambda_{trace} = 62.88$ is smaller than the 90% critical value 64.74. The next step is to test the number of cointegrating relations using the λ_{max} statistic. $\lambda_{max} = 21.11$ implies that H_0: 3 cointegrating relations exist is rejected against H_A: 4 cointegrating relations exist. However, $\lambda_{max} = 16.66$ suggests not to reject H_0: 4 cointegrating relations exist against H_A: 5 cointegrating relations exist because $\lambda_{max} = 16.66 < 17.14 = \lambda_{max}(90)$. Again, it is difficult to determine the number of cointegrating relations since the two tests give different results. Here, it can be concluded that there exist 3 cointegrating relations among these 8 variables.[14] Although the cointegrating relations $\beta' X$ can be interpreted as long-run relations in economic sense, if more than one cointegrating relation exists, their interpretations are not necessarily obvious and easy, as seen in the Japan-U.S. case.

Table 7.22 contains the estimates of the matrix β, the estimated long-run parameters. The matrix $\hat{\beta}$ is an (8×3) matrix since 3 cointegrating relations were found in the cointegration analysis. Each column presents a long-run relation among the 8 variables. Note that in all 3 cointegrating relations the coefficient of German exchange rate is normalized, i.e., its coefficient is set to one.

The estimates of the adjustment coefficients, $\hat{\alpha}$, and their associated t-values are found in Table 7.23. The matrix $\hat{\alpha}$ is an (8×3) matrix. The adjustment coefficients in α are interpreted as the speed of moving back to long-run relations once variables move away from the long-run equilibrium. Most of these numbers are small, indicating adjustment speed is slow in the long-run once the system deviates from the long-run equilibria.

The existence of weakly exogenous variables in the system is verified by testing on the rows of α matrix. If the entire row of α is 0, then the corresponding variable will be treated as weakly exogenous.

[14] In comparing results from the 2 cointegration case and the 3 cointegration case, no major changes in the results for the preliminary investigation were found.

Table 7.22 The Estimates of the Long-Run Parameters: β

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
EG	1.000	1.000	1.000
EJ	-1.278	-2.991	-0.779
MG	-1.060	-0.604	1.412
MJ	1.217	4.412	0.406
MUS	0.186	1.432	-2.777
GG	0.248	1.363	-7.472
GJ	0.576	-5.203	2.870
GUS	1.080	0.455	8.456

Table 7.23 The Estimates of the Adjustment Parameters: α

	$\hat{\alpha}_1$	t-values for $\hat{\alpha}_1$	$\hat{\alpha}_2$	t-values for $\hat{\alpha}_2$	$\hat{\alpha}_3$	t-values for $\hat{\alpha}_3$
EG	0.037	0.670	0.005	0.255	0.017	0.559
EJ	0.141	2.712	0.039	2.145	-0.004	-0.143
MG	0.060	3.146	0.002	0.378	0.021	1.957
MJ	-0.128	-6.456	-0.019	-2.837	0.009	0.834
MUS	-0.024	-1.277	0.027	4.145	-0.008	-0.770
GG	0.005	0.342	0.000	0.041	0.030	3.926
GJ	-0.089	-7.198	0.008	1.958	0.015	2.124
GUS	-0.005	-0.517	-0.003	-0.732	-0.013	-2.198

Before performing formal tests, it is suspected that at least German exchange rate (EG) and U.S. GNP (GUS) can be weakly exogenous since all three of their t-values are small. In fact, the formal χ^2 test indicates these two variables can be treated as weakly exogenous ($\chi^2 = 0.67$ and $p-value = 0.88$ for EG and $\chi^2 = 3.39$ and $p-value = 0.33$ for GUS). The third and sixth row of α, corresponding to German money supply (MG) and German GNP (GG), could also be 0 since $\chi^2 = 10.95$ and $p-value = 0.01$ and $\chi^2 = 7.79$ and $p-value = 0.05$. The other variables can not be treated as weakly exogenous.[15] When the four variables, EG, MG, GG and GUS, were tested simultaneously, the hypothesis that all four variables can be weakly exogenous ($\chi^2 = 20.57$ and $p-value = 0.06$) was accepted.

To reformulate the full system model into the partial system model, EG, MG, GG and GUS are used as weakly exogenous variables. In the model (7.3) the endogenous variables x_t will consist of the four variables and the exogenous variables z_t contains the four variables.

After reformulating to the partial model, it is possible to test for the number of cointegrations in the

[15] The second row of α, corresponding to Japanese exchange rate, could be weakly exogenous since $\chi^2 = 8.21$ and $p-value = 0.04$. However, when testing simultaneously with other variables, the hypothesis was rejected.

Table 7.24 The Results of Testing Cointegrating Relations in the Partial System: Germany-Japan-U.S.

Eigenvalues	$Trace$	$H_0 : r = h$	p_z	$p_y - r$	$Trace(90)$
0.5539	143.13	0	4	4	76.4
0.4182	78.55	1	4	3	52.4
0.2260	35.22	2	4	2	32.3
0.1681	14.72	3	4	1	15.7

partial system setting. The results are shown in Table 7.24. This table is the same as Table 7.5. The table confirms that there exist three cointegrating relations among the variables since $Trace = 14.72 < 15.7 = Trace(90)$.

Table 7.25 is the result of the estimates of β in the partial system. Note that the order of the variables are different because this is the partial system model and the coefficient of Japanese exchange rate is normalized. The first column of β shows that the following relation will exist among the variables:

$$\begin{aligned} EJ &= 0.834MJ + 0.136MUS + 0.501GJ + 0.774EG \\ &\quad -0.776MG + 0.369GG + 0.577GUS \end{aligned} \tag{7.14}$$

In all three long-run relations, it is noted that U.S. money supply, German exchange rate and GNP are positively related to the Japanese exchange rate. For the other variables, the signs of coefficients can be both positive and negative. To investigate long-run relations more thoroughly, more structures need to be imposed on the long-run relations.

Will these 3 cointegrating relations among the 8 variables be explained by the theoretical relations presented here? Recall that the model derived in the previous chapter found the following 5 theoretical

Table 7.25 The Estimates of the Long-Run Parameters in the Partial System: β

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
EJ	1.000	1.000	1.000
MJ	-0.834	-1.501	4.766
MUS	-0.136	-0.555	-2.173
GJ	-0.501	1.790	1.212
EG	-0.774	-0.308	-0.755
MG	0.776	0.194	-0.655
GG	-0.369	-0.603	-1.419
GUS	-0.577	0.198	-2.766

long-run relations, presented below for the reader's convenience:

$$-\frac{1}{\beta}M + \frac{1}{\beta^*}M^* + \frac{\alpha}{\beta}Y - \frac{\alpha^*}{\beta^*}Y^* = 0 \tag{7.15}$$

$$-\frac{1}{\beta}M + \frac{1}{\beta^{**}}M^{**} + \frac{\alpha}{\beta}Y - \frac{\alpha^{**}}{\beta^{**}}Y^{**} = 0 \tag{7.16}$$

$$-\delta E_1 - \sigma E_2 + \frac{\lambda}{\beta}M + (\phi - 1 - \frac{\lambda\alpha}{\beta})Y = 0 \tag{7.17}$$

$$(\delta^* + \sigma^*)E_1 - \sigma^* E_2 + \frac{\lambda^*}{\beta^*}M^* + (\phi^* - 1 - \frac{\lambda^*\alpha^*}{\beta^*})Y^* = 0 \tag{7.18}$$

$$-\sigma^{**}E_1 + (\delta^{**} + \sigma^{**})E_2 + \frac{\lambda^{**}}{\beta^{**}}M^{**} + (\phi^{**} - 1 - \frac{\lambda^{**}\alpha^{**}}{\beta^{**}})Y^{**} = 0 \tag{7.19}$$

The first 2 relations (7.15) and (7.16) are the same relations that were derived in the two-country case, the money market equilibria in the two countries. The reason why the money market equilibrium conditions do not contain the third country's variables is due to the assumption that the money demand functions do not directly include any foreign variables. See (4.20), (4.21) and (4.22). The relations (7.17), (7.18) and (7.19) are also similar to (7.9) and (7.10) although they include two exchange rates in the relations unlike (7.9) and (7.10). Table 7.26[16] lists the possible signs of the coefficients predicted by the model (7.15) - (7.19).

To map these theoretical long-run relations to the empirical long-run relations that were discovered in the cointegration analysis, the procedure of interpreting the long-run relations that will apply to this case are the same as in the two-country case. To do this, a series of restriction matrices Rs

Table 7.26 The Possible Signs of Coefficients

Equation	E_1	E_2	M^*	M^{**}	M	Y^*	Y^{**}	Y
(7.15)			+		-	-		+
(7.16)				+	-		-	+
(7.17)	+	+			-			?
(7.18)	+	-	+			?		
(7.19)	-	+		+			?	

[16] Note that all the variables are on left-hand side.

corresponding to the above 5 relations are used, written as follows[17]:

$$R_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{7.20}$$

$$R_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{7.21}$$

$$R_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{7.22}$$

$$R_4 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.23}$$

$$R_5 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{7.24}$$

For instance, the relation (7.15) excludes 4 variables, two exchange rates and the second foreign country's variables, from the relation. Exclusion restrictions are imposed on these parameters. The columns correspond to the order of E_1, E_2, M^*, M^{**}, M, G^*, G^{**} and G. The 1 in the third row and fourth column in R_1, for example, implies exclusion of money supply of the second foreign country (Japan). The rest of the 4 matrices Rs are interpreted in a similar fashion. Note, however, that no restrictions are imposed on the signs of the coefficients, only implementing exclusion restrictions on the coefficients.

[17] In the partial system model the order of the variables are EJ, MJ, MUS, GJ, EG, MG, GG and GUS.

Table 7.27 The Estimates of the Long-Run Parameters: German-Japan-U.S. imposing the relation (7.19)

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
EJ	1.000	1.000	1.000
MJ	-0.189	-1.900	-2.886
MUS			
GJ	-0.956	1.386	5.849
EG	-0.636	-0.305	-2.170
MG			
GG			
GUS			

The first question to be examined is whether each of the theoretical long-run relations (7.15) - (7.19) will be supported by the cointegrating space. The relation (7.15) is imposed on all three vectors and tested, obtaining $\chi^2 = 78.36$ and $p - value = 0.00$. Therefore, the hypothesis that the first relation (7.15) is supported by the cointegrating space is not accepted. The other 4 relations are also tried and, except for the fifth relation (7.19), give the same results. In all the 3 cases, $p-value = 0.00$ and causing the rejection of the hypotheses that these 3 relations are supported by the cointegrating space. For the relation (7.19), the result is $\chi^2 = 23.01$ and $p - value = 0.03$. The estimates are found in Table 7.27. In comparison with Table 7.26, the coefficients of MJ in all three vectors have a wrong sign. The signs of EG are as predicted.

When an attempt was made to impose three different relations on the three long-run vectors, there were so many combinations of relations on the three vectors[18] that only some of the results can be reported here.

Having already seen some unexpected signs of the coefficients in Table 7.25, for instance, when a negative relation between EJ and MJ was expected, a positive relation between these variables was obtained. One way to interpret this is to assume that each vector represents a linear combination of some relations. In other words, some different relations are embedded together in each vector. Hence, simply looking at the coefficients of the two variables does not make the relation between these two variables clear.

In considering the relations (7.16), (7.17) and (7.19), 6 different combinations were attempted,

[18] There are 60 (3 out of 5) possible cases, however all 60 cases were not attempted because intuitive information from the estimates in Table 7.25 was used to limit the attempts.

Table 7.28 The Estimates of the Long-Run Parameters: German-Japan-U.S. imposing the relation (7.19) on β_1, (7.17) on β_2 and (7.16) on β_3

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
EJ	-2.311	-1.616	
MJ	2.428		1.000
MUS		0.496	-4.983
GJ	-0.384		12.284
EG	1.000	1.000	
MG			
GG			
GUS		1.326	-5.614

depending on which relations are mapped on which vectors. All 6 cases are accepted although p-values vary from 0.05 to 0.15. However, none of these 6 cases gives the expected sign patterns of the coefficients. For instance, Table 7.28 presents the coefficient estimates of one of the 6 cases. The relation (7.19) is imposed on β_1, (7.17) on β_2 and (7.16) on β_3. The results are $\chi^2 = 9.43$ and $p - value = 0.15$.

Notice that the sign of MJ is incorrect in β_1 again, while the sign of EJ is as predicted. In β_2, both the sign of EJ and MUS are incorrect. Finally, the signs of GJ and GUS are wrong in β_3. This is a typical result from the above 6 cases. None of the combinations satisfy the predicted sign patterns. In fact, this is what happened in other cases. In one more group of combinations that was accepted as the result of hypothesis testing of the combination of the relations (7.17), (7.18) and (7.19). Again, 6 cases can be considered depending on how the relations are mapped. In all 6 cases, the results are the same ($\chi^2 = 11.74$ and $p - value = 0.07$). However, it is impossible to find any combination that satisfy the sign patterns. For instance, consider the relation (7.18) on β_1, (7.19) on β_2 and (7.17) on β_3. Table 7.29 shows the coefficient estimates of this case.

Observe that the signs of both EJ and MG are wrong in β_1. However, all the signs in β_2 turn out to be correct. Finally, the signs of EJ and MUS are incorrect. For the other 5 cases the results are similar.

In sum, although the existence of some linear combination was found among the variables, their relations are not what the model predicted. In particular, the sign of Japanese money supply is wrong in most of the cases. The two cases presented here, the relations (7.17) and (7.19), where there is a linear combination among the variables contained in these relations which are stationary. However, the model does not predict these relations.

Table 7.29 The Estimates of the Long-Run Parameters: German-Japan-U.S. imposing the relation (7.19) on β_1, (7.18) on β_2 and (7.17) on β_3

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
EJ	-17.649	-2.772	-1.860
MJ		1.660	
MUS			0.608
GJ		1.941	
EG	1.000	1.000	1.000
MG	-16.516		
GG	3.169		
GUS			1.650

8 SHORT-RUN DYNAMIC ANALYSIS

The previous chapter discussed long-run relations among the variables in the partial system model. The error-correction (full system) model, the idea of rank of Π, weak exogeneity and the partial system model were all applied to the data. Although the results indicated the existence of some long-run relations among the variables, these long-run relations could not be well explained by the theoretical long-run relation. It could be interpreted that these long-run relations among the variables are linear combinations of several relations.

In this chapter, attention is focused on the short-run dynamics of the model. It will analyze the short-run dynamics of the model based on the model found in the previous chapter and will report the results for variance decompositions and impulse response functions for the Germany-U.S., Japan-U.S. and Germany-Japan-U.S. cases.

8.1 Variance Decomposition Analysis

8.1.1 Germany-U.S.

This section reports the results from variance decomposition analysis for the Germany-U.S. case. Table 8.1 reports the variance decomposition analysis for four different forecasting horizons, giving the results for the full system model. To impose more restrictions on variance structure, a Choleski decomposition is applied.[1] The order of the variables is as follows; German money (MG) $\rightarrow$ German GNP (GG) $\rightarrow$ U.S. GNP (GUS) $\rightarrow$ German exchange rate (EG) $\rightarrow$ U.S. money (MUS).[2] It can be observed, as Dibooglu (1993) pointed out, that the money supplies and GNPs innovation explain the preponderance of their forecast error variance.

The MG innovation accounts for approximately 74% of its own forecast error variance. The EG innovation also explains 17% of the error variance. Approximately 85% of German GNP forecast error variance is explained by its own innovation. The MG and MUS innovation also account for a small

[1] Recall that restrictions on variance structure were mentioned when identification issue in VAR was discussed.
[2] This order of the variables is determined based on the partial model as will be seen later.

Table 8.1 Variance Decomposition for Full System: Germany-U.S.

Variable	Steps	MG	GG	GUS	EG	MUS
MG	1	100.00	0.00	0.00	0.00	0.00
	4	74.05	5.04	0.32	16.20	4.38
	8	73.97	5.05	0.35	16.25	4.38
	12	73.97	5.05	0.35	16.25	4.38
GG	1	3.25	96.75	0.00	0.00	0.00
	4	4.99	84.45	0.64	3.63	6.29
	8	4.99	84.40	0.64	3.67	6.29
	12	4.99	84.40	0.64	3.67	6.29
GUS	1	0.24	1.16	98.60	0.00	0.00
	4	2.06	1.44	91.78	2.82	1.90
	8	2.08	1.48	91.54	3.00	1.90
	12	2.08	1.48	91.54	3.00	1.90
EG	1	2.16	1.47	0.35	96.01	0.00
	4	2.61	3.22	1.43	91.15	1.58
	8	2.61	3.22	1.44	91.14	1.58
	12	2.61	3.22	1.44	91.14	1.58
MUS	1	4.48	2.79	5.01	8.78	78.95
	4	3.77	3.91	3.99	25.38	62.96
	8	3.77	3.91	3.99	25.39	62.94
	12	3.77	3.91	3.99	25.39	62.94

portion of the GG error variance.

The EG innovation accounts for almost 92% of its own variance error and only 10% of the forecast error variance is explained by the other variable innovations in this model. The other variables do not have much explanatory power for the EG variance. The portion explained by the other variables is much smaller than what Dibooglu found.[3]

Finally, the MUS forecast error variance is mainly explained by the MUS and EG innovation (63% and 26% respectively). In this full system model, it is observed that both money supply forecast variances are, to some extent, explained by German exchange rate. In other words, the explanatory power of the German exchange rate for the two money supplies can not be ignored.

In Table 8.2, the results for variance decomposition from the partial system model are found. The order of the variables is the same as in Table 8.1. The order is determined as follows: the weakly exogenous variables are followed by the endogenous variables and ordered by money supply → GNP → exchange rate in each category.

[3]This difference will be discussed when the problems of the innovation analysis are referred to later in the chapter. Recall also that Dibooglu analyzed French Francs and Italian Lira.

Table 8.2 Variance Decomposition for Partial System: Germany-U.S.

Variable	Steps	MG	GG	GUS	EG	MUS
MG	1	100.00	0.00	0.00	0.00	0.00
	4	73.07	6.40	0.25	17.95	2.33
	8	72.99	6.41	0.27	18.00	2.33
	12	72.99	6.41	0.27	18.00	2.33
GG	1	2.45	97.55	0.00	0.00	0.00
	4	4.23	88.25	0.65	2.50	4.38
	8	4.23	88.21	0.65	2.52	4.38
	12	4.23	88.21	0.65	2.52	4.38
GUS	1	0.14	1.38	98.48	0.00	0.00
	4	1.67	1.61	92.10	2.25	2.37
	8	1.69	1.64	91.89	2.40	2.37
	12	1.69	1.64	91.89	2.40	2.37
EG	1	2.06	1.56	0.34	96.03	0.00
	4	2.45	3.48	1.31	91.13	1.63
	8	2.45	3.48	1.32	91.12	1.63
	12	2.45	3.48	1.32	91.12	1.63
MUS	1	3.44	5.32	5.54	8.43	77.27
	4	3.03	5.37	4.73	22.99	63.88
	8	3.03	5.37	4.73	23.00	63.87
	12	3.03	5.37	4.73	23.00	63.87

Recall that only the U.S. money supply is treated as endogenous and the rest of the variables are weakly exogenous in the model. The results from the partial system model are essentially the same as the results from the full system model, that is, no drastic changes are observed. It still can be observed that the money and GNP innovations explain the preponderance of their forecast error variances. Even after the fourth quarter, moneys and GNPs explain more than 60% of their forecast error variances (MUS:64%, MG:73%, GUS:92% and GG:88%). Except for MG, the portion of the error variance explained by its own variable innovation slightly increases from the full system model to the partial system model. For instance, the portion of the GG forecast error variance explained by its own innovation increases from 84% to 88%.

The EG innovation attributes to at most 18% and 23% of the forecast error variances of MG and MUS. EG still has some explanatory power for both money supplies, however, the exchange rate innovation does not account for much of the variance of either GNP. In the variance decomposition analysis, the order of the variables is important, thus when a different order is adopted different results may be obtained.

In general, if the correlations among the variables are small, the order of the variables is not important while if the correlations are high, the order becomes important (Enders (1995)). This is one of the reasons the results differ from what Dibooglu has found.

8.1.2 Japan-U.S.

The results of the same analysis for the Japan-U.S. case is reported here. Table 8.3 presents the results for the full system model. The variable were ordered in this way; GJ → EJ → MUS → MJ → GUS. Recall that GJ and EJ are treated as weakly exogenous variables and MUS, MJ and GUS are treated as endogenous variables. Although it may be difficult to justify the comparison of the above results in Table 8.3 with the results in Table 8.1 in the previous section, since a different set of variables is treated as weakly exogenous and the order of the variables is not the same, it still can be observed from the preponderance of money supplies and GNPs that the largest portion of the forecast variance comes from its own innovation. The GJ and EJ innovation each account for the largest part of their own forecast error variance, approximately 96% for each. The MUS innovation explains 80% of its own forecast error variance. In the long-run (after 1 year), EJ also explains 18% of the MUS forecast variance. In the German-U.S. case, EG was the second important component in explaining the MUS forecast variance (Table 8.1). Again, the exchange rate has some explanatory power for the MUS forecast variance.

The MJ innovation accounts for 72% of its own forecast variance. Unlike the MUS case, the GJ innovation, not the EJ innovation, accounts for the second largest portion of the variance which is as much as 12% of the forecast variance. The GUS forecast variance explained by its own innovation is even lower, 66%. Approximately 27% of its own variance is explained by both GJ and EJ, which are treated as weakly exogenous variables.

Table 8.4 presents the results for the partial system model. The over-all results are essentially the same as the results for the full system model in Table 8.3. The portion of the forecast variance explained by its own innovation becomes slightly smaller in all forecast error variances. In other words, the other variable innovations explain slightly larger portions of the forecast variance and gain more explanatory power.

The GJ and EJ innovation still account for more than 90% of its own forecast error variance respectively. For MUS, the portion of forecast variance explained by the EJ innovation increases as much as the portion accounted for by its own innovation decreases (approximately 4.5%). The portions of the MUS variance accounted for by the other variables do not change. The MJ forecast variance is mainly

Table 8.3 Variance Decomposition for Full System: Japan-U.S.

Variable	Steps	GJ	EJ	MUS	MJ	GUS
GJ	1	100.00	0.00	0.00	0.00	0.00
	4	95.63	0.53	2.33	0.72	0.80
	8	95.59	0.56	2.33	0.73	0.80
	12	95.59	0.56	2.33	0.73	0.80
EJ	1	2.15	97.84	0.00	0.00	0.00
	4	2.06	96.36	0.11	0.76	0.72
	8	2.06	96.33	0.11	0.76	0.74
	12	2.06	96.33	0.11	0.76	0.74
MUS	1	0.51	2.85	96.64	0.00	0.00
	4	1.53	18.08	80.19	0.14	0.07
	8	1.53	18.17	80.07	0.14	0.08
	12	1.53	18.17	80.07	0.14	0.08
MJ	1	11.65	0.51	7.76	80.07	0.00
	4	12.60	1.96	8.52	72.88	4.02
	8	12.61	2.07	8.52	72.78	4.02
	12	12.61	2.07	8.52	72.78	4.02
GUS	1	9.93	1.23	3.68	0.30	84.86
	4	10.45	16.09	5.31	1.14	67.01
	8	10.37	16.79	5.28	1.14	66.43
	12	10.37	16.79	5.28	1.14	66.42

explained by MJ and GJ innovations. The MUS innovation also accounts for approximately 10% of the variance. The portion of the GUS variance explained by Japanese variables, mainly GJ and EJ (weakly exogenous), becomes even larger (approximately 35%).

8.1.3 Germany-Japan-U.S.

This section will examine the results for variance decomposition for the three-country case; the Germany-Japan-U.S. case. Table 8.5 gives the results for variance decomposition from the full system model, with the following order of the variables; MG → GG → GUS → EG → MUS → MJ → GJ → EJ.[4] The order of the variables was determined as in the previous two-country cases.

A large portion of the MG forecast error variance is explained by the MG and EG innovation, a result similar to that of the two-country (Germany-U.S.) case in Table 8.1, however; the size of the portion itself decreases.[5] The portion of the MG variance accounted for by the MUS innovation slightly

[4] The first 4 variables are being treated as weakly exogenous variables and the latter 4 variables are endogenous variables in the partial system model.

[5] The three-country case uses more variables than the two-country case and the order of the variables has changed. It may not be appropriate to compare the numbers from the different cases directly.

Table 8.4 Variance Decomposition for Partial System: Japan-U.S.

Variable	Steps	GJ	EJ	MUS	MJ	GUS
GJ	1	100.00	0.00	0.00	0.00	0.00
	4	91.88	2.70	4.72	0.15	0.55
	8	91.70	2.86	4.73	0.15	0.56
	12	91.70	2.86	4.73	0.15	0.56
EJ	1	2.60	97.40	0.00	0.00	0.00
	4	2.50	95.67	1.05	0.54	0.24
	8	2.51	95.64	1.06	0.54	0.25
	12	2.51	95.64	1.06	0.54	0.25
MUS	1	0.95	4.87	94.18	0.00	0.00
	4	1.66	22.27	75.88	0.12	0.08
	8	1.66	22.37	75.76	0.12	0.09
	12	1.66	22.37	75.76	0.12	0.09
MJ	1	17.20	0.77	7.51	74.52	0.00
	4	15.39	2.44	9.80	69.41	2.97
	8	15.38	2.58	9.79	69.30	2.97
	12	15.38	2.58	9.79	69.30	2.97
GUS	1	10.95	0.85	5.04	0.42	82.73
	4	17.31	15.91	6.60	1.69	58.49
	8	17.16	16.82	6.62	1.68	57.73
	12	17.15	16.83	6.62	1.68	57.73

increases from 4.83% to 7.63%. EJ, which is the third-country variable, also accounts for 6.50% of the forecast variance. Although this portion is small, it still suggests that the third-country variables can not be totally ignored.

The largest component of the GG forecast variance comes from the GG innovation. This portion is now larger than in the two-country case (84.40% → 90.59%). However, the EG innovation does not explain as much as in the two-country case. The third-country variables, including EJ, do not have much explanatory power for the GG variance.

The GUS innovation accounts for 82.05% of its own forecast error variance. It is also noted that German variables, MG, GG and EG, also explain approximately 13% of the variance while Japanese variables do not attribute to the forecast variance as much.

92.50% of the EG forecast error variance is explained by its own innovation. The other variable innovations, especially the third-country variable innovations, do not have much explanatory power for the EG variance (less than 1%).

For MUS, 65.3% of its forecast error variance is accounted for by the MUS innovation. German

Table 8.5 Variance Decomposition for Full System Model: Germany-Japan-U.S.

Variable	Lags	MG	GG	GUS	EG	MUS	MJ	GJ	EJ
MG	1	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4	65.26	5.30	0.29	13.44	7.66	1.42	0.31	6.33
	8	65.00	5.27	0.41	13.42	7.62	1.43	0.35	6.50
	12	64.99	5.27	0.41	13.42	7.63	1.43	0.35	6.50
GG	1	1.44	98.56	0.00	0.00	0.00	0.00	0.00	0.00
	4	1.92	90.74	0.15	2.16	1.26	0.32	0.86	2.59
	8	1.94	90.59	0.18	2.17	1.26	0.32	0.87	2.68
	12	1.94	90.59	0.18	2.17	1.26	0.32	0.87	2.68
GUS	1	2.33	4.64	93.04	0.00	0.00	0.00	0.00	0.00
	4	5.78	4.58	82.62	2.64	0.25	0.70	1.58	1.85
	8	5.86	4.55	82.06	2.78	0.25	0.70	1.58	2.21
	12	5.86	4.55	82.05	2.78	0.25	0.70	1.58	2.22
EG	1	1.67	0.73	0.28	97.32	0.00	0.00	0.00	0.00
	4	1.74	2.08	1.13	92.59	1.64	0.01	0.75	0.06
	8	1.74	2.08	1.14	92.57	1.64	0.01	0.76	0.06
	12	1.74	2.08	1.14	92.57	1.64	0.01	0.76	0.06
MUS	1	3.55	9.42	2.84	9.33	74.86	0.00	0.00	0.00
	4	3.40	8.22	2.58	16.45	65.42	0.29	0.80	2.86
	8	3.41	8.20	2.61	16.43	65.30	0.30	0.81	2.93
	12	3.41	8.20	2.62	16.43	65.30	0.30	0.81	2.93
MJ	1	4.11	8.67	0.01	0.24	1.26	85.72	0.00	0.00
	4	2.79	6.40	2.17	2.46	3.09	67.67	5.69	9.72
	8	2.79	6.39	2.19	2.46	3.08	67.48	5.87	9.73
	12	2.79	6.39	2.19	2.46	3.08	67.48	5.87	9.73
GJ	1	3.00	0.41	4.46	0.11	0.81	2.74	88.47	0.00
	4	2.42	1.88	3.71	5.52	2.37	2.08	76.05	5.97
	8	2.43	1.88	3.74	5.52	2.37	2.08	75.98	5.99
	12	2.43	1.88	3.74	5.52	2.37	2.08	75.98	6.00
EJ	1	4.28	2.88	0.26	37.31	0.58	0.98	0.42	53.28
	4	6.65	2.26	4.34	28.38	0.44	1.38	1.26	55.35
	8	6.78	2.23	4.71	28.07	0.44	1.30	1.25	55.21
	12	6.79	2.23	4.72	28.06	0.44	1.30	1.25	55.21

variables also have some explanatory power and explain approximately 28% of the MUS variance. Compared with the portion explained by these German variable innovations, the portion by the Japanese variable innovations is much smaller (only 4%). This asymmetry in explanatory powers of German and Japanese variable innovations comes from the fact that all German variables are ordered before the Japanese variables.[6]

The largest portion (67.5%) of the MJ forecast variance is explained by the MJ innovation. Approximately 10% of the forecast variance is attributed tc by the EJ innovation. The GJ innovation does not have much explanatory power. The three German variable innovations can not be ignored, although the portion accounted for by the three innovations is small (11%) and none of the individual innovation contributes much. In general, the portion of the Japanese forecast variance explained by the German variable innovations is larger than the portion of the German forecast variance explained by the Japanese variable innovations, as will be seen in the rest of the two Japanese forecast variances.

The results are similar for the GJ forecast variance, GJ attributes most to its variance (76%). The German innovations explain approximately 10% of the GJ variance, while the EJ innovation explains only 55% of its own forecast variance. This portion is much smaller than any other portions explained by the own innovations. Interestingly, the EG innovation has some explanatory power for EJ. It accounts for 28% of the EJ forecast variance.

The other variable innovations do not attribute to the EJ forecast variance as much as the EG innovation. The third-country variable, EG, is important in explaining the variability of EJ.

A similar analysis was performed for the partial system model. The results for the analysis are found in Table 8.6. Recall that the difference between the full system model and the partial system model is that, in the partial system model, the first 4 variables are treated as weakly exogenous variables and the last 4 as endogenous variables .

Meanwhile, there is no such a distinction among the variables and all the variables are treated equally in the full system model.

First, it is noted that the over-all results do not change drastically. It is still true that the largest component of the forecast error variance comes from its own innovation (see the diagonal of Table 8.6.). In many cases, the portion explained by its own innovation seems to be smaller in the partial system model than in the full system model. For some forecast error variances, the portions explained by German variable (weakly exogenous) innovations become larger and the portions explained by Japanese variable (endogenous) innovations are smaller in the partial system model than in the full system model,

[6] When a different order of the variables is applied, the results change. We discuss this problem later in this chapter.

Table 8.6 Variance Decomposition for Partial System Model: Germany-Japan-U.S.

Variable	Lags	MG	GG	GUS	EG	MUS	MJ	GJ	EJ
MG	1	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	4	71.73	5.43	0.40	17.37	3.42	1.11	0.53	0.01
	8	71.56	5.45	0.44	17.42	3.41	1.11	0.54	0.09
	12	71.56	5.45	0.44	17.42	3.41	1.11	0.54	0.09
GG	1	3.77	96.23	0.00	0.00	0.00	0.00	0.00	0.00
	4	5.61	78.63	0.38	2.87	7.54	0.44	3.88	0.65
	8	5.61	78.50	0.40	2.93	7.52	0.44	3.92	0.67
	12	5.61	78.50	0.40	2.93	7.52	0.44	3.92	0.67
GUS	1	0.45	0.93	98.62	0.00	0.00	0.00	0.00	0.00
	4	2.20	0.85	88.67	1.11	1.06	1.02	0.37	4.73
	8	2.23	0.85	88.49	1.13	1.06	1.02	0.37	4.85
	12	2.23	0.85	88.49	1.13	1.06	1.02	0.37	4.85
EG	1	2.45	0.82	0.28	96.45	0.00	0.00	0.00	0.00
	4	2.77	2.77	1.38	90.92	1.06	0.01	0.71	0.37
	8	2.77	2.79	1.43	90.80	1.06	0.01	0.72	0.44
	12	2.77	2.79	1.43	90.80	1.06	0.01	0.72	0.44
MUS	1	2.51	5.80	5.11	9.88	76.70	0.00	0.00	0.00
	4	2.39	5.10	4.30	17.77	64.27	0.85	1.95	3.36
	8	2.39	5.10	4.30	17.77	64.25	0.86	1.97	3.37
	12	2.39	5.10	4.30	17.77	64.25	0.86	1.97	3.37
MJ	1	3.56	7.75	0.31	0.10	2.41	85.87	0.00	0.00
	4	2.59	6.45	1.46	2.04	2.64	69.11	5.07	10.63
	8	2.59	6.45	1.46	2.05	2.69	69.01	5.12	10.62
	12	2.59	6.46	1.46	2.05	2.69	69.01	5.12	10.62
GJ	1	3.64	1.39	2.80	0.21	1.33	2.16	88.47	0.00
	4	3.04	2.66	2.26	6.45	1.57	1.72	74.94	7.35
	8	3.04	2.68	2.29	6.46	1.58	1.72	74.86	7.38
	12	3.04	2.68	2.29	6.46	1.58	1.72	74.86	7.38
EJ	1	7.14	2.67	0.62	38.97	0.11	2.11	0.31	48.07
	4	9.03	2.60	1.41	34.38	0.44	1.85	2.35	47.94
	8	9.03	2.60	1.42	34.37	0.45	1.85	2.35	47.94
	12	9.03	2.60	1.42	34.37	0.45	1.85	2.35	47.94

however, this is not a clear cut phenomenon. For the U.S. variable forecast variances, the portion accounted for by the Japanese variable innovations increases as much as the portion explained by German variable innovations increases.

For the EG forecast variance, the portion explained by the EJ innovation slightly increases (0.06% $\rightarrow$ 0.44%), but it is still a small portion. The three Japanese variable innovations have little explanatory power for the EG forecast variance, a little over 1%. The portion of the EJ forecast variance explained by the EG innovation increases from 28.06% to 34.37%. The German variable innovations account for a larger portion of the EJ variance in the partial system model than in the full system model (approximately 37% $\rightarrow$ 46%). In the partial system model, the third-country variables, in this case German variables, become more important explanatory variables for EJ. Meanwhile, the EJ innovation attributes less to its own variance (55.21% $\rightarrow$ 47.94%). It is also observed that the contributions of all three Japanese variable innovations to the EJ variance slightly decreases from 58% to 52%. Since all German variables are treated as weakly exogenous and all Japanese variables as endogenous, the contribution of German variable innovations to Japanese forecast variances is larger than the contribution of Japanese variable innovations to German forecast variances.

8.2 Impulse Response Analysis

Long-run equilibrium was estimated in the previous chapter. However, as Lütkpohl (1991) points out, it is not appropriate to interpret the coefficients in the long-run equilibrium equations as the long-run effect of a unit increase in one variable on the other since this ignores all the other relations among the variables summarized in the system. For instance, in the long-run equilibrium equation (7.7) the coefficient of German money supply is -0.988. This should not be interpreted as an increase in the German exchange rate of 0.988 when German money supply increases by one unit, since the other variables are held fixed. In the long-run, all the variables in the system move so that impulse response analysis is more appropriate in order to investigate the long-run relations among the variables. It is appropriate to apply impulse response analysis to error correction model where all the effects are on the first-order difference of the variables. Hence, when a positive effect on a variable is observed, it is possible to conclude that the variables moves in a positive direction, that is, the variables increases. The larger effect implies the larger change in the variables or a faster speed of change in the variable.

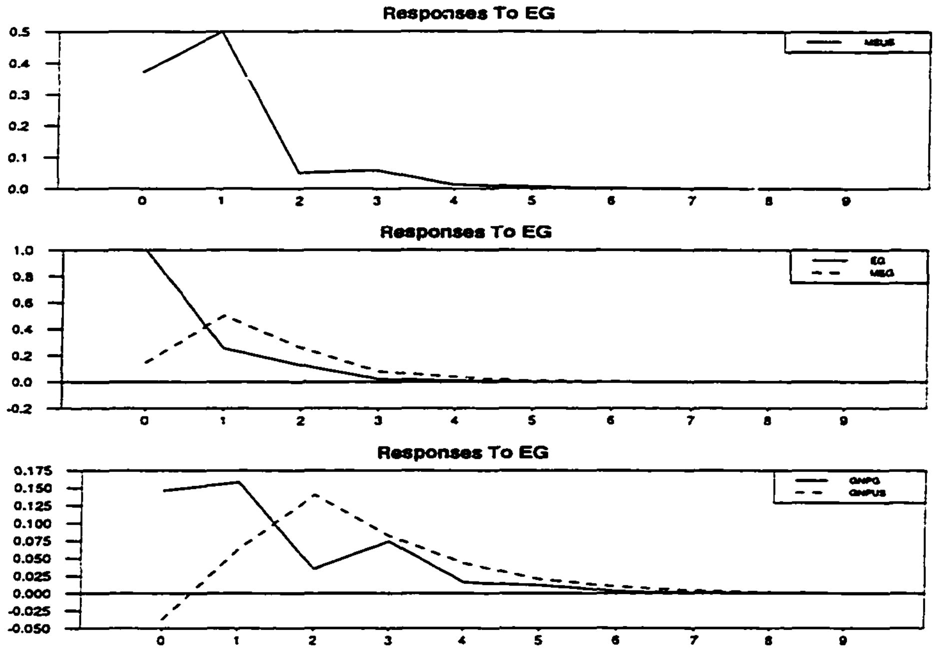

Figure 8.1 Responses to German Exchange Rate: Germany-U.S.

8.2.1 Germany-U.S.

Impulse response functions for the Germany-U.S. case are reported in Figure 8.1 - 8.5. Five different figures are given for the impulse response functions from the partial system model since there are 5 variables in the system. Figures are given in the order of the variables used in the variance decomposition analysis discussed in the previous section. Figure 8.1 shows the responses of 5 variables, including German real exchange rate itself, to the German real exchange rate shock. The German real exchange rate shock will cause both the German and the U.S. money supply to increase immediately. The exchange rate affects U.S. money supply more than German money supply at Q_0. The effect on U.S. and German money supply will last for approximately 6 quarters and, then die down. Since the effects on both money supplies are positive throughout time, this implies that the effect of the German exchange rate shock on both money supplies will be permanent. The effects on German and U.S. GNP are in opposite directions at Q_0. In the U.S., GNP responds negatively while the German GNP responds positively to the German exchange rate shock at Q_0. However, the effect on U.S. GNP immediately turns positive at Q_1. Both effects are also positive, lasting for 2 years.

One standard deviation of the German money supply shock (Figure 8.2) does not induce German real exchange rate to move immediately. There is no contemporaneous effect on the real exchange rate.

The German exchange rate slightly appreciates for the first 2 quarters and, then the effect will die down. Since the effect on the real exchange rate is positive over time, the effect will be permanent. The German money supply shock has an immediate positive effect on German GNP. As the German money supply shock diminishes, the effect on GNP will decrease in Germany and lasts for 3 quarters. Again, the shock is positive over time (it turns to be slightly negative at Q_3), and is considered to be permanent. It is apparent that German money supply shock has positive effects on the U.S. money supply and GNP at Q_0. The U.S. money supply decreases in Q_1 while U.S. GNP keeps increasing for 6 quarters. The effect on U.S. money supply may be considered temporary, since it fluctuates in both negative and positive direction, but the effect on U.S. GNP is permanent. Therefore, the German money supply increases both German and U.S. GNP in this model.

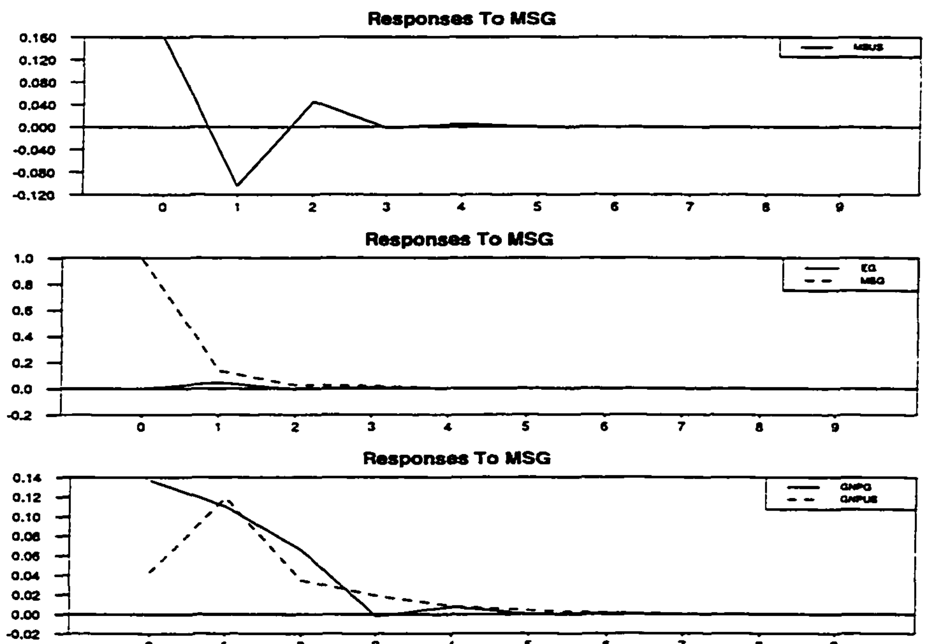

Figure 8.2 Responses to German Money Supply: Germany-U.S.

In Figure 8.3 similar results can be observed. The U.S. money supply shock does not have a contemporaneous effect on exchange rate and it will lead to exchange rate appreciation at Q_1. Although the effect goes down to almost zero at Q_2, over time the effect on the exchange rate is positive and seems to be permanent. German money supply follows the exchange rate and it seems that it moves in order to offset the exchange rate appreciation. The U.S. money supply shock is negatively related to the German GNP. German GNP does not move contemporaneously and decreases at Q_1 when the U.S. money supply shock occurs. At Q_2, German GNP moves in a positive direction and offsets the negative effect at Q_1, but the overall effect may not be as large as the positive offset. The U.S. GNP does not initially move and increases at Q_1. The effect will continue for approximately 1 year and is considered to be a permanent effect. The effect on both money supplies and GNPs will die down after the sixth quarters.

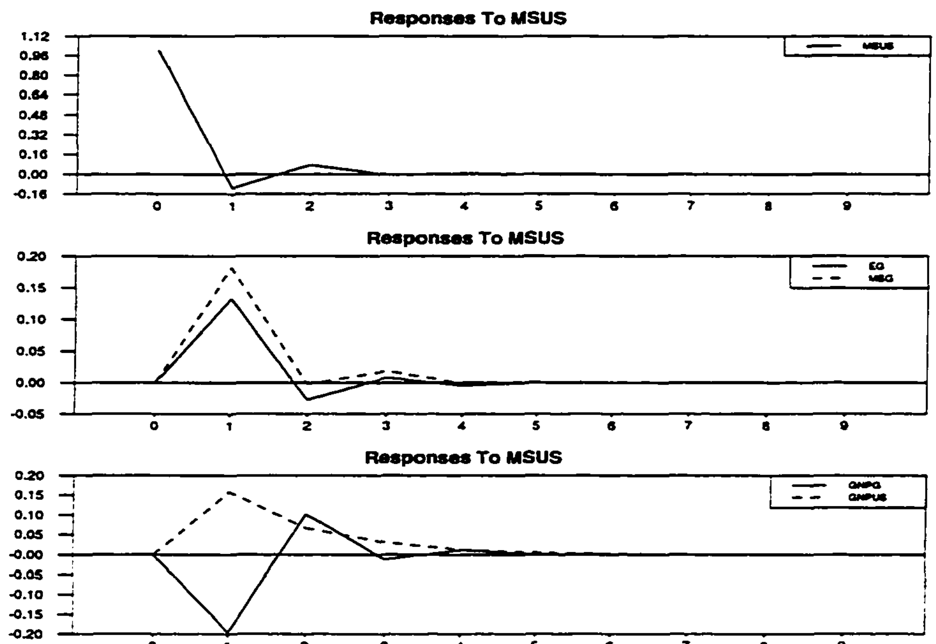

Figure 8.3 Responses to U.S. Money Supply: Germany-U.S.

The German GNP shock (Figure 8.4) does not have contemporaneous effects on the exchange rate or the German money supply. The effect on the German exchange rate is a rise at Q_1, which will not diminish for approximately 2 years, a permanent effect. The German GNP shock does not immediately increase German money supply. As the exchange rate appreciates, the German monetary authority intervenes in the market in order to slow the fluctuations of the exchange rate down. After the money supply increases for approximately 4 quarters, it stops increasing when the monetary authority does some fine tuning. The monetary authority acts passively, responding to the GNP shock and the exchange rate. The effects on the German money supply are also permanent. The U.S. money supply increases contemporaneously at Q_0 and the effect is positive over time. The German GNP shock on U.S. GNP is also positive. It will have a small positive effect on U.S. GNP. This effect continues for approximately 4 quarters and is considered to be permanent.

The U.S. GNP shock (Figure 8.5) does not have contemporaneous effects on the exchange rate, the German money supply or the German GNP but it does have some effect on the U.S. money supply. The

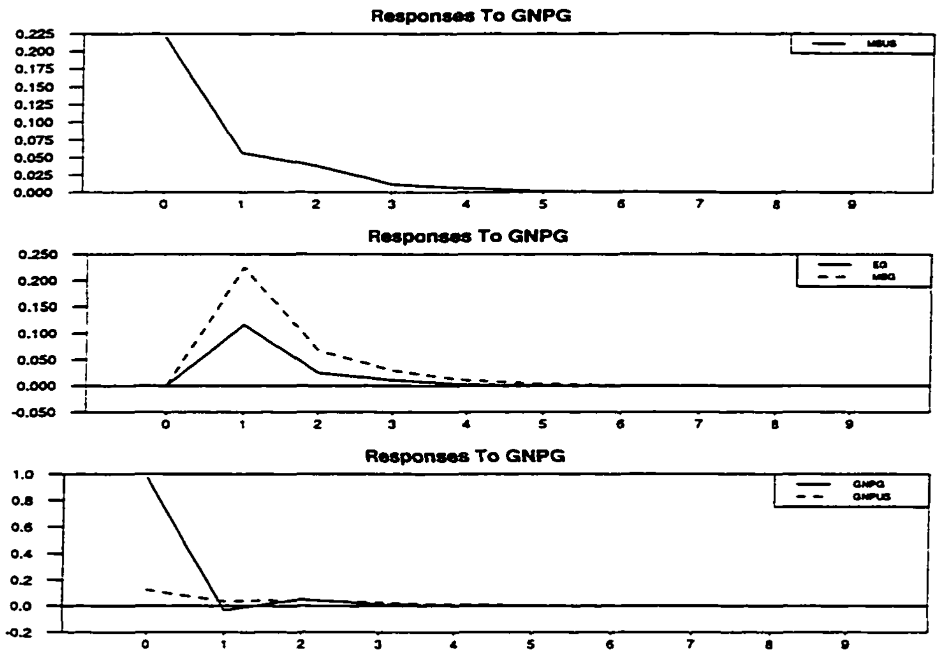

Figure 8.4 Responses to German GNP: Germany-U.S.

U.S. money supply responds positively to U.S. GNP shock when the U.S. money authority passively reacts to the GNP shock by increasing the money supply in the economy. The German real exchange rate depreciates at Q_1 and the German money authority decreases the German money supply to induce the exchange rate to appreciate. Due to their effort, the speed of the depreciation decreases after Q_1. Since the effect on the exchange rate is negative over time, the U.S. GNP shock has a permanent negative effect on the exchange rate, similarly German money supply will decrease in the long run.

Table 8.7 summarizes the above results for impulse responses. The arrows indicate long-run permanent effects of the shocks. It indicates that the effects of the shocks are temporary rather than permanent.

Table 8.7 Summary of Impulse Responses: Germany-U.S.

	EG	MG	MUS	GG	GUS
EG	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$
MG	$\nearrow$	$\nearrow$	-	$\nearrow$	$\nearrow$
MUS	$\nearrow$	$\nearrow$	$\nearrow$	$\searrow$	$\nearrow$
GG	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$
GUS	$\searrow$	$\searrow$	$\nearrow$	$\nearrow$	$\nearrow$

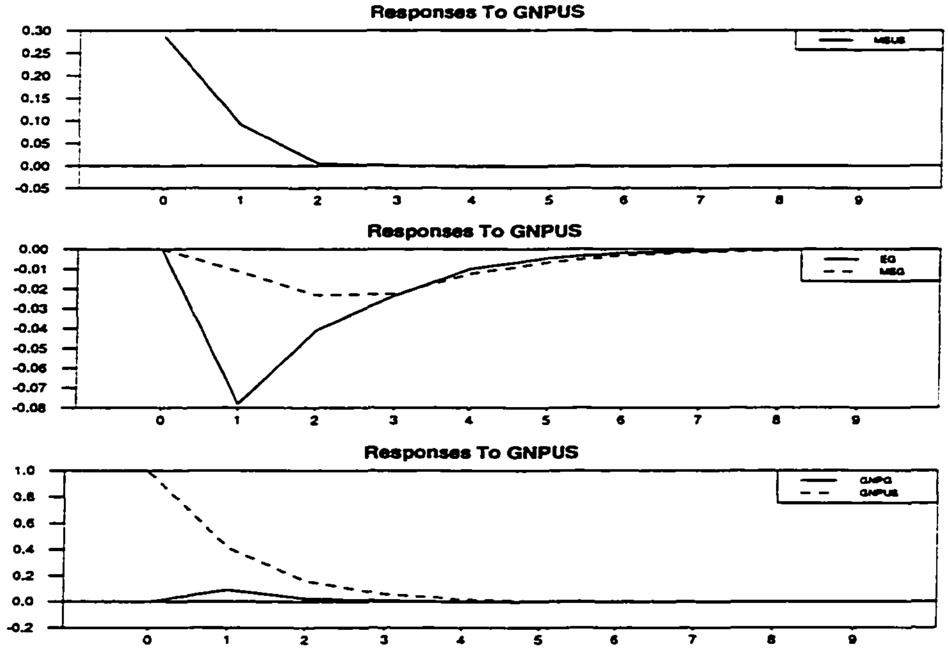

Figure 8.5 Responses to U.S. GNP: Germany-U.S.

8.2.2 Japan-U.S.

Figures 8.6 - 8.10 present the results for impulse response functions of the Japan-U.S. case, also taken from the partial system model. In the long-run, the Japanese exchange rate shock (Figure 8.6) will have a positive effect on the Japanese exchange rate itself since the negative change at Q_0 is offset by the positive change initiated at Q_2. The Japanese exchange rate shock has contemporaneous effects on both the U.S. and the Japanese money supply. The U.S. money supply increases immediately while the Japanese money supply decreases at Q_0. However, Japanese money supply also increases after Q_1 and continues to increase thereafter. The Japanese exchange rate shock has a permanent positive effect on U.S. money supply since the overall effect is positive over the time. Japanese GNP responds negatively to the Japanese exchange rate shock. The appreciation of the Japanese Yen dampens Japanese exports and decreases Japanese GNP. As the speed of the appreciation of the Yen slows down, Japanese GNP recovers and after Q_1, it increases. On the other hand, U.S. GNP responds positively to the Japanese exchange rate shock. The effect on U.S. GNP is permanently positive and dies down after 2 years.

The Japanese money supply increases contemporaneously responding to its own shock (Figure 8.7) as expected, however, at Q_1, it decreases, and, then, after Q_3, Japanese money supply does not change any longer. The Japanese money supply shock has positive effects on both the U.S. money supply and

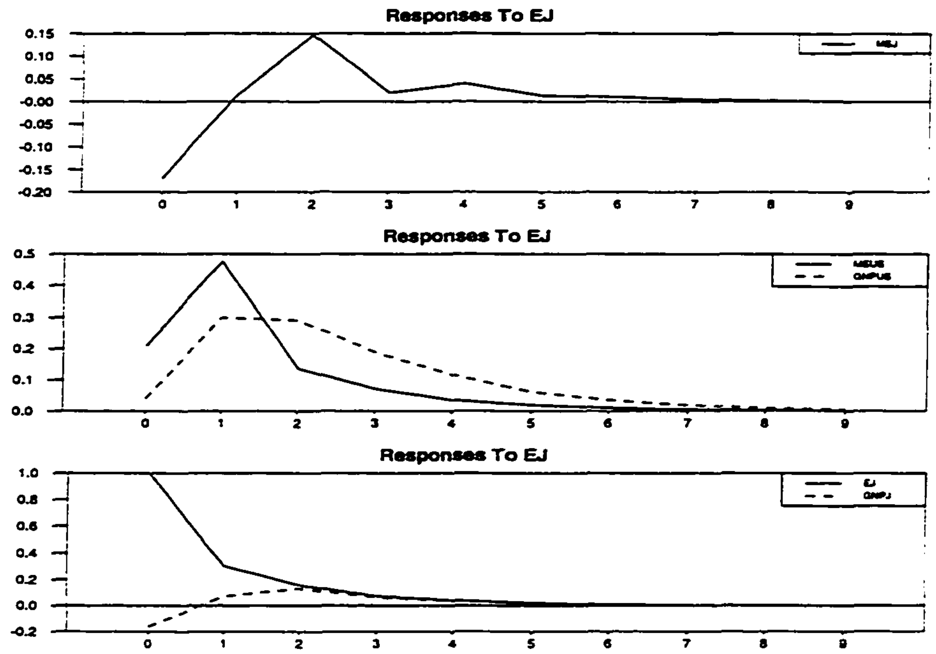

Figure 8.6 Responses to Japanese Exchange Rate: Japan-U.S.

GNP. In both cases, the positive changes are larger than the negative changes, so the Japanese money supply shock has positive effects on both variables. However, the Japanese money supply shock does not have any contemporaneous effects on the Japanese exchange rate or Japanese GNP. The exchange rate decreases at Q_1. Then, at Q_2, although it increases, the positive effect is smaller than the negative effect at Q_1. The Japanese money supply shock will have a negative effect on the exchange rate, as theory predicts. An increase in Japanese money supply leads to depreciation of the Japanese yen, so Japanese GNP increases at Q_1 and decreases at Q_2. Yet the overall effect seems to be small due to the offset of the two effects and, in the end, all the effects of the Japanese money supply shock will die down within 2 years.

The U.S. money supply shock (Figure 8.8) does not have contemporaneous effects on the Japanese money supply, GNP or exchange rate, but U.S. GNP positively responds to the U.S. money supply shock. It is also observed that the U.S. money shock has a permanent effect on U.S. GNP. The Japanese money supply starts to respond positively at Q_1 and decreases at Q_2. The positive changes in the Japanese money supply seem to be larger than the negative changes, so that the long-run U.S. money supply effect is positive. At Q_1, both the Japanese exchange rate and GNP increase, this positive change in exchange rate is as expected: an increase in U.S. money supply induces the Japanese yen to appreciate, mainly due to the fact that the U.S. money supply increases faster than the Japanese money supply.

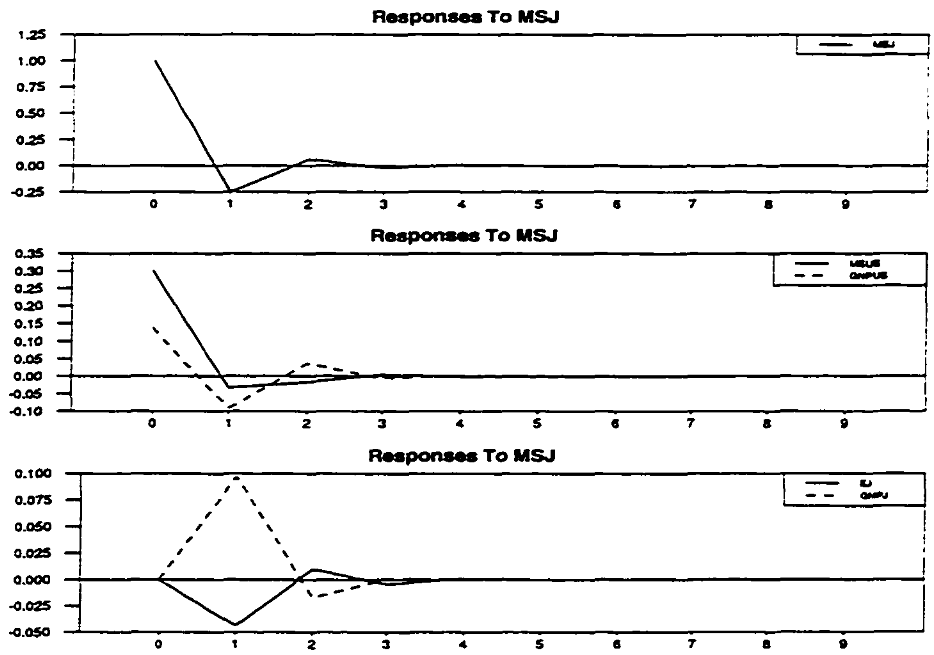

Figure 8.7 Responses to Japanese Money Supply: Japan-U.S.

The change in Japanese GNP is not so large as the changes in the other variables. Although the speeds of the changes in Japanese exchange rate and GNP decrease at Q_2, the two variables continue to increase and the effects on both variables are permanently positive. The effects on all 5 variables will diminish after Q_6.

The Japanese GNP shock (Figure 8.9) immediately affects variables other than the Japanese exchange rate. The exchange rate receives a positive shock at Q_1. At Q_2, the change in the Japanese exchange rate starts slowing down, then it appreciates for the three quarters and its effect dies out quickly. Here, both Japanese and U.S. money supply move in the same direction, but the change in the Japanese money supply is larger than the change in the U.S. money supply.

Now, the Japanese monetary authority will try to reduce money supply to slow down the economy and induce the Japanese yen to depreciate after Q_0. At Q_0, the Japanese GNP shock has a positive effect on U.S. GNP and its effect is permanently positive. In fact, the effect on U.S. GNP will last longer than the effect on Japanese GNP.

The U.S. GNP shock (Figure 8.10) will not contemporaneously affect any other variables than itself. It has an effect on Japanese exchange rate at Q_1, which is permanently positive and continues for approximately 2 years. When the U.S. economy expands, it will help the Japanese economy to expand by increasing Japanese exports to the U.S., thus Japanese GNP also responds positively to the U.S.

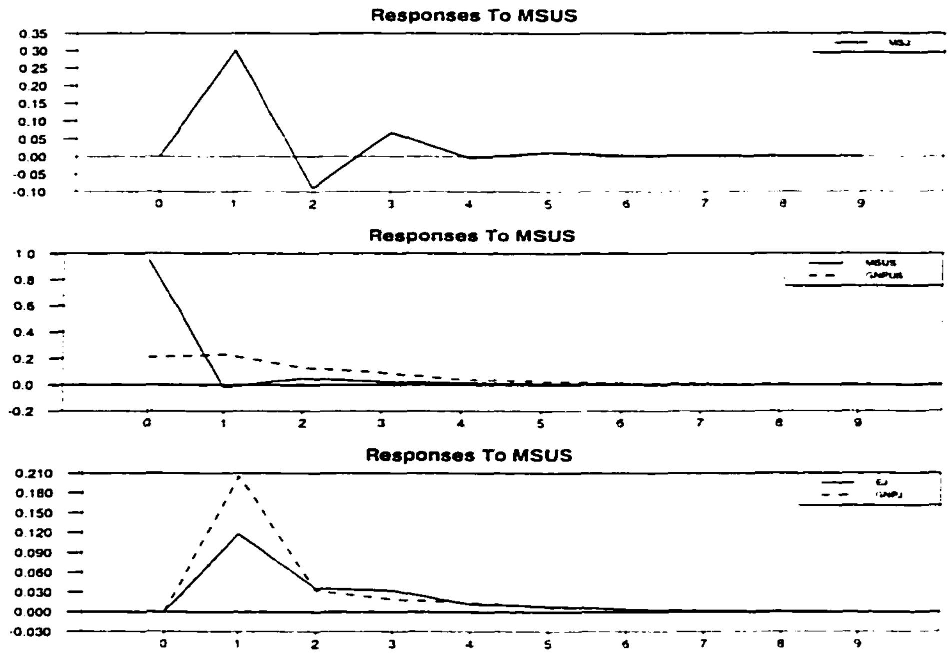

Figure 8.8 Responses to U.S. Money Supply: Japan-U.S.

GNP shock. The effect on the U.S. money supply is very small but positive. As far as the Japanese money supply is concerned, it also increases at Q_1, but at Q_2, it decreases slightly. The effects on Japanese exchange rate and Japanese GNP seem to continue slightly longer than the effects on the other variables.

Table 8.8 summarizes the above results for impulse responses. Table will be read in the same fashion as Table 8.7.

Table 8.8 Summary of Impulse Responses: Japan-U.S.

	EJ	MJ	MUS	GJ	GUS
EJ	↗	-	↗	-	↗
MJ	↘	↗	↗	↗	↗
MUS	↗	↗	↗	↗	↗
GJ	↘	↗	-	↗	↗
GUS	↘	↘	-	↗	↗

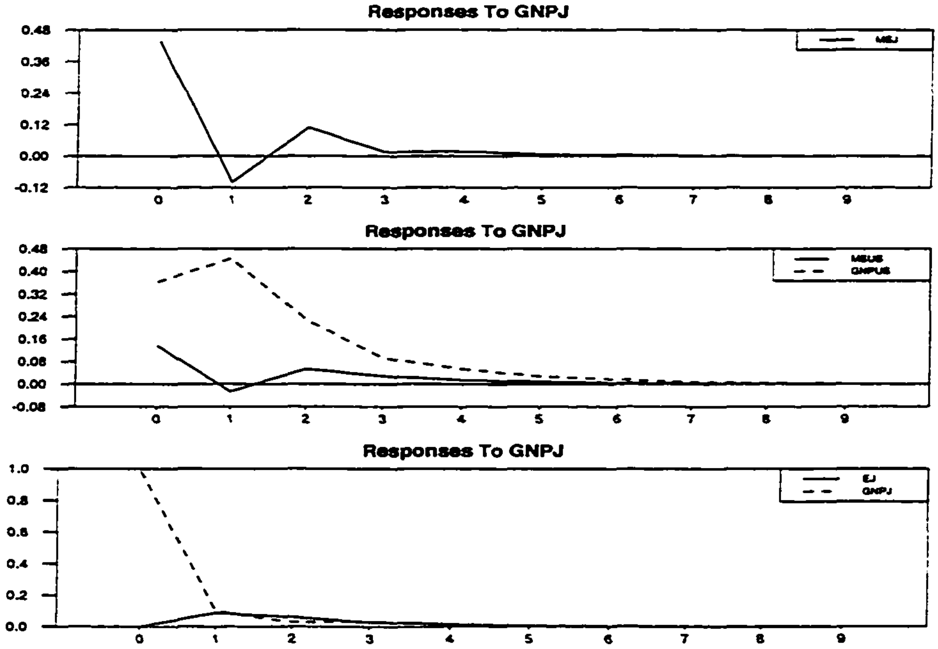

Figure 8.9 Responses to Japanese GNP: Japan-U.S.

8.2.3 Germany-Japan-U.S.

Figure 8.11 - 8.18 are impulse response functions for the Germany-Japan-U.S. case. In this case, there are 8 variables in the system, therefore 8 different impulse response functions. In this three-country case, interest naturally lies in the third country effects.

The German exchange rate shock (Figure 8.11) will have contemporaneous effects on all the variables in the system, although the effects on Japanese and U.S. GNP are minor. The German exchange rate shock will induce both the German Mark and the Japanese Yen to appreciate against the U.S. dollar. It is interesting to note that the Japanese exchange rate responds to the German exchange rate shock as much as the German exchange rate does. While the speed of appreciation of both exchange rates slows down after Q_1, the German exchange rate shock will have permanent effects on both currencies. All three money supplies respond positively to the German exchange rate shock, here, the effect on the Japanese money supply is the smallest among all the three. However consistently positive the reactions of the three money supplies, the effects on GNPs are various. The effects on German GNP are consistently positive and permanent and will die down after one year. The effects on Japanese and U.S. GNP are initially negative and small, however, at Q_1, both effects turn positive and continue to be positive until the effects die down.

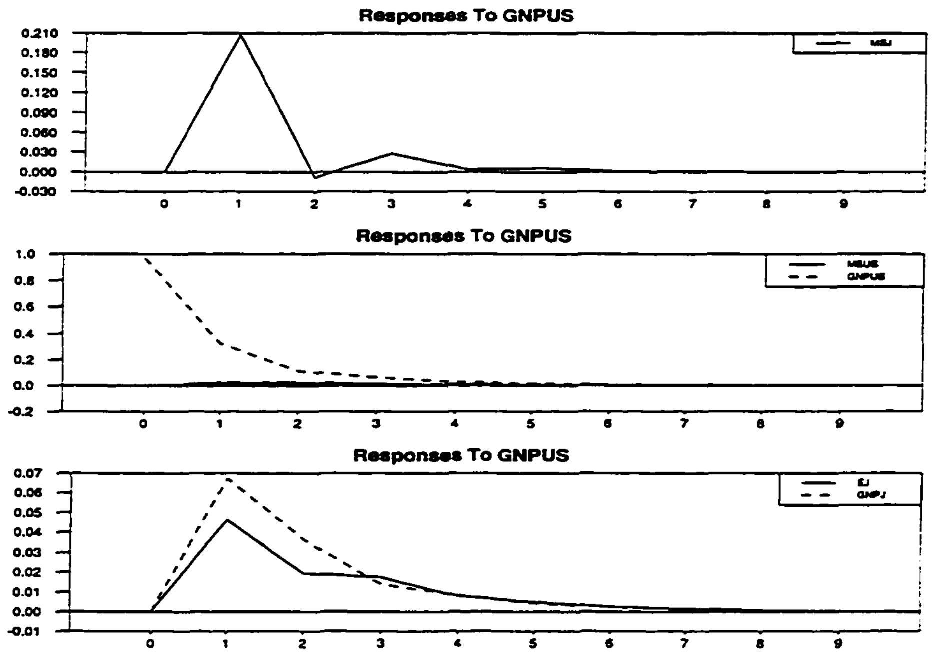

Figure 8.10 Responses to U.S. GNP: Japan-U.S.

The Japanese exchange rate shock (Figure 8.12) has different effects on the variables from the German exchange rate shock because Japanese exchange rate is treated in a different way (as weakly exogenous variable) in the system. It has immediate effects on weakly exogenous variables such as Japanese exchange rate itself, Japanese money supply, U.S. money supply and Japanese GNP, but no contemporaneous effects exist on the other endogenous variables. The Japanese exchange rate shock will have a positive and permanent effect on its own variable. However, it will have a negative effect on the German exchange rate, unlike the previous case, and this effect is consistently negative, the German mark will depreciate due to the Japanese exchange rate shock. In this model, the effect of the Japanese exchange rate shock on the German exchange rate is opposite of the effect of the German exchange rate shock on the Japanese exchange rate. The effect on the Japanese money supply is initially negative, but fluctuates after Q_2. The overall effect seems to be negative, although this negative effect is partially offset by the positive effect.

The effect on U.S. money supply initially moves in a positive direction and turns positive at Q_1, when the effect on German money supply is negative. The effect on German money supply continues longer than the effects on the two other money supplies. The Japanese exchange rate shock has a negative effect on Japanese GNP, caused by the appreciation of Japanese Yen which decreases Japanese exports to the other countries. On the other hand, its effect on U.S. GNP is consistently positive, while

German GNP will fall over time, but the magnitude of this effect is not as large as those of the other effects.

The German money supply shock (Figure 8.13) immediately induces the Japanese exchange rate to appreciate. The German exchange rate also appreciates after Q_2, which contradicts the predictions from the theoretical model. Interestingly, the effect on the German exchange rate is smaller than the effect on the Japanese exchange rate and it also dies down more quickly. Both U.S. and Japanese money supply increase at Q_0 and immediately decrease at Q_1, then, the effects on both money supplies quickly die down after Q_3. Since the effects fluctuate over time, they are temporary effects. On the other hand, the effect on the German money supply itself is consistently positive and permanent. The effects on all GNPs are contemporaneously positive and continue for one year, and, hence, the effects on all GNPs are permanently positive.

The Japanese money supply shock (Figure 8.14) does not have contemporaneous effects on either the Japanese exchange rate or the German exchange rate. In fact, German exchange rate hardly responds to the Japanese money supply shock. The Japanese exchange rate will slightly appreciate, but the effect quickly dies down. Both the Japanese and the U.S. money supply immediately increase and fluctuate, though the effect on U.S. money supply seems temporary. The overall effect on Japanese money supply is positive, while some of this positive effect is offset as time passes. The German money supply does not immediately respond to the Japanese money supply shock, but it starts increasing at Q_1 and fluctuates thereafter, moving in the opposite direction of the U.S. and the Japanese money supplies over time, so the effect on the German money supply is temporary. Japanese GNP increases at Q_0, but the effect is not large and the initial positive effect will be offset by some negative effects later. Both the U.S. and German GNPs move in the same direction, the effect on German GNP being larger than the effect on U.S. GNP, but the effects on both GNPs are temporary.

The U.S. money supply shock (Figure 8.15) has contemporaneous effects only on the U.S. money supply itself and the Japanese money supply. The two exchange rates do not respond to the U.S. money supply shock at Q_0, but the German exchange rate appreciates and the Japanese exchange rate depreciates at Q_1. It seems that the German exchange rate is more responsive to the U.S. money supply shock than the Japanese exchange rate, though the effects fluctuate over time and are considered to be temporary. Both the German and the Japanese money supply increase at Q_1. Although the effect on German money supply dies down after Q_2, the effect on Japanese money supply will continue for 6 quarters. The U.S. money supply shock has a positive effect on U.S. GNP and the effect is permanently positive. On the other hand, the shock negatively affects both German and Japanese GNP initially, but as time progresses, both GNPs also increase due to the expansion of U.S. economy. Both the negative and positive effects on the two GNPs are approximately the same, so they tend to offset each other and the overall effect may not be large.

While German exchange rate does not immediately respond to the German GNP shock (Figure 8.16), Japanese exchange rate contemporaneously responds in a positive way. Due to the fluctuation, the overall effect on the Japanese exchange rate is small, but permanently positive. All three money supplies increase due to the German GNP shock, but once again, the German money supply does not immediately respond. Even though the effects on the U.S. and the Japanese money supplies are sometimes negative, overall, they are positive. The German GNP shock will increase money supplies in the three countries and it also has positive effects on all three GNPs.

The Japanese GNP shock (Figure 8.17) does not have contemporaneous effects on any variables other than Japanese GNP, but the Japanese and the German exchange rates appreciate at Q_1, and both exchange rates fluctuate thereafter. The positive effects on both exchange rates are larger than the negative effects, so the overall effect on the exchange rates will be positive and the two exchange rates appreciate due to the Japanese GNP shock. The three money supplies respond in the same direction, decreasing at Q_1 and fluctuating over time. The Japanese GNP shock seems to have a negative effect on Japanese money supply, while it it has only temporary effects on German and U.S. money supply. It is also noted that the Japanese GNP shock induces the three GNPs, including Japanese GNP, to increase over time, therefore; in this system, the Japanese GNP shock will positively contribute to the GNPs in all three countries.

Finally, the U.S. GNP shock (Figure 8.18) induces German exchange rate to depreciate and Japanese exchange rate to appreciate, both effects being permanent. This appreciation is not what the model predicts. It also increases both the Japanese and the U.S. money supply and decreases the German money supply. The U.S. GNP shock will induce German and Japanese GNPs to increase as well as sparking an increase in the U.S. GNP itself, so in the end it has a positive effect on all three GNPs.

The summary of the above results is presented in Table 8.9.

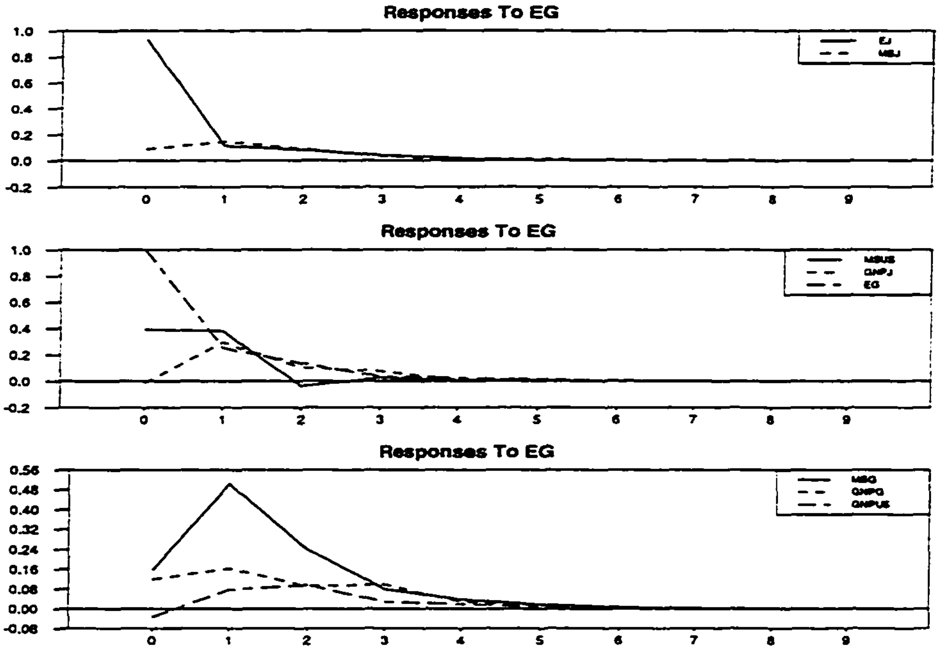

Figure 8.11 Responses to German Exchange Rate: Germany-Japan-U.S.

Table 8.9 Summary of Impulse Responses: Germany-Japan-U.S.

	EG	EJ	MG	MJ	MUS	GG	GJ	GUS
EG	↗	↗	↗	↗	↗	↗	↗	↗
EJ	↘	↗	↘	↘	↗	↘	↘	↗
MG	↗	↗	↗	↗	-	↗	↗	↗
MJ	-	↗	↗	↗	-	-	-	-
MUS	↗	↗	↗	↗	↗	↘	-	↗
GG	↗	↗	↗	↗	↗	↗	↗	↗
GJ	↗	↗	-	↘	-	↗	↗	↗
GUS	↘	↗	↘	↗	↗	↗	-	↗

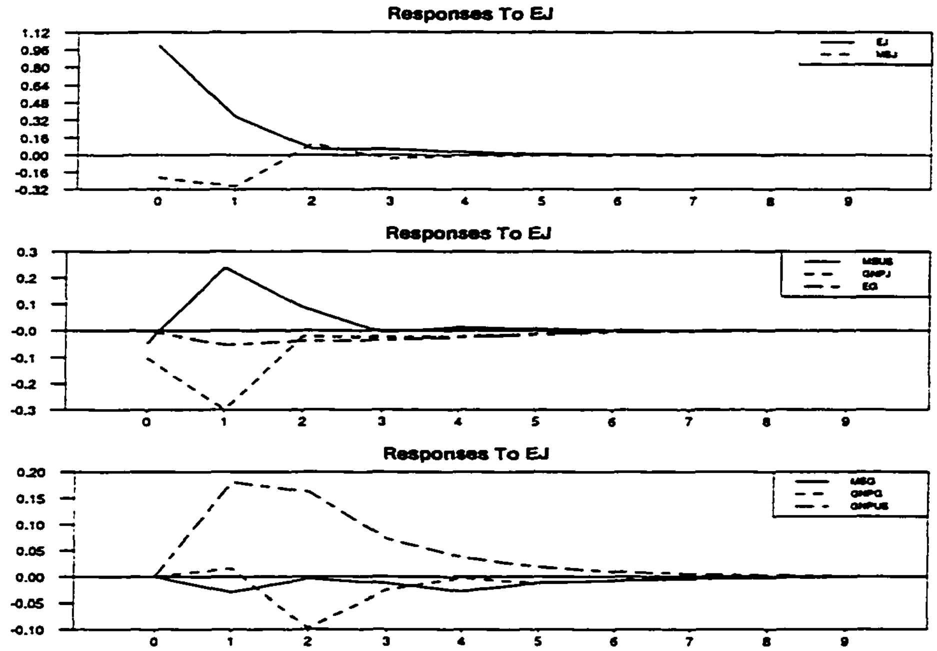

Figure 8.12 Responses to Japanese Exchange Rate: Germany-Japan-U.S.

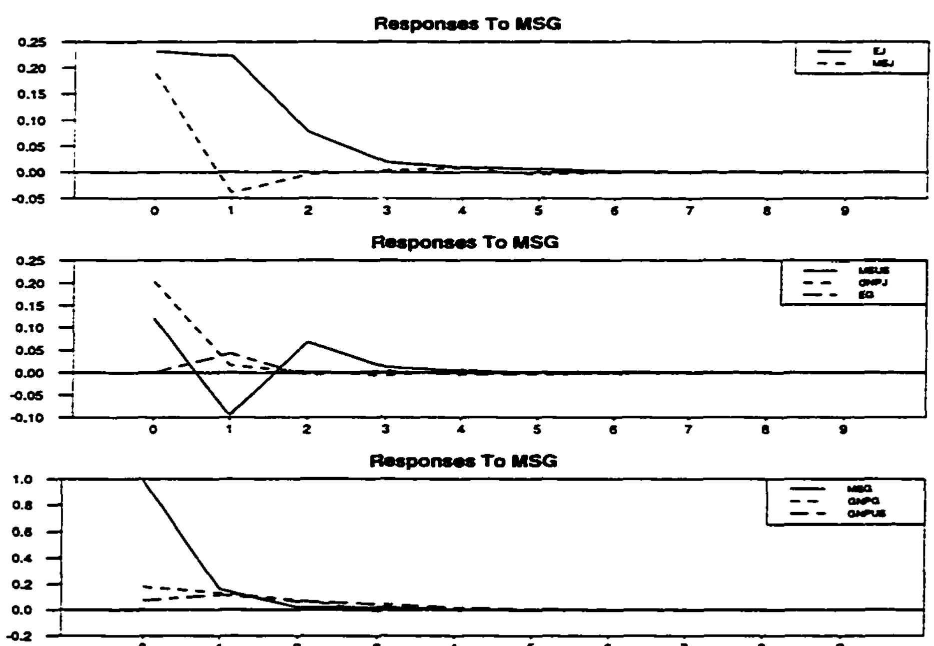

Figure 8.13 Responses to German Money Supply: Germany-Japan-U.S.

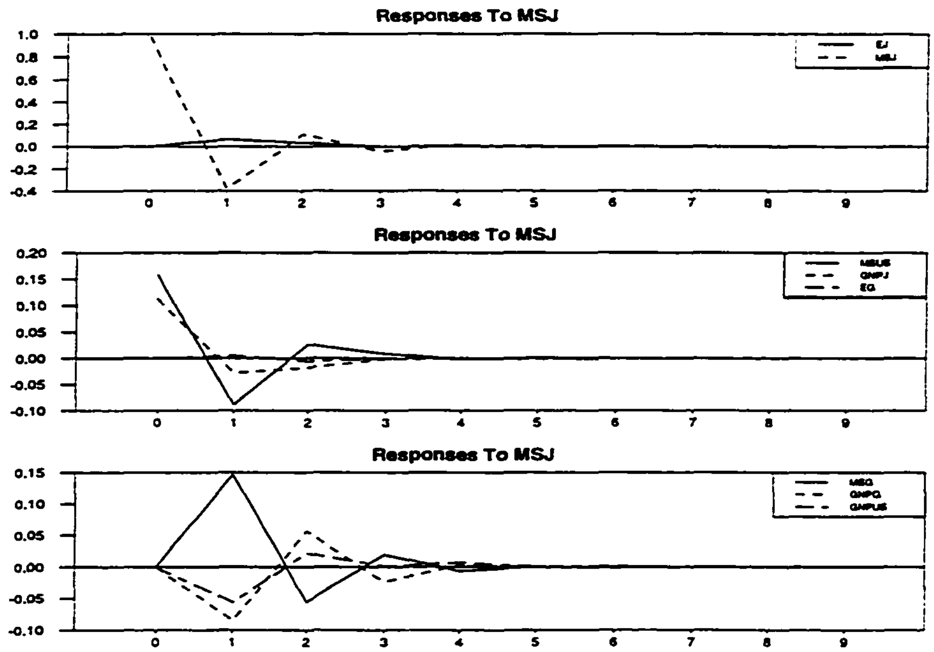

Figure 8.14 Responses to Japanese Money Supply: Germany-Japan-U.S.

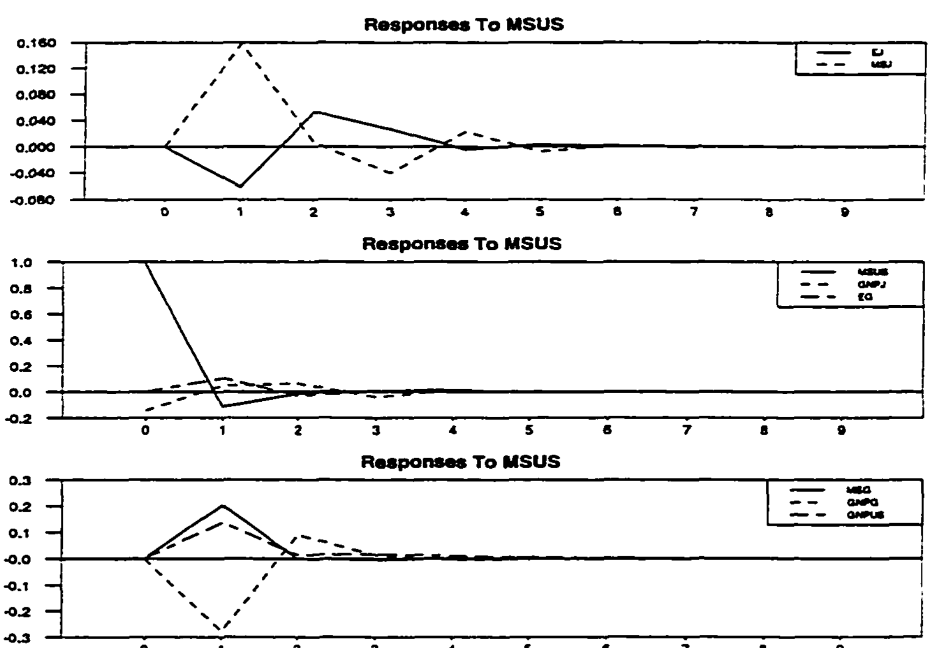

Figure 8.15 Responses to U.S. Money Supply: Germany-Japan-U.S.

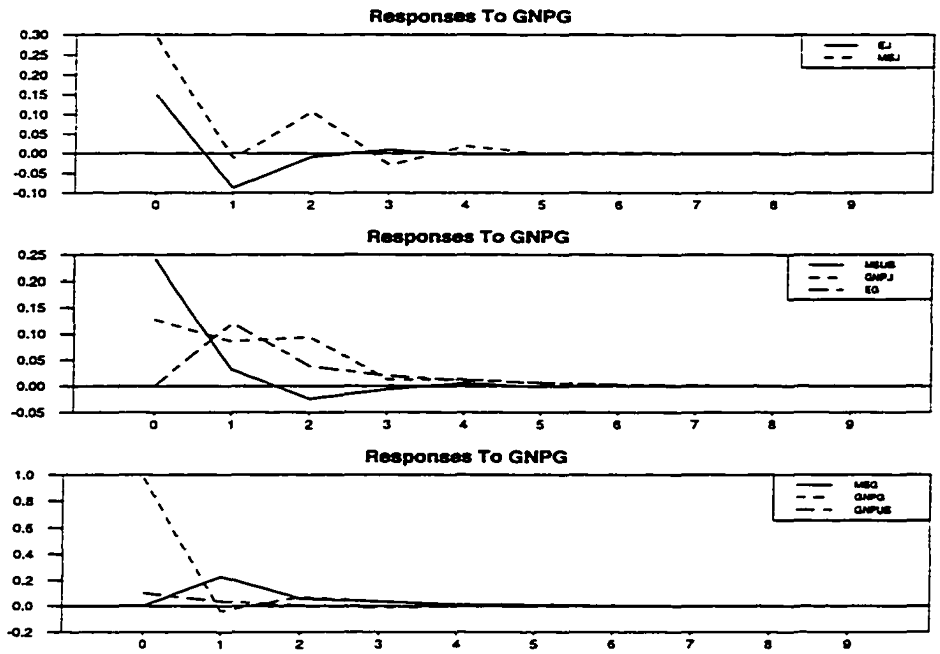

Figure 8.16 Responses to German GNP: Germany-Japan-U.S.

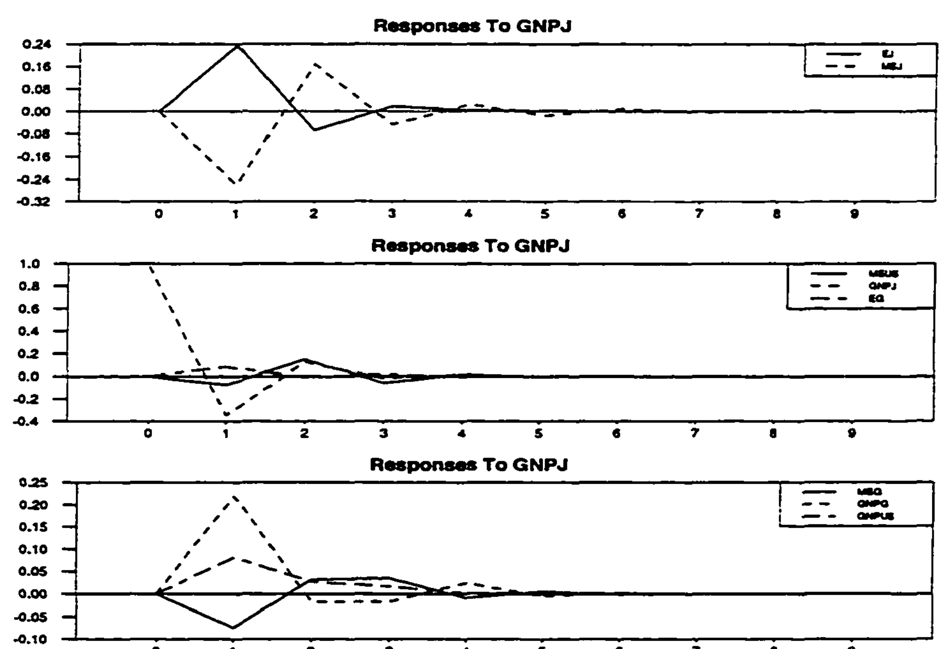

Figure 8.17 Responses to Japanese GNP: Germany-Japan-U.S.

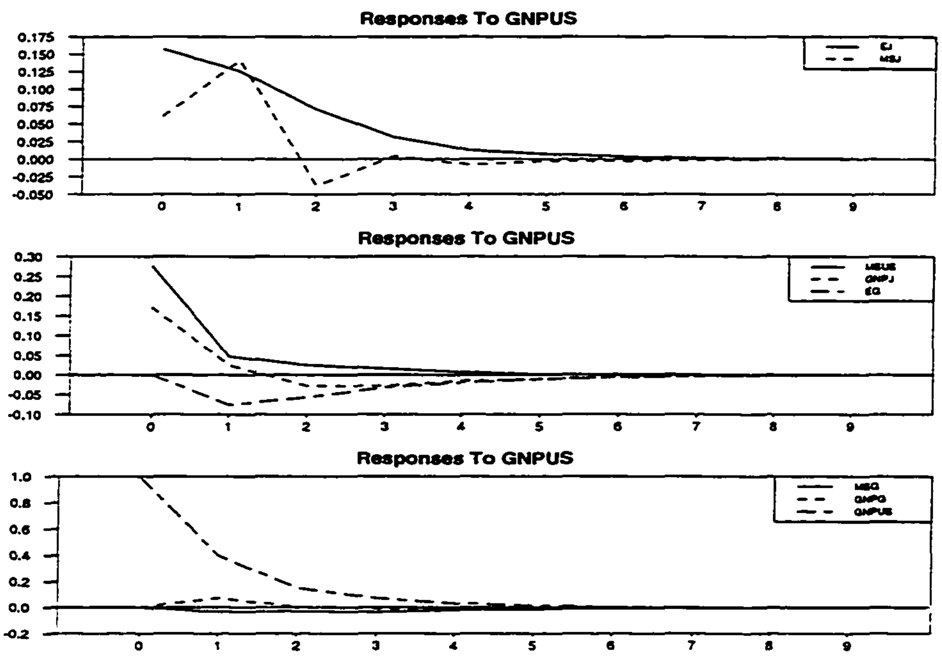

Figure 8.18 Responses to U.S. GNP: Germany-Japan-U.S.

9 CONCLUSION

In this part, a multivariate statistical model, the partial system model, was applied to a data set consisting of exchange rates, money supplies and GNPs. This work was initially inspired by the work done by Dibooglu (1993). This part extended his work by applying Johansen's partial system model, instead of the full system model Dibooglu applied to his data set, which treats some variables in the system as weakly exogenous and the others as endogenous. This made it possible to deal with more variables than the full system model, since the number of parameters to be estimated is fewer than in the full system model. When the third country's variables were added to the system, there were a total of 8 variables in the model; 2 exchange rates, 3 money supplies and 3 GNPs.

First, the existence of unit roots in the time series data was investigated by using the Dickey-Fuller augmented unit root test. This series of tests showed that all the variables in the data set are integrated of order one, i.e., all the variables contain a unit root.

Secondly, the full system model was investigated using the error correction model, and then, the weakly exogenous and endogenous variables in this full system model were determined. In the Germany-U.S. case, 4 weakly exogenous variables in the system were found, while 2 weakly exogenous variables were identified in the Japan-U.S. case. Finally, in the Germany-Japan-U.S. model, 4 weakly exogenous variables were found. As discussed in the text, weak exogeneity is not the same as causality. After identifying weakly exogenous variables, the system was reformulated into the partial system model, and then, the number of cointegrating relations among the variables under the partial system examined. These cointegrating relations were tested by applying the rank and maximum test. As it has often been pointed out, the analysis of cointegrating relations is very sensitive because the distributions of statistics are not ordinary distributions, they depend on nuisance parameters, and the critical values are derived from simulations. It could be argued that the results derived from the model are not robust and some researchers are even sceptical about the procedures. However, research on cointegration analysis under the partial system model has just started. This is one area which promises to be fertile ground for research in the future.

Conintegrating relations are interpreted as long-run equilibrium. The theoretical model is based on Dornbusch's sticky-price model and assumed that all countries in the model are large countries which endogenizes all prices in the system. The results do not completely match the theoretical long-run relations, most notably, the relations between exchange rates and some of the money supplies are not what theories predict. However, the relations between exchange rates and GNPs are as expected. In particular, the third country's variables were tested to see if they have some effect on exchange rate and other variables, since hypothesis testing shows that the effects of the third country's variables on other countries' variables can not be ignored. On occasion, the results for these effects are not consistent with what the theory predicts, and, sometimes, the signs of the coefficients do not agree with the theoretical signs.

To investigate short-run dynamics and long-run effects of the system, impulse response analysis is more appropriate. The coefficients in the long-run equilibrium equation should not be interpreted as the elasticity, which indicates the change in one variable caused by a unit of change in the other variable, and impulse response analysis accounts for changes in all other variables in the system. Chapter 6 presented the results for variance decomposition and impulse response analysis. The results for impulse response analysis are summarized in Tables 8.7-8.9. Because some of the coefficients for money supplies in the long-run equilibrium equations were opposite to the predicted signs, there were similar problems with the relations between exchange rates and some money supplies, i.e., some of the relations between exchange rates and money supplies were not the predicted relations. While evidence indicates that the third country's variables have some explanatory power concerning changes in the variables of the other two countries, in most of the cases, it is difficult to interpret the results when the third country's variables are included.

There are some critiques of impulse response analysis. As discussed in Chapter 6, restrictions were imposed on the system by specifying the ordering of the variables when variance decomposition and impulse response analysis was performed. The ordering that was adopted is one of many possible orderings. In other words, there are other orderings of the variables and different results corresponding to these orderings. This makes some researchers sceptical of the above analysis. In fact, when the orderings were changed, different results (not shown in this part) were obtained. There are other problems that render the interpretation of impulse response analysis difficult. If the model has important variables missing, it may lead to major distortions in the analysis and make the analysis worthless for structural interpretations, although the model may still be useful for predictions. Additional problems result from measurement errors and the use of seasonally adjusted or temporally or contemporaneously aggregated variables.

Variance decomposition analysis is subject to the same criticism as impulse response analysis. First, variance decomposition is not unique since it depends on the choice of transformation. If other possible orderings of the variables were chosen, i.e., another choice of transformation, it would be possible to obtain different results. Although Choleski decomposition was applied in this part, there are other types of decomposition, for instance, Blanchard-Quah decomposition, that might have been considered. Hossain applied and compared two types of decomposition in his paper. The variance decomposition is conditional on the system under consideration, so the results may change if the system is changed by adding or deleting some of the variables from the system. However, here, the results are not so sensitive to the choice of models, i.e., full system model versus partial system model. Measurement errors, seasonal adjustment and the aggregate variables may affect the results for variance decomposition.

The theoretical model is based on Dornbusch's sticky price model with modified assumptions, including an unconditional interest parity assumption which enables interest rates to be removed from the system. This assumption makes the model simpler since the variables are now only prices, money supplies, GNPs and exchange rates, however; this makes comparison of results with others which include interest rates in the system more difficult. There are also some empirical results which refute Dornbusch's sticky price model. In further research, some of the other theoretical models that were reviewed earlier, such as monetary model, portfolio model and currency substitution model, might be extended to a three-country model. As the world economy becomes more interdependent, a particular country's policy will have greater effect on variables in other countries. It will be increasingly important to expand the model while, at the same time, keeping it as simple as possible.

PART II

APPLICATION OF REGIME-SWITCHING STOCHASTIC VOLATILITY MODEL TO EXCHANGE RATES

10 INTRODUCTION

In financial economics, and some areas of econometrics, the volatility of financial assets, including foreign exchange rates, draws researchers' attention. Many researchers have established empirical regularities of financial asset volatility. These regularities are well summarized in Bollerslev, Engle and Nelson (1994). For instance, Bollerslev et al. refer to (1) thick tails, (2) volatility clustering, (3) leverage effect, (4) volatility and serial correlation and (5) co-movement in volatilities. Thick tails refers to the fact that researchers often find the distribution of asset returns tends to have fat tails. The leverage effect refers to the tendency for changes in stock prices to be negatively correlated with volatility. Of course, foreign exchange rates do not necessarily satisfy all of the above regularities, often exhibiting time-varying volatility. This part will pursue the issues of changing volatility over time and volatility clustering. Volatility clustering is described by Mandelbrot (1963) as:

> *large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes.*

For example, as Figure 13.1 shows, one of the characteristics of the foreign currency exchange rates is its time-varying volatility, that is, the phenomenon that a tranquil period is followed by a volatile period.[1]

Volatility in exchange rates is not constant but varies over time. For the last 10 years, many efforts have been made to model the volatility of financial assets including foreign exchange rates. In econometrics, the ARCH (Autoregressive Conditional Heteroscedasticity) and the GARCH (Generalized Autoregressive Conditional Heteroscedasticity) model and their variations have been extensively considered. More recently, the stochastic volatility model and its variations have been considered.

This part will attempt to model time-varying volatility by adopting a switching-regime stochastic volatility model which is a variation of the stochastic volatility model. In Chapter 11, the two basic classes of models will be summarized: ARCH-type model including GARCH model and stochastic volatility model. Chapter 12 will introduce the switching-regime stochastic volatility model which will

[1] This phenomenon could be observed more often and more clearly in other financial markets such as stock markets.

be applied to the data. Chapter 13 will present empirical results and conclusions will be discussed in Chapter 14.

11 TIME-VARYING VARIANCE MODELS

This chapter will summarize three models to describe time-varying variance: ARCH, GARCH and stochastic volatility. For each of these, basic modeling and estimation methods will be illustrated.

11.1 ARCH Model

Since Engle (1982) introduced an ARCH (Autoregressive Conditional Heteroscedasticity) model to model changing variance of the time series over time, the ARCH and GARCH (Generalized Autoregressive Conditional Heteroscedasticity) models have been among the most popular models in econometrics and financial economics to capture time-varying conditional variance.

In this section, a basic ARCH model followed by GARCH and a stochastic volatility model, are outlined, in particular, to illustrate differences between the stochastic volatility model and the ARCH and GARCH model. Extensive discussions on ARCH and GARCH models can be found in Bollerslev et al. (1992), Bollerslev et al. (1994) and Enders (1994). The key idea to capturing the time-varying volatility and volatility clustering is the distinction between the unconditional variance and the conditional variance. The idea is that the conditional variance depends on the information of the past periods and varies over time while the unconditional variance is time-invariant.

Consider the following simple model that sketches the essence of the ARCH model. Suppose that the $\{y_t\}$ process follows an AR(p) process:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{11.1}$$

where $\{\varepsilon_t\}$ is a white noise:

$$\mathrm{E}[\varepsilon_t] = 0 \qquad \text{for all } t \tag{11.2}$$

$$\mathrm{E}[\varepsilon_t \varepsilon_\tau] = \begin{cases} \sigma_\varepsilon^2 & \text{for } t = \tau, \\ 0 & \text{otherwise} \end{cases} \tag{11.3}$$

The $\{y_t\}$ process is covariance-stationary, if all the roots of

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0 \tag{11.4}$$

are assumed to lie outside the unit circle. The mean of the process $\{y_t\}$ is:

$$\mathrm{E}[y_t] = \alpha/(1 - \phi_1 - \phi_2 - \cdots - \phi_p) \tag{11.5}$$

Suppose that the square of $\{\varepsilon_t\}$ itself also follows AR(q):

$$\varepsilon_t^2 = \beta + \theta_1 \varepsilon_{t-1}^2 + \theta_2 \varepsilon_{t-2}^2 + \cdots + \theta_q \varepsilon_{t-q}^2 + \eta_t \tag{11.6}$$

where $\{\eta_t\}$ is also a white noise process:

$$\mathrm{E}[\eta_t] = 0 \qquad \text{for all } t \tag{11.7}$$

$$\mathrm{E}[\eta_t \eta_\tau] = \begin{cases} \sigma_\eta^2 & \text{for } t = \tau, \\ 0 & \text{otherwise.} \end{cases} \tag{11.8}$$

When (11.6), (11.7) and (11.8) hold, the process $\{\varepsilon_t\}$ is said to follow an ARCH(q) process and this will be denoted as $\varepsilon_t \sim ARCH(q)$. A further restriction is required for the ARCH process, the assumption that all the roots of $(1 - \theta_1 z - \theta_2 z^2 - \cdots - \theta_q z^q) = 0$ are outside the unit circle. If this holds, then the unconditional variance of $\{\varepsilon_t\}$ is calculated as:

$$\mathrm{var}[\varepsilon_t] = \mathrm{E}[\varepsilon_t^2] = \beta/(1 - \theta_1 - \theta_2 - \cdots - \theta_q) \tag{11.9}$$

On the other hand, using the assumption that $\{\varepsilon_t\}$ is a white noise process, the conditional variance of $\{\varepsilon_t\}$ based on the observation of time $t-1$ is expressed as:

$$\mathrm{var}[\varepsilon_t | I_{t-1}] = \mathrm{E}[\varepsilon_t^2 | I_{t-1}] = \beta + \theta_1 \varepsilon_{t-1}^2 + \theta_2 \varepsilon_{t-2}^2 + \cdots + \theta_q \varepsilon_{t-q}^2 \tag{11.10}$$

where I_{t-1} is an information set of time $t-1$ or the observations at time $t-1$. It can be seen from (11.9) that the ARCH model is still consistent with the assumption that the unconditional variance is constant.

The unconditional mean and variance of $\{y_t\}$ are the same as (11.5) and (11.9). The conditional mean and variance are still the same as previously:

$$\mathrm{E}[y_t | I_{t-1}] = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} \tag{11.11}$$

$$\mathrm{var}[y_t | I_{t-1}] = \mathrm{var}[\varepsilon_t | I_{t-1}] = \beta + \theta_1 \varepsilon_{t-1}^2 + \theta_2 \varepsilon_{t-2}^2 + \cdots + \theta_q \varepsilon_{t-q}^2 \tag{11.12}$$

It should be noted that the conditional variance is a function of the past realizations.

Some authors use alternative representations; see, for instance, Bollerslev et al. (1992) and Harvey (1993). Following Harvey (1993), some of the other properties of the ARCH model are better illustrated, using the alternative representation.

Suppose that the process $\{y_t\}$ is, instead, expressed as follows:

$$y_t = \sigma_t u_t \tag{11.13}$$

$$\sigma_t^2 = \gamma + \alpha y_{t-1}^2 \tag{11.14}$$

where $\gamma > 0$, $\alpha \geq 0$ and $\{u_t\}$ is n.i.d.(0,1). Two conditions are needed, $\gamma > 0$ and $\alpha \geq 0$, so that σ_t^2 is always nonnegative. Note that the model is conditionally Gaussian and $y_t|y_{t-1} \sim N(0, \sigma_t^2)$. Firstly, the ARCH model is a Martingale Difference (MD) and its unconditional mean is zero and it is serially uncorrelated.[1] If $0 < \alpha < 1$, the unconditional variance of $\{y_t\}$ can be written as:

$$\text{var}[y_t] = \text{E}[y_t^2] = \gamma/(1-\alpha) \tag{11.15}$$

Therefore, the ARCH process is a white noise though it is not a strict white noise. Although it is conditionally Gaussian, the process is not unconditionally Gaussian. It is also noted that the kurtosis, $3(1-\alpha^2)/(1-3\alpha^2)$, is greater than 3 if $3\alpha^2 < 1$. This implies that the data distribution has heavier tails than the normal distribution whose kurtosis is 3. Hence, the ARCH model can take into account another regularity that many of the financial data show, leptkurtosis. In other words, the ARCH model can explain the data which are generated by a fat-tailed distribution. Using (11.1) and (11.2), it can be shown that the squared observations, $\{y_t^2\}$, actually follow an AR(1) process. The ACF of $\{y_t^2\}$ is written as:

$$\rho(\tau, y_t^2) = \sigma^\tau \qquad \tau = 0, 1, 2, \cdots \tag{11.16}$$

The MSE of the prediction under the alternative model is:

$$\text{MSE}(\tilde{y}_{T+1|T}) = \gamma(1 + \alpha + \alpha^2 + \alpha^3 + \cdots + \alpha^{l-1}) + \alpha^l y_T^2 \tag{11.17}$$

Note that as $l \to \infty$, the expression of (11.17) will tend to that of (11.15) since $0 < \alpha < 1$. When the value of l is finite and small, the two expressions are different.

[1] The $\{y_t\}$ process is called an MD when $\{y_t\}$ satisfies:

$$\text{E}[y_t|I_{t-1}] = 0.$$

11.2 Generalized ARCH Model

A natural extension of the ARCH model is the Generalized ARCH model or GARCH(p,q), which was first introduced by Bollerslev (1986). The model assumes that the conditional variance follows an ARMA(p,q) process instead of an AR process. The conditional variance $\{h_t\}$, will be written as follows:

$$h_t = \beta + \theta_1 \varepsilon_{t-1}^2 + \theta_2 \varepsilon_{t-2}^2 + \cdots + \theta_p \varepsilon_{t-p}^2 + \phi_1 h_{t-1} + \cdots + \phi_q h_{t-q} \tag{11.18}$$

or

$$h_t = \Phi(L)^{-1}\beta + \Phi(L)^{-1}\Theta(L)\varepsilon_t^2 \tag{11.19}$$

where $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_p L^p$ and $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_q L^q$. L is a lag operator. All the roots of $\Phi(z) = 0$ are assumed to lie outside the unit circle. Restrictions must be imposed on the parameters so that the conditional variance is nonnegative. In a simple GARCH(1,1) model, the restriction is equivalent to both θ_1 and ϕ_1 being nonnegative. To determine the orders of p and q, the usual ACF/PACF techniques will be applied to the residuals. Hamilton (1994) shows that if $\{\varepsilon_t\}$ follows a GARCH(p,q) process, then $\{\varepsilon_t^2\}$ is described by an ARMA(m,p), where $m = \max(p,q)$. By observing ACF and PACF of $\{\varepsilon_t^2\}$, the range of the possible orders of p and q can be narrowed down. Bollerslev et al. (1992) point out that $p = q = 1$ is sufficient in most of the empirical cases.

11.3 Estimation Methods

There are three principal methods to estimate the ARCH and GARCH model; the maximum likelihood method, the quasi-maximum likelihood method and the method of moments. See Hamilton (1994), Bollerslev et al. (1992) and Bollerslev (1994) for the detailed discussions on these three methods. Here, only basic ideas are illustrated.

11.3.1 Maximum Likelihood Method

To explain the maximum likelihood method, consider the following model:

$$y_t = x_t'\beta + u_t \tag{11.20}$$

where x_t is a vector of explanatory variables and β is a vector of coefficients. Suppose the error term $\{u_t\}$ follows an ARCH process:

$$u_t = \sqrt{h_t}\nu_t \tag{11.21}$$

$$E[\nu_t] = 0, \qquad E[\nu_t^2] = 1 \qquad \text{for all } t \tag{11.22}$$

The conditional variance, $E[u_t^2|u_{t-1}, u_{t-2}, \cdots] \equiv h_t$, is assumed to evolve as follows:

$$h_t = \zeta + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_m u_{t-m}^2 \tag{11.23}$$

This indicates that u_t follows an ARCH(m) process. By conditioning on the first m observations, T numbers of observations are used to estimate parameters. There will be, at time t, the following vector of the observations:

$$z_t \equiv (y_t, y_{t-1}, \cdots, y_0, \cdots, y_{-m+1}; x_t, x_{t-1}, \cdots, x_0, \cdots, x_{-m+1}) \tag{11.24}$$

If it is assumed that u_t has a Gaussian distribution $N(0,1)$ and is independent of both x_t and z_{t-1}. Then, a joint distribution of y_t can be written as:

$$f(y_t|x_t, z_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} \exp\left\{\frac{-(y_t - x_t'\beta)^2}{2h_t}\right\} \tag{11.25}$$

where

$$h_t = \zeta + \alpha_1(y_{t-1} - x_{t-1}'\beta)^2 + \cdots + \alpha_m(y_{t-m} - x_{t-m}'\beta)^2 \tag{11.26}$$

Hence, the log likelihood function conditional on the first m observations will be:

$$\begin{aligned} l(\theta) &= \sum_{t=1}^{T} \log f(y_t|x_t, z_{t-1}; \theta) \\ &= -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(h_t) - \frac{1}{2}\sum_{t=1}^{T}(y_t - x_t'\beta)^2/h_t \end{aligned} \tag{11.27}$$

To maximize the log likelihood function in (11.27), several techniques are available such as the method of scoring (Engle (1982)), or the BHHM algorithm (Berndt, Hall, Hall, and Hausman (1974) and Bollerslev (1986)).

Some researchers have attempted to extend the above model to incorporate the empirical regularity that many financial data come from the fat-tailed distribution. For instance, Bollerslev (1987) considers a non-Gaussian distribution case. He assumes that u_t has a t-distribution with κ degrees of freedom and a scale parameter, M_t, which is also a parameter to be estimated by maximum likelihood method. In this case, the density function is written as:

$$f(u_t) = \frac{\Gamma[(\kappa+1)/2]}{(\pi\kappa)^{1/2}\Gamma(\kappa/2)} M_t^{-1/2}[1 + \frac{u_t^2}{M_t\kappa}]^{-(\kappa+1)/2} \tag{11.28}$$

where $\Gamma[\cdot]$ is the gamma function. The t-distribution is symmetric around zero and its kurtosis is $3(\kappa-2)/(\kappa-4)$ which is greater than 3 if $\kappa > 4$. The conditional variance, then, is:

$$E[u_t^2] = \frac{M_t\kappa}{(\kappa - 2)} \tag{11.29}$$

If $M_t = h_t(\kappa - 2)/\kappa$, then the density becomes:

$$f(u_t) = \frac{\Gamma[(\kappa+1)/2]}{(\pi)^{1/2}\Gamma(\kappa/2)}(\kappa-2)^{-1/2}h_t^{-1/2}[1 + \frac{u_t^2}{h_t(\kappa-2)}]^{-(\kappa+1)/2} \tag{11.30}$$

Using (11.30) instead of (11.28), the following log likelihood function conditional on the first m observations is obtained:

$$\begin{aligned} l(\theta) &= \sum_{t=1}^{T} \log f(y_t|x_t, z_{t-1}; \theta) \\ &= T \log \frac{\Gamma[(\kappa+1)/2]}{\pi^{1/2}\Gamma(\kappa/2)}(\kappa-2)^{-1/2} - \frac{1}{2}\sum_{t=1}^{T}\log(h_t) \\ &\quad - \frac{\kappa+1}{2}\sum_{t=1}^{T}[1 + (y_t - x_t'\beta)^2/h_t(\kappa-2)] \end{aligned} \tag{11.31}$$

where

$$\begin{aligned} h_t &= \zeta + \alpha_1(y_{t-1} - x_{t-1}'\beta)^2 + \cdots + \alpha_m(y_{t-m} - x_{t-m}'\beta)^2 \\ &= [w_t(\beta)]'\delta \end{aligned} \tag{11.32}$$

where

$$[w_t(\beta)]' \equiv \left[\begin{array}{cccc} 1 & (y_{t-1} - x_{t-1}'\beta)^2 & \cdots & (y_{t-m} - x_{t-m}'\beta)^2 \end{array}\right]' \tag{11.33}$$

$$\delta \equiv \left[\begin{array}{cccc} \zeta & \alpha_1 & \cdots & \alpha_m \end{array}\right]' \tag{11.34}$$

Again, by using the available methods, the maximum likelihood estimates can be found numerically. For other distributions than t-distribution, Jorion (1988) proposes a normal-Poisson mixture distribution. Baillie and Bollerslev (1989) considers power exponential distribution and Hsieh (1989) uses normal-log normal mixture.

11.3.2 Quasi-Maximum Likelihood Method

Weiss (1984, 1986), Bollerslev and Woodridge (1992), and Glosten, Jagannathan and Runkle (1989) pointed out that the maximum likelihood method discussed in the above will provide consistent estimates even when u_t has a non-Gaussian distribution, if it is assumed:

$$E[\nu_t|x_t, z_{t-1}] = 0 \tag{11.35}$$

and

$$E[\nu_t^2|x_t, z_{t-1}] = 1 \tag{11.36}$$

They showed that, under certain regularity conditions, the following will hold:

$$\sqrt{T}(\hat{\theta}_T - \theta) \xrightarrow{d} N(0, D^{-1}SD^{-1}) \tag{11.37}$$

where $\hat{\theta}_T$ is the estimate and θ is the true value. S and D in (11.37) are:

$$S = \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T}[s_t(\theta)]\cdot[s_t(\theta)] \tag{11.38}$$

where $s_t(\theta)$ is a score vector calculated by[2]:

$$s_t(\theta) = \frac{\partial \log f(y_t|x_t, z_{t-1};\theta)}{\partial\theta} \tag{11.39}$$

and

$$D = \operatorname*{plim}_{T\to\infty} T^{-1}\sum_{t=1}^{T} -E\left\{\frac{\partial s_t(\theta)}{\partial\theta'}|x_t, z_{t-1}\right\} \tag{11.40}$$

S and D are consistently estimated by:

$$\hat{S}_T = T^{-1}\sum_{t=1}^{T}[s_t(\hat{\theta}_T)]\cdot[s_t(\hat{\theta}_T)]' \tag{11.41}$$

$$\hat{D}_T = T^{-1}\sum_{t=1}^{T}\{\frac{1}{2\hat{h}_t^2}\begin{bmatrix} \sum_{j=1}^{m} -2\hat{\alpha}_j\hat{u}_{t-j}x_{t-j} \\ w_t(\beta) \end{bmatrix}$$

$$\times\begin{bmatrix} \sum_{j=1}^{m} -2\hat{\alpha}_j\hat{u}_{t-j}x'_{t-j} & w_t(\hat{\beta})' \end{bmatrix} + \frac{1}{\hat{h}_t}\begin{bmatrix} x_t x_t' & 0 \\ 0 & 0 \end{bmatrix}\} \tag{11.42}$$

Note that if the data were generated from Gaussian distribution, then $S = D$ holds.

11.3.3 Generalized Method of Moments

The third method to estimate the parameters is generalized method of moments. To apply this method, two conditions must be satisfied. The first one, from (11.20), is that the residual in the regression is orthogonal with the explanatory variables, x_t:

$$\mathrm{E}[u_t x_t] = 0 \tag{11.43}$$

The second condition is the implicit error in forecasting that the squared residual is orthogonal with lagged squared residuals:

$$\mathrm{E}[(u_t^2 - h_t)w_t] = 0 \tag{11.44}$$

[2]Derivations are in Hamilton (1994).

To minimize, $\theta = (\beta', \delta,)'$ is chosen:

$$g(\theta : z_t)' \hat{s}_T' g(\theta : z_t) \tag{11.45}$$

where

$$g(\theta : z_t) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T}(y_t - x_t'\beta_t)x_t \\ \frac{1}{T}\sum_{t=1}^{T}\{(y_t - x_t'\beta_t)^2 - w_t(\beta)'\delta\}w_t(\beta) \end{bmatrix} \tag{11.46}$$

After deriving the first order conditions from (11.45), estimates of the parameters can be found numerically. Further discussion on generalized method of moments can be found in Hamilton (1994).

11.4 Stochastic Volatility Model

The stochastic volatility model is another way to capture the time-varying volatility of the time series data. Although the model imposes less restrictions and fits in a theoretical framework more naturally than the ARCH and GARCH model, it is very difficult to obtain the exact likelihood function for the stochastic model and to estimate by maximum likelihood method, since the likelihood function is an N-dimensional integral, where N is the number of observations. Thus, its empirical application has been limited. While the ARCH and GARCH model assume that the conditional variance is a function of the past variance and the squares of the past observations, this approach assumes that variance is an unobservable variable that follows some stochastic process, for example, an AR process. Another advantage of the stochastic volatility model is that the extension to multivariate models is more natural: see Harvey et al. (1994). Following Harvey (1993) and Harvey et al. (1994), this section discusses a simple univariate stochastic volatility model.

Consider the following simple univariate model:

$$y_t = \exp\left\{\frac{\alpha_t}{2}\right\}\varepsilon_t \qquad t = 1, 2, \cdots, T \tag{11.47}$$

where $\varepsilon_t \sim NID(0, 1)$ and α_t is assumed to follow a stochastic process, say, an AR(1) process:

$$\alpha_t = \gamma + \phi\alpha_{t-1} + \eta_t \tag{11.48}$$

where $\eta_t \sim NID(0, \sigma_\eta^2)$. It is also assumed that the processes $\{\varepsilon_t\}$ and $\{\eta_t\}$ are independent of each other for all t. If $|\phi| < 1$, then the process $\{\alpha_t\}$ is stationary with mean and variance:

$$\mathrm{E}[\alpha_t] = \gamma_\alpha = \frac{\gamma}{(1-\phi)} \tag{11.49}$$

$$\mathrm{Var}[\alpha_t] = \sigma_\alpha^2 = \frac{\sigma_\eta^2}{(1-\phi^2)} \tag{11.50}$$

Harvey et al. (1994) point out that the restrictions necessary to ensure the stationarity of the process $\{y_t\}$ are the ones to ensure the stationarity of the process $\{\alpha_t\}$ because the process $\{y_t\}$ is a product of two stationary processes. Since the processes $\{\varepsilon_t\}$ and $\{\eta_\tau\}$ are independent of each other for all t and τ, the process $\{y_t\}$ is a white noise process. Its mean and autocovariance are:

$$\mathrm{E}[y_t] = \mathrm{E}\left[\exp\left\{\frac{\alpha_t}{2}\right\}\right]\mathrm{E}[\varepsilon_t] = 0 \qquad \forall t \tag{11.51}$$

and

$$\begin{aligned}
\mathrm{E}[y_t y_\tau] &= \mathrm{E}\left[\varepsilon_t \exp\left\{\frac{\alpha_t}{2}\right\}\varepsilon_\tau \exp\left\{\frac{\alpha_\tau}{2}\right\}\right] \\
&= \mathrm{E}[\varepsilon_t\varepsilon_\tau]\mathrm{E}\left[\exp\left\{\frac{(\alpha_t+\alpha_\tau)}{2}\right\}\right] \\
&= \begin{cases} \exp\left\{(\gamma_\alpha + \frac{1}{2}\sigma_\alpha^2)\right\} & t=\tau, \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{11.52}$$

The odd moments of the process $\{y_t\}$ are all zero because of the symmetry of $\{\varepsilon_t\}$. The even moments are derived by using the properties of log-normal distribution, $\exp\{\alpha_t\}$:

$$\mathrm{E}[\exp\{j\alpha_t\}] = \exp\left\{j\gamma_\alpha + \frac{1}{2}j^2\sigma_\alpha^2\right\} \tag{11.53}$$

Most importantly, the fourth moment exists and it is:

$$\begin{aligned}
\mathrm{E}[y_t^4] &= \mathrm{E}[\varepsilon_t^4]\mathrm{E}[\exp\{2\alpha_t\}] \\
&= 3\exp\{2\gamma_\alpha + 2\sigma_\alpha^2\}
\end{aligned} \tag{11.54}$$

The Kurtosis is, then, calculated:

$$\begin{aligned}
m_3 &= \frac{\mathrm{E}[y_t^4]}{\{\mathrm{E}[y_t^2]\}^2} \\
&= \frac{3\exp\{2\gamma_\alpha + 2\sigma_\alpha^2\}}{[\exp\{\gamma_\alpha + \frac{1}{2}\sigma_\alpha^2\}]^2} \\
&= 3\exp\{\sigma_\alpha^2\}
\end{aligned} \tag{11.55}$$

So, if σ_α^2 is positive then the kurtosis is greater than 3, which describes a fat-tailed distribution. It is sometimes useful to use a transformed process, $\{\log y_t^2\}$, rather than the process $\{y_t\}$ to capture the properties of the dynamics. From (11.47):

$$\log(y_t^2) = \alpha_t + \log(\varepsilon_t^2) \tag{11.56}$$

Since $\{\varepsilon_t\}$ has a standard normal distribution, $\log(\varepsilon_t^2)$ has the mean -1.27 and the variance $\pi^2/2 = 4.93$. Then, (11.56) can be written as:

$$\begin{aligned} \log(y_t^2) &= \alpha_t + \log(\varepsilon_t^2) - 1.27 + 1.27 \\ &= -1.27 + \alpha_t + \varepsilon_t^* \end{aligned} \tag{11.57}$$

where $\varepsilon_t^* = \log(\varepsilon_t^2) + 1.27$. Hence, $\log(y_t^2)$ is the sum of an AR(1) process and a white noise. That is, $\log(y_t^2)$ is an ARMA(1,1) process with autocorrelation function:

$$\rho_\tau = \frac{\phi^\tau}{1 + 4.93/\sigma_\alpha^2}. \tag{11.58}$$

The model can be generalized by assuming that the process $\{\alpha_t\}$ follows any stationary ARMA(p, q) process. Then, the process $\{y_t\}$ still follows a stationary process.

Another direction of the generalization is to assume a non-normal distribution for $\{\varepsilon_t\}$ as the ARCH model is generalized by using t-distribution.

Suppose that the process $\{\varepsilon_t\}$ has a t-distribution. The t-distribution is:

$$t = \frac{z}{\sqrt{v}} \tag{11.59}$$

where $z \sim N(0,1)$ and $\nu v \sim \chi^2(\nu)$. z and v are independent. Hence, $\{\varepsilon_t\}$ can be written as:

$$\varepsilon_t = \frac{\zeta_t}{\sqrt{\kappa_t}} \tag{11.60}$$

where $\zeta_t \sim N(0,1)$ and $\nu\kappa_t \sim \chi^2(\nu)$, ν degrees of freedom. Then, from (11.60):

$$\log \varepsilon_t^2 = \log \zeta_t^2 - \log \kappa_t \tag{11.61}$$

where $\log \kappa_t$ is a log of $\frac{\chi^2(\nu)}{\nu}$ and its expectation and variance are, respectively:

$$\mathrm{E}[\log \kappa_t] = \Psi\left(\frac{\nu}{2}\right) - \log\left(\frac{\nu}{2}\right) \tag{11.62}$$

$$\mathrm{Var}[\log \kappa_t] = \Psi'\left(\frac{\nu}{2}\right) \tag{11.63}$$

where $\Psi(\cdot)$ is the digamma and $\Psi'(\cdot)$ is the trigamma function. Substituting these results into (11.57) gives:

$$\begin{aligned} \log(y_t^2) &= -1.27 + \alpha_t + \varepsilon_t^* - \mathrm{E}[\log \kappa_t] + \mathrm{E}[\log \kappa_t] \\ &= -1.27 - \left\{\Psi\left(\frac{\nu}{2}\right) - \log\left(\frac{\nu}{2}\right)\right\} + \alpha_t + \varepsilon_t^{**} \end{aligned} \tag{11.64}$$

where $\varepsilon_t^{**} = \varepsilon_t^{*} + E(\log \kappa_t)$. The expectation and variance of the process $\{\varepsilon_t^{**}\}$ are:

$$\mathrm{E}[\varepsilon_t^{**}] = 0 \tag{11.65}$$

$$\mathrm{Var}[\varepsilon_t^{**}] = 4.93 + \Phi'\left(\frac{\nu}{2}\right) \tag{11.66}$$

Again, $\log(y_t^2)$ is a sum of the AR(1) process and the white noise. The ACF has the following form:

$$\rho_\tau = \frac{\phi^\tau}{\{1 + [\Psi'(\frac{\nu}{2}) + 4.93]/\sigma_\alpha^2\}} \qquad \tau = 1, 2, \cdots \tag{11.67}$$

11.5 Estimation Methods

There are mainly three methods of estimating a stochastic volatility model: method of moments, quasi-maximum likelihood method, and Bayesian approach. Since the Bayesian approach is applied to the model in the next chapter, it will be discussed there.

11.5.1 Method of Moments

There is some work on parameter estimation based on the method of moments; see Wiggins (1987), and Melino and Turnbull (1990). Melino and Turnbull point out that the work done by the other three has found the sensitivity of the parameters to the moments they fitted but that they could not test whether the different parameters they obtained are due to sampling error. This section illustrates the generalized method of moments procedure used by Melino and Turnbull. In their paper, Melino and Turnbull estimated a U.S.-Canada daily exchange rate with about 3,000 observations using a stochastic volatility model. Their data are unevenly spaced. The estimated equations are:

$$S(t_i) = ah_i + (1 + bh_i)S(t_{i-1}) + v(t_{i-1})S(t_{i-1})^{\beta/2}h_i^{1/2}e(t_i) \tag{11.68}$$

and

$$\ln v(t_i) = \alpha h + (1 + \delta h)\ln v(t_i - h) + \gamma h^{1/2}u(t_i) \tag{11.69}$$

where $S(t_i)$ is a spot exchange rate at time t_i, $v(t_i)$ a level of a volatility, $h_i = t_i - t_{i-1}$, and $h = min\{h_i\}$. Two error terms are assumed:

$$\begin{pmatrix} e(t_i) \\ u(t_i) \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right) \tag{11.70}$$

They demonstrate that if $\delta < 0$ and the appropriate initial conditions are met, the even spacing discrete time approximation will lead to the stationary volatility and that:

$$\ln v_t \sim N(\mu_v, \sigma_v^2) \tag{11.71}$$

where $\mu_v = -\frac{\alpha}{\delta}$ and $\sigma_v^2 = \frac{h\gamma^2}{[1-(1+\delta h)^2]}$. Melino and Turnbull define $\theta = (a, b, \alpha, \delta, \gamma, \rho; \beta)$ and $w_i(\theta)$ by:

$$w_i(\theta) = \frac{S(t_i) - ah_i - (1 + bh_i)S(t_{i-1})}{[h_i S^{\beta}(t_{i-1})]^{1/2}} \tag{11.72}$$

where $w_i(\theta)$ represents the normalized one-observation-ahead forecast errors. In general the expectation of functions of w_i will be functions of θ. The method of moments estimates the parameters θ by equating the computed sample moments of these functions to their population moments. They consider the following functions in reference to the three criteria; familiarity, identification, and efficiency. See Melino and Turnbull for a detailed discussion:

$$w_i^m(\theta) \qquad m = 1, 2, 3, \cdots \tag{11.73}$$

$$|w_i^m(\theta)| \qquad m = 1, 2, 3, \cdots \tag{11.74}$$

$$w_i(\theta) w_{i-j}(\theta) \qquad j = 1, 2, 3, \cdots \tag{11.75}$$

$$|w_i(\theta) w_{i-j}(\theta)| \qquad j = 1, 2, 3, \cdots \tag{11.76}$$

$$w_i^2(\theta) w_{i-j}^2(\theta) \qquad j = 1, 2, 3, \cdots \tag{11.77}$$

$$|w_i(\theta)| w_{i-j}(\theta) \qquad j = 0, \pm 1, \pm 2, \pm 3, \cdots \tag{11.78}$$

Melino and Turnbull provide the unconditional expectation of these functions in the appendix. Then, they follow Hansen's (1982) general framework. $f_i(\theta) \in R^p$ denotes a vector whose components are function of w_i, and $g_n(\theta)$ is defined to be:

$$g_n(\theta) = \frac{1}{n} \sum_{i=1}^{n} f_i(\theta) \tag{11.79}$$

Then, an optimal θ will be chosen as:

$$\hat{\theta}_n = \operatorname*{argmin}_{\theta \in \Theta} g_n'(\theta) \hat{W}_n g_n(\theta) \tag{11.80}$$

where Θ is the permissible parameter space and $\hat{W}_n$ is a positive definite weighting matrix. Under certain regularity conditions, $\hat{\theta}_n$ is consistent and asymptotically normal:

$$n^{-1/2}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, V_n) \tag{11.81}$$

where V_n can be consistently estimated by:

$$\hat{V}_n = (D_n' \hat{W}_n D_n)^{-1} D_n' \hat{W}_n \hat{\Sigma}_n \hat{W}_n D_n (D_n' \hat{W}_n D_n)^{-1} \tag{11.82}$$

where $D_n'(\theta) = dg_n(\theta)/d\theta$, that is, the Jacobian matrix. Melino and Turnbull estimated $\hat{\Sigma}_n$ by using the Newey-West method and set $\hat{W}_n = \hat{\Sigma}_n^{-1}$ for the simplicity. The results are presented in section 4.2 in their paper.

11.5.2 Quasi-Maximum Likelihood Method

This method has been discussed in many papers; see Harvey, Ruiz, and Shepard (1994), Ruiz (1994), Kim and Shepard (1994), Jacquier, Polson, and Rossi (1994) and Breidt and Carriquiry (1996). In this section, a framework of the method based on the above papers is presented. As Ruiz (1994) and Jacquier et al. (1994) point out, the method of moments estimates are inefficient and show poor performances over repeated samplings relative to likelihood-based estimates. Jacquier et al. show that this problem is particularly serious in a stochastic volatility case because it is difficult to choose moments to be computed without the help of the score function. Harvey et al., Ruiz, and Kim et al. transform the SV model to a linear model in a state-space model and use the Kalman filter to estimate the unobservable volatility and a quasi-maximum likelihood function to obtain the parameters. In the simple model used in this part, (11.48) can be considered to be the transition equation and (11.57) can be seen as the measurement equation. Harvey and Shepard (1992) showed that η_t in (11.48) and ε_t^* in (11.57) are uncorrelated even if η_t in (11.48) and ε_t in (11.47) are correlated. As seen in the above, $\varepsilon_t^* = \log \varepsilon_t^2 + 1.27$ does not have a Gaussian distribution. In other words, if the Kalman filter is applied, the estimates are the MMSLE (Minimum Mean Square Linear Estimator), but not the MMSE (Minimum Mean Square Estimator). An exact likelihood function cannot be obtained from the Kalman filter because the model does not have a conditional Gaussian distribution. However, the model can be treated as if it had a Gaussian distribution and the quasi-maximum likelihood function can be maximized instead of the exact likelihood function. Ruiz (1992) points out that the assumption that ε_t is a Gaussian will not improve the efficiency even if it is true while Harvey states that if the distribution of ε_t is not specified, the level of volatility is not identified because E[$\log \varepsilon_t{}^2$] is unknown. If the distribution of ε_t^* is assumed to be a t-distribution, then ν can be obtained from (11.68). Then, E[$\log \varepsilon_t{}^2$] can be computed by (11.61) and (11.62). Breidt and Carriquiry propose another transformation that is called the robustified transformation instead of a square-log transformation that we considered in (11.56) and apply the quasi-maximum likelihood method.

12 REGIME-SWITCHING STOCHASTIC MODEL

This section discusses the principal model; the regime-switching stochastic volatility (RSSV) model, two different version of the model. The first model is an extension of Schmidt's model (1996), different only in that four regimes are used here while there are two regimes in Schmidt's model. The second model is a mean model which considers an explicit relation between exchange rate and interest rate and assumes that an error term will explain volatility in exchange rates. Finally, the structural model, derived from the interest parity condition, will be assumed throughout.

Discussion for the first model in this section will follow Schmidt's discussion with some modifications.

12.1 An Extension of Schmidt's Model

The first model, a simple extension of Schmidt's model, will be expressed as follows:

$$\begin{aligned} y_t &= \exp\left\{\frac{\alpha_t}{2}\right\}\zeta_t \\ \alpha_t &= \beta_{s_t} + \sigma_{s_t}\eta_t \end{aligned} \tag{12.1}$$

where the two errors, η_t and ζ_t, are assumed to have normal distributions with mean zero and variance one:

$$\begin{aligned} &\zeta_t \ iid\ N(0,1) \\ &\eta_t \ iid\ N(0,1) \end{aligned}$$

As Schmidt points out, a mean of zero in η_t is logical since the mean in the β_{s_t} term can be accounted for and σ_{s_t} can account for the variance process, as many researchers have found[1], the expected change of the exchange rate is assumed to be zero and the assumption that ζ_t has a mean zero is also valid.

While Schmidt discusses the case where α_t follows an AR(1) process; $\alpha_t = \beta_{s_t} + \phi_{s_t}\alpha_{t-1} + \sigma_{s_t}\eta_t$, this part simply applies the case where α_t has a constant term and an error term in order to keep the model simple. The state of the economy is represented by s_t. There are four states of the economy in this model, so that s_t takes four values; $s_t = 1, 2, 3, 4$. The four economic states will be determined

[1] Many researchers have found that short-run exchange rates such as daily exchange rates follow a random walk process. For instance see Meese and Rogoff (1983).

by the following two factors; (a) two observable economic states which is a change in the difference of the interest rates in the two countries, and (b) two unobservable economic states A and B. If capital mobility is assumed, the interest parity condition holds and the difference of the two interest rates is the expected appreciation (depreciation) of domestic currency, then it is possible to examine how the higher expected appreciation or depreciation will affect a change in exchange rate. In other words, whether higher expected appreciation (depreciation) will induce exchange rate to be more volatile or less volatile. Defining x_t as the difference between the foreign and U.S. interest rate, say, French interest rate – U.S. interest rate, then the four economic are defines as follows:

state 1: $|x_t| \geq k$ and unobservable state A

state 2: $|x_t| \geq k$ and unobservable state B

state 3: $|x_t| < k$ and unobservable state A

state 4: $|x_t| < k$ and unobservable state B

where k is some fixed number. The economy is in state 1 if the interest rate differential is greater than or equal to some fixed value k and the economy is in unobservable state A. If the economy is in state 1, the change in exchange rate on day t, then y_t, will be modeled as:

$$y_t = \exp\left\{\tfrac{1}{2}(\beta_1 + \sigma_1\eta_t)\right\}\zeta_t$$

Similarly if the economy is in state 2 where the interest rate differential is greater than or equal to some fixed value k and the economy is in unobservable state B, then y_t will be modeled as:

$$y_t = \exp\left\{\tfrac{1}{2}(\beta_2 + \sigma_2\eta_t)\right\}\zeta_t.$$

To simplify the model assumes that σ differs depending only on the unobservable states. In other words, it is assumed that $\sigma_1 = \sigma_3$ and $\sigma_2 = \sigma_4$. In Table 12.1, y_t in all four states is summarized.

Table 12.1 States of the Economy and y_t

	Unobservable State A	Unobservable State B
$\lvert x_t\rvert \geq k$	1.$\exp\left\{\frac{1}{2}(\beta_1 + \sigma_1\eta_t)\right\}\zeta_t$	2.$\exp\left\{\frac{1}{2}(\beta_2 + \sigma_2\eta_t)\right\}\zeta_t$
$\lvert x_t\rvert < k$	3.$\exp\left\{\frac{1}{2}(\beta_3 + \sigma_1\eta_t)\right\}\zeta_t$	4.$\exp\left\{\frac{1}{2}(\beta_4 + \sigma_2\eta_t)\right\}\zeta_t$

To describe the process of switching states from 1 or 3 to 2 or 4, or vice versa, a Markov chain model is applied. As in Schmidt, a fixed transition probability matrix is assumed in this Markov chain:

$$P(s_t = 1 \text{ or } 3 | s_{t-1} = 2 \text{ or } 4) = \varepsilon_1$$

$$P(s_t = 2 \text{ or } 4 | s_{t-1} = 1 \text{ or } 3) = \varepsilon_2 \qquad (12.2)$$

Equation (12.2) above indicates that the probability that the state of economy shifts from the state 2 or 4 to the state 1 or 3 is ε_1 and the probability that the state of economy moves from the state 1 or

3 to the state 2 or 4 is ε_2. The probability that the state of economy remains the same in each case is $1-\varepsilon_1$ and $1-\varepsilon_2$, respectively. The transition probability matrix is written as:

$$P = \begin{pmatrix} 1-\varepsilon_1 & \varepsilon_1 \\ \varepsilon_2 & 1-\varepsilon_2 \end{pmatrix}. \tag{12.3}$$

The first row and column represent the states 1 and 3 and the second row and column represent the states 2 and 4. If the regularity condition for the Markov chain is assumed, then:

$$P'\pi = \pi \tag{12.4}$$

where $\pi = (\pi_1 \ \pi_2)'$ is a (2×1) vector. π is called the limiting probability distribution and can be solved in terms of ε_1 and ε_2:

$$\begin{aligned} \pi_1 &= \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2} \\ \pi_2 &= \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} \end{aligned} \tag{12.5}$$

The Gibbs sampler technique will approximate the posterior distribution of all unknown model parameters. The joint and conditional distributions used in the Gibbs sampler technique follow.

Consider the observed data $\boldsymbol{y} = (y_1, \cdots, y_n)'$ and

$$\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\varepsilon}, \boldsymbol{s})' \tag{12.6}$$

where

$$\begin{aligned} \boldsymbol{\beta} &= (\beta_1, \beta_2, \beta_3, \beta_4)', \\ \boldsymbol{\sigma} &= (\sigma_1, \sigma_2)', \\ \boldsymbol{\alpha} &= (\alpha_1, \cdots, \alpha_n)', \\ \boldsymbol{\varepsilon} &= (\varepsilon_1, \varepsilon_2)', \\ \boldsymbol{s} &= (s_1, \cdots, s_n)'. \end{aligned}$$

The joint posterior distribution needed for the analysis is:

$$P(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\varepsilon}, \boldsymbol{s}|\boldsymbol{y}) \quad \propto \quad P(\boldsymbol{y}|\boldsymbol{\alpha})P(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{s})P(\boldsymbol{s}|\boldsymbol{\varepsilon})P(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\varepsilon}) \tag{12.7}$$

However, this joint posterior distribution is difficult to obtain analytically, as Schmidt observes. Instead of directly using this joint posterior distribution, the Gibbs sampler technique draws samples from the joint posterior by sequentially drawing subvectors of $\boldsymbol{\theta}$ from their conditional distributions. Suppose the parameter vector $\boldsymbol{\theta}$ is divided into d (in our case, five) subvectors. Each iteration of the Gibbs sampler cycles through the subvectors of $\boldsymbol{\theta}$, drawing each subset conditionally on the value of all the others and on y. There are d steps in iteration t. At iteration t, an order of the d subvectors of θ is selected and each subvector is conditionally updated, given all the other components of θ:

$$P(\theta_i|\theta_{-i}^{t-1}, y)$$

where θ_{-i}^{t-1} is all the components of θ, except for θ_i, at their current values:

$$\theta_{-i}^{t-1} = (\theta_1^t, \cdots, \theta_{i-1}^t, \theta_{i+1}^{t-1}, \cdots, \theta_d^{t-1})$$

To apply the Gibbs sampler technique, the following conditional distributions are needed: the conditional distribution of the transition probability, $P(\varepsilon_i|s)$, the conditional distribution of the state vector s_t; $P(s_t|y, s_{-t}, \beta, \sigma, \alpha, \varepsilon)$, the conditional distribution of β; $P(\beta|\alpha, \varepsilon, s, y)$, the conditional distribution of σ; $P(\sigma|\beta, \alpha, \varepsilon, s, y)$ and the conditional distribution of α_t; $P(\alpha_t|\beta, \alpha_{-t}, \sigma, \varepsilon, s, y)$. The conditional probabilities of β and σ need some modifications due to the increase in the number of regimes. The other three are the same as in Schmidt.

Before examining the above conditional distributions, it is necessary to define the following indicator functions:

$$\begin{aligned}
I_{1t} &= 1_{\{s_t=1\}}, \\
I_{2t} &= 1_{\{s_t=2\}}, \\
I_{3t} &= 1_{\{s_t=3\}}, \\
I_{4t} &= 1 - I_{1t} - I_{2t} - I_{3t}, \\
\hat{I}_{1t} &= I_{1t} + I_{3t}, \\
\hat{I}_{2t} &= I_{2t} + I_{4t}.
\end{aligned}$$

The last two indicator functions, $\hat{I}_{1t}$ and $\hat{I}_{2t}$, mean that:

$$\hat{I}_{1t} = \begin{cases} 1 & \text{if state} = 1 \text{ or } 3, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\hat{I}_{2t} = \begin{cases} 1 & \text{if state} = 2 \text{ or } 4, \\ 0 & \text{otherwise.} \end{cases}$$

12.1.1 Conditional Distribution of the Transition Probability

As noted earlier, the states 1 and 3 and the states 2 and 4 are treated the same in terms of transition probabilities. In other words, in terms of transition probabilities, there exist only two exactly the same way as in Schmidt. A subscript i will denote 1 if the economy is in state 1 or 3 and will also denote 2 if the economy is in state 2 or 4. From (12.7), the conditional distribution of the transition probability,

ε_i, depends only on the states and the prior distribution:

$$P(\varepsilon_i|\cdot) \propto P(s|\varepsilon)P(\varepsilon) \qquad i = 1, 2 \tag{12.8}$$

Following Schmidt, independent beta prior distributions for the ε_i; $\varepsilon_i \sim Beta(\gamma_{i1}, \gamma_{i2})$ are applied. The probability density function is:

$$f(\varepsilon_i) = \frac{\Gamma(\gamma_{i1} + \gamma_{i2})}{\Gamma(\gamma_{i1})\Gamma(\gamma_{i2})}\varepsilon_i^{\gamma_{i1}-1}(1-\varepsilon_i)^{\gamma_{i2}-1} \qquad i = 1, 2 \tag{12.9}$$

This application will assign the value of one to both γ_{i1} and γ_{i2}, so that $f(\varepsilon_i)$ can be treated as a uniform distribution. The conditional distribution of the states, given the transition probabilities, can be expressed as:

$$\begin{aligned} P(s_1, s_2, \cdots, s_n|\varepsilon) \quad \propto \quad & [\varepsilon_1^{\hat{I}_{1n}}(1-\varepsilon_1)^{1-\hat{I}_{1n}}]^{1-\hat{I}_{1n-1}}[\varepsilon_2^{1-\hat{I}_{1n}}(1-\varepsilon_2)^{\hat{I}_{1n}}]^{\hat{I}_{1n-1}} \\ & \times[\varepsilon_1^{\hat{I}_{1n-1}}(1-\varepsilon_1)^{1-\hat{I}_{1n-1}}]^{1-\hat{I}_{1n-2}}[\varepsilon_2^{1-\hat{I}_{1n-1}}(1-\varepsilon_2)^{\hat{I}_{1n-1}}]^{\hat{I}_{1n-2}} \\ & \qquad \vdots \\ & \times[\varepsilon_1^{\hat{I}_{12}}(1-\varepsilon_1)^{1-\hat{I}_{12}}]^{1-\hat{I}_{11}}[\varepsilon_2^{1-\hat{I}_{12}}(1-\varepsilon_2)^{\hat{I}_{12}}]^{\hat{I}_{11}} \end{aligned} \tag{12.10}$$

The simplified conditional distribution results from the following counts of numbers used in a designated set:

$$\begin{aligned} k_1 &= \#\{t : s_t = 1 \text{ or } 3,\ s_{t+1} = 2 \text{ or } 4,\ 1 \le t < n\}, \\ k_2 &= \#\{t : s_t = 2 \text{ or } 4,\ s_{t+1} = 1 \text{ or } 3,\ 1 \le t < n\}, \\ n_1 &= \#\{t : s_t = 1 \text{ or } 3,\ 1 \le t < n\}, \\ n_2 &= \#\{t : s_t = 2 \text{ or } 4,\ 1 \le t < n\}. \end{aligned}$$

and

$$P(s_1, s_2, \cdots, s_n|\varepsilon) \propto \varepsilon_1^{k_1}(1-\varepsilon_1)^{n_1-k_1}\varepsilon_2^{k_2}(1-\varepsilon_2)^{n_2-k_2} \tag{12.11}$$

For instance, k_1 is the number of counts for the current state, being either 1 or 3, and the future state, being 2 or 4. See Schmidt for the detailed derivation of (12.10) and (12.11). Hence, multiplying (12.9) and (12.11), yields:

$$P(\varepsilon_i|s) \sim Beta(\gamma_{i1} + k_i, \gamma_{i2} + n_i - k_i) \qquad i = 1, 2 \tag{12.12}$$

12.1.2 Conditional Distribution of the State Vector

Let s_{-t} be the state vector with the current state, s_t, deleted. Then, the conditional distribution of s_t will be expressed as:

$$P(s_t|y, s_{-t}, \beta, \alpha, \sigma, \varepsilon) \propto P(y, s, \beta, \alpha, \sigma, \varepsilon) \tag{12.13}$$

Furthermore:

$$P(y, s, \beta, \alpha, \sigma, \varepsilon) \propto P(\alpha_t | s_t, \beta, \alpha_{-t}, \sigma) P(s_{t+1} | s_t, \varepsilon) P(s_t | s_{t-1}, \varepsilon) \tag{12.14}$$

The first term on the right hand side in (12.14), the conditional distribution of α_t, has normal distribution with mean β_{s_t} and variance σ_{s_t}:

$$P(\alpha_t | s_t, \beta, \alpha_{-t}, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2_{s_t}}} \exp\left\{ \frac{-(\alpha_t - \beta_{s_t})^2}{2\sigma^2_{s_t}} \right\} \tag{12.15}$$

The second and third terms on the right hand side of (12.14) are respectively expressed as:

$$\begin{aligned} P(s_{t+1} | s_t, \varepsilon) &= [\varepsilon_1^{\tilde{I}_{1t+1}} (1 - \varepsilon_1)^{1 - \tilde{I}_{1t+1}}]^{1 - \tilde{I}_{1t}} [\varepsilon_2^{1 - \tilde{I}_{1t+1}} (1 - \varepsilon_2)^{1 - \tilde{I}_{1t+1}}]^{\tilde{I}_{1t}} \\ P(s_t | s_{t-1}, \varepsilon) &= [\varepsilon_1^{\tilde{I}_{1t}} (1 - \varepsilon_1)^{1 - \tilde{I}_{1t}}]^{1 - \tilde{I}_{1t-1}} [\varepsilon_2^{1 - \tilde{I}_{1t}} (1 - \varepsilon_2)^{1 - \tilde{I}_{1t}}]^{\tilde{I}_{1t-1}} \end{aligned} \tag{12.16}$$

By multiplying all three:

$$\begin{aligned} P(s_t | s_{-t}, \beta, \alpha, \sigma, \varepsilon, y) \quad \propto \quad & \frac{1}{\sqrt{2\pi\sigma^2_{s_t}}} \exp\left\{ \frac{-(\alpha_t - \beta_{s_t})^2}{2\sigma^2_{s_t}} \right\} \\ & \times [\varepsilon_1^{\tilde{I}_{1t+1}} (1 - \varepsilon_1)^{1 - \tilde{I}_{1t+1}}]^{1 - \tilde{I}_{1t}} [\varepsilon_2^{1 - \tilde{I}_{1t+1}} (1 - \varepsilon_2)^{1 - \tilde{I}_{1t+1}}]^{\tilde{I}_{1t}} \\ & \times [\varepsilon_1^{\tilde{I}_{1t}} (1 - \varepsilon_1)^{1 - \tilde{I}_{1t}}]^{1 - \tilde{I}_{1t-1}} [\varepsilon_2^{1 - \tilde{I}_{1t}} (1 - \varepsilon_2)^{1 - \tilde{I}_{1t}}]^{\tilde{I}_{1t-1}} \end{aligned} \tag{12.17}$$

Note that this is a discrete distribution.

12.1.3 Conditional Distribution of β

Let $\beta' = (\beta_1,\ \beta_2,\ \beta_3,\ \beta_4)$. This produces:

$$P(\beta | \alpha, \sigma, \varepsilon, s, y) \propto P(\alpha | \beta, \sigma, s) P(\beta) \tag{12.18}$$

Now, if the prior distribution of β is multivariate normal with mean β_0 and covariance Σ, then:

$$\begin{aligned} \beta_0' &= (\beta_{10},\ \beta_{20},\ \beta_{30},\ \beta_{40}) \\ \Sigma_0 &= \Delta^{-1} diag(\sigma_0^2,\ \sigma_1^2,\ \sigma_0^2,\ \sigma_1^2) \end{aligned}$$

$\Delta^{-1} = diag(\delta_1,\ \delta_2,\ \delta_3,\ \delta_4)$ and δ_j is a specified positive number. If Z and X may defined as $Z = (\alpha_1, \cdots, \alpha_n)'$, $x_t' = (I_{1t},\ I_{2t},\ I_{3t},\ I_{4t})$ and $X = [x_t']_{t=1}^n$, then the diagonal matrix $X'X$ is:

$$X'X = \begin{pmatrix} \sum_{t=1}^n I_{1t} & 0 & 0 & 0 \\ 0 & \sum_{t=1}^n I_{2t} & 0 & 0 \\ 0 & 0 & \sum_{t=1}^n I_{3t} & 0 \\ 0 & 0 & 0 & \sum_{t=1}^n I_{4t} \end{pmatrix} \tag{12.19}$$

The conditional distribution, $P(\beta|\alpha,\ \sigma,\ s)$, is obtained which is multivariate normal with mean vector:

$$(\Delta + X'X)^{-1}(\Delta\beta_0 + X'Z),$$

and covariance matrix:

$$diag(\sigma_1^2,\ \sigma_2^2,\ \sigma_1^2,\ \sigma_2^2)(\Delta + X'X)^{-1}.$$

12.1.4 Conditional Distribution of σ

The conditional distribution of σ is expressed as follows:

$$P(\sigma|\beta,\alpha,\epsilon,s,y) \propto P(\alpha|\beta,\sigma,s)P(\beta|\sigma)P(\sigma) \tag{12.20}$$

Recall that the foregoing assumed that states 1 and 3 share the same variance and states 2 and 4 share the same variance. Now, the last term on the right, $P(\sigma)$, is assumed to be the product of independent inverse gamma distributions:

$$P(\sigma_j^2) \propto (\sigma_j^2)^{-\nu-1}\exp\left\{\frac{-1}{\lambda\sigma_j^2}\right\} \qquad j = 0,1 \tag{12.21}$$

where $\nu = \nu_{0j}$ and $\lambda = 2/(\nu_{0j}s_{0j}^2)$, and ν_{0j} and s_{0j}^2 are prespecified positive numbers. This yields:

$$\begin{aligned} P(\sigma_j^2|\alpha,\beta,s) &\propto (\sigma_j^2)^{-\frac{1}{2}\sum\limits_{t=1}^{n}\hat{I}_{it}} \exp\left\{-\frac{1}{2\sigma_j^2}\sum_{t=1}^{n}(\alpha_t-\beta_j)^2\hat{I}_{it}\right\} \\ &\quad \times(\sigma_i^2)^{-1}\exp\left\{\frac{(\beta_j-\beta_{0j})'\Delta_j(\beta_j - \ beta_{0j})}{2\sigma_i^2}\right\} \\ &\quad \times(\sigma_i^2)^{-\nu_{0i}-1}\exp\left\{-\frac{\nu_{0i}s_{0i}^2}{2\sigma_i^2}\right\} \\ &= (\sigma_i^2)^{-\nu_i-1}\exp\left\{\frac{-1}{\lambda_i\sigma_i^2}\right\} \end{aligned} \tag{12.22}$$

where

$$\nu_i = \tfrac{1}{2}\textstyle\sum_{t=1}^{n}\hat{I}_{it} + \nu_{0i} + 1$$
$$\lambda_i = 2\{\textstyle\sum_{t=1}^{n}(\alpha_t-\beta_j)^2\hat{I}_{jt} + (\beta_j-\beta_{0j})'\Delta_j(\beta_j-\beta_{0j}) + \nu_{0i}s_{0i}^2\}^{-1}$$

for $i = 1,2$ and $j = 1,2,3,4$.

12.1.5 Conditional Distribution of α_t

If α_{-t} is the vector α with α_t deleted, then:

$$P(\alpha_t|\alpha_{-t},\beta,\sigma,\epsilon,s,y) \propto P(y_t|\alpha_t)P(\alpha_t|,s_t,\beta,\sigma) \tag{12.23}$$

The conditional distribution of α is much simpler than that of Schmidt because α does not follow AR(1) process, as it does in Schmidt. α_t has a constant term and an error term as in (12.1).

12.2 Mean Model

This section considers the mean model of the regime-switching volatility model, beginning with the following:

$$\begin{cases} z_t = \rho x_t + y_t \\ y_t = \exp\left\{\frac{\alpha_t}{2}\right\}\zeta_t \\ \alpha_t = \beta_{s_t} + \sigma_{s_t}\eta_t \end{cases} \tag{12.24}$$

where the two errors, η_t and ζ_t, are assumed to have normal distributions with mean zero and variance one:

$$\zeta_t \; iid \; N(0,1)$$
$$\eta_t \; iid \; N(0,1)$$

z_t and x_t in (12.24) are defined as follows:

$$z_t = e_{t+1} - e_t$$
$$x_t = (i_{us,t} - i_{f,t})e_t = i_t e_t$$

where e_t is exchange rate at time period t and $i_{us,t}$ is U.S. interest rate at time period t and $i_{f,t}$ is foreign interest rate at time period t. The first equation in (12.24) is based on the interest parity condition. The interest parity condition is:

$$i_{us,t} = i_{f,t} - (E_t e_{t+1} - e_t)/e_t \tag{12.25}$$

Assuming that rational expectations, $E_t e_{t+1} = e_{t+1}$ hold, the following equation results:

$$e_{t+1} - e_t = (i_{us,t} - i_{f,t})e_t \tag{12.26}$$

The first equation in (12.24) was constructed from (12.26) above and indicates that z_t consists of structural components and an error component, which is characterized by a stochastic volatility. Thus, this equation can be used to test whether the interest parity condition holds.

As in the previous section, α_t is assumed to have a constant term and an error term in order to keep the model simple. The state of the economy is represented by s_t and assumes only two unobservable states in this model, so that s_t takes two values; $s_t = 0, 1$. If the economy is in an unobservable state 0, the change in exchange rate, z_t, will be modeled as:

$$z_t = \rho x_t + \exp\left\{\frac{1}{2}(\beta_0 + \sigma_0\eta_t)\right\}\zeta_t$$

Similarly, if the economy is in unobservable state 1, then z_t will be modeled as:

$$z_t = \rho x_t + \exp\left\{\tfrac{1}{2}(\beta_1 + \sigma_1\eta_t)\right\}\zeta_t.$$

Note that the coefficient of x_t, ρ, does not depend on unobservable states.

To describe the process of switching states from 0 to 1, or vice versa, a Markov chain model, which assumes a fixed transition probability matrix, is used:

$$\begin{aligned} P(s_t = 0|s_{t-1} = 1) &= \varepsilon_0 \\ P(s_t = 1|s_{t-1} = 0) &= \varepsilon_1 \end{aligned} \tag{12.27}$$

(12.27) gives exactly the same interpretation as in the previous section. The probability that the state of economy shifts from 1 to the state 0 is ε_0 and the probability that the state of economy moves from 0 to the state 1 is ε_1. A transition probability matrix is also written as in (12.4):

$$P = \begin{pmatrix} 1-\varepsilon_0 & \varepsilon_0 \\ \varepsilon_1 & 1-\varepsilon_1 \end{pmatrix} \tag{12.28}$$

Assuming the regularity condition for the Markov chain, the limiting probability distribution can be solved in terms of ε_0 and ε_1:

$$\begin{aligned} \pi_0 &= \frac{\varepsilon_1}{\varepsilon_0 + \varepsilon_1} \\ \pi_1 &= \frac{\varepsilon_0}{\varepsilon_0 + \varepsilon_1} \end{aligned} \tag{12.29}$$

A Gibbs sampler technique will be applied to estimate parameters. The Gibbs sampler technique uses joint and conditional distributions to consider the observed data $\boldsymbol{z} = (z_1, \cdots, z_n)'$ and $\boldsymbol{x} = (x_1, \cdots, x_n)'$ and parameters to be estimated in the model:

$$\boldsymbol{\theta} = (\rho, \boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\varepsilon}, \boldsymbol{s})' \tag{12.30}$$

where

$$\begin{aligned} \boldsymbol{\beta} &= (\beta_0, \beta_1)', \\ \boldsymbol{\sigma} &= (\sigma_0, \sigma_1)', \\ \boldsymbol{\alpha} &= (\alpha_1, \cdots, \alpha_n)', \\ \boldsymbol{\varepsilon} &= (\varepsilon_0, \varepsilon_1)', \\ \boldsymbol{s} &= (s_1, \cdots, s_n)'. \end{aligned}$$

The joint posterior distribution, that is needed for this analysis is:

$$P(\rho, \boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\alpha}, \boldsymbol{\varepsilon}, \boldsymbol{s}|\boldsymbol{x}, \boldsymbol{z}) \propto P(\boldsymbol{z}|\rho, \boldsymbol{x}, \boldsymbol{\alpha})P(\boldsymbol{\alpha}|\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{s})P(\boldsymbol{s}|\boldsymbol{\varepsilon})P(\boldsymbol{\beta}, \boldsymbol{\sigma}, \boldsymbol{\varepsilon}) \tag{12.31}$$

Therefore, this model requires the following conditional distributions: the conditional distribution of the transition probability ε; $P(\varepsilon_i|s)$, the conditional distribution of the state vector s_t; $P(s_t|\rho, s_{-t}, \beta, \sigma, \alpha, \varepsilon, x, z)$, the conditional distribution of β; $P(\beta|\rho, \alpha, \varepsilon, s, x, z)$, the conditional distribution of σ; $P(\sigma|\rho, \beta, \alpha, \varepsilon, s, x, z)$, the conditional distribution of α_t; $P(\alpha_t|\rho, \beta, \alpha_{-t}, \sigma, \varepsilon, s, x, z)$ and the conditional distribution of ρ; $P(\rho|\beta, \alpha, \sigma, \varepsilon, s, x, z)$.

Among the conditional probabilities listed above, the conditional distribution of the transition probability, $P(\varepsilon_i|\cdot)$, and the conditional distribution of the state vector, $P(s_t|\rho, s_{-t}, \beta, \sigma, \alpha, \varepsilon, x, z)$, are exactly the same as in the previous section. The conditional distribution of ρ, $P(\rho|\beta, \alpha, \sigma, \varepsilon, s, x, z)$, is newly introduced, which requires some discussion. The other three conditional distributions need some modifications.

12.2.1 Conditional Distribution of β

Let $\beta' = (\beta_0,\ \beta_1)$, this gives:

$$P(\beta|\phi, \alpha, \sigma, \varepsilon, s, x, z) \propto P(\alpha|\beta, \sigma, s)P(\beta) \tag{12.32}$$

Now, when the prior distribution of β is taken to be multivariate normal with mean β_0 and covariance Σ:

$$\beta_0' = (\beta_{00},\ \beta_{10})$$
$$\Sigma_0 = \Delta^{-1} diag(\sigma_0^2,\ \sigma_1^2)$$

where $\Delta^{-1} = diag(\delta_0,\ \delta_1)$ and δ_j is a specified positive number. Now,if the following indicator functions; $I_{0t} = 1_{\{s_t=0\}}$, and $I_{1t} = 1_{\{s_t=1\}} = 1 - I_{0t}$ are defined and, also, $L = (\alpha_1, \cdots, \alpha_n)'$, $m_t' = (I_{1t},\ I_{2t})$ and $M = [m_t']_{t=1}^n$. Then, the diagonal matrix $M'M$ is:

$$M'M = \begin{pmatrix} \sum_{t=1}^n I_{1t} & 0 \\ 0 & \sum_{t=1}^n I_{2t} \end{pmatrix} \tag{12.33}$$

The resulting conditional distribution, $P(\beta|\alpha,\ \sigma,\ s)$, is multivariate normal with mean vector:

$$(\Delta + M'M)^{-1}(\Delta\beta_0 + M'L),$$

and covariance matrix:

$$diag(\sigma_0^2,\ \sigma_1^2)(\Delta + M'M)^{-1}.$$

12.2.2 Conditional Distribution of σ

The conditional distribution of σ is expressed as follows:

$$P(\sigma|\phi,\beta,\alpha,\varepsilon,s,x,z) \propto P(\alpha|\beta,\sigma,s)P(\beta|\sigma)P(\sigma) \tag{12.34}$$

Now, the last term on the right, $P(\sigma)$, is assumed to be the product of independent inverse gamma distributions:

$$P(\sigma_i^2) \propto (\sigma_i^2)^{-\nu-1}\exp\left\{\frac{-1}{\lambda\sigma_i^2}\right\} \qquad i=0,1 \tag{12.35}$$

where $\nu=\nu_{0i}$ and $\lambda=2/(\nu_{0i}s_{0i}^2)$. ν_{0i} and s_{0i}^2 are prespecified positive numbers, this gives:

$$\begin{aligned} P(\sigma_i^2|\alpha,\beta,s) &\propto (\sigma_i^2)^{-\frac{1}{2}\sum_{t=1}^{n} I_{it}} \exp\left\{-\frac{1}{2\sigma_i^2}\sum_{t=1}^{n}(\alpha_t-\beta_i)^2 I_{it}\right\} \\ &\quad \times(\sigma_i^2)^{-1}\exp\left\{\frac{(\beta_i-\beta_{0i})'\Delta_i(\beta_i-\beta_{0i})}{2\sigma_i^2}\right\} \\ &\quad \times(\sigma_i^2)^{-\nu_{0i}-1}\exp\left\{-\frac{\nu_{0i}s_{0i}^2}{2\sigma_i^2}\right\} \\ &= (\sigma_i^2)^{-\nu_i-1}\exp\left\{\frac{-1}{\lambda_i\sigma_i^2}\right\} \end{aligned} \tag{12.36}$$

where

$$\begin{aligned} \nu_i &= \tfrac{1}{2}\textstyle\sum_{t=1}^{n} I_{it}+\nu_{0i}+1 \\ \lambda_i &= 2\{\textstyle\sum_{t=1}^{n}(\alpha_t-\beta_i)^2 I_{it}+(\beta_i-\beta_{0i})'\Delta_i(\beta_i-\beta_{0i})+\nu_{0i}s_{0i}^2\}^{-1} \end{aligned}$$

for $i=0,1$.

12.2.3 Conditional Distribution of α_t

Conditional distribution of α_t can be obtained as follows:

$$\begin{aligned} P(\alpha_t|\phi,\beta,\sigma,\varepsilon,s,x,z) &\propto P(y_t|\alpha_t)P(\alpha_t|s_t,\beta,\sigma) \\ &= P(z_t-\theta x_t|\alpha_t)P(\alpha_t|s_t,\beta,\sigma) \end{aligned} \tag{12.37}$$

12.2.4 Conditional Distribution of ρ

First the conditional distribution of ρ can be written as:

$$P(\rho|\alpha,\beta,\sigma,\varepsilon,s,x,z) \propto P(z|\rho,\alpha,x)P(\rho) \tag{12.38}$$

Because the conditional distribution of z is a normal distribution with mean of ρx_t and variance of e^{α_t}, the distribution can be written as:

$$P(z|\rho,\alpha,x) = \Pi_{t=1}^{n} \frac{1}{\sqrt{2\pi e^{\alpha_t}}} \exp\left\{\frac{-(z_t-\rho x_t)^2}{2e^{\alpha_t}}\right\} \tag{12.39}$$

Assuming that the prior distribution of ρ is also a normal distribution with mean ρ_0 and variance $\sigma_{\rho_0}^2$:

$$\frac{1}{\sqrt{2\pi\sigma_{\rho_0}^2}} \exp\left\{\frac{-(\rho-\rho_0)^2}{2\sigma_{\rho_0}^2}\right\} \tag{12.40}$$

Hence, the conditional distribution of ρ is:

$$\begin{aligned}
&P(\rho|\alpha,\beta,\sigma,\varepsilon,s,x,z)\\
&= \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi e^{\alpha_t}}} \exp\left\{\frac{-(z_t-\rho x_t)^2}{2\exp\{\alpha_t\}}\right\} \times \frac{1}{\sqrt{2\pi\sigma_{\rho_0}^2}} \exp\left\{\frac{-(\rho-\rho_0)^2}{2\sigma_{\rho_0}^2}\right\}\\
&= \frac{1}{(\sqrt{2\pi})^n} \frac{1}{\sqrt{2\pi\sigma_{\rho_0}^2 \exp\{\sum_{t=1}^{n}\alpha_t\}}} \exp\left\{-\sum_{t=1}^{n} \frac{(z_t-\rho x_t)^2}{2\exp\{\alpha_t\}} - \frac{(\rho-\rho_0)^2}{2\sigma_{\rho_0}^2}\right\}
\end{aligned} \tag{12.41}$$

When the exponential part of (12.41) expanded and the square completed with respect to ρ:

$$\begin{aligned}
-\sum_{t=1}^{n} \frac{(z_t-\rho x_t)^2}{2\exp\{\alpha_t\}} - \frac{(\rho-\rho_0)^2}{2\sigma_{\rho_0}^2} &= -\frac{A\rho^2 - 2B\rho + C}{2\sigma_{\rho_0}^2 \exp\{\sum_{t=1}^{n}\alpha_t\}}\\
&= -\frac{(\rho-\frac{B}{A})^2}{\frac{2}{A}\sigma_{\rho_0}^2 \exp\{\sum_{t=1}^{n}\alpha_t\}} - \frac{C-\frac{B^2}{A}}{2\sigma_{\rho_0}^2 \exp\{\sum_{t=1}^{n}\alpha_t\}}
\end{aligned} \tag{12.42}$$

where

$$\begin{aligned}
A &\equiv \sigma_{\rho_0}^2 \sum_{j=1}^{n} \exp\left\{\sum_{t=1}^{n}\alpha_t - \alpha_j\right\} x_j^2 + \exp\left\{\sum_{t=1}^{n}\alpha_t\right\}\\
B &\equiv \sigma_{\rho_0}^2 \sum_{j=1}^{n} \exp\left\{\sum_{t=1}^{n}\alpha_t - \alpha_j\right\} x_j z_j + \rho_0 \exp\left\{\sum_{t=1}^{n}\alpha_t\right\}\\
C &\equiv \sigma_{\rho_0}^2 \sum_{j=1}^{n} \exp\left\{\sum_{t=1}^{n}\alpha_t - \alpha_j\right\} z_j^2 + \rho_0^2 \exp\left\{\sum_{t=1}^{n}\alpha_t\right\}
\end{aligned}$$

Therefore:

$$\begin{aligned}
&P(\rho|\alpha,\beta,\sigma,\varepsilon,s,x,z)\\
&= \frac{1}{(\sqrt{(2\pi)^n}\sqrt{A}} \exp\left\{\frac{-C+\frac{B^2}{A}}{2\sigma_{\rho_0}^2 \exp\left\{\sum_{t=1}^{n}\alpha_t\right\}}\right\}\\
&\times \frac{1}{\sqrt{\frac{2\pi}{A}\sigma_{\rho_0}^2 \exp\left\{\sum_{t=1}^{n}\alpha_t\right\}}} \exp\left\{\frac{(\rho-\frac{B}{A})^2}{\frac{2}{A}\sigma_{\rho_0}^2 \exp\left\{\sum_{t=1}^{n}\alpha_t\right\}}\right\}
\end{aligned} \tag{12.43}$$

The last two terms (the third line) in (12.43) indicate that ρ will be drawn from a normal distribution with mean of B/A and variance of $\sigma^2_{\rho_0} \exp\left\{\sum_{t=1}^{n} \alpha_t\right\}/A$ in Gibbs sampling: $N\left(B/A, \sigma^2_{\rho_0} \exp\left\{\sum_{t=1}^{n} \alpha_t\right\}/A\right)$, where A and B are as previously defined.

The next chapter will apply the data set to the model generated above, and also report some of the results.

13 EMPIRICAL RESULTS

13.1 Data Descriptions

Here the data set is applied to the switching-regime stochastic volatility model. The data sets consist of three daily foreign exchange rates and four daily interest rates for four countries from Jan. 1 1975 to Dec. 31 1993.[1] The countries are France, Germany, United Kingdom and United States. Daily call money rates are used for daily interest rates. In Table 13.1 and Table 13.2, summary statistics for each time series are provided.

In Table 13.1, the data summary for the three daily exchange rates is presented. The second column (N) in the table is the number of observations. Basic statistics such as mean and standard deviation are given in the table. The exchange rates are defined as the foreign currency prices of U.S. dollar for the rest of the chapter.[2] For example, EF, which stands for French exchange rate, is the franc price of the U.S. dollar. EG and EUK are the German exchange rate and the British exchange rate, respectively. Δ is the first difference operator:

$$\Delta EF_t = EF_t - EF_{t-1} \qquad \forall t$$

Here, the means of the first differences are essentially zero and the standard deviations of the first differences are smaller than those of the original series. The first difference series also have large kurtoses (greater than 3), implying that the first order series have fat-tailed distributions relative to the normal.

Figure 13.1, Figure 13.2 and Figure 13.3 present the plots of exchange rates (left) and their first differences (right). From the plots of the original series, it it noted that U.S. dollars appreciated through the first half of the 1980s and depreciated in the second half of the 1980s against the German Mark and the French Franc. U.S. dollars actually depreciated against British pounds in the mid-80s. In all cases, the first differences of exchange rates exhibit time-changing volatility, especially around the middle of the 1980s (observations 2000-2500), where they exhibit larger volatilities.

[1] The data sets were kindly provided by Mr. Patrick Decker of the Federal Reserve Bank of Washington, D.C..

[2] In the previous chapters, the exchange rates were defined as the dollar price of the foreign currencies.

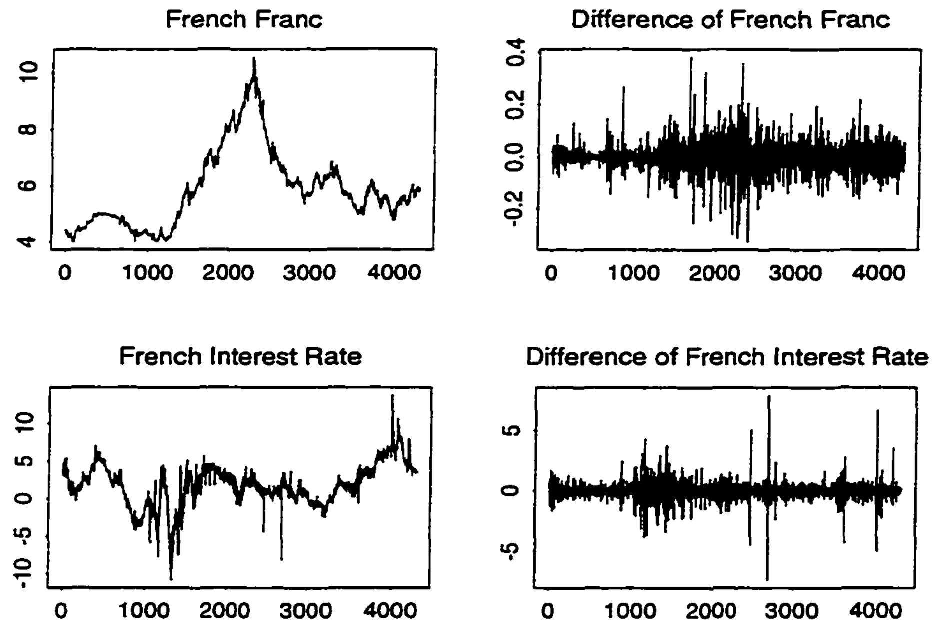

Figure 13.1 French Exchange Rate and Interest Rate

Table 13.1 Data Summary for Daily Exchange Rates

Variables	N	Mean	Standard Error	Skewness	Kurtosis
EF	4289	5.88	1.40	0.97	0.46
ΔEF	4289	0.00	0.04	-0.09	7.56
EG	4294	2.11	0.43	0.57	-0.52
ΔEG	4294	0.00	0.01	-0.26	5.01
EUK	4316	1.75	0.29	0.37	-0.16
ΔEUK	4316	0.00	0.01	-0.33	3.70

Table 13.2 gives the data summary for the three daily interest rates. Daily interest rates are defined as the difference between foreign call money rate and the U.S. call money rate. $INTF$ is, for instance, defined as the difference between French call money rate and U.S. call money rates:

$$INTF \equiv \text{French call money rate} - \text{U.S. call money rate.}$$

The other two interest rates are similarly defined. In the bottom of Figure 13.1, Figure 13.2 and Figure 13.3, interest rates (left) and their first differences (right) are plotted. These plots show that foreign interest rates were relatively lower than the U.S. interest rate during the 1980s. On the other hand, foreign interest rates were higher than the U.S. interest rate in the 1990s. They also show that volatility in interest rates changes over time.

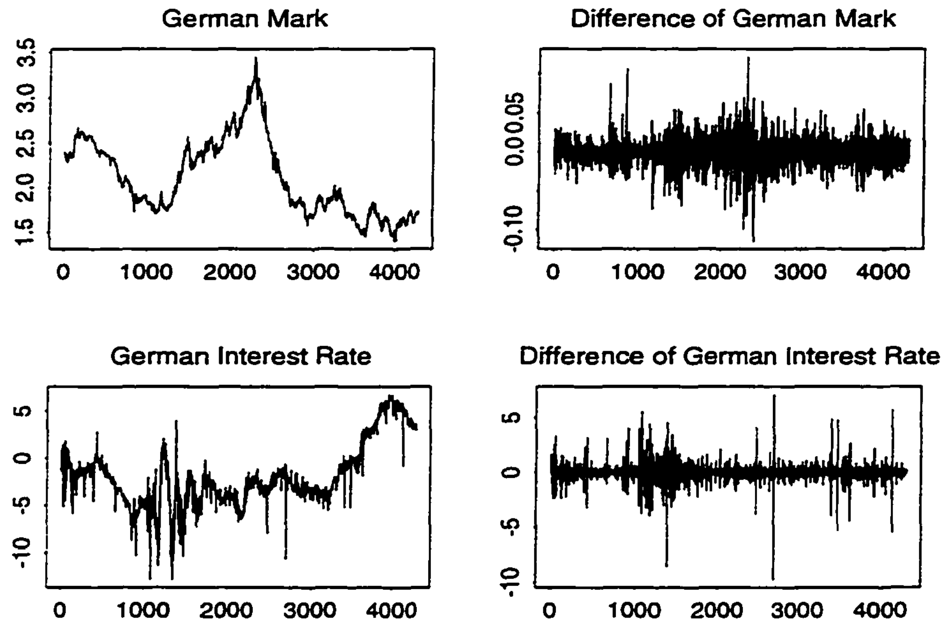

Figure 13.2 German Exchange Rate and Interest Rate

Table 13.2 Data Summary for Daily Interest Rates

Variables	N	Mean	Standard Error	Skewness	Kurtosis
$INTF$	4289	1.78	2.81	-0.39	1.57
$\Delta INTF$	4289	0.00	0.52	-0.45	27.34
$INTG$	4294	-1.95	3.43	0.65	0.34
$\Delta INTG$	4294	0.00	0.59	-0.92	34.03
$INTUK$	4316	2.84	3.15	-0.64	1.16
$\Delta INTUK$	4316	0.00	0.61	-0.56	15.74

This chapter examines the question of whether or not the interest rate differential explains movements of the exchange rate including the volatility. At this moment, there is no clear relationships between movement of exchange rates and that of interest rates. At the beginning of the sample (up to the 2000th observation), exchange rates and interest rates move in a similar fashion.

The purpose in this part is to construct a model that relates exchange rate volatility to movement of interest rates to see if movements of exchange rates depend on the size of the interest rate differential. First, it is necessary to divide the whole data set depending on whether or not the size of difference of two interest rates is greater than and equal to some positive number (k). Here 3%, 4% and 5% have

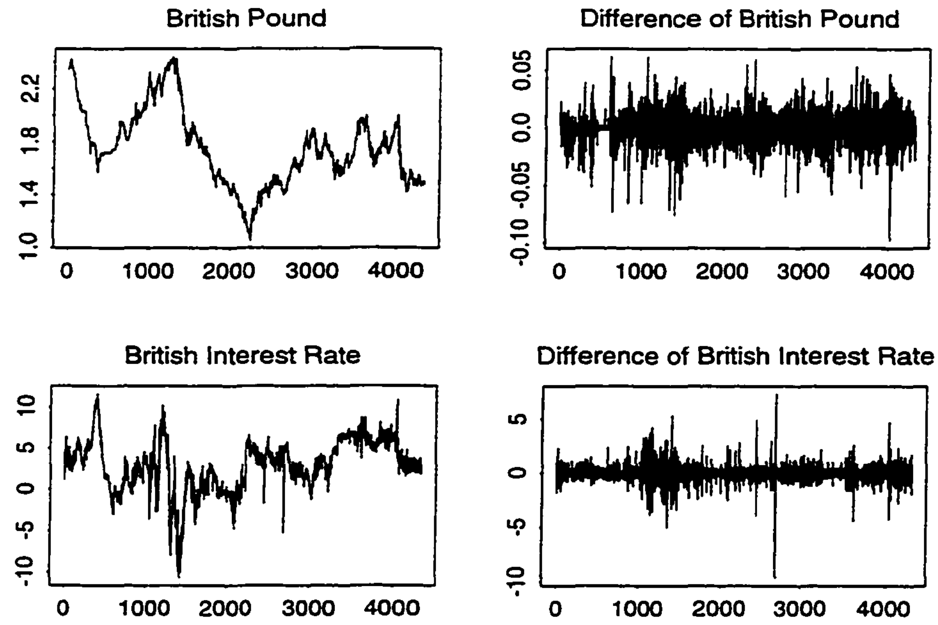

Figure 13.3 British Exchange Rate and Interest Rate

been chosen for k.[3] In other words, the data is split into two parts depending on whether the interest rate differential stays inside the bounds or moves outside the bounds of the prespecified interest rate differential.

Table 13.3 presents the data summary of exchange rates with $k = 3$. The subscript G indicates that the data correspond to difference of interest rates greater than or equal to 3% while the subscript L implies the data corresponding to difference of interest rates less than 3%. Note that there is a higher kurtosis in each case of the exchange rate with subscript L, which implies a higher volatility.

13.2 Empirical Results I

The model used to estimate is described in (12.1):

$$y_t = \exp\left\{\frac{\alpha_t}{2}\right\}\zeta_t$$
$$\alpha_t = \beta_{s_t} + \sigma_{s_t}\eta_t \tag{13.1}$$

The following eight parameters need to be estimated in the model: β_1, β_2, β_3, β_4, σ_1, σ_2, ε_1, and ε_2. The Gibbs sampler technique was used to estimate the parameter values. A burn-in period of 5,000 Gibbs iterates was chosen and 10,000 observations were used in the analysis. All the results are based

[3]This section will report the results for $k = 3$. The results for $k = 4, 5$ will be presented in the Appendix.

Table 13.3 Data Summary for Daily Exchange Rates: $k = 3$

Variables	N	Mean	Standard Error	Skewness	Kurtosis
EF_G	1599	5.54	1.01	1.00	0.74
EF_L	2690	6.08	1.55	0.76	-0.17
ΔEF_G	1599	0.00	0.04	0.26	6.64
ΔEF_L	2690	0.00	0.04	-0.22	7.79
EG_G	2490	2.06	0.43	0.74	-0.54
EG_L	1804	2.17	0.42	0.37	-0.28
ΔEG_G	2490	0.00	0.01	-0.08	4.85
ΔEG_L	1804	0.00	0.01	-0.59	5.22
EUK_G	2347	1.77	0.30	0.39	-0.26
EUK_L	1969	1.73	0.26	0.28	-0.21
ΔEUK_G	2347	0.00	0.01	-0.35	3.13
ΔEUK_L	1969	0.00	0.01	-0.32	4.53

Table 13.4 Data Summary for Daily Interest Rates: $k = 3$

Variables	N	Mean	Standard Error	Skewness	Kurtosis
$INTF_G$	1599	3.27	3.79	-1.44	1.82
$INTF_L$	2690	0.89	1.39	-0.66	-0.30
$INTG_G$	2490	-2.52	4.25	0.92	-0.24
$INTG_L$	1804	-1.17	1.43	1.13	0.94
$INTUK_G$	2347	4.48	3.25	-2.18	5.91
$INTUK_L$	1969	0.89	1.49	-0.52	-0.70

on 5,000 observations after a burn-in period of 5,000. The estimated marginal posterior distributions of the parameters for each country are shown in Figure 13.4, Figure 13.5, and Figure 13.6. The figures in the middle and the bottom are estimated marginal posterior distributions of σs and εs, respectively. For figures of distribution of βs, the solid line represents the distribution of β_1. The dotted line is the distribution of β_2. Finally, the lighter broken line is the distribution of β_3 and the heavy broken line represents the distribution of β_4. For figures of σs and εs, the solid line represents an unobservable state A (states 1 and 3) and the dotted line is an unobservable state B (states 2 and 4).

The estimated posterior means of the parameters are given in Table 13.5. The numbers in parenthesis are variances, the figure in the top, the estimated marginal posterior distribution of βs.

In the French case, there is a distinction in β between unobservable state A and unobservable state B. State A represents larger values of β which implies that the larger change in exchange rate, y_t. However, within state A there is not much distinction in the distributions of βs between state 1 and 3. β_1 and β_3 seem to have similar distributions, although distribution of β_3 is slightly rightward to

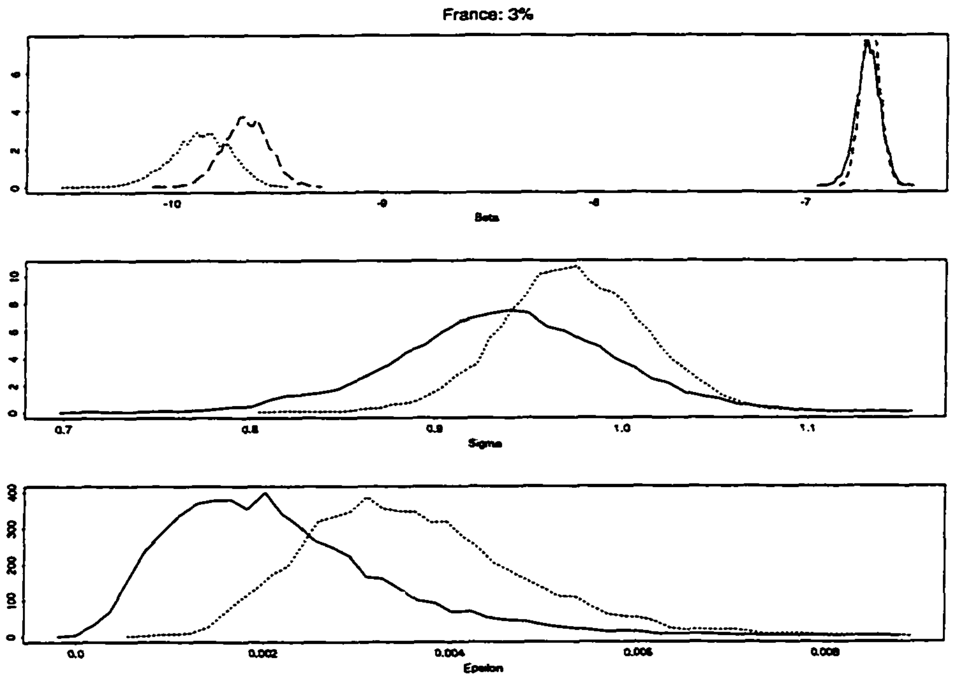

Figure 13.4 Estimated Marginal Posterior Distribution: France

distribution of β_1. In other words, it seems that once the economy enters unobservable state A, whether the interest rate differential or the expected appreciation (depreciation) stays inside or outside the 3% bounds does not make much difference.

To see whether the two parameters β_1 and β_3 are different, in particular, if the posterior probability that β_1 is larger than β_3, $P(\beta_1 > \beta_3)$, the posterior probability was computed. The results of comparisons of βs by posterior probabilities are given in Table 13.6. The result for β_1 and β_3 is 0.44. Looking at Figure 13.4, it is difficult to distinguish β_1 and β_3. On the other hand, the difference between β_2 and β_4 is more visible. From Figure 13.4, β_2 seems to take on smaller values than β_4.

The posterior probability that β_2 is greater than β_4, $P(\beta_2 > \beta_4)$, is 0.09. Less than 10% of the pairs of β_2 and β_4 satisfy $\beta_2 > \beta_4$, so it can be concluded that β_2 is likely to be smaller than β_4. If the economy is in state B, the interest rate differential seems to make some difference. If the absolute value of the interest rate differential is larger than 3%, the value of β (β_2) tends to be smaller, which implies that the change in the exchange rate is more likely to be smaller. If the interest rate differential is within the 3% bounds, the change in the exchange rate tends to be larger.

The variability parameters σ_1 and σ_2 do not have distinct distributions. The posterior probability, $P(\sigma_1 > \sigma_2)$, is 0.37. It is not clear whether σ_1 is smaller than σ_2. It may be concluded that the values of σs do not depend on the two unobservable states very much.

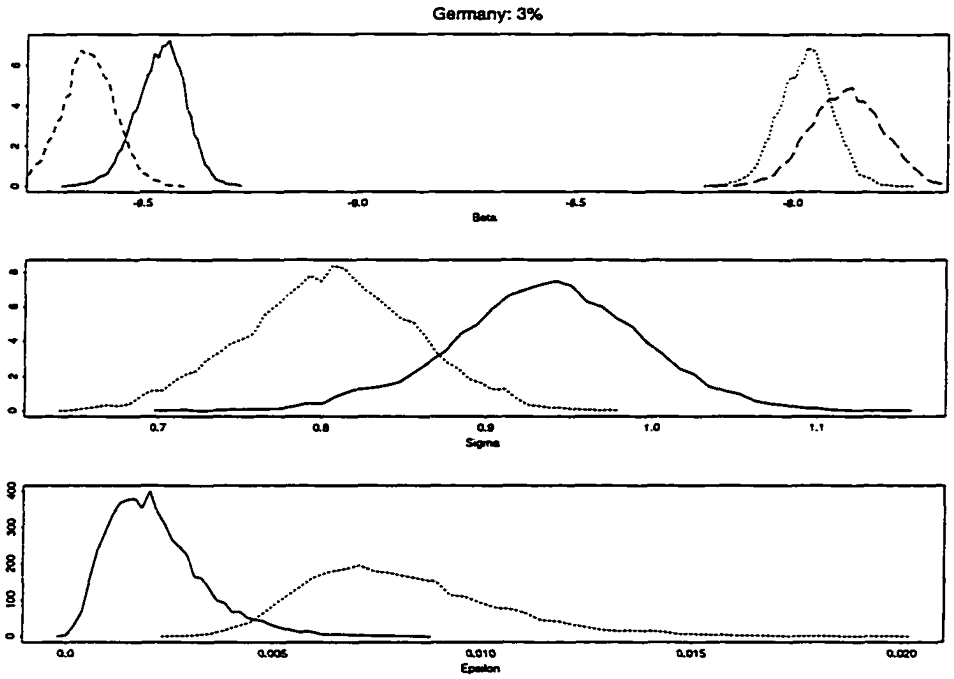

Figure 13.5 Estimated Marginal Posterior Distribution: Germany

The means of the state values drawn at each time period are plotted in Figure 13.7, along with the data of both exchange rates and interest rates. In the French case, the economy appears not to change states often. From observation 1 to approximately observation 1300 (November 7 1980), most of the time, it stays in unobservable state B, which represents a state of lower volatility of exchange rate. After about the observation 1300, it switches to state A, which is state of higher volatility of exchange rate, and stays in state A. This is approximately one year after the Reagan administration took the office. It is well known that during the Reagan administration the exchange rate was allowed to move relatively freely.

In the German case, the results seem to give clearer implications. The two unobservable states, A and B, make more differences in β. Clearly, β in state A (state 1 and 3) is smaller than β in state B (state 2 and 4). It is also noted that the variability parameter, σ, in state A is larger than in state B. Note that unobservable states A and B are reversed, compared with the French case, since the parameters β_1 and β_3 in state A are smaller than β_2 and β_4 in state B. Hence, state A implies a state of lower volatility and state B represents a state of higher volatility. Unobservable states A and B are named to provide a convenient distinction. Schmidt discusses the issue of identifiability in more detail. Interest rate differential makes some distinction between the values of β within each unobservable state. More specifically, if the economy is in state A and the interest rate differential is outside the 3% bounds,

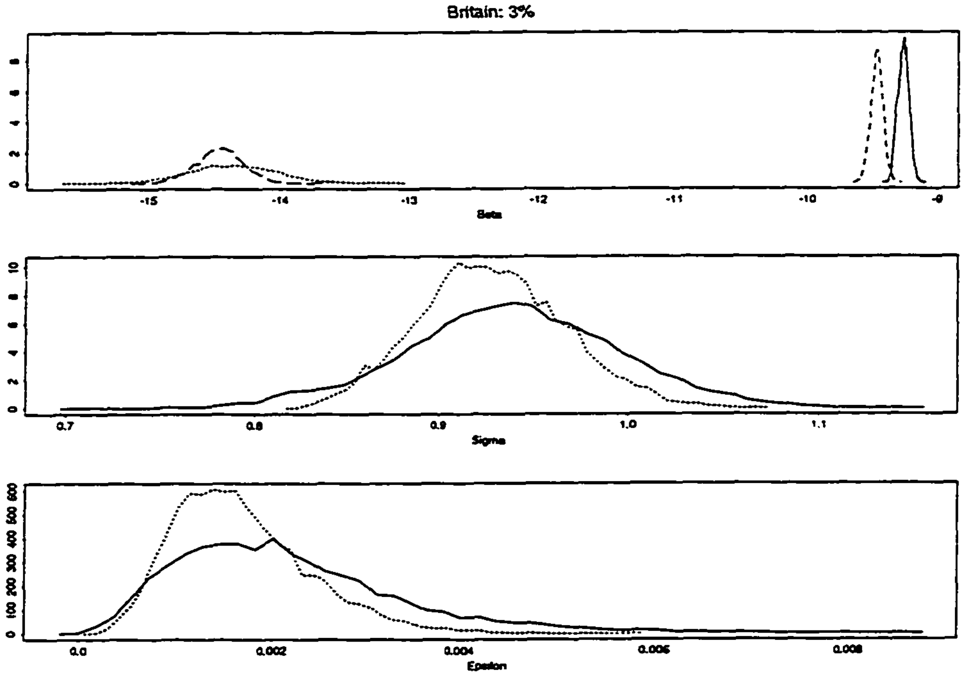

Figure 13.6 Estimated Marginal Posterior Distribution: Britain

then the change in the exchange rate tends to be larger than otherwise. In other words, the value of β_1 is likely to be larger than the value of β_3. Note also that the posterior distribution $P(\beta_1 > \beta_3)$ is 0.99 from Table 13.6. In state B, however, an interest rate differential greater than the 3% bounds leads to smaller change in exchange rate. The value of β_2 tends to be smaller than the value of β_4. The computed posterior distribution $P(\beta_2 > \beta_4)$ is only 0.16. Less than 20% of the pairs of β_2 and β_4 satisfies the relation of $\beta_2 > \beta_4$.

The variability parameters, σ, also appear to depend on the two unobservable states. The variability parameter in state A, σ_1, is likely to be larger than σ_2 in state B. The posterior probability $P(\sigma_1 > \sigma_2)$ is 0.98, which conforms to the observation.

The means of the state values are plotted with the data in Figure 13.8. It is evident that there is more often change in the state of economy. The state A will be interpreted as a state of low volatility and the state B will be a state of high volatility, that is, the first 1300 observations and the last 1500 observations are likely to stay in state A and the middle 1500 observations tend to stay in state B although the state frequently changes. In the data, the 1300th observation is dated October 17, 1980 and the observation 2800 is June 15, 1987. These observations correspond to what was observed in the exchange rate movement during the 1980s.

In the British case, βs are different, depending on which unobservable state the economy is in. βs

Table 13.5 Estimated Posterior Means of the Parameters: $k = 3$

	β_1	β_2	β_3	β_4	σ_1	σ_2	ε_1	ε_2
France	-6.70	-9.87	-6.69	-9.65	0.97	1.01	0.004	0.02
	(0.003)	(0.019)	(0.002)	(0.012)	(0.002)	(0.008)	(0.000[a])	(0.000)
Germany	-9.46	-7.97	-9.63	-7.88	0.80	0.63	0.008	0.02
	(0.003)	(0.004)	(0.004)	(0.007)	(0.003)	(0.004)	(0.000)	(0.000)
Britain	-9.26	-14.37	-9.47	-14.47	0.93	0.59	0.001	0.05
	(0.002)	(0.117)	(0.002)	(0.032)	(0.001)	(0.039)	(0.000)	(0.000)

[a]This does not mean that the variance is zero. The value is very small (5.290017e-06).

Table 13.6 Comparisons of βs: Posterior Probability: $k = 3$

	$P(\beta_1 > \beta_3)$	$P(\beta_2 > \beta_4)$	$P(\sigma_1 > \sigma_2)$
France	0.44	0.09	0.37
Germany	0.99	0.16	0.98
Britain	1.00	0.39	0.94

in state A are larger than βs in state B. State A can be interpreted as a state of a larger change in exchange rate and state B, a state of a smaller change. In Figure 13.9, the state means will verify our interpretations of the unobservable states. The interest rate differential or expected appreciation (depreciation) will be important in state A. If the interest rate differential is outside the 3% bounds, then β takes even larger values and if the interest rate differential is within the 3% bounds, β will be slightly smaller, that is,β_1 tends to be larger than β_3. Note, also, that the estimated posterior probability, $P(\beta_1 > \beta_3)$, is 1.00. In state B, however, the interest rate differential does not seem to play a role. The two marginal posterior distributions of β_2 and β_4 overlap very much. The estimated posterior probability, $P(\beta_2 > \beta_4)$, is 0.39 which makes it difficult to separate β_2 from β_4.

The variability parameter seems to be independent of unobservable states since its marginal posterior distributions overlap. However, the posterior probability of $P(\sigma_1 > \sigma_2)$ is as high as 0.94 which says that, most of the time, the variability parameter in state A, σ_1, is larger than the variability parameter in state B, σ_2.

In Figure 13.9, state A, a high volatility state, is the most permanent and it seldom changes state from state A to state B. In the British case, exchange rate seems to have already been in the state of higher volatility around the observation 600 (October 26, 1977). Since then, the exchange rate stays in the highly volatile state.

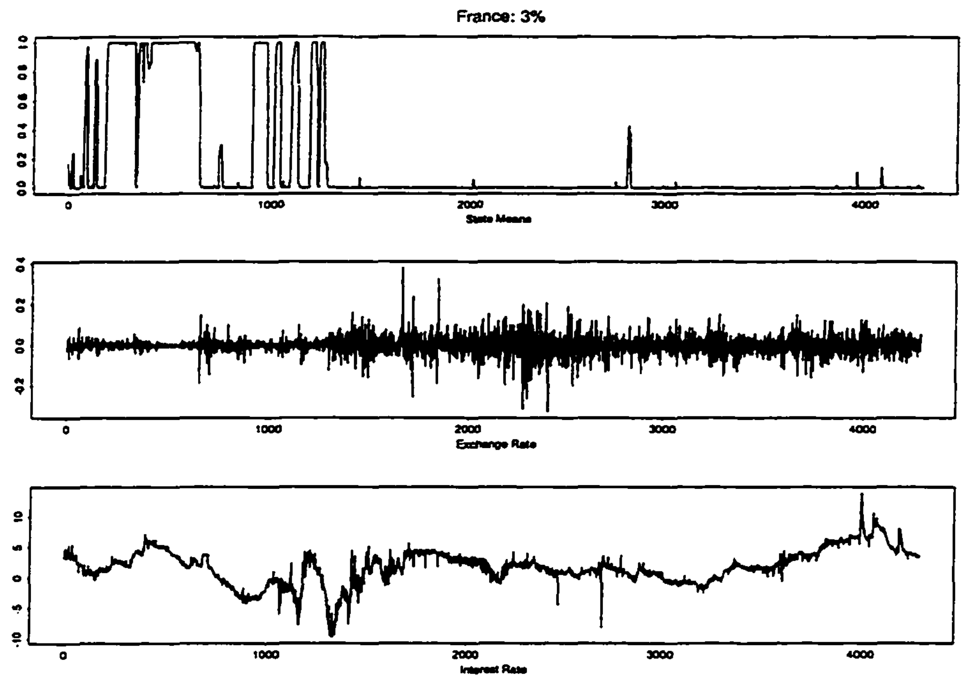

Figure 13.7 State Means and Data: France

13.3 Empirical Results II: Mean Model

The previous section examined interest rates to see if they have some explanatory power for different regimes of exchange rate, however the results were not so promising. It seems that interest rates do not contribute to separating regimes in the model previously specified. To further investigate a relationship between exchange rate and interest rate, it is necessary to estimate a mean model (12.24).

This section will present the results for the mean model (12.24) as derived from the interest parity condition. Here, the focus lies in the coefficient of x_t, ρ. If ρ equals one, then the interest parity condition holds. If ρ is zero, then exchange rate follows a random walk and interest rates do not explain the movement of exchange rates.

In this model, as previously discussed, it is necessary to estimate the following parameters; ρ, β_0, β_1, σ_0, σ_1, ε_0 and ε_1.

Table 13.7 reports the estimated posterior means of the seven parameters for each country and shows that the estimated parameter ρ is essentially zero in all 3 cases. This implies that all three exchange rates follow a random walk process since (12.24) becomes:

$$
\begin{aligned}
z_t &= y_t = \exp\left\{\frac{\alpha_t}{2}\right\}\zeta_t \\
\alpha_t &= \beta_{s_t} + \sigma_{s_t}\eta_t
\end{aligned}
\qquad (13.2)
$$

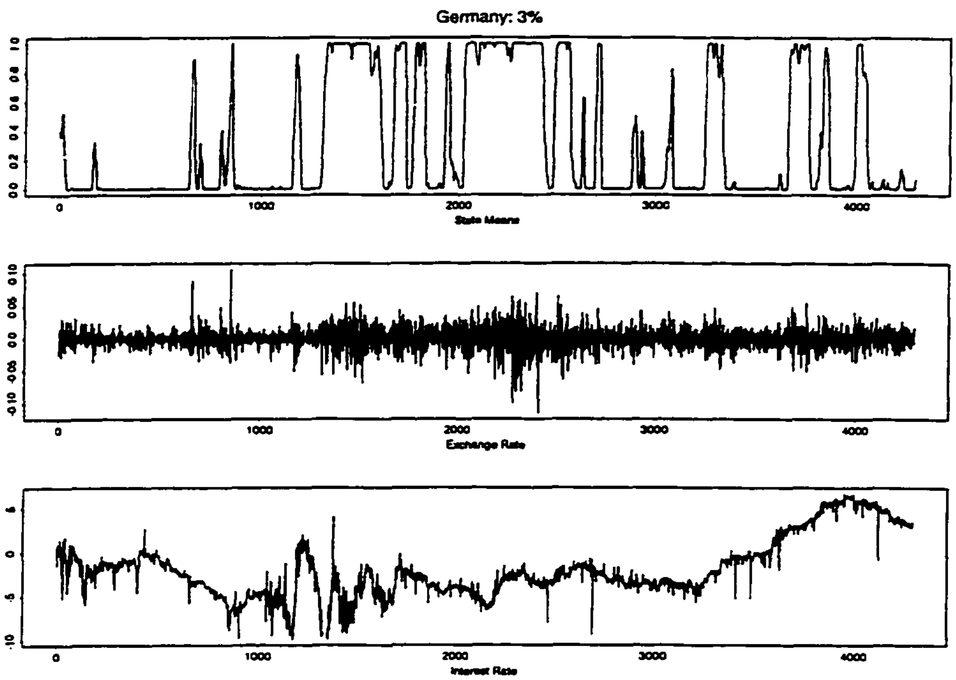

Figure 13.8 State Means and Data: Germany

This conclusion can also be confirmed by checking Figure 13.10, Figure 13.12 and Figure 13.14. At the top of each Figure is the marginal posterior distribution of ρ. In all three cases, the distribution of ρ is mound-shaped with mean of approximately zero and the distribution does not include one. Therefore, the null hypothesis, ρ equals one, is rejected. This result indicates that the data not only reject the interest parity condition but also implies that exchange rates follows a random walk process. This result also means that a relationship between exchange rates and interest rates does not exist for the daily data in these three countries.

In the French case, it is obvious that β_0 takes on larger values than β_1. State 0 is considered to capture a state of larger volatility, while state 1 represents a state of smaller volatility. σ_0 is more likely to take on smaller values than σ_1. Figure 13.11 gives the plots of the parameters. β_0 and β_1 take on distinct values. Figure 13.16 reports a comparison of state means and data of first difference of exchange rate. Again, state 0 corresponds to a state of larger volatility and state 1 represents a state of smaller volatility. As was seen previously, after the observation 1300, the state stays in a high volatility state most of the time.

In the German case, the state 0 implies a state of smaller variability since β_0 takes on smaller values than β_1. There is little observable difference in the values of σ and ε, in particular, the two distributions of ε overlap very much. The same results are found in the plot of the parameters in

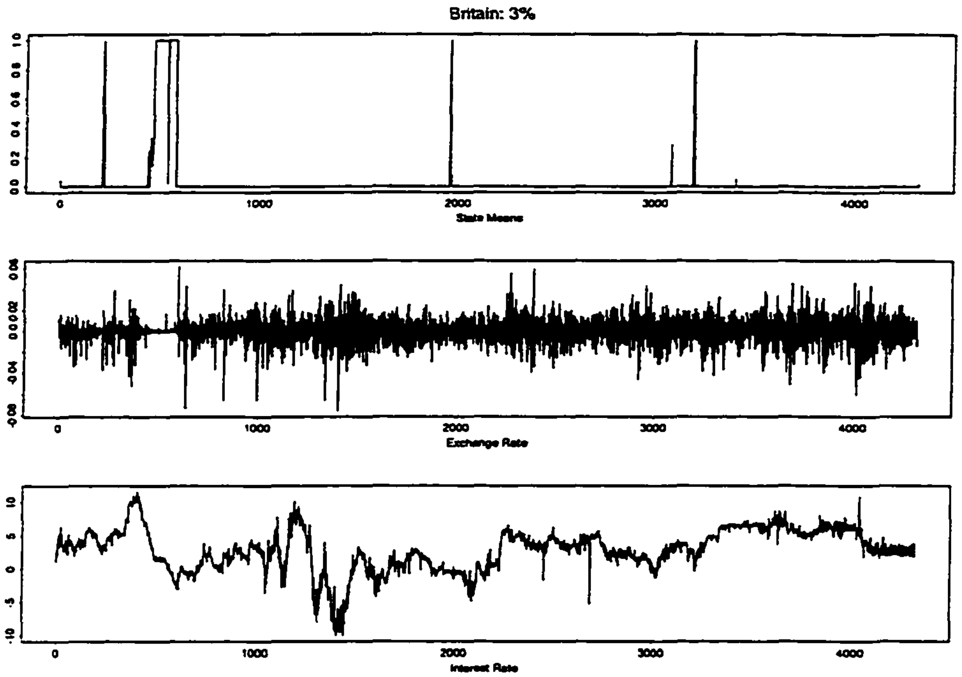

Figure 13.9 State Means and Data: Britain

Figure 13.13. Figure 13.17 shows that state 0 corresponds to a state of smaller volatility and state 1 represents a state of larger volatility, and states more often switch between 0 and 1, when compared with the French case.

In the British case, a state 0 represents a state of small variability, since the distribution of β_0 is flatter than the distribution of β_1. This is also implied in the second figure in Figure 13.15, by the fact that β_0 varies more than β_1. σ_0 takes on larger values but it also has a larger variance. Figure 13.18 indicates that state 0 is a state of lower volatility and state 1 is a higher volatility state. In the British case, state stays in a higher volatility state most of the time after the observation 700.

Table 13.7 Estimated Posterior Means of the Parameters

	ρ	β_0	β_1	σ_0	σ_1	ε_0	ε_1
France	-0.000007	-6.50	-9.11	0.89	1.09	0.01	0.01
	(0.00)[a]	(0.000)	(0.09)	(0.003)	(0.008)	(0.004)	(0.004)
Germany	-0.00001	-9.78	-8.16	0.79	0.66	0.01	0.02
	(0.000)[b]	(0.016)	(0.014)	(0.003)	(0.004)	(0.001)	(0.002)
Britain	0.00001	-12.78	-9.23	1.59	0.84	0.04	0.004
	(0.000)[c]	(1.030)	(0.037)	(0.058)	(0.001)	(0.002)	(0.003)

[a]This does not mean that the variance is zero. The value is very small (3.61922e-10).
[b]5.230819e-10.
[c]1.940401e-10.

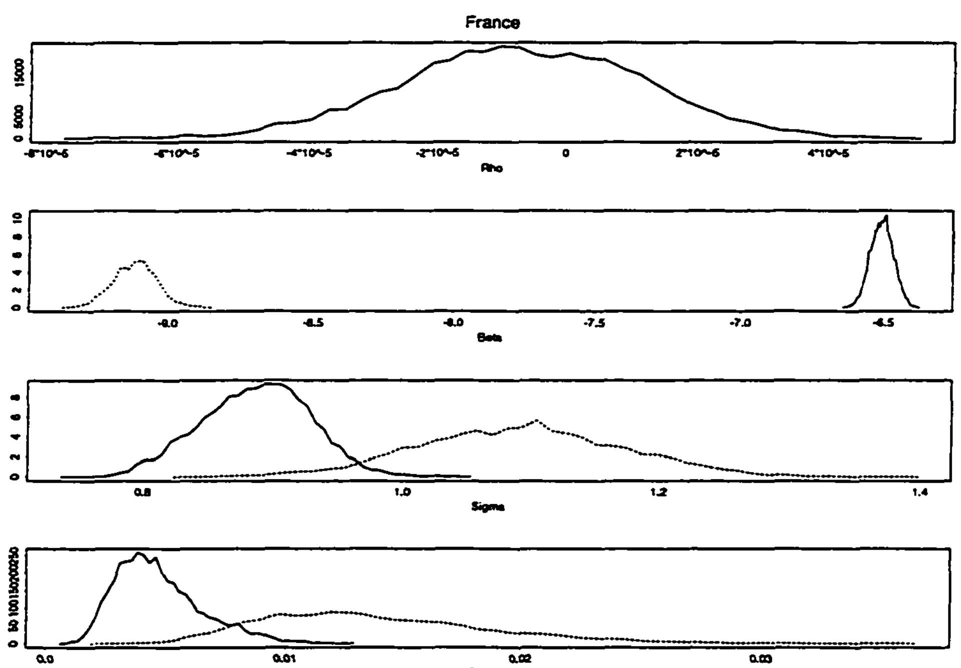

Figure 13.10 Estimated Marginal Posterior Distribution: France

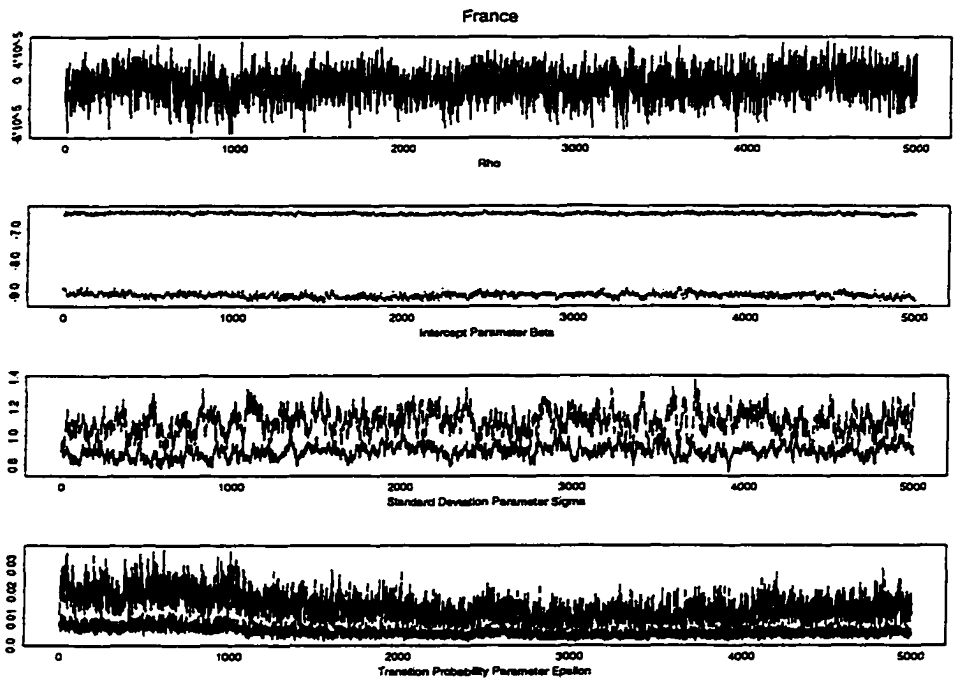

Figure 13.11 Estimated Parameters: France

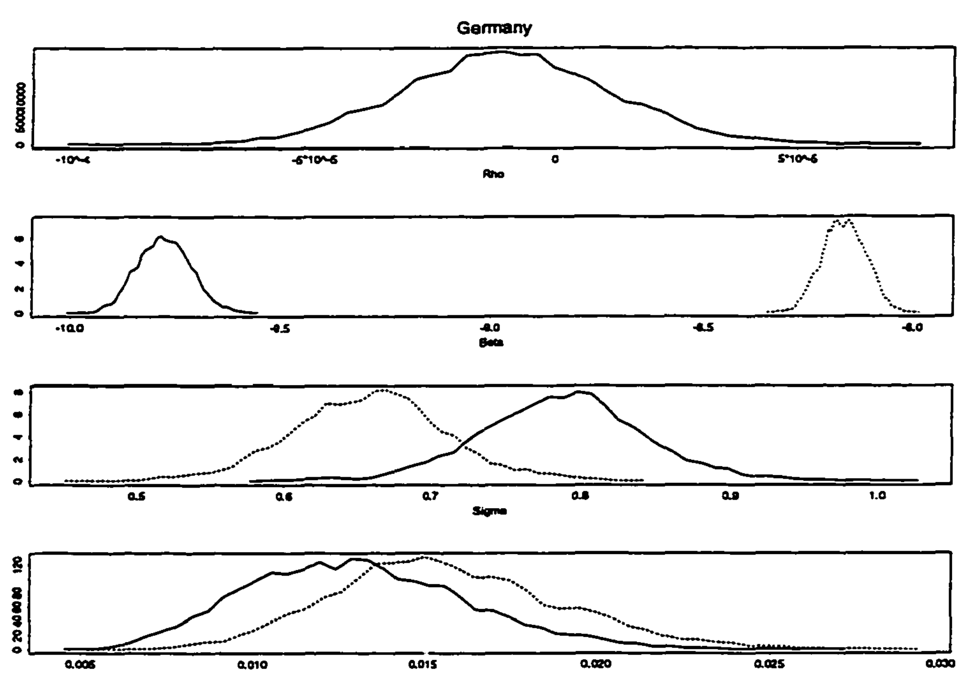

Figure 13.12 Estimated Marginal Posterior Distribution: Germany

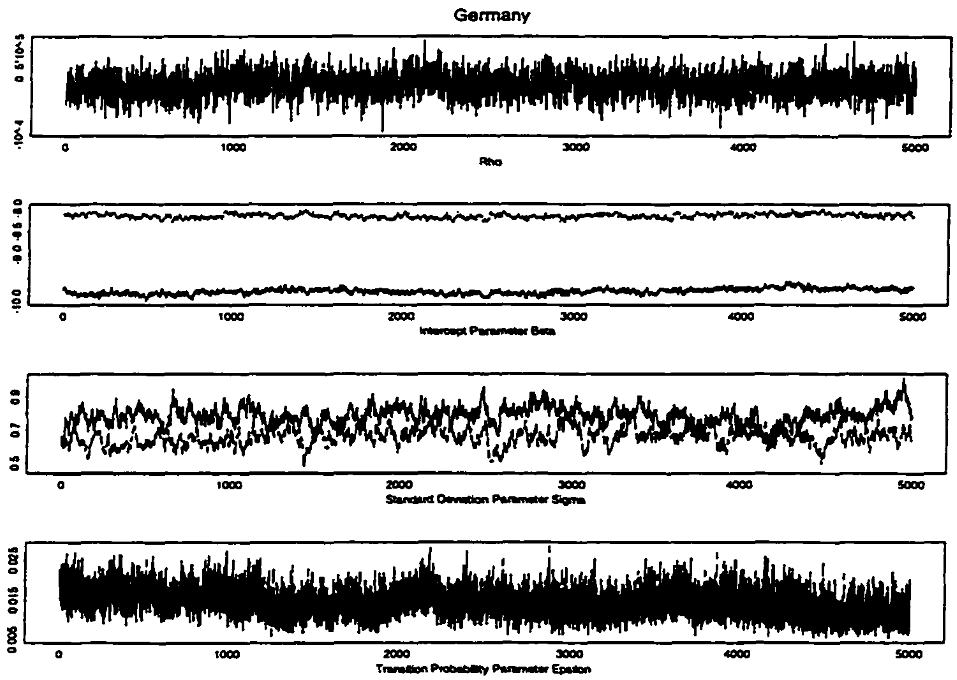

Figure 13.13 Estimated Parameters: Germany

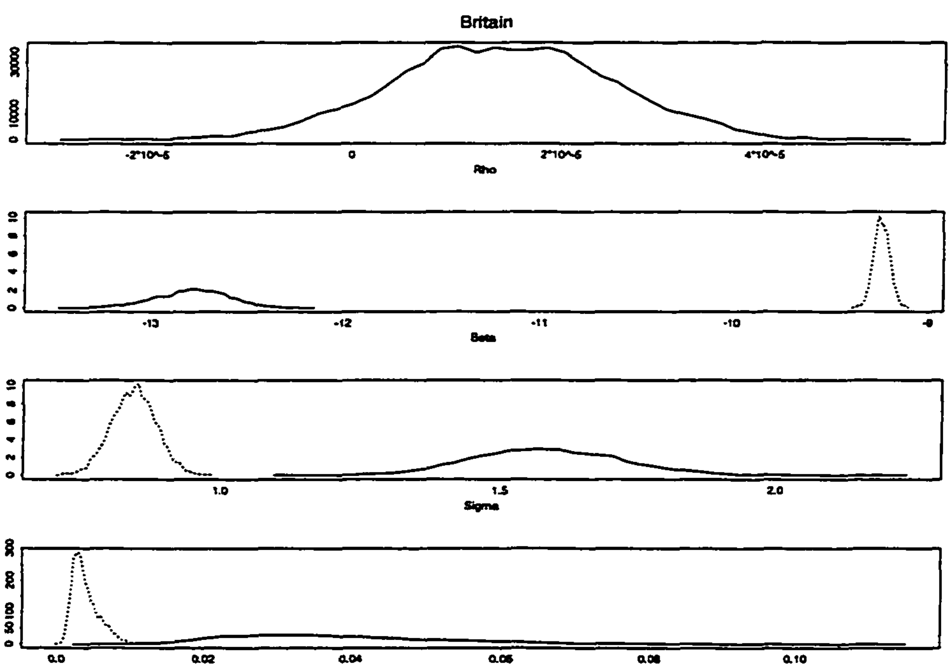

Figure 13.14 Estimated Marginal Posterior Distribution: Britain

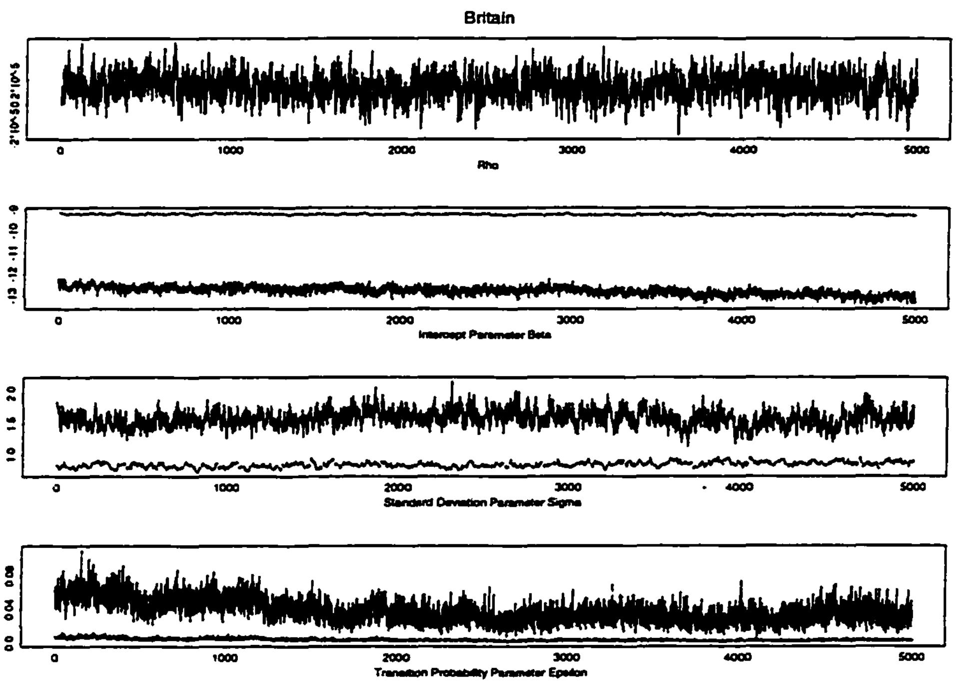

Figure 13.15 Estimated Parameters: Britain

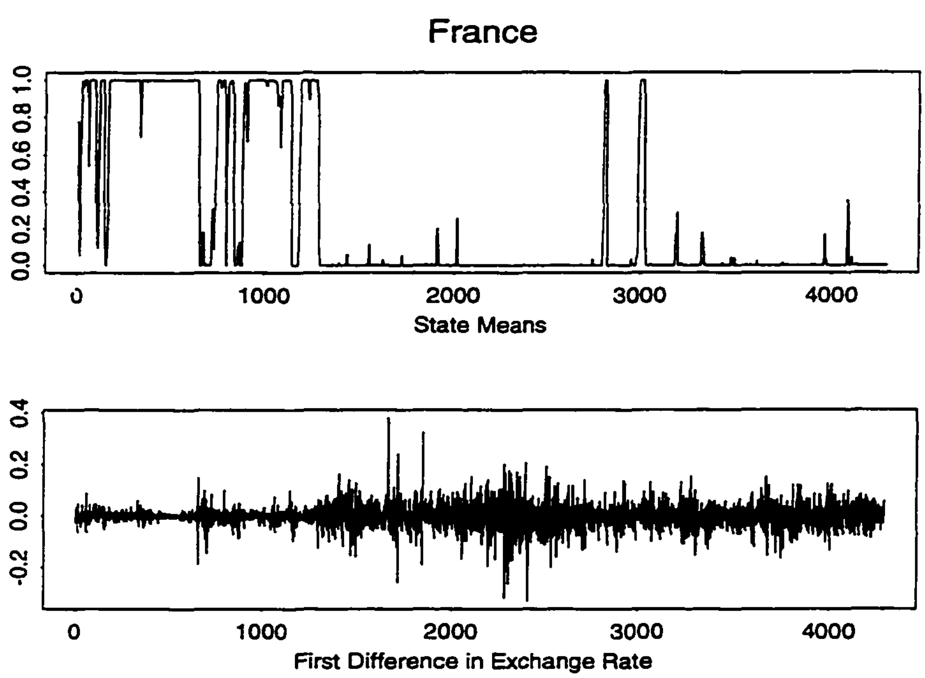

Figure 13.16 State Means and Data: France

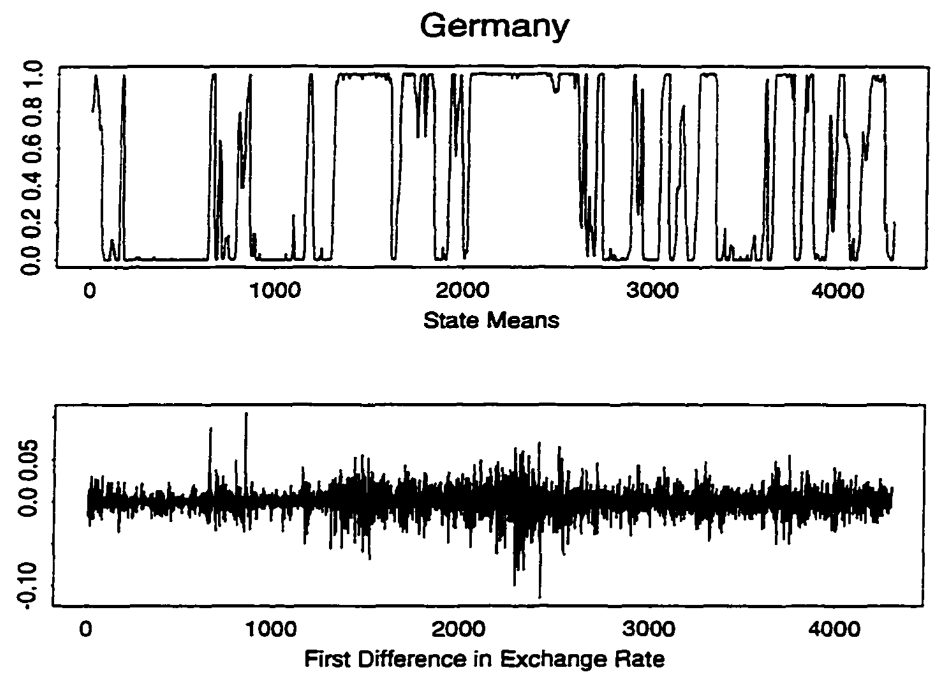

Figure 13.17 State Means and Data: Germany

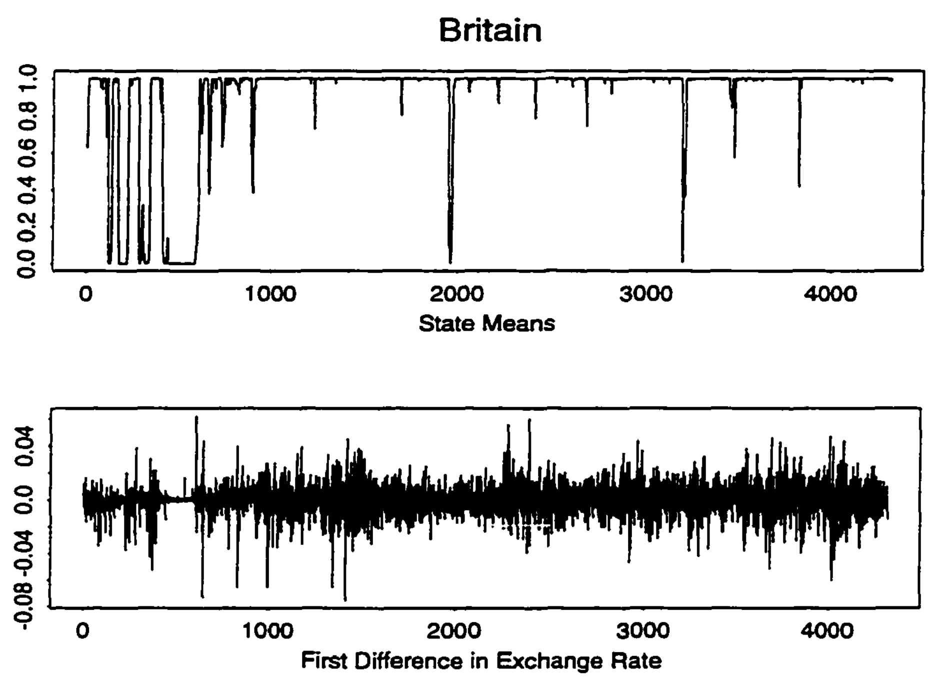

Figure 13.18 State Means and Data: Britain

14 CONCLUSION

This part attempted to model the volatility of exchange rates, applying a regime-switching stochastic volatility model to the exchange rate data to examine the volatility of exchange rates. The model used, the four-regime-switching stochastic volatility model, was an extension of Schmidt's two-regime-switching stochastic volatility model. Observable states, depending on interest rate differential along with unobservable states, were introduced for the modification, specifically, the assumption that the interest rate differential is equal to expected appreciation or depreciation, if the interest parity condition holds. Introduction of another set of unobservable states would make the model more complicated, however, by introducing a set of observable states, the model was extensively simplified. In terms of the model parameters, there were four different βs; β_1 to β_4. To avoid further complication, it was assumed that the variability parameter, σ, depends only on unobservable states. In other words, there are only two variability parameters in the model, including only two transition probability parameters, ε_1 and ε_2. As Schmidt pointed out, the primary advantage for the model is its ability to allow for the possibility of multiple states and this has been achieved by our model. In all cases, the French, German and British cases, it is clear that there exist two distinct unobservable states. This can be seen from the fact that there are always two distinct sets of parameters βs. In some cases (for example, the German case), the variability parameters σs also depend on these unobservable states. Regardless of the issue of identifiability, these two states can be interpreted as the high volatility state and the low volatility state. These results correspond to the results obtained in Schmidt. Using the value-weighted market index from the Center for Research in Security Prices (CRSP), she also found that two unobservable states play a rather important role for the volatility in the market index.

On the other hand, the observable states introduced here do not play as crucial a role as the unobservable states. This is partly because the choice of variable, the interest rate, may not be a good one. For instance, in the German case, the interest rate plays a relatively important role in the model, as evidenced by the fact that the values of β_1 and β_3, and β_2 and β_4 are more discernible. We can see the same implication from the computed posterior probabilities $P(\beta_1 > \beta_3)$ and $P(\beta_2 > \beta_4)$ in Table 13.6.

However, in the French and British case, the role of the observable states is not clear, except for the state of high volatility in Britain, as can be seen by the overlaps in the marginal posterior distributions of βs and σs. After changing the value of k, which is the interest rate bound, the results do not appear to change dramatically. These results are reported in the Appendix. In some cases, for instance, the French case with the 4% bounds, the role of the interest rate seems to become more important. In the other cases tested, the results are similar.

It is also observed that the exchange rate very often causes the state to switch between unobservable state A and B in the German case, while the exchange rate seldom causes this switch between states in the French and British cases. Both the French and British exchange rates stay in a high volatility state most of the time. Introduction of interest rates as observable states did not give clear results.

To examine the relationship between exchange rate and interest rates along with exchange rate volatility, a mean model of regime-switching volatility model, derived from the interest parity condition, was introduced. The results indicate that interest rates do not have explanatory power for exchange rates. So, it is concluded that exchange rates simply follow a random walk process. This result was observed in all three countries. On the other hand, two distinct states in exchange rates were found to exist. They can be interpreted as state of a high volatility and state of a low volatility.

Forecasts using the above regime-switching stochastic volatility model are very possible. Here is a rough sketch of the prediction procedures. The procedure will start with the following predictive posterior distribution:

$$P(y_{t+1}|\boldsymbol{y}) = \int P(y_{t+1}|\alpha_{t+1})P(\alpha_{t+1}|s_{t+1},\theta)P(s_{t+1}|\theta)P(\boldsymbol{\theta}|\boldsymbol{y})\,d\alpha_{t+1}\,ds_{t+1}\,d\boldsymbol{\theta} \tag{14.1}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\sigma}, \boldsymbol{\epsilon}, \boldsymbol{s})$. This will be approximated as follows:

$$P(y_{t+1}|\boldsymbol{y}) \approx \frac{1}{K}\sum_{k=1}^{K} P(y_{t+1}^{(k)}|\alpha_{t+1}^{(k)})P(\alpha_{t+1}^{(k)}|s_{t+1}^{(k)},\theta^{(k)})P(s_{t+1}^{(k)}|\theta^{(k)}) \tag{14.2}$$

where k indicates the kth iteration of Gibbs sampling. It is possible to simulate state variable s_{t+1} and continue to find all other parameters, and then, finally y_{t+1}. Finding s_{t+1}, it is then possible to find $\beta_{s_{t+1}}$, $\sigma_{s_{t+1}}$ and $\alpha_{s_{t+1}}$. Therefore, the predictive distribution of y_{t+1} will be derived based on these values. Updating state variable $t+2, t+3, \cdots$ makes it possible to forecast further y_{t+2}, y_{t+3}.

The forecasting issue may be approached in a similar way for the mean model. However, not only parameters are needed but also the value of x_{t+1} to forecast y_{t+1}. It is necessary to model the process $\{x_t\}$.

15 GENERAL CONCLUSION

Each of the two statistical models, the cointegration partial model and the stochastic volatility model, that were discussed and applied to the data set in this thesis, investigated different exchange rate related questions. The first part of the thesis applied the cointegration partial system model to a set of monthly data that included exchange rates, money supplies, and GNPs. The goal here was to investigate the third country effects on the exchange rate determination, and, indeed, adding the third country's variables drastically increases the number of parameters in the model. The partial system model solves this problem by adopting the concept of weak endogeneity. While the results indicate some evidence that the third country's effects can not be ignored, their interpretations are not obvious. In particular, the three-country theoretical model, based on Dornbusch sticky price model, could not explain the empirical results well, since, in the end, the signs of coefficients did not follow the signs predicted by the model.

For further research, the model can be modified by introducing other assumptions, particularly, the interest parity condition that was assumed by the model. As some past research has reported an inability of the condition, relaxing the interest parity condition may yield superior results . Similarly, the model might also be extended using one of the other exchange rate determination models discussed in the first part, rather than Dornbusch's sticky price model, which has served as the base for the model presented here.

In the second part, the regime-switching stochastic volatility model was applied to the daily exchange rate data in order to investigate volatility of the exchange rates and, simultaneously to examine the relation between daily interest rates and daily exchange rates. Here, the results did not find any relations between interest rate and exchange rate, which implies that the daily exchange rate follows a random walk process. However, the model successfully captured the two different regimes; the highly volatile state and the less volatile state.

For further study, this model can be extended in many directions. Different economic assumptions will create more and different structural assumptions that may be imposed on the model. Also, the a priori structures imposed on the relation between interest rate and exchange rate might have been unrealistic, that is, the interest rate may be influenced by the exchange rate. If this is indeed the case, the interest rate should be endogenized in the model. Finally, prediction of the exchange rate using the model is also an intriguing topic for further research.

APPENDIX

In this appendix we will report the results for $k = 4$ and $k = 5$. In other words, these are the results for setting the interest differential bounds to be 4% and 5%.

Table A.1 Data Summary for Daily Exchange Rates: $k = 4$

Variables	N	Mean	Standard Error	Skewness	Kurtosis
EF_G	811	5.45	0.75	1.14	1.68
EF_L	3408	5.99	1.50	0.79	-0.07
ΔEF_G	811	0.00	0.04	0.69	6.33
ΔEF_L	3408	0.00	0.04	-0.25	7.76
EG_G	1451	2.04	0.43	0.67	-0.72
EG_L	2843	2.14	0.43	0.54	-0.38
ΔEG_G	1451	0.00	0.01	0.11	3.61
ΔEG_L	2843	0.00	0.01	-0.48	5.70
EUK_G	1671	1.77	0.29	0.33	0.22
EUK_L	2645	1.74	0.28	0.40	-0.43
ΔEUK_G	1671	0.00	0.01	-0.39	3.00
ΔEUK_L	2645	0.00	0.01	-0.29	4.24

Table A.2 Data Summary for Daily Interest Rates: $k = 4$

Variables	N	Mean	Standard Error	Skewness	Kurtosis
$INTF_G$	811	4.16	4.26	-1.87	2.84
$INTF_L$	3408	1.16	1.85	-0.66	-0.11
$INTG_G$	1451	-2.67	5.09	0.80	-0.86
$INTG_L$	2843	-1.59	2.04	1.15	0.60
$INTUK_G$	1671	5.05	3.53	-2.57	6.86
$INTUK_L$	2645	1.44	1.82	-0.66	-0.34

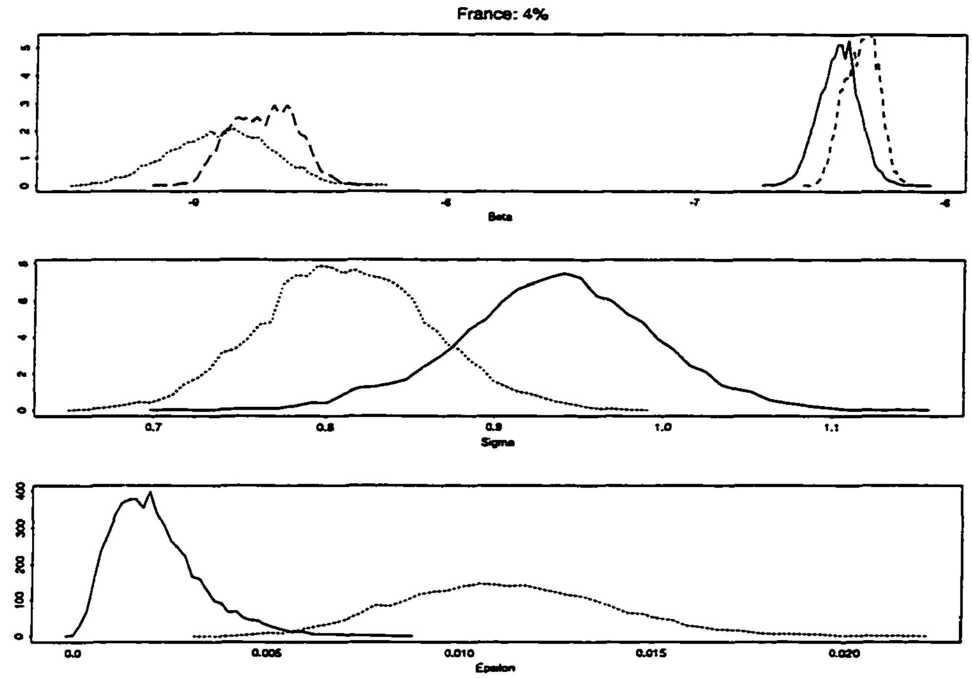

Figure A.1 Estimated Marginal Posterior Distribution: France, $k = 4$

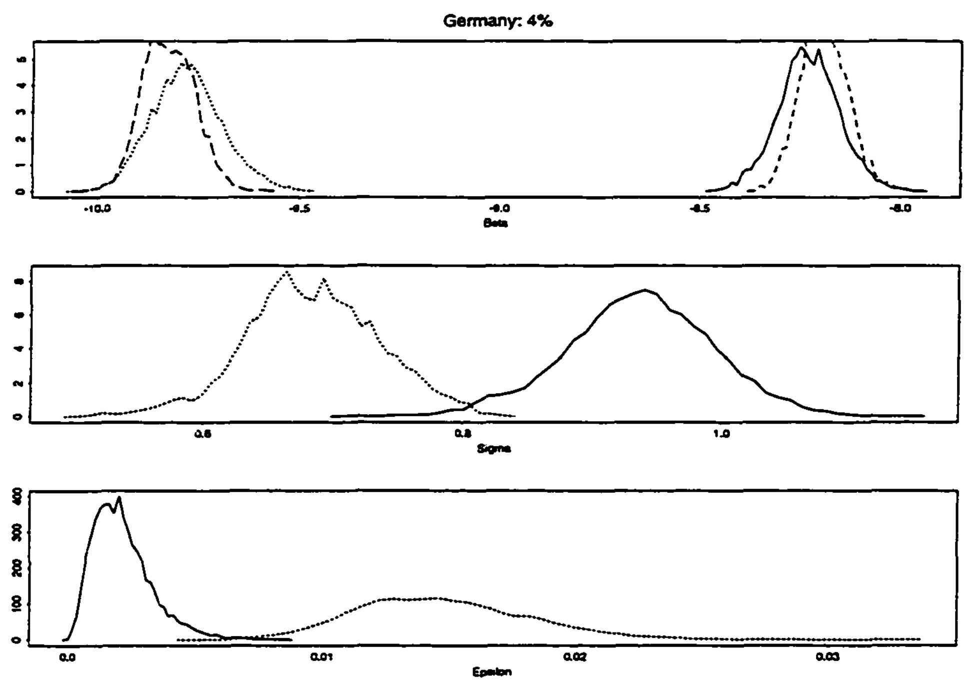

Figure A.2 Estimated Marginal Posterior Distribution: Germany, $k = 4$

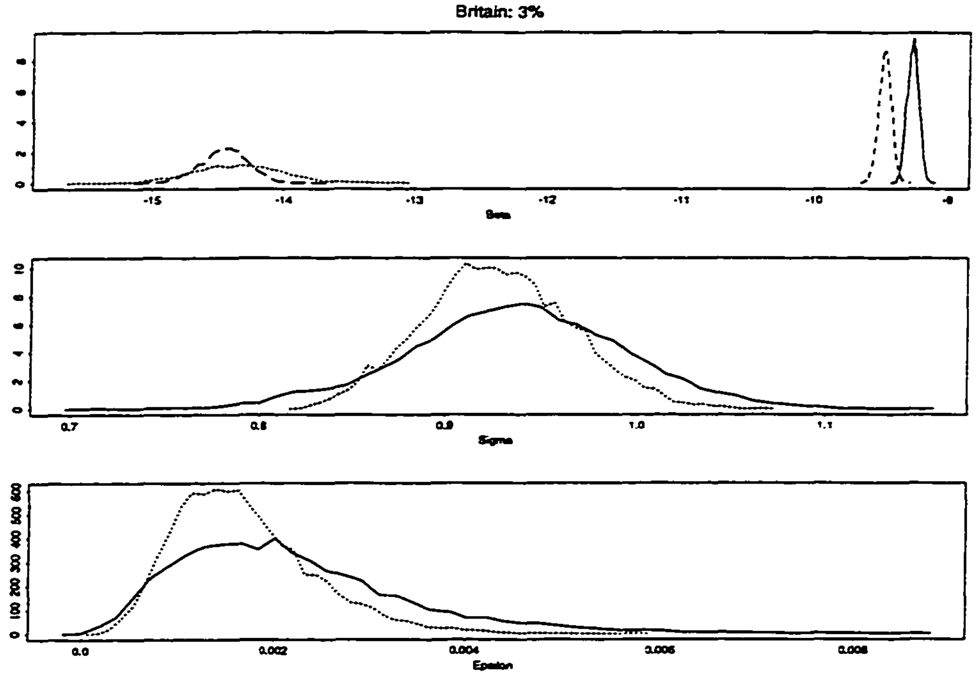

Figure A.3 Estimated Marginal Posterior Distribution: Britain, $k = 4$

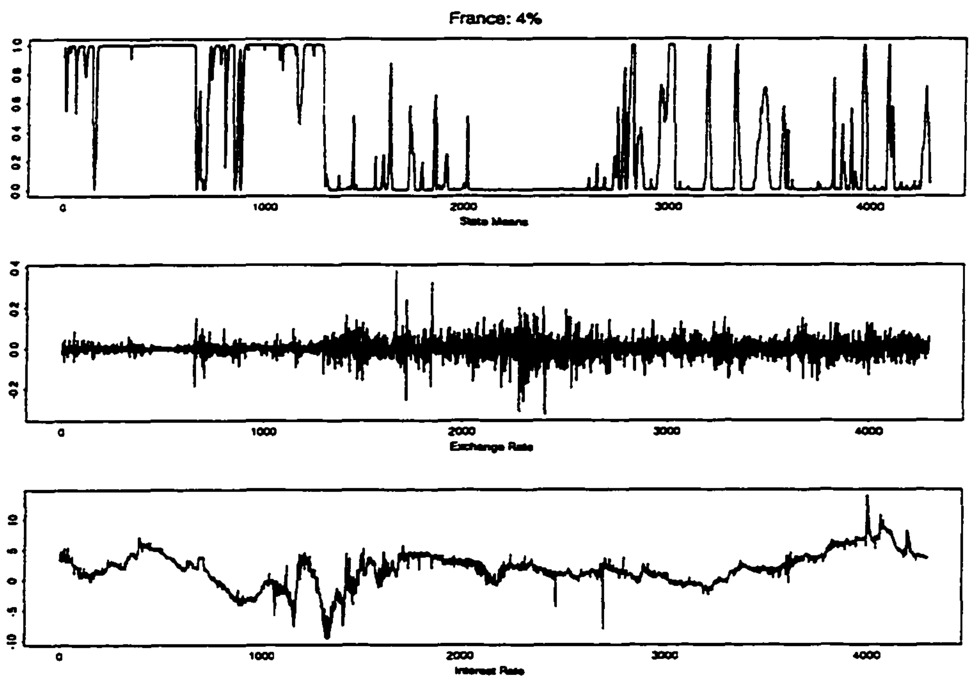

Figure A.4 State Means and Date: France, $k = 4$

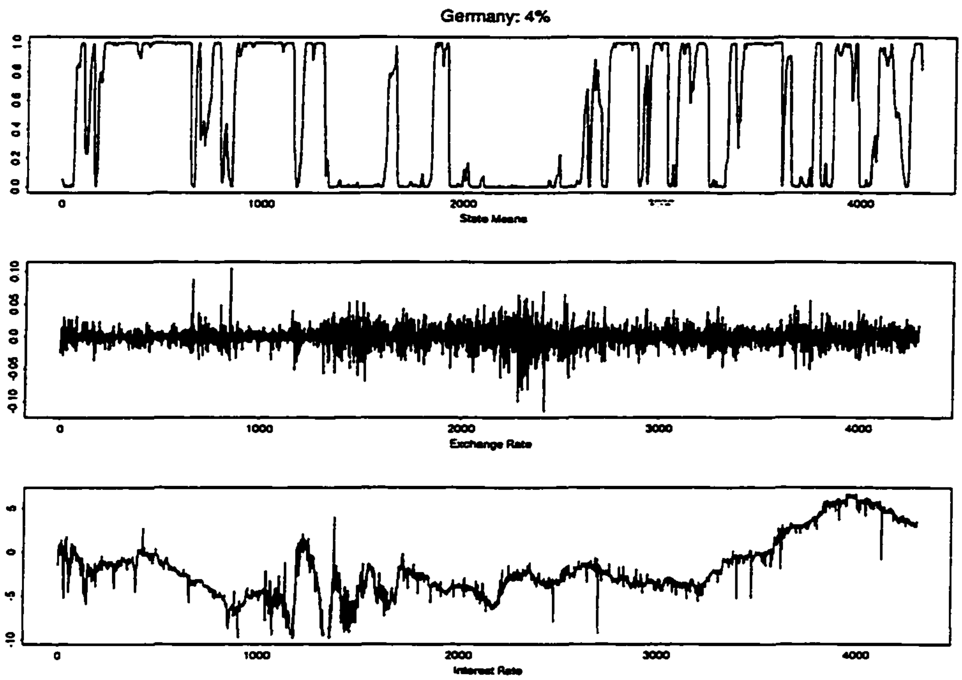

Figure A.5 State Means and Data: Germany, $k = 4$

Table A.3 Estimated Posterior Means of the Parameters: $k = 4$

	β_1	β_2	β_3	β_4	σ_1	σ_2	ε_1	ε_2
France	-6.42	-8.90	-6.34	-8.73	0.81	1.11	0.01	0.02
	(0.006)	(0.038)	(0.004)	(0.016)	(0.003)	(0.006)	(0.000)[a]	(0.000)
Germany	-8.23	-9.78	-8.19	-9.82	0.68	0.78	0.01	0.01
	(0.006)	(0.008)	(0.004)	(0.004)	(0.003)	(0.003)	(0.000)	(0.000)
Britain	-9.26	-14.37	-9.47	-14.47	0.93	0.59	0.001	0.05
	(0.002)	(0.117)	(0.002)	(0.032)	(0.001)	(0.039)	(0.000)	(0.000)

[a]Again, this does not mean that the variance is zero. The value is very small (7.332668e-06).

Table A.4 Comparisons of βs : Posterior Probabilities: $k = 4$

	$P(\beta_1 > \beta_3)$	$P(\beta_2 > \beta_4)$	$P(\sigma_1 > \sigma_2)$
France	0.16	0.14	0.01
Germany	0.31	0.68	0.10
Britain	1.00	0.39	0.94

Table A.5 Data Summary for Daily Exchange Rates: $k = 5$

Variables	N	Mean	Standard Error	Skewness	Kurtosis
EF_G	551	5.22	0.41	0.45	1.91
EF_L	3738	5.98	1.47	0.80	0.03
ΔEF_G	551	0.00	0.04	0.33	1.60
ΔEF_L	3738	0.00	0.04	-1.33	8.17
EG_G	880	1.98	0.43	0.88	-0.27
EG_L	3414	2.14	0.43	0.52	-0.51
ΔEG_G	880	0.00	0.02	0.19	4.10
ΔEG_L	3414	0.00	0.01	-0.42	5.26
EUK_G	1250	1.81	0.27	0.20	1.01
EUK_L	3066	1.73	0.26	0.28	-0.21
ΔEUK_G	1250	0.00	0.12	-0.46	3.57
ΔEUK_L	3066	0.00	0.01	-0.26	3.63

Table A.6 Data Summary for Daily Interest Rates: $k = 5$

Variables	N	Mean	Standard Error	Skewness	Kurtosis
$INTF_G$	551	4.54	4.97	-1.83	2.07
$INTF_L$	3738	1.37	2.04	-0.61	-0.06
$INTG_G$	880	-2.52	5.93	0.58	-1.39
$INTG_L$	3414	-1.81	2.38	1.13	0.59
$INTUK_G$	1250	5.36	3.89	-2.60	6.25
$INTUK_L$	3066	1.81	2.05	-0.62	-0.18

Table A.7 Estimated Posterior Means of the Parameters: $k = 5$

	β_1	β_2	β_3	β_4	σ_1	σ_2	ε_1	ε_2
France	-6.48	-9.41	-6.47	-9.01	0.87	1.10	0.01	0.02
	(0.009)	(0.056)	(0.002)	(0.006)	(0.002)	(0.007)	(0.000)	(0.000)
Germany	-9.87	-8.26	-9.88	-8.27	0.74	0.70	0.02	0.01
	(0.013)	(0.008)	(0.007)	(0.004)	(0.004)	(0.002)	(0.000)	(0.000)
Britain	-9.13	-13.22	-9.33	-13.44	0.86	1,43	0.003	0.04
	(0.003)	(0.092)	(0.001)	(0.039)	(0.002)	(0.023)	(0.000)	(0.000)

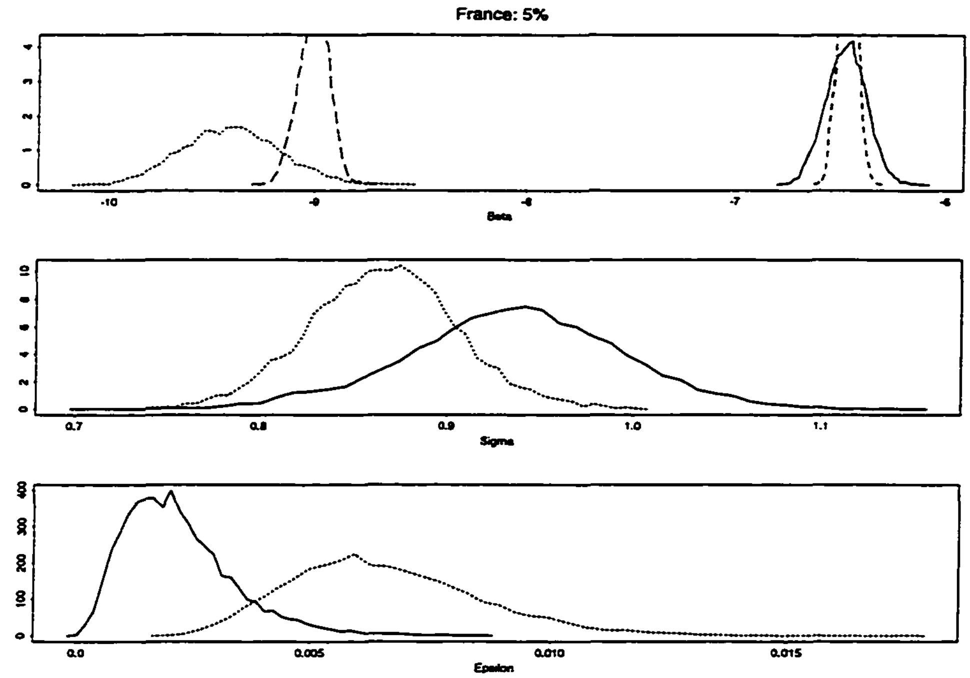

Figure A.6 Estimated Marginal Posterior Distribution: France, $k = 5$

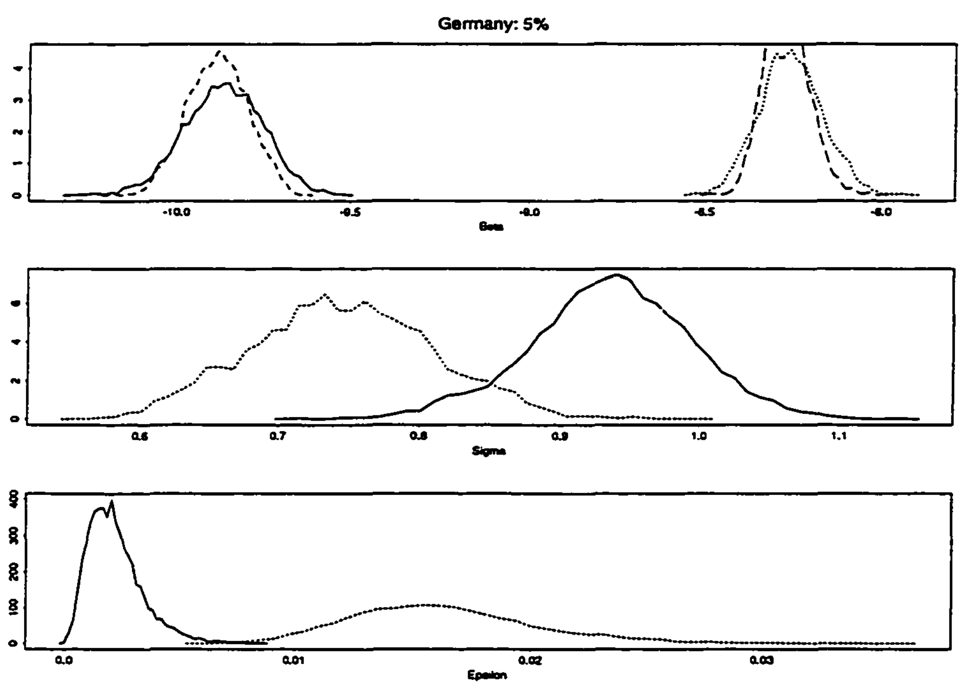

Figure A.7 Estimated Marginal Posterior Distribution: Germany, $k = 5$

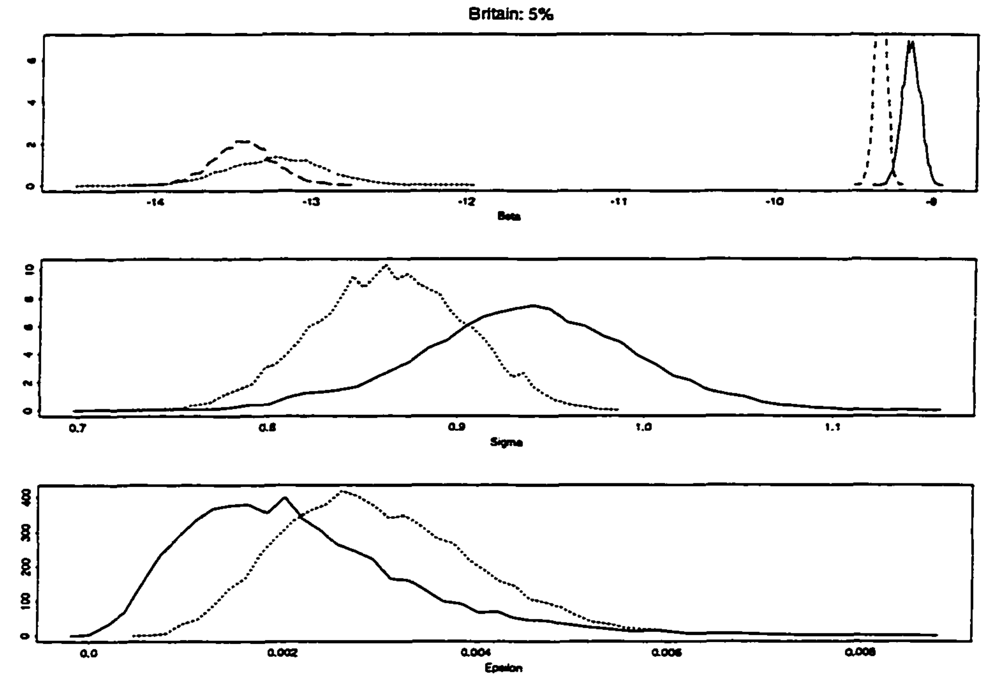

Figure A.8 Estimated Marginal Posterior Distribution: Britain, $k = 5$

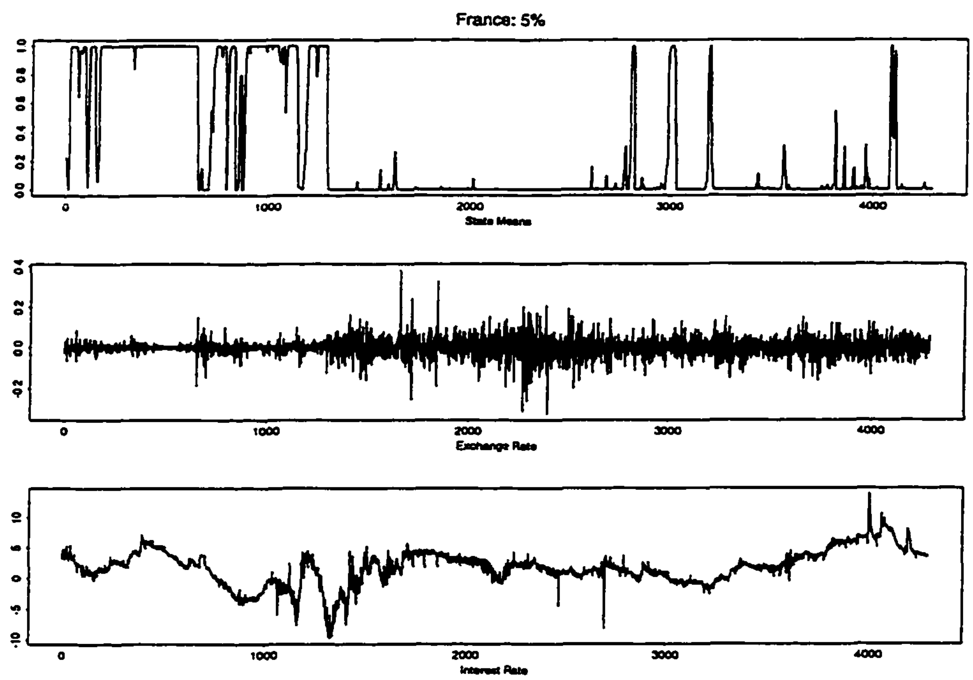

Figure A.9 State Means and Data: France, $k = 5$

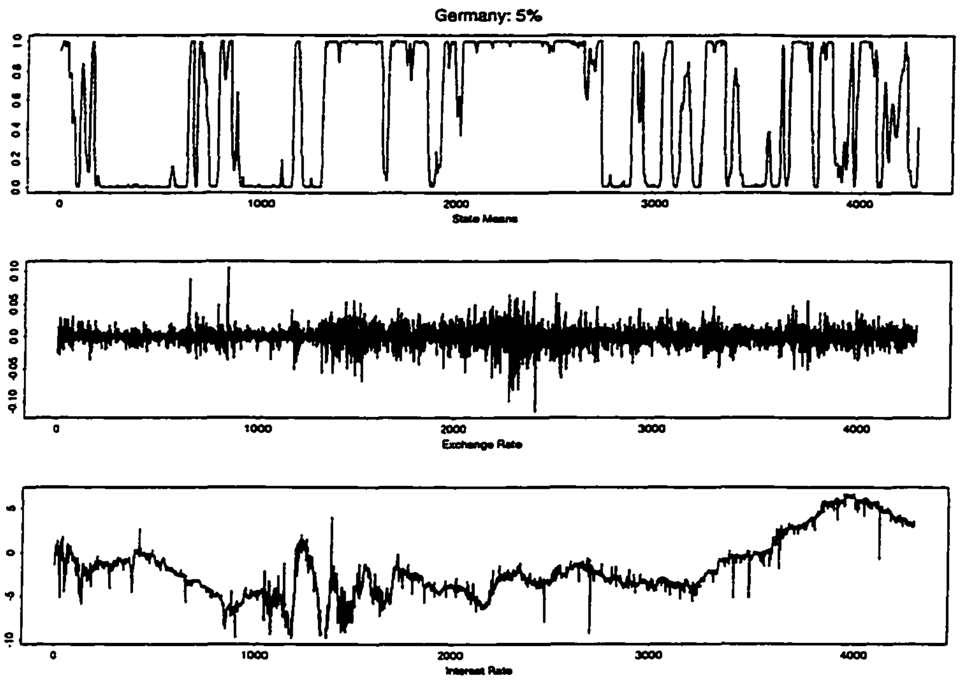

Figure A.10 State Means and Data: Germany, $k = 5$

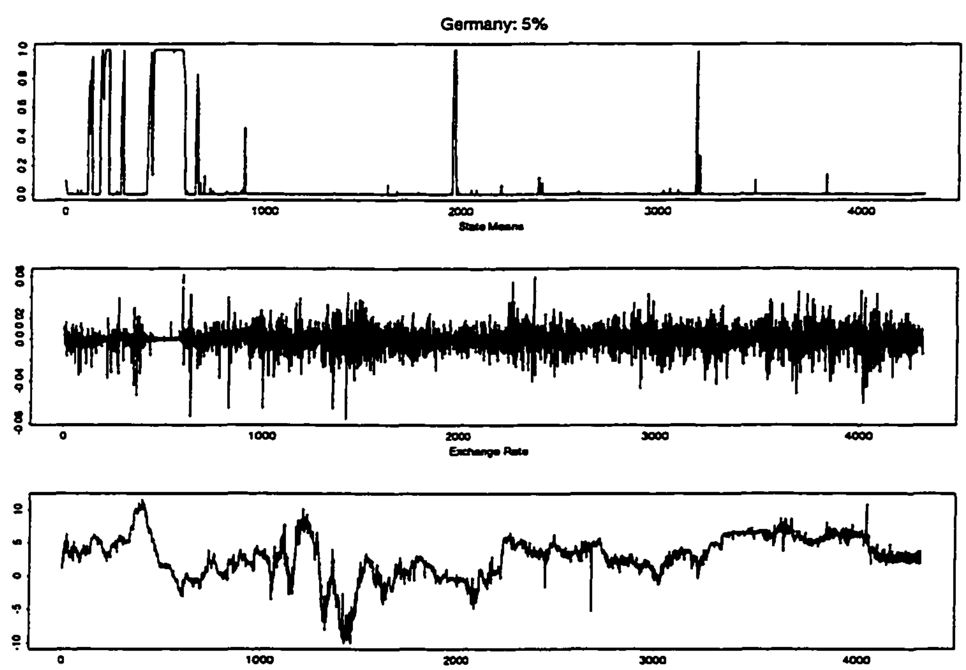

Figure A.11 State Means and Data: Britain, $k = 5$

BIBLIOGRAPHY

Abramovitz, M. and N. C. Stegun. (1964) *Handbook of Mathematical Functions with Formulas, Graphs, and Tables.* National Bureau of Standard Applied Mathematics Series 55, Washington.

Baillie, R. T. and Selover, D. D. (1987) "Cointegration and models of exchange rate determination." *International Journal of Forecasting,* v3: 43-51.

Baillie, R. T. and McMahon, P. C. (1989) *The Foreign Exchange Market: Theory and Econometric Evidence.* Cambridge University Press, Cambridge.

Baillie, R. T. and Bollerslev, T. (1989) "The message in daily exchange rates: A conditional-variance tale." *Journal of Business and Economic Statistics,* v7: 297-305.

Bain, L. J. and Engelhardt, M. (1990) *Introduction to Probability and Mathematical Statistics.* PWS-KENT Publishing Company, Boston.

Banerjee, Anindya, Dolado, Juan, Galbraith, John W., and Hendry, David F. M. (1993) *Co-integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford University Press, Oxford.

Berndt, E. K., Hall, B. H., Hall, R. E. and Hausman, J. A. (1974) "Estimation and inference in nonlinear structural models." *Annals of Economic and Social measurement,* 3: 653-65.

Bilson, John F. O. (1978a) "Rational expectations and the exchange rate" *The Economics of Exchange Rates: Selected Studies,* ed. by J. A. Frenkel and H. G. Johnson, Addison-Wesley, Reading, Mass.

Bilson, John F. O. (1978b) "The monetary approach to the exchange rate: Some empirical evidence." *IMF Staff Papers,* 25: 48-75.

Blanchard, Oliver, and Quah, Danny. (1989) "The dynamic effects of aggregate demand and aggregate supply disturbances." *American Economic Review,* 79: 655-673.

Blanchard, Oliver, and Watson, Mark. (1986) "Are business cycle all alike?" *The American Business Cycle,* ed. by Robert J. Gordon, University of Chicago Press, Chicago.

Bollerslev, T. (1986) "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics,* 51: 309-324.

Bollerslev, T. (1987) "A conditionally heteroskedastic time series model for speculative prices and rates of return." *Review of Economics and Statistics,* 69: 542-547.

Bollerslev, T. (1990) "Modelling the coherence in short-run nominal exchange rate: A multivariate generalized ARCH approach." *Review of Economics and Statistics,* 72: 498-505.

Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992) "ARCH modeling in finance: A review of the theory and empirical evidence." *Journal of Econometrics,* 52: 5-59.

Bollerslev, T., and Wooldridge, J. M. (1992) "Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances." *Econometric Reviews,* v11: 143-172.

Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994) "ARCH models." *Handbook of Econometrics,* vol.IV, ed. by R. F. Engle and D. L. McFadden, North-Holland, Amsterdam: 2959-3038.

Branson, W. H. (1968) *Financial Capital Flows in the US Balance of Payments.* North-Holland, Amsterdam.

Branson, W. H. (1975) "Comment on Whitman", *Brookings Papers on Economic Activity,* Brookings Institution, Washington: 537-541.

Breidt, F. J. and Carriquiry, A. L. (1996) "Improved quasi-maximum likelihood estimation for stochastic volatility models." *Modelling and Prediction: Honoring Seymour Geisser,* ed. by Jack C. Lee, Wesley O. Johnson and Arnold Zellner, Springer-Verlag, New York: 228-247.

Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods.* Springer-Verlag, New York.

Calvo, Guillermo and Rodriguez, Carlos A. (1977) "A model of exchange rate determination under currency substitution and rational expectations." *Journal of Political Economy,* 85: 617-625.

Campbell, John J., and Perron, Pierre. (1991) "Pitfalls and opportunities: What macroeconomists should know about Unit Roots." *Macroeconomics Annual, 1991,* NBER, Cambridge, Mass.: 141-201.

Clark, P. K. (1973) "A subordinate stochastic process model with finite variance for speculative prices." *Econometrica,* 41: 135-155.

Dibooglu, S. (1993) *Multiple cointegration and structural models: applications to exchange rate determination,* Iowa State University thesis, Ames, IA.

Dibooglu, S, and Enders, W. (1994) "Multiple cointegrating vectors and structural economic models: An application to the French Franc/U.S. Dollar Exchange Rate." *Southern Economic Journal,* 61: 1098-1116.

Dickey, D. A., and Fuller, W. A. (1979) "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association,* 74: 427-431.

Dickey, D. A., and Fuller, W. A. (1981) "Likelihood ratio statistics for autoregressive time series with a unit root." *Econometrica,* 49: 1057-1072.

Dornbusch, Rudigar. (1976a) "Expectations and exchange rate dynamics." *Journal of Political Economy,* 84: 1161-1176.

Dornbusch, Rudigar. (1976b) "Exchange rate expectations and monetary policy." *Journal of International Economics,* 6: 231-244.

Dornbusch, Rudigar. (1980a) "Exchange rate economics: Where do we stand ?" *Brookings Papers on Economic Activity,* 1, Brookings Institution, Washington: 143-185.

Dornbusch, Rudigar. (1980b) *Open Economy Macroeconomics,* Basic, New York.

Enders, W. (1995) *Applied econometric time series,* Wiley, New York.

Engle, R. F. (1982) "Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation." *Econometrica,* 50: 987-1007.

Engle, R. F., Hendry, D. F., and Richard, J. F. (1983) "Exogeneity." *Econometrica,* 51: 277-304.

Engle, R. F. and Bollerslev, T. (1986) "Modelling the persistence of conditional variance." *Econometrics Review,* 5: 1-50.

Frankel, Jeffrey A. (1979) "On the Mark: A theory of floating exchange rates based on real interest differentials." *American Economic Review,* 20: 407-417.

Frankel, Jeffrey A. (1983) "Monetary and Portfolio-Balance Models of Exchange Rate Determination." *Economics Interdependence and Flexible Exchange Rates,* ed. by J. Bhandari and B. Putnam, MIT Press, Cambridge, Mass.

Frankel, Jeffrey A. (1993) *On Exchange Rates,* MIT Press, Cambridge, Mass.

Frenkel, Jacob A. (1976) "A monetary approach to the exchange rate: Doctrinal aspects and empirical evidence." *Scandinavian Journal of Economics,* 78: 200-204.

Frenkel, Jacob A. (1978) "Purchasing power parity: Evidence from the 1920s." *Journal of International Economics,* 8: 161-191.

Frenkel, Jacob A and Mussa, M. L. (1985) "Asset markets, exchange rates, and the balance of payments." *Handbook of International Economics,* vol.II, ed. by R. W. Jones and P. B. Kenen, North-Holland, Amsterdam: 679-747.

Frenkel, Jacob A. and Rodriguez, C. A. (1982) "Exchange rate dynamics and the overshooting hypothesis." *IMF Staff Papers*: 1-30.

Fuller, W. A. (1996) *Introduction to Statistical Time Series,* John Wiley, New York.

Gardeazabal, J. and Regulez, M. (1992) *The monetary model of exchange rates and cointegration: estimation, testing and prediction,* Springer-Verlag, New York.

Geweke, John. (1982) "Measurement of linear dependence and feedback between multiple time series." *Journal of American Statistical Association,* 77: 304-314.

Girton, R. and Roper, D. (1986) "Theory and implications of currency substitution." *The monetary approach to international adjustment.* ed. by Bluford H. Putnam and D. Sykes Wilford, Praeger, New York.

Glosten, L. R., Jogannathan, R. and Runkle, D. (1989) "Relationship between the expected value and the volatility of the nominal excess return on stocks." Mimeo, Northwestern University, IL.

Godfrey, L. G. (1988) *Misspecification tests in econometrics: The Lagrange Multiplier principle and other approaches,* Cambridge University Press, Cambridge.

Greene, W. H. (1990) *Econometric Analysis,* Macmillan, New York.

Hamilton, J. (1994) *Time Series,* Princeton University Press, Princeton, New Jersey.

Hansen, L. P. (1982) "Large sample properties of generalized method of moments estimators." *Econometrica,* v50: 1029-1054.

Hansen, H. and Juselius, K. (1995) *CATS in RATS: Cointegration Analysis of Time Series,* Estima, Evanston.

Harboe, I., Johansen, S., Nielsen, B. G., and Rahbek, A. C. (1995) "Test for cointegrating rank in partial systems." Discussion Paper, University of Copenhagen, Denmark.

Harvey, A. C. (1990) *The Econometric Analysis of Time Series,* MIT Press, Cambridge, Mass.

Harvey, A. C. (1991) *Forecasting, Structural Time Series Models and the Kalman Filters,* Cambridge University Press, Cambridge.

Harvey, A. C. (1993) *Time Series Models,* MIT Press, Cambridge, Mass.

Harvey, A. C. and Shephard, N. G. (1992) "Structural time series models." *Handbook of Statistics,* vol.11, ed. by Rao, C. R. and Maddala, G. S., North-Holland, Amsterdam.

Harvey, A. C. and Ruiz, E. and Shephard, N. G. (1994) "Multivariate stochastic variance models." *Review of Economic Studies,* v61: 247-64.

Hatanaka, M. (1996) *Time-Series-Based Econometrics: Unit Roots and Co-Integrations,* Oxford University Press, Oxford.

Hendry, M. (1995) *Dynamic Econometrics,* Oxford University Press, Oxford.

Hodrick, R. J. (1978) "An empirical analysis of the monetary approach to the determination of the exchange rate." *The Economics of Exchange Rates: Selected Studies,* ed. by J. A. Frenkel and H. G. Johnson. Addison-Wesley, Reading, Mass.

Hossain, F. (1995) *Current account determination in the intertemporal framework: an empirical analysis,* Iowa State University thesis, Ames, IA.

Hsieh, D. A. (1989) "Modeling heteroscedasticity in daily foreign-exchange rates." *Journal of Business and Economic Statistics,* 7: 307-317.

International Financial Statistics, (various issues, 1973-1994), International Monetary Fund, Washington, D.C.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994) "Bayesian analysis of stochastic volatility models." *Journal of Business & Economic Statistics,* vol.12.

Johansen, S. (1988) "Statistical analysis of cointegration vectors." *Journal of Economic Dynamics and Control.* 12: 231-54.

Johansen, S. (1991a) "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models." *Econometrica*, 59: 1554-80.

Johansen, S. (1991b) "Determination of cointegration rank in the presence of a linear trend." *Oxford Bulletin of Economics and Statistics*, 54: 383-97.

Johansen, S. (1992) "Testing weak exogeneity and the order of cointegration in UK money demand data." *Journal of Policy Modelling*, 14: 313-35.

Johansen, S. (1992) "Cointegration in partial systems and the efficiency of single equation analysis." *Journal of Econometrics*, 52: 389-402.

Johansen, S. (1995) *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

Johansen, S., and Juselius, K. (1990) "Maximum likelihood estimation and inference on cointegration: with applications to the demand for money." *Oxford Bulletin of Economics and Statistics*, 52: 169-210.

Johansen, S., and Juselius, K. (1992) "Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK." *Journal of Econometrics*, 53: 211-44.

Johnston, J.J. (1984) *Econometric methods*, McGraw-Hill, New York.

Jorion, P. (1988) "On jump processes in the foreign exchange and stock markets." *Review of Financial Studies*, 1: 427-45.

Juselius, K. (1992) "Domestic and foreign effects on prices in an open economy: the case of Denmark." *Journal of Policy Modelling*, 14: 401-28.

Kim, S. and Shepard, N. (1994) "Stochastic volatility: likelihood inference and comparison with ARCH Models." Working Paper, Nuffield College, Oxford.

King D. T., Putnam B. H. and Wilford, D. S. (1986) "A Currency portfolio approach to exchange rate determination: Exchange rate stability and the interdependence of monetary policy." *The Monetary approach to international adjustment*. ed. by Bluford H. Putnam and D. Sykes Wilford, Praeger, New York.

King, R. G., Plosser, C. I., Stock, J. H. and Watson, M. W. (1991) "Stochastic trends and economic fluctuations." *American Economic Review*, 81: 819-840.

Kouri, P. J. K. (1976) "The Exchange rate and the balance of payments in the short run and in the long run: a Monetary approach." *Scandinavian Journal of Economics*, 78, 2: 280-304.

Kouri, P. J. K. and de Macedo. (1978) "Exchange rates and the international adjustment process." *Brookings Papers on Economic Activity*, 1, Brookings Institution, Washington: 111-150.

Levich, R. M. (1985) "Empirical studies of exchange rates: Price behavior, rate determination and market efficiency." *Handbook of International Economics, vol.II* ed. by R. W. Jones and P. B. Kenen, North-Holland, Amsterdam: 979-1040.

Lütkepohl, H. (1991) *Introduction to Multiple Time Series Analysis*, Springer-Verlag, New York.

MacDonald, R. (1988) *Floating Exchange Rates: Theories and Evidence,* Unwin Hyman, London.

Mahieu, R. and Schotman, P. (1994) "Stochastic volatility and the distribution of exchange rate news." Discussion Paper, 96, Institute for Empirical Macroeconomics, University of Minnesota, Minnesota.

Mahieu, R. and Schotman, P. (1994) "Neglected common factors in exchange rate volatility." *Journal of Empirical Finance*: 279-311.

Mandelbrot, B. (1963) "The variation of certain speculative prices." *Journal of Business,* 36: 394-419.

McKinnon, R. I. and Oates, W. J. (1966) "The implications of economic integration for monetary, fiscal and exchange rate policy." *Princeton Studies in International Finance,* No.16.

McKinnon, R. I. (1969) "Portfolio balance and international payments adjustment." *Monetary Problems of the International Economy,* ed. by R. A. Mundell and A. K. Swoboda, The University of Chicago, Chicago.

McKinnon, R. I. (1982) "Currency substitution and instability in the world dollar standard." *American Economic Review,* 72: 320-333.

Meese, R. A. and Rogoff, K. (1983) "Empirical exchange rate model of the seventies: Do they fit out of sample?." *Journal of International Economics,* 14: 3-24.

Meese, R. A. and Rogoff, K. (1983) "The out-of-sample failure of empirical exchange rate models: Simply error or misspecification." *Exchange Rate and International Macroeconomics,* ed. by J. A. Frankel, University of Chicago, Chicago.

Melino, A. and Turnbull, S. M. (1990) "Pricing foreign currency options with stochastic volatility." *Journal of Econometrics,* v45: 239-65.

Mills, T. C. (1992) *Time Series Techniques for Economists,* Cambridge University Press, Cambridge.

Mussa, M. L. (1976) "The exchange rate, the balance of payments and monetary and fiscal policy under a regime of controlled floating." *Scandinavian Journal of Economics,* 78: 229-254.

Mussa, M. L. (1979) "Empirical regularities in the behavior of exchange rates and theories of the foreign markets." *Policies for Employment, Prices and Exchange Rates, Carnegie-Rochester Conference Series on Public Policy,* ed. by K. Brunner and A. H. Metzler.

Mussa, M. L. (1984) "The theory of exchange rate determination." *Exchange Rate Theory and Practice,* ed. by John F. O. Bilson and Richard C. Marston, The University of Chicago Press, Chicago.

Nelson, C. R., and Plosser, C. I. (1982) "Trends and random walks in macroeconomic time series: Some evidence and implications." *Journal of Monetary Economics,* 10: 139-162.

Phillips, P. C. B. (1987) "Multiple Regression with Integrated Time Series." *Yale Cowles Foundation Discussion Paper,* 852, 41.

Phillips, P. C. B., and Perron, P. (1988) "Testing for a unit root in time series regression." *Biometrika,* 75: 335-346.

Richard, J. F. (1980) " Models with several regimes and changes in exogeneity." *Review of Economic Studies*, 47: 1-20.

Ruiz, E. (1993) "Stochastic volatility versus autoregressive conditional heteroscedasticity." Working Paper, Universidad Carlos III de Madrid, Spain.

Ruiz, E. (1994) "Quasi-maximum likelihood estimation of stochastic volatility models." *Journal of Econometrics*, 289-306.

Schmidt, P. (1996) *Bayesian Analysis of a Regime-Switching Volatility Model via the Gibbs Sampler*, Iowa State University creative component (Unpublished), Ames, IA.

Shenton, L. R. and Bowman, K. O. (1977) "A bivariate model for the distribution of $\sqrt{b_1}$ and b_2." *Journal of the American Statistical Association*, 72: 206-211.

Sims, C. A. (1980) "Macroeconomics and reality." *Econometrica*, 48: 1-48.

Sims, C. A. (1986) "Are forecasting models usable for policy analysis?" *Quarterly Review of the Federal Reserve Bank of Minneapolis*, Winter: 2-16.

Sims, C. A., Stock, J. H., and Watson, M. W. (1990) "Inference in linear time series models with some unit roots." *Econometrica*, 58: 113-44.

Taylor, S. (1986) *Modeling Financial Time Series*, John Wiley & Sons, New York.

Taylor, M. P. and McMahon, P. C. (1988) "Long-run purchasing power parity in the 1920s," *European Economic Review*, v32: 179-197.

Urbain, J. P. (1988) *The Econometric Analysis of Import Demand Functions: an application of cointegration analysis and exogeneity tests*, Unpublished master thesis, Université Catholique de Louvain-La-Neuve, Belgium.

Urbain, J. P. (1992) "On weak exogeneity in error correction models." *Oxford Bulletin of Economics and Statistics*, 52: 187-202.

Urbain, J. P. (1993) *Exogeneity in error correction models*, Springer-Verlag, New York.

Watson, M. (1986) "Univariate detrending methods with stochastic trends." *Journal of Monetary Economics*, 18: 49-75.

Watson, M. (1994) "Vector autoregressions and cointegration." *Handbook of Econometrics*, vol.IV, ed. by R. F. Engle and D. L. McFadden, North-Holland, Amsterdam: 2843-2915.

Weiss, A. A. (1984) "ARMA models with ARCH errors." *Journal of Time Series Analysis*, 5: 129-43.

Weiss, A. A. (1986) "Asymptotic theory for ARCH models: Estimation and testing." *Econometric Theory*, 2: 107-31.

Wiggins, J. B. (1987) "Option values under stochastic volatility: Theory and empirical estimates." *Journal of Financial Economics*, v19: 351-372.

ACKNOWLEDGMENTS

I am indebted to my major professors, Dr. Stefano Athanasoulis and Dr. F. Jay Breidt whose guidance, encouragement, enthusiasm, and availability made both my dissertation and graduate study much more enjoyable and rewarding than I had anticipated.

I would also like to thank Dr. Harvey Lapan, who served as my major professor during an earlier stage of my graduate program. I appreciate his help and his patience with me.

I am grateful to my committee members: Dr. Alicia Carriquiry, Dr. John Schroeter, and Dr. Young Kihl. Their professionalism and expertise contributed greatly to the success of my program.

I would also like to express my gratitude to two Iowa State alumni, Dr. Selahattin Dibooglu and Dr. Ferdaus Hossain for their advice and encouragement. I am indebted to them for the idea of my dissertation and the methodology.

Finally, I would like to acknowledge the sacrifice and support of my parents and brothers. Without their understanding and support, I would never have been able to finish my PhD program.

I would like to dedicate this dissertation to my family.