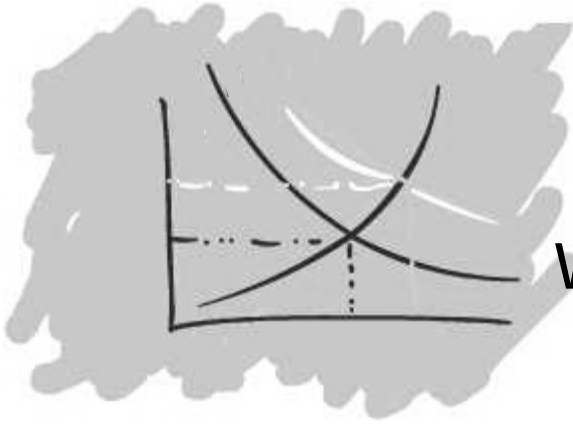# Economic Analysis

Working Paper Series

Long swings in Japan's current account

and in the yen

♦

Nikolas A. Müller-Plantenberg

Working Paper 8/2012

UAM

UNIVERSIDAD AUTONOMA
DE MADRID

Departamento de Análisis Económico:
Teoría Económica e Historia Económica

Long swings in Japan's current account and in the yen[*]

Nikolas A. Müller-Plantenberg[†]

Abstract

The yen has experienced several big swings over recent decades. This paper argues that the fluctuations of the Japanese exchange rate resulted mainly from corresponding movements in the current account, which affected the demand for yen relative to other currencies. The paper builds a vector error correction model for the exchange rate and the current account, based on the idea that the exchange rate and its economic fundamental do not move too far apart over time. In addition, the model allows for a Markov-switching stochastic trend in the current account. Regime changes occur at uncertain dates, possibly in response to exchange rate changes. Bayesian estimation proceeds using an innovative Gibbs-sampling procedure. The empirical results suggest that recurrent structural breaks in the yen's fundamentals account for the large fluctuations of the Japanese exchange rate.

---

1 Introduction

Along with its rise in the post-war period to one of the world's largest economies, Japan has experienced a sustained appreciation of its currency, the yen. Less often noticed, but just as remarkable, is the fact that the yen's value has fluctuated widely over the years, both in nominal and real terms, since it started to float in the early 1970s.

Exchange rate fluctuations.   Some numbers help to illustrate the magnitudes involved. After the breakdown of the Bretton Woods system and the yen's first years of floating, Japan's nominal effective exchange rate appreciated by 45% in less than two years, namely from 1976Q4 to 1978Q3. The currency then fell and rose again by a quarter of its value before entering a period of relative tranquility in the early 1980s, marked by a strong and rising current account surplus. Compared to other industrial countries, Japan liberalized its capital markets at a relatively late stage. In the first half of the 1980s, the opening of its financial account induced strong capital outflows, which helped to contain the yen's strength for some time. Once the initial wave of capital outflows was over, however, the nominal exchange rate shot up by 61% between 1985Q3 and 1988Q4. It subsequently weakened and recovered, to return to almost the 1988 level in 1992. Yet from 1992Q3 to 1995Q2, the yen's trade-weighted value rose once more by more than a half, only to fall and rise again by as much in the following two years.

The sizeable and prolonged swings of Japan's nominal exchange rate translated into very similar movements of the real exchange rate throughout the floating period. However, the yen appreciated less in real terms than in nominal terms. From 1980Q1 to 2000Q1, Japan's annual inflation rate stayed 1.7% below the weighted inflation rates of its trading partners, with little variation. By contrast, Japan's nominal effective exchange rate rose 4.5% per year on average, implying a substantial real appreciation of the yen over the years. Clearly, purchasing power parity has been the exception rather than the rule in Japan.

What explains these massive exchange rate movements? This paper is based on the belief that conventional theories of exchange rate determination cannot explain the yen's performance in a satisfactory way. First, the actual fluctuations of the Japanese currency appear quite simply too large for many models. It is difficult to think, for instance, that international return differentials or differences in money growth could account for the yen's double-digit movements, which occurred often within a matter of months. Nor can productivity-based theories, which in principle might help to understand the yen's long-term appreciation, explain episodes during which the yen depreciated.[1]

---

[1] The Balassa-Samuelson effect is one important example. For many economists, Japan is the showcase for the Balassa-Samuelson effect that predicts an appreciation of the real exchange rate in countries with

Trade and capital flows.   This paper tells a different story. It argues that it was in the first place the Japanese current account that caused the large fluctuations of the Japanese currency over the past three decades as well as its sustained appreciation. In this interpretation, Japan's large surpluses during the 1980s and 1990s influenced the demand for yen relative to other currencies and helped to strengthen the currency. The variability of the current account thus implied large changes in the currency's value.[2]

Yet capital flows mattered, too, in Japan. In particular, the massive purchases of foreign debt securities by the Japanese as a result of their strong export performance tended to ease the pressure on the yen and thus allowed for a more gradual and delayed adjustment of the exchange rate to the movements of the current account (Müller-Plantenberg 2006). As will be shown in section 2.2, a simple model of international adjustment can explain the long swings of Japan's external balances and of its exchange rate.

The main objective of this paper is, however, an econometric one, namely to model the cyclical movements of the yen and its main economic fundamental within a multivariate nonlinear time series framework and to demonstrate the feasibility of Bayesian inference upon the model with the aid of the Gibbs sampler. It is this multivariate perspective that sets the paper apart from many recent studies whose aim it has been to assess the validity of purchasing power parity by examining the univariate nonlinear adjustment behaviour of the real exchange rate (see, for example, Michael et al. 1997, Obstfeld & Taylor 1997).

---

high productivity growth (Balassa 1964, Samuelson 1964). This theory has the potential to explain large long-run movements of the real exchange rate. It is indeed the case that Japan experienced both strong productivity growth and a sustained increase in the real value of its currency over the past fifty years. However, other implications of the theory are less well matched by the data. According to the theory, the real exchange rate movements are brought about by movements in the ratio of nontraded versus traded goods prices. However, Engel (1999) recently found that relative prices of nontraded goods account for almost none of the movements in real exchange rates in major industrial countries, irrespective of the time horizon (the method he used was based on a decomposition of the mean squared error of real exchange rate changes). As Engel points out, relative prices of nontraded goods in Japan increased by about 40 percent since the 1970s but the appreciation of the real exchange rate was 90 percent. However, since the relative price of nontradables in the United States closely mirrored the relative price of nontradables in Japan, the Balassa-Samuelson effect was effectively neutralized. As a final argument, there have not been any large downswings in the relative price of nontradable goods in Japan but, as we have seen, there have been several instances of a dramatic depreciation of the yen.

[2]The argument that the current account matters for the exchange rate reminds of the recent attempts in the microstructure literature on exchange rates to link exchange rate movements to order flows in the foreign exchange market (Evans & Lyons 2002, 2003). It is also closely related to the balance of payments (BOP) flow approach to exchange rate determination in macroeconomics (see, for example, chapter 12 in Kenen 2000). For a long time, the conventional way of analysing exchange rate behaviour was to monitor the flow supplies of, and flow demands for, foreign currencies in the foreign exchange markets (Rosenberg 1996). However, the BOP flow approach lost much of its appeal to academic economists in the 1970s when more stock-oriented exchange rate theories came into vogue. Rosenberg (1996, page 69) has pointed out, however, that despite the sceptical view of academics, "most market participants today probably still rely on some variant of the BOP flow approach in their analysis of exchange rate movements and in their formulation of international investment strategy".

Specifically, the model developed here consists of the current account and the real effective exchange rate. While individually nonstationary, the two variables turn out to be cointegrated, suggesting a vector error correction model as a natural starting point. However, the model is modified in an important respect in that the intercept of the current account equation is allowed to switch between two states, or regimes, according to an unobservable Markov process, so as to take account of the recurrent swings in the current account and the gradual response of the exchange rate to those swings.

Outline.   The paper is organized as follows. Section 2 discusses the time series evidence and builds a theoretical model to explain the cyclical movements of Japan's current account, debt balance and exchange rate. Section 3 sets out the empirical model. Section 4 describes how Gibbs sampling can be used to carry out Bayesian inference on the model. Section 5 discusses the empirical results. Section 6 provides conclusions.

2   Japan's external performance and the demand for yen

Figure 1:   **Japanese current account and exchange rate (1980s and 1990s).** Japanese current account (left scale, in trillions of yen) and nominal effective exchange rate (right scale, in logarithms), period from 1977Q1 to 2001Q1. *Source: International Financial Statistics (IMF).*

Key for understanding the dynamics of the Japanese currency during recent decades are, according to this paper, the strong surges and occasional declines in the demand for Japanese goods and services. Before studying this hypothesis theoretically and empirically, it is helpful to start by examining the available time series evidence.

4

2.1 Time series evidence

Figure 1 plots Japan's current account and nominal effective exchange rate for the period from 1977Q1 to 2001Q4. As in the rest of this paper, the nominal exchange rate is defined as the foreign-currency price of the domestic currency; that is, a rise in the nominal exchange rate implies an appreciation of the domestic currency. One can observe that the current account went through four big swings. The nominal exchange rate followed these movements quite closely. It similarly experienced large, protracted swings related to those of the current account. The statistical part of this paper will show that the Japanese current account and its real exchange rate, which mirrored the movements of its nominal counterpart, are cointegrated over the period for which data are available.

The relationship is less clear only after 1981, when the yen suddenly weakened for several quarters, despite a rising current account. As mentioned in the introduction, large capital outflows occurred at that time due to the liberalization of Japan's capital account. (Note that the US dollar experienced a sharp appreciation during the pre-1985 period even after US interest rates had fallen from their record levels of the early 1980s.)

The exchange rate movements appear to have followed the current account movements with a substantial lag. Another way to look at it is to note that the yen rose most strongly in those years in which the current account hit its peaks, namely in 1978, 1986, 1992 and 1998 (as well as in 1971, as we are about to see).
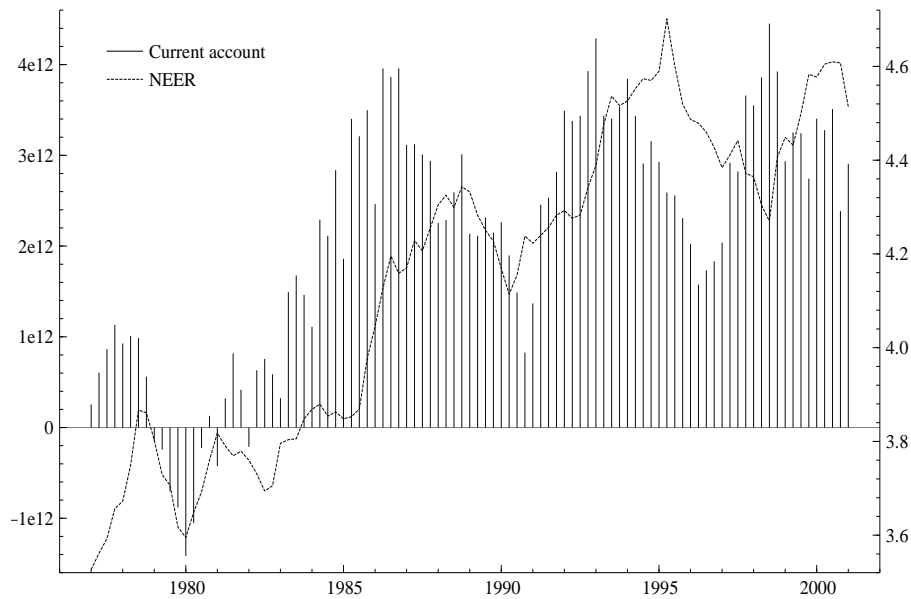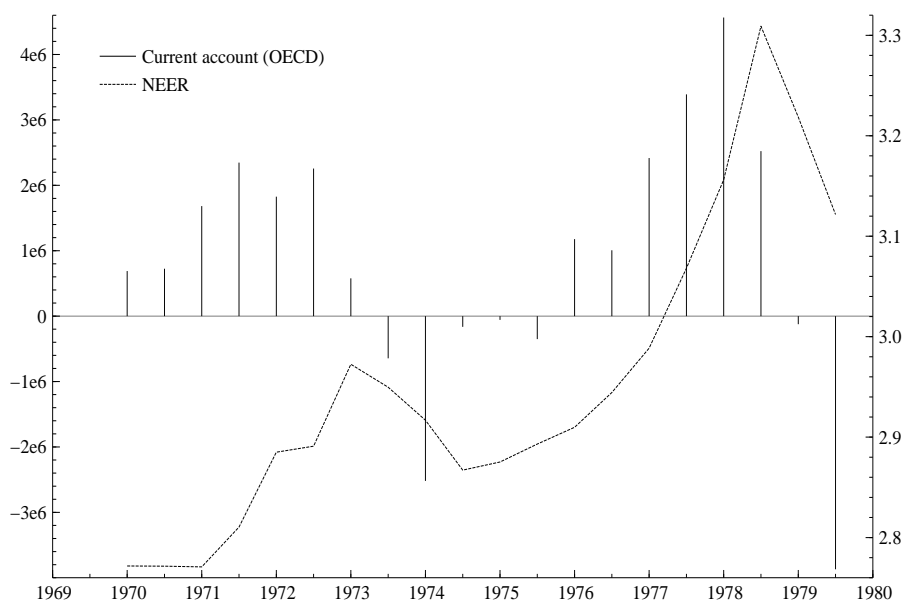

Figure 2: **Japanese current account and exchange rate (1970s).** Japanese current account (left scale, in millions of yen) and nominal effective exchange rate (right scale, in logarithms), period from 1970H1 to 1979H2. *Source: Economic Outlook (OECD).*

Quarterly data are available only from 1977. Before that year, data for both variables exist only at a biannual frequency. Consider figure 2 which again plots the same variables as figure 1, this time for the period from 1970 until the end of 1979, allowing for a little time overlap in both plots. Figure 2 shows another swing of the current account in the first half of the 1970s, with a corresponding rise and fall of the exchange rate. Again, one can observe a short lag between both variables. As in figure 1, the yen appreciates very strongly at a time when the current account reaches a temporary peak, namely in the years 1971 and 1972. Both figures therefore suggest that the current account provided a good indicator of the changing desire of economic agents to obtain and spend Japanese currency.[3]
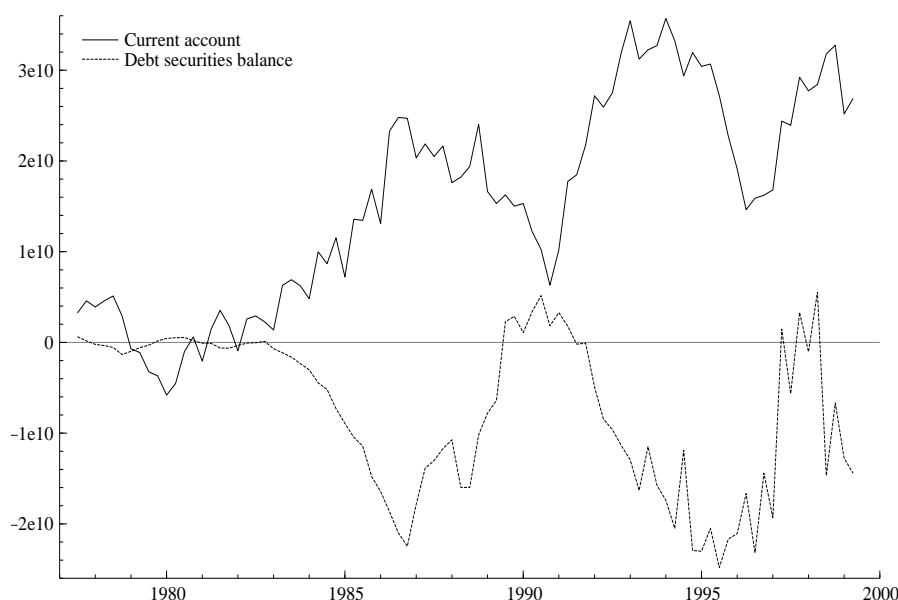
Figure 3:   **Japan's lending abroad.** Japanese current account and debt securities balance. Both variables in US-$. Debt securities balance as a moving average with two leads and lags. Period from 1977Q3 to 1999Q2. *Source: International Financial Statistics (IMF).*

Turning to the capital side of the Japanese balance of payments, let us now consider figure 3, which plots the Japanese current account together with the debt securities balance. Debt securities are part of the portfolio investment balance and they arguably represent the most important item in Japan's financial account. From figure 3 it is evident that Japan's purchases of foreign debt securities mirrored the movements of its current account during many years. Over time, though, net foreign lending increased less strongly

---

[3]The interpretation seems to be confirmed by Brooks et al. (2001) who find that the yen exchange rate has remained closely tied to the current account over recent years; portfolio flows appear to have been less relevant for the yen-dollar exchange rate than, say, for the euro-dollar exchange rate.

than foreign exports (becoming even negative in the early 1990s), as foreigners started gradually to pay off their debt to the Japanese.

What this suggests is that Japan's heavy external lending provided a temporary buffer against the exchange rate pressure from its external surplus. By doing so, it contributed to the lagged responses of the Japanese currency to the movements in net exports.[4]

2.2 Theoretical model

The argument presented above will now be formalized in terms of a simple theoretical model. The model is capable of explaining the recurrent swings in Japan's external balances and in its nominal and real exchange rate. No attempt is made to account for the yen's long-term appreciation. Instead, the analysis will focus on Japan's economic fluctuations at medium frequencies. As with business cycles of economic activity, there is no doubt that cyclical variations of the balance of payments and of the exchange rate with periods of, say, a couple of quarters or years are of strong interest to anyone concerned with economic performance.

Suppose there are two countries: a home country indexed by 1 and a foreign country indexed by 2. The nominal exchange rate between the home country and the foreign country, $S(t)$, can conveniently be thought of as the product of the ratio of the foreign and domestic price levels, $P_2(t)$ and $P_1(t)$, and the ratio of the net demands for the domestic and foreign currencies, $\tilde{C}_1(t)$ and $\tilde{C}_2(t)$:

$$S(t) = \frac{P_2(t)}{P_1(t)} \times \frac{\tilde{C}_1(t)}{\tilde{C}_2(t)}. \tag{1}$$

The first factor in this formula represents the Casselian hypothesis and says that the relative price of two currencies should reflect their relative purchasing powers. The second factor accounts for the fact that purchasing power parity seldom holds and assumes therefore that deviations from purchasing power parity are determined by the supply and demand conditions in the foreign exchange market.

The variables $\tilde{C}_1(t)$ and $\tilde{C}_2(t)$ can be interpreted as normalized indices measuring the net demand for the domestic and foreign currency, respectively. Both variables depend on

---

[4]Finally, it should be noted that Japan has been accumulating vast reserves of foreign exchange over recent decades. Purchases of foreign exchange appear to have been particularly heavy during those years in which the exchange rate appreciated most strongly. Although these purchases could never fully offset the appreciation of the yen, it appears that they did have a moderating impact (Müller-Plantenberg 2006). Altogether, reserves played more of an endogenous role, reacting whenever economic fundamentals put too strong upward pressure on the exchange rate. This view is also in line with Girton & Roper (1977) who suggest that both exchange rate adjustments and reserve changes serve as indicators of exchange market pressure. In the multivariate model of section 3, I choose not to include foreign exchange reserves as an additional variable in order to preserve a parsimonious setup.

the net flow of money the home and foreign countries receive in return for their balance of payments transactions, or in other words on each country's cash flow vis-à-vis the rest of the world:

$$\tilde{C}_1(t) = \exp\left(\frac{\xi}{2P_1(t)} \int_{-\infty}^{t} c_{21}(\tau) - (S(\tau))^{-1} c_{12}(\tau)\, d\tau\right), \tag{2a}$$

$$\tilde{C}_2(t) = \exp\left(\frac{\xi}{2P_2(t)} \int_{-\infty}^{t} c_{12}(\tau) - S(\tau)\, c_{21}(\tau)\, d\tau\right). \tag{2b}$$

Here, $c_{ij}(t)$ represents the flow of money from country i to country j (denominated in country j's currency by assumption, to keep things simple), which increases the net supply of country i's currency in the foreign exchange market and reduces that of country j's currency. It is important to note that even if all currency exchanges to which balance of payments transactions give rise are carried out by participants of the foreign exchange market, the market as a whole may not have to clear. On the contrary, except when exports and imports as well as capital inflows and outflows exactly net each other out exactly, the demands and supplies of currencies in the foreign exchange market are likely to differ, leading to foreign exchange imbalances that accumulate over time. It is the cumulative net demand of a currency that thus determines its value. Balance of payments flows, on the other hand, measure only incremental changes in that demand and should therefore be associated with currency appreciations or depreciations.

Even though the foreign exchange market is thought to drive the nominal exchange rate in this model, it should be noted that the model implies a one-to-one relationship between currency flows and the real exchange rate. This is evident from equation (1), which may be rearranged as follows:

$$Q(t) = \frac{\tilde{C}_1(t)}{\tilde{C}_2(t)}. \tag{3}$$

In conjunction with equations (2a) and (2b), equation (3) implies that movements in prices and in the nominal exchange rate affect the response of the real exchange rate to movements in international cash flow in possibly quite complex ways. However, linearization can yield a simple relationship between country 1's net cash inflow and its real exchange rate provided country 1 and country 2 are treated symmetrically. Under such conditions, proportional changes in the real exchange rate depend positively on the net cash flow country 1 receives:

$$\dot{q}(t) = -\xi c(t), \tag{4}$$

8

where

$$c(t) := c_{12}(t) - c_{21}(t).$$

Equation (4) specifies in a simple way how the exchange rate is set in the foreign exchange market. To complete the model, it is necessary to make a number of assumptions regarding the interaction of the current account, the debt balance (as the empirically most relevant item within the financial balance) and the monetary balance.

First, we adopt the familiar assumption that the Marshall-Lerner conditions holds and that changes in the current account are determined by the real exchange rate:

$$\dot{z}(t) = -\phi_1 z(t) - \phi_2 q(t). \tag{5}$$

The purpose of the first term on the right-hand side of equation (5) is to keep balance of payments imbalances in check and to ensure the dynamic stability of the model.

Second, we assume that current account liabilities are paid off gradually. The debt balance is thus the difference between newly incurred liabilities, which are equal to the current account, and the previously incurred debt that is now becoming due. Formally, the debt balance is defined as follows:

$$d(t) = -z(t) - \gamma \int_0^t e^{-\gamma(t-\tau)} d(\tau) d\tau, \tag{6}$$

where an exponentially declining maturity structure of foreign debt is chosen for convenience, so as to facilitate the application of Laplace transforms in the solution of the model.

Finally, the balance of payments identity implies that international cash flow is determined by the current account and the debt balance since these are the only balance of payments components considered in the model. Therefore:

$$c(t) = -z(t) - d(t). \tag{7}$$

The model can be solved with the method of Laplace transforms (see appendix A). The solution for $z(t)$ and $d(t)$ is:

$$z(t) = A_z\, e^{-B_z t} + \frac{-C_z E_z + D_z}{F_z}\, e^{-E_z t}\, \sin(F_z t) + C_z\, e^{-E_z t}\, \cos(F_z t), \tag{8a}$$

$$d(t) = A_d\, e^{-B_d t} + \frac{-C_d E_d + D_d}{F_d}\, e^{-E_d t}\, \sin(F_d t) + C_d\, e^{-E_d t}\, \cos(F_d t), \tag{8b}$$

where the A_z, B_z, ..., F_z and A_d, B_d, ..., F_d are constants that depend on the model's parameters and initial conditions. For our purposes, it is important to note the following: First, the constants B_z and B_d are always positive so that the first term in equation (8a) as well as that in equation (8b) vanish as t increases. Second, allowing for the debt-financing of current account imbalances implies that the model is potentially unstable. Specifically, if the parameter limiting large current account imbalances, ϕ_1, is zero or takes on a small value, the constants E_z and E_d are negative, implying ever greater oscillations in $z(t)$ and $d(t)$ over time. Last but not least, the frequency of the balance of payments swings depends on the maturity of foreign debt: a higher value of γ implies that debt amortization is anticipated, resulting in a quicker response of the exchange rate to current account movements and thus a shortening of the cycle period of $z(t)$ and $d(t)$. Since it is difficult in general to know the maturity structure of a country's overall foreign debt, the empirical model of the next section will allow for the data to determine the frequency with which Japan's current account and exchange rate revert their trends.
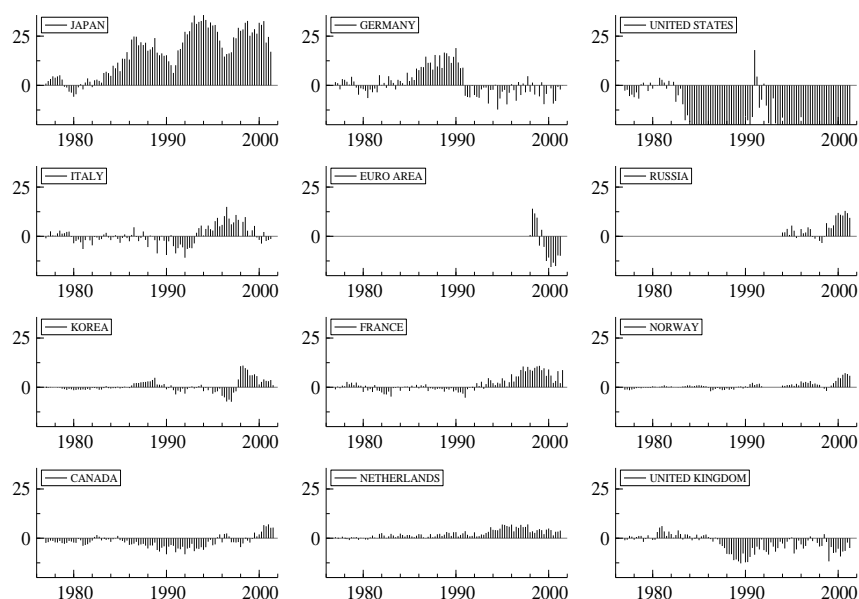
Figure 4: **Large current account surpluses.** Current account balances of countries with large current account surpluses (in billions of US dollar). Countries are selected and ordered according to the highest current account balance they have achieved in any single quarter in the period from 1977Q1 to 2001Q3. *Source: International Financial Statistics (IMF).*

2.3 Why study Japan?

What is the merit of studying the exchange rate performance of Japan which is but a single economy? Wouldn't we need to observe similar time series patterns as in Japan

elsewhere as well in order to make the above a convincing theory? To answer these questions, the first thing to note is that Japan's export boom in the 1980s and 1990s has been truly exceptional. As figure 4 demonstrates, Japan's current account surplus has been far greater than the surplus of any other country; the changes in the balance of payments have also been much more sizeable than those in other countries. As for the capital account, other advanced economies, such as the United States or the European economies, have traditionally been more integrated into world financial markets than Japan with its well-documented home bias, which can explain why the Japanese current account had such a stable impact on the yen (Brooks et al. 2001) and why the capital account was mainly accommodating Japan's current account imbalances.

Upon closer inspection, however, one also finds many episodes in other countries where exchange rates were influenced by current account movements in much the same way as in Japan. Currency crises are the case in point. In many of the recent exchange rate crises, the sequence of events, other factors notwithstanding, has been quite similar. The countries affected were usually running large current account deficits prior to their crises. In many cases, they received large capital inflows which helped to keep their exchange rate from depreciating. However, their national currencies got into trouble as soon as capital inflows dried up or reversed in sign. Eichengreen (2003, chapter 8) and Bussiere & Mulder (1999), for instance, have shown that a small set of variables—including the current account as a percentage of GDP, export growth, international reserves and short-term foreign debt relative to reserves—do a very good job in predicting the EMS crisis in 1992–1993, the Mexican crisis in 1994–1995 as well as the Asian crisis in 1997.

Japan allows us to study the impact of trade and capital flows on exchange rate behaviour under more stable conditions. Since Japan has been running large export surpluses for a long time, the country only needed to decide on how to fend off the upward pressure on its exchange rate and how to invest its export revenues. Conditions were similar for Germany in the 1980s, before German unification, when it also ran a substantial current account surplus (see figure 4). During these years, the German mark responded to movements in Germany's current account in much the same way as the yen did with respect to the Japanese balance of payments (figure 5). For deficit countries, the situation is quite different since these countries are confronted with issues such as current account sustainability and finiteness of reserves that need to be resolved in one way or the other, for instance by attracting capital flows from abroad. This might explain why in the United States, the country with the largest current account deficit in the world and with very large and open financial markets, exchange rate movements have been more sensitive to capital flow determinants such as international return differentials (Brooks et al. 2001, Eichengreen 1996).
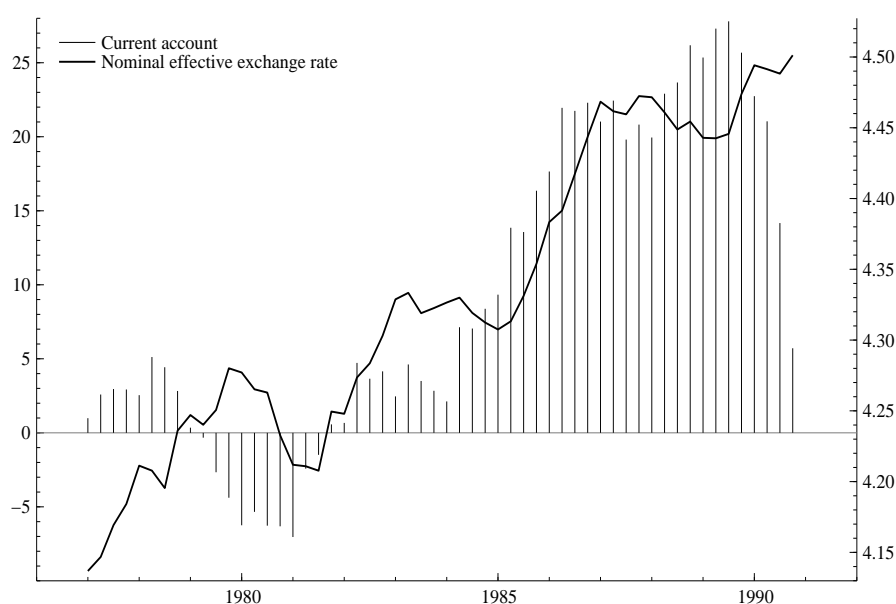
Figure 5: **German current account and nominal exchange rate in the 1980s.** German current account (left scale, in German mark) and nominal effective exchange rate (right scale, in logarithms), period from 1977Q1 to 1990Q4. *Source: International Financial Statistics (IMF).*

3 Empirical modelling

This section presents an empirical model based on the close relationship between the current account and the real exchange rate in Japan. The model allows for large swings in the Japanese current account. At the same time, it takes into account the repercussions of these swings on the performance of the yen.

3.1 Cointegration of exchange rate and economic fundamental

The statistical analysis of Japan's real effective exchange rate as well as its current account reveal that both variables are nonstationary. Table 1 reports the results of the augmented Dickey-Fuller tests for both variables. The null hypothesis of a unit root cannot be rejected for either variable.[5]

While individually I(1), I find the Japanese current account and real effective exchange rate to be cointegrated, that is, linear combinations of both variables exist that are I(0). To test for cointegration, I apply the methodology of Johansen (1988). Table 2 gives the results. I can reject the null hypothesis of no cointegration at the 1% significance level (applying critical values as computed by PcGive).

---

[5]For the general question whether the current account can be regarded as a nonstationary variable, see Lau et al. (2006) and references therein.

Constant	Trend	Lags	Coefficient	Std. error	t statistic	5% crit.
Real effective exchange rate (1978Q3–2001Q3)						
No	No	4	1.0001	0.04496	0.05417	-1.94
Yes	No	3	0.94170	0.04460	-1.926	-2.89
Yes	Yes	3	0.85491	0.04316	-3.300	-3.46
Current account (real volume) (1978Q3–2001Q3)						
No	No	6	0.99017	0.005188	-0.4861	-1.94
Yes	No	5	0.91713	0.005106	-2.068	-2.89
Yes	Yes	4	0.82356	0.005041	-3.456	-3.46

Table 1: **Unit root tests.** *Augmented Dickey-Fuller tests. Lag length selection based on Akaike information criterion.*

H_0	H_A	Trace test	p-value
$r = 0$	$r > 0$	19.143	[0.003] **
$r \leq 1$	$r > 1$	3.5246	[0.070]

Table 2: **Testing for cointegration.** *Testing for the number of distinct cointegrating vectors, using 5 lags. Double asterisks (**) mark significance at the 1% level.*

Given that both variables are cointegrated, their joint dynamic behaviour can conveniently be represented by a vector error correction model. Estimation can proceed along traditional lines using maximum likelihood methods.

3.2 A Markov-switching vector error-correction model

Motivated by the inspection of the data in section 2.1, I seek to analyse a modified vector error correction model. I make two alterations to the conventional setup. First, the intercept of the current account equation is allowed to switch between two unknown values according to a two-state Markov process with constant, unknown transition probabilities. While the current account is nonstationary, the time series in figures 1 and 2 suggest that it is subject to two kinds of drifts, an upward drift in some periods and a downward drift in other periods.

Second, seasonal dummies are present only in the current account equation, not in the exchange rate equation. This seems justified since the exchange rate, unlike the current

account, does not appear to exhibit any seasonality. From an economic point of view, it is also hard to conceive why the exchange rate, by itself, should be fluctuating seasonally.

The model I analyse takes the following form:

$$
\begin{bmatrix} \Delta q_t \\ \Delta z_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \psi_{z,1} & \psi_{z,2} & \psi_{z,3} \end{bmatrix} \begin{bmatrix} d_{1,t} \\ d_{2,t} \\ d_{3,t} \end{bmatrix}
$$
$$
+ \begin{bmatrix} \pi_{0,1} \\ \pi_{0,2} \end{bmatrix} + \begin{bmatrix} 0 \\ \nu_z \end{bmatrix} R_t + \sum_{i=1}^{h-1} \begin{bmatrix} \pi_{i,11} & \pi_{i,12} \\ \pi_{i,21} & \pi_{i,22} \end{bmatrix} \begin{bmatrix} \Delta q_{t-i} \\ \Delta z_{t-i} \end{bmatrix} \qquad (9)
$$
$$
+ \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} q_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{q,t} \\ \varepsilon_{z,t} \end{bmatrix},
$$

where

$$
\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} = \alpha\beta' \equiv \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix},
$$
$$
\nu_z > 0,
$$
$$
\text{Prob}\left(R_t = 1 | R_{t-1} = 1\right) = p, \quad \text{Prob}\left(R_t = 0 | R_{t-1} = 0\right) = q.
$$

In this representation, q_t is the real effective exchange rate of Japan and z_t is the Japanese current account (in the domestic currency). The vector $[d_{1,t}, d_{2,t}, d_{3,t}]'$ contains seasonal dummies.

In period t, the system is in one of two regimes, 0 or 1, according to the random variable R_t. The variable R_t follows a Markov process with transition probabilities p and q. The regime affects the intercept of the current account equation, which switches between a lower level, $\pi_{0,2}$, and a higher level, $\pi_{0,2} + \nu_z$. The current account is drifting downward whenever the system is in regime 0; it is drifting upward whenever the system is in regime 1. The transition probabilities are assumed constant here in order to avoid a too complex model, as is done in most applications of Markov-switching models.

By collecting both variables in a vector $y_t = [q_t, z_t]'$, the model can be written as:

$$
\Delta y_t = \Psi d_t + \pi_0 + \nu R_t + \sum_{i=1}^{h-1} \Pi_{t-i} \Delta y_{t-i} + \Pi y_{t-1} + \varepsilon_t
$$
$$
= \Psi d_t + \pi_0 + \nu R_t + \sum_{i=1}^{h-1} \Pi_{t-i} \Delta y_{t-i} + \alpha\beta' y_{t-1} + \varepsilon_t \qquad (10)
$$
$$
= \Psi d_t + \pi_0 + \nu R_t + \sum_{i=1}^{h-1} \Pi_{t-i} \Delta y_{t-i} + \alpha\eta_{t-1} + \varepsilon_t,
$$

14

$$\varepsilon_t \sim \text{ i.i.d. N}(0, \Sigma).$$

As equation (10) indicates, the matrix of long-run responses, Π, is the product of the feedback vector, $\alpha = [\alpha_1, \alpha_2]'$, and the cointegrating vector, $\beta' = [\beta_1, \beta_2]$. The vector α is called feedback vector since it measures the system's response to the error from the long-run equilibrium relation. This error is given by $\eta_{t-1} = \beta' y_{t-1}$.

The model may be written still more compactly as:

$$\begin{aligned} Y &= D\Psi' + R\nu' + X\Gamma + Z\beta\alpha' + E \\ &= D\Psi' + R\nu' + WB + E, \end{aligned} \tag{11}$$

where

$$Y = \begin{bmatrix} \Delta y_1' \\ \Delta y_2' \\ \vdots \\ \Delta y_T' \end{bmatrix}, \quad D = \begin{bmatrix} d_1' \\ d_2' \\ \vdots \\ d_T' \end{bmatrix}, \quad R = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_T \end{bmatrix},$$

$$X = \begin{bmatrix} 1 & \Delta y_0' & \Delta y_{-1}' & \cdots & \Delta y_{2-h}' \\ 1 & \Delta y_1' & \Delta y_0' & \cdots & \Delta y_{3-h}' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \Delta y_{T-1}' & \Delta y_{T-2}' & \cdots & \Delta y_{T-h+1}' \end{bmatrix}, \quad Z = \begin{bmatrix} y_0' \\ y_1' \\ \vdots \\ y_{T-1}' \end{bmatrix}, \quad E = \begin{bmatrix} \varepsilon_1' \\ \varepsilon_2' \\ \vdots \\ \varepsilon_T' \end{bmatrix},$$

$$\Gamma = \begin{bmatrix} \pi_0' \\ \Pi_1' \\ \Pi_2' \\ \vdots \\ \Pi_{h-1}' \end{bmatrix}, \quad W = \begin{bmatrix} X & Z\beta \end{bmatrix}, \quad B = \begin{bmatrix} \Gamma \\ \alpha' \end{bmatrix}.$$

4   Bayesian inference with Gibbs sampling

The model introduced in the previous section is nonlinear and difficult to estimate by classical statistical methods. I will now demonstrate how Bayesian inference, and in particular the simulation tool of Gibbs sampling, can be used to estimate the model. The key of the approach taken here is that the unobserved regimes can be treated as additional unknown parameters (see Albert & Chib 1993). They can then be analysed along with the model's unknown parameters via the simulation method of Gibbs sampling.

Applying the method of Gibbs sampling to time series has become increasingly popular in the literature in recent years. Examples can be found in Albert & Chib (1993) and Kim & Nelson (1998). Kim & Nelson (1999) survey the literature and provide many illustrations. However, I have not so far encountered a *Bayesian* treatment of a model such as the one presented in this paper that combines a vector error correction model with Markov-switching elements. I show that inference becomes feasible once the model's equations are made independent by making use of the Choleski decomposition of their covariance matrix since univariate inference can then be carried out directly on

any equation-specific components of the model. This method is simple to use and it is not difficult to think of applications to other economic problems involving cointegrated systems that are subject to influences affecting variables individually (shifts in regimes, changing volatilities etc.).

For further use, define $\tilde{y}_t \equiv [y_1, y_2, \ldots, y_t]$ and $\tilde{R}_t \equiv [R_1, R_2, \ldots, R_t]$ for $t = 1, 2, \ldots, T$. The sample size is T, so $\tilde{y}_T$ is the vector of all observations and $\tilde{R}_T$ is the vector of all regimes, or states. Let θ denote the vector of all parameters of the model.

4.1 Introducing the Gibbs sampler

Bayesian inference about θ is based on the posterior distribution:

$$g(\theta|\tilde{y}_T) \propto L(\theta|\tilde{y}_T)g(\theta), \tag{12}$$

where $g(\theta)$ is the prior and $L(\theta|\tilde{y}_T)$ is the likelihood function. Direct inference would not be practical here due to the difficulties involved in computing the likelihood function. The alternative is to simulate the posterior density using the Markov chain Monte Carlo (MCMC) method referred to as the Gibbs sampler. For an introduction to MCMC methods and the Gibbs sampler, see for example Gelman et al. (1995) and Kim & Nelson (1999).

Suppose the vector of all parameters, θ, is partitioned into vector components, $\theta_1, \theta_2, \ldots, \theta_k$. There is thus a complete set of conditional distributions, $g(\theta_1|\tilde{y}_T, \theta_2, \theta_3, \ldots, \theta_k)$, $g(\theta_2|\tilde{y}_T, \theta_1, \theta_3, \ldots, \theta_k)$, ..., $g(\theta_k|\tilde{y}_T, \theta_1, \theta_2, \ldots, \theta_{k-1})$. Given some arbitrary starting values for the parameters, $\theta_1^0, \theta_2^0, \ldots, \theta_k^0$, the Gibbs algorithm involves iterating through the following cycle (with the current iteration denoted by i):

Step 1   Draw θ_1^i from $g(\theta_1|\tilde{y}_T, \theta_2^{i-1}, \ldots, \theta_k^{i-1})$.

Step 2   Draw θ_2^i from $g(\theta_2|\tilde{y}_T, \theta_1^i, \theta_3^{i-1}, \ldots, \theta_k^{i-1})$.

$\ldots$

Step k   Draw θ_k^i from $g(\theta_k|\tilde{y}_T, \theta_1^i, \theta_2^i, \ldots, \theta_{k-1}^i)$.

Steps 1 through k can be iterated T times, so as to obtain a full set of simulated parameters for every iteration. Under regularity conditions, the distribution of θ^i converges to the distribution of θ as T goes to infinity (see references quoted in Albert & Chib 1993, Kim & Nelson 1999). This suggests setting $T = N + M$, so that when N initial simulations are discarded, the remaining M drawings of all parameters can be used as an approximate simulated sample from $g(\theta_1, \theta_2, \ldots, \theta_k)$.

4.2 Details of the Gibbs sampling procedure

My objective is to find a complete set of conditional distributions of all the parameters on which the Gibbs sampling scheme can be run. It turns out that this task is facilitated once I treat the regimes $\tilde{R}_T$ as additional unknown parameters and analyse them jointly with θ, the vector of all other parameters. Given $\tilde{R}_T$, conditional inference on θ is basically equivalent to inference on a vector error correction model. Given θ on the other hand, procedures are available that help us to retrieve the conditional distribution of the regimes.

I thus obtain a tractable conditional structure as a basis of the Gibbs simulations. Gibbs sampling proceeds by iteratively drawing from the following sequence of conditional distributions:

- $[\,\tilde{R}_T\,|\,\tilde{y}_T, \Psi, \nu, B, \beta, \Sigma, p, q\,]$,

- $[\,p, q\,|\,\tilde{R}_T\,]$,

- $[\,\Psi, \nu\,|\,\tilde{y}_T, \tilde{R}_T, B, \beta, \Sigma\,]$,

- $[\,\Sigma\,|\,\tilde{y}_T, \tilde{R}_T, \Psi, \nu, B, \beta\,]$,

- $[\,B\,|\,\tilde{y}_T, \tilde{R}_T, \Psi, \nu, \beta\,]$,

- $[\,\beta\,|\,\tilde{y}_T\,]$.

I now turn to the details of the Gibbs sampling procedure.

4.2.1 Generating the regimes, $\tilde{R}_T$

To generate the regimes, $\tilde{R}_T$, I employ multi-move Gibbs sampling (see Kim & Nelson 1999). Multi-move Gibbs sampling refers to the simulation of all the regimes as a block from the joint conditional distribution:

$$g(\tilde{R}_T|\theta_{-\tilde{R}_T}, \tilde{y}_T),$$

where $\theta_{-\tilde{R}_T}$ refers to all the parameters of the model other than $\tilde{R}_T$ (which is treated here as a vector of parameters). It follows from the Markov property of S_t that the joint conditional density can be factorized as follows:

$$g(\tilde{R}_T|\theta_{-\tilde{R}_T}, \tilde{y}_T) = g(R_T|\tilde{y}_T) \prod_{t=1}^{T-1} g(R_t|R_{t+1}, \tilde{y}_t). \tag{13}$$

Equation (13) shows that $g(\tilde{R}_T|\theta_{-\tilde{R}_T}, \tilde{y}_T)$ can be evaluated once $g(R_T|\tilde{y}_T)$ as well as $g(R_t|R_{T+1}, \tilde{y}_t)$, $t = T-1, T-2, \ldots, 1$, are known. This suggests employing a two-step procedure in order to generate these conditional densities.

Step 1 Run the basic filter of Hamilton (1989) to get $g(R_t|\tilde{y}_t)$, $t = 1, 2, \ldots, T$. That is, iterate on the following pair of equations:

$$g(R_t = i|\tilde{y}_t) = \frac{g(R_t = i, y_t|\tilde{y}_{t-1})}{g(y_t|\tilde{y}_{t-1})}$$
$$= \frac{g(y_t|R_t = i, \tilde{y}_{t-1})\, g(R_t = i|\tilde{y}_{t-1})}{\sum_{j=1}^2 g(y_t|R_t = j, \tilde{y}_{t-1})\, g(R_t = j|\tilde{y}_{t-1})}, \quad i = 0, 1,$$

$$\begin{bmatrix} g(R_{t+1} = 0|\tilde{y}_t) \\ g(R_{t+1} = 1|\tilde{y}_t) \end{bmatrix} = \begin{bmatrix} q & 1-p \\ 1-q & p \end{bmatrix} \begin{bmatrix} g(R_t = 0|\tilde{y}_t) \\ g(R_t = 1|\tilde{y}_t) \end{bmatrix}.$$

The last iteration of the filter provides us with $g(R_T|\tilde{y}_T)$, from which R_T is generated.

Step 2 The following result, which follows from the Markov property of R_t, provides us with the smoothed conditional densities $g(R_t|R_{T+1}, \tilde{y}_t)$:

$$g(R_t|R_{T+1}, \tilde{y}_t) = \frac{g(R_{t+1}|R_t, \tilde{y}_t)\, g(R_t|\tilde{y}_t)}{g(R_{t+1}|\tilde{y}_t)}$$
$$= \frac{g(R_{t+1}|R_t)\, g(R_t|\tilde{y}_t)}{g(R_{t+1}|\tilde{y}_t)}$$
$$\propto g(R_{t+1}|R_t)\, g(R_t|\tilde{y}_t),$$

where $g(R_{t+1}|R_t)$ is the transition probability and $g(R_T|\tilde{y}_T)$ is saved from step 1. The result in equation (4.2.1), can be used to recursively generate R_t from $g(R_t|R_{T+1}, \tilde{y}_t)$, for $t = T-1, T-2, \ldots, 1$.

4.2.2 Generating the transition probabilities, p and q

Conditional on $\tilde{R}_T$, the transition probabilities p and q are independent of the data set, $\tilde{y}_T$, and the model's other parameters. The conditional distribution $p, q|\tilde{R}_T$ can be obtained from standard Bayesian results on Markov chains. Given, $\tilde{R}_T$, the transitions n_{ij} from state i to j, with $i, j = 0, 1$, provide sufficient statistics for p and q. The likelihood function for p and q is given by:

$$L(p, q|\tilde{R}_T) = p^{n_{11}}(1-p)^{n_{10}}q^{n_{00}}(1-q)^{n_{01}}.$$

The form of the likelihood suggests the use of the beta distribution as a conjugate prior for the transition probabilities.

Prior   Assuming independent beta distributions for the priors of p and q, the prior is given by:

$$p \sim \mathrm{beta}(u_{11}, u_{10}),$$
$$q \sim \mathrm{beta}(u_{00}, u_{01}),$$

with

$$g(p, q) \propto p^{u_{11}-1}(1-p)^{u_{10}-1}q^{u_{00}-1}(1-q)^{u_{01}-1},$$

where u_{ij}, $i, j = 0, 1$, are the hyperparameters of the prior.

Combining the prior and the likelihood, the following posterior is obtained:

Posterior

$$
\begin{aligned}
g(p, q | \tilde{R}_T) &= g(p, q) L(p, q | \tilde{R}_T) \\
&\propto \quad p^{u_{11}-1}(1-p)^{u_{10}-1}q^{u_{00}-1}(1-q)^{u_{01}-1} \\
&\qquad p^{n_{11}}(1-p)^{n_{10}}q^{n_{00}}(1-q)^{n_{01}} \\
&= \quad p^{u_{11}+n_{11}-1}(1-p)^{u_{10}+n_{10}-1}q^{u_{00}+n_{00}-1}(1-q)^{u_{01}+n_{01}-1}.
\end{aligned}
\tag{14}
$$

I have estimated the model using a non-informative, flat, prior as well as an informative prior. For the non-informative prior, I set $u_{11} = u_{10} = u_{01} = u_{00} = 1$. Using the formulas for the expected value and variance of a beta random variable, I obtain:

$$\mathrm{E}(p) = \mathrm{E}(q) = \frac{1}{2}, \qquad \mathrm{Var}(p) = \mathrm{Var}(q) = \frac{1}{12}.$$

Based on data prior to our sample, it appears that the upswings and downswings of the Japanese current account have an average duration of roughly ten to twelve quarters. This suggests choosing an informative prior for the transition probabilities with the following first and second moments:

$$\mathrm{E}(p) = \mathrm{E}(q) = \frac{10}{11}, \qquad \mathrm{Var}(p) = \mathrm{Var}(q) = \frac{1}{2000}.$$

Notice that the expected duration of regime 1, say, equals $1/(1-p)$. The implied hyper-parameters of the prior are:

$$u_{11} = u_{00} = 149, \qquad u_{10} = u_{01} = 14.9.$$

4.2.3 Generating ν and Ψ

Conditional on all other parameters, $\psi_{z,1}$, $\psi_{z,2}$, $\psi_{z,3}$ and ν_z may be generated as follows. Define C to be the Choleski decomposition of Σ. Since Σ is given, C may be evaluated and the equation (10) may be multiplied through with the inverse of C. After rearranging, I obtain:

$$y^* = \Psi^* d_t + \nu^* R_t + \varepsilon_t^*, \tag{15}$$

where

$$y^* \equiv C^{-1} \left(\Delta y_t - \pi_0 - \sum_{i=1}^{h-1} \Pi_{t-i} \Delta y_{t-i} - \Pi y_{t-1} \right),$$

$$\Psi^* \equiv C^{-1} \Psi, \quad \nu^* \equiv C^{-1} \nu, \quad \varepsilon_t^* \equiv C^{-1} \varepsilon_t.$$

Notice that as $\varepsilon_t^* \sim \mathrm{N}(0, I)$, equation (15) represents a system of two independent equations. With all elements of y^* given, the second equation—which contains the parameters that are of interest to us—may be estimated in isolation. Consider therefore the Bayesian estimation of the following linear regression model:

$$Y^* \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} D & R \end{bmatrix} \begin{bmatrix} \psi_{z,1} \\ \psi_{z,2} \\ \psi_{z,3} \\ \nu_z \end{bmatrix} + E^* \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

where $E^* [0, 1]' \sim \mathrm{N}(0, I_T)$.

The natural conjugate prior for $[\psi_{z,1}, \psi_{z,2}, \psi_{z,3}, \nu]'$ is the normal density:

Prior of $[\psi_{z,1}, \psi_{z,2}, \psi_{z,3}, \nu]'$

$$\begin{bmatrix} \psi_{z,1} & \psi_{z,2} & \psi_{z,3} & \nu \end{bmatrix}' \sim \mathrm{N}(\gamma_0, \Omega_0),$$

where γ_0 and Ω_0 can be chosen appropriately (a non-informative prior may be obtained by choosing Ω_0 very large).

20

The posterior distribution for $[\psi_{z,1}, \psi_{z,2}, \psi_{z,3}, \nu]'$ is given by the following normal distribution (Kim & Nelson 1999, page 174):

Posterior of $[\psi_{z,1}, \psi_{z,2}, \psi_{z,3}, \nu]'$

$$\begin{bmatrix} \psi_{z,1} & \psi_{z,2} & \psi_{z,3} & \nu \end{bmatrix}' \sim \mathrm{N}(\gamma_1, \Omega_1),$$

where

$$\gamma_1 = \left[\Omega_0^{-1} + R'D'DR\right]^{-1} \left[\Omega_0^{-1}\gamma_0 + R'D'Y^* \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right],$$

$$\Omega_1 = [\Omega_0^{-1} + R'D'DR]^{-1}.$$

4.2.4 Generating Σ and B

Consider now the estimation of Σ and B conditional on all the other parameters. The system in equation (10) then becomes:

$$
\begin{aligned}
y_t^{**} &\equiv y_t - \Psi d_t - \nu R_t \\
&= \pi_0 + \sum_{i=1}^{h-1} \Pi_{t-i}\Delta y_{t-i} + \alpha\beta' y_{t-1} + \varepsilon_t, \\
\varepsilon_t &\sim \text{ i.i.d. } \mathrm{N}(0, \Sigma).
\end{aligned}
\tag{16}
$$

Since β as well as all the parameters contained in y_t^{**} are given, I can treat equation (16) as a multivariate linear regression model. Recalling the notation of equation (11), this model becomes:

$$
\begin{aligned}
Y^{**} &\equiv Y - D\Psi' - R\nu' \\
&= WB + E, \\
\mathrm{vec}(E) &\sim \mathrm{N}_{Tn}(0, \Sigma \otimes I_T).
\end{aligned}
\tag{17}
$$

This regression can be estimated using Theorem 9.1 (or Theorem 9.3) in Bauwens et al. (1999).

Prior of B and Σ   A non-informative prior is applied:

$$g(B, \Sigma) \propto |\Sigma|^{-(n+1)/2}.$$

Posterior of B and Σ Let k equal the number of columns of W and Mt and IW denote the matricvariate Student and inverted Wishart distributions, respectively. The posterior densities of B and Σ are given by:

$$B \sim \mathrm{Mt}_{k\times n}(\hat{B}, W'W, S, T-k),$$
$$\Sigma \sim \mathrm{IW}_n(S, T),$$

where

$$\hat{B} = (W'W)^{-1}W'Y,$$
$$S = (Y - W\hat{B})'(Y - W\hat{B}).$$

4.2.5 Generating β using griddy Gibbs sampling

I now turn to the problem of estimating the cointegrating vector β. Let $g(\beta)$ be the prior for β. Following Bauwens et al. (1999, Theorem 9.3), the kernel of β is given by:

$$g(\beta|\tilde{y}_T) \propto g(\beta)|\beta'V_0\beta|^{l_0}/|\beta'V_1\beta|^{l_1}, \tag{18}$$

where

$$V_0 = Z'M_X Z,$$
$$V_1 = Z'M_Y[I_T - X(X'M_Y X)^{-1}X']M_Y Z$$
$$\quad = Z'M_X[I_T - Y(Y'M_X Y)^{-1}Y']M_X Z,$$
$$M_X = I_T - X(X'X))^{-1}X',$$
$$M_Y = I_T - Y(Y'Y))^{-1}Y',$$
$$l_0 = (T - k - n)/2,$$
$$l_1 = (T - k)/2.$$

Note that when the first element of β is normalized to one, the task reduces to sampling from the univariate conditional posterior distribution of β_2. Some analytical results are available. Bauwens et al. (1999, Corollary 9.4) show that under a non-informative prior for β—which is the simplest case—, the posterior density of β is a 1-1 poly-t density. However, the simulation of a 1-1 poly-t density is by no means easy and involves a substantial fixed cost in terms of programming (for an algorithm, see Bauwens & Richard 1985).

For this reason, I apply the griddy Gibbs sampler as proposed by Ritter & Tanner (1992). The griddy Gibbs sampler is an attractive device whenever it is difficult to directly

sample from $g(\theta_i|\tilde{y}_T, \theta_{-i})$, the posterior density of a parameter θ_i conditional on the data and all other parameters of the model. The idea is to form a simple approximation to the inverse cdf of this density based on the evaluation of $g(\theta_i|\tilde{y}_T, \theta_{-i})$ on a grid of points. The procedure consists of the following steps (letting j denote the current Gibbs iteration):

Step 1 Specify a grid for θ_i, say $\theta_{i,1}, \theta_{i,2}, \ldots, \theta_{i,n}$.

Step 2 Evaluate $g(\theta_i|\tilde{y}_T, \theta_{-i})$, using the most recently simulated values for θ_{-i}, to obtain $w_1, w_2, \ldots, w_n$.

Step 3 Use $w_1, w_2, \ldots, w_n$ to obtain an approximation to the empirical inverse cdf of $g(\theta_i|\tilde{y}_T, \theta_{-i})$. Denote the approximate inverse cdf as $\hat{F}^{-1}(\cdot)$.

Step 4 Draw $\zeta^{(j)}$, a uniformly distributed random variable on $[0, 1]$, to obtain $\theta_i^{(j)} = \hat{F}^{-1}(\zeta^{(j)})$.

To simulate β, I use a piecewise linear approximation to the empirical inverse cdf based on the posterior density of β given in equation (18). Ritter & Tanner (1992) discuss a number of possible enhancements to the procedure, regarding for example the adjustment of the grid or the approximation to the inverse cdf. However, I find that already a simple approximation and a uniformly spaced, stable grid do a satisfactory job.

4.2.6 Design of the Gibbs sampler

The Gibbs sampling scheme is run over 5000 iterations, of which the first 2500 simulations are discarded.

To start the sampling algorithm, initial values for the parameters are specified. As an starting value for the cointegrating vector, β, I take the estimate from a standard vector error correction model with constant, non-switching intercepts. I further choose 0.5 as the initial value for the transition probabilities. Σ is set to equal the identity matrix. All other parameters are set to zero. The estimation results appear robust to changes in these specifications.

5 Empirical results

5.1 Parameter estimates

The model is estimated with two different assumptions regarding the prior for the transition probabilities. The exact way of choosing the prior is discussed in appendix 4.2.2. The estimations were carried out using a lag length, h, of 4.

5.1.1 Using a non-informative prior

I first use a flat prior for p and q. The problem of using a non-informative prior is that during some periods of the sampling, the draws for p and q are unreasonably low. Even though most of the density mass of the posterior density for these parameters lies in the region between 0.8 and 0.9, there is still substantial mass in the lower tail of the density. This indicates that the transition probabilities occasionally pick up seasonal and short-term fluctuations. Figure 6 plots the estimated probability of being in regime 1. However, the graph shows that even with a non-informative prior, the shifts of the model from one regime to another are clearly discernible.
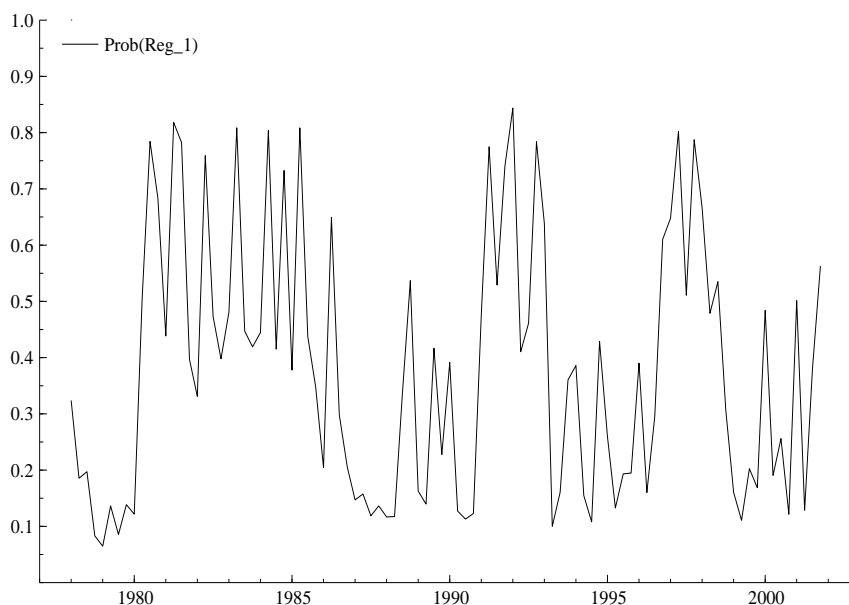

Figure 6: **Regime probabilities (non-informative prior).** Probability of regime with high current account intercept, using a non-informative prior for transition probabilities.

5.1.2 Using an informative prior

Problems with non-informative priors similar to the one described here also arise in the literature that applies Bayesian Markov-switching models to business cycle data (see for instance Kim & Nelson 1998). The usual approach in that literature is to specify informative priors for the parameters, in particular the transition probabilities, incorporating what is known or believed about the average duration of business cycle phases. I proceed in a similar way here by choosing priors for p and q that give a low weight to

Parameter	Posterior		
	Mean	Median	95% interval
β_2	-7.9984	-8.0668	-12.986, -2.7074
ν_z	0.0085137	0.0085855	0.0056633, 0.011192
$\psi_{z,3}$	0.00091097	0.00094097	-0.0011817, 0.0031250
$\pi_{0,1}$	0.60585	0.61299	0.24555, 0.95397
$\pi_{0,2}$	-0.0044141	-0.0048382	-0.047484, 0.043137
$\pi_{1,11}$	0.29403	0.29911	0.13980, 0.45853
α_1	-0.13231	-0.13553	-0.20944, -0.052209
α_2	-0.00017333	5.8164e-005	-0.010508, 0.0094079
σ_s^2	0.0021557	0.0020229	0.0015517, 0.0031886
σ_{sz}	1.0751e-005	9.8425e-006	-3.2738e-005, 5.1592e-005
q	0.91140	0.91119	0.87648, 0.94534
p	0.90679	0.90708	0.87458, 0.93724

Table 3: **Parameter estimates.** *Posterior estimates of selected parameters. Mean, median and 95% interval of the Gibbs simulations for each parameter, after discarding initial simulations (see section 4.2.6).*

transition probabilities that would imply very frequent current account reversals (details in Appendix 4.2.2).[6]

Table 3 presents the estimation results for a subset of the parameters using these priors. Figure 7 plots the Gibbs simulations for several parameters, along with a correlogram indicating the degree of serial correlation of the drawings. Also drawn are the simulated posterior densities of the parameters, which all appear well-shaped. Note that the estimate for the cointegrating vector, β, $[1.0, -0.25018]'$, is relatively close to the estimate from an ordinary, non-switching vector error correction model, which is $[1.0, -0.26805]'$. Finally, figure 8 plots the estimated probability of being in regime 1. Periods when the current account, and thus the exchange rate, are strengthening (regime 1), are again clearly distinguished from the other periods when the current account is downward-drifting (regime 0).

5.2 Exchange rate simulations with and without current account shifts

I have carried out some simulations of the estimated model in order to assess the significance of its regime-switching component for the variability of the current account and the exchange rate. I simulate the model twice (with an identical sequence of random shocks hitting the variables in both cases). The simulation period is sixty years, not taking into account an initial warming-up period. In the first simulation, I generate artificial time

---

[6]Alternatively, one could smooth the current account to remove its short-term fluctuations and then estimate the model using a non-informative prior.
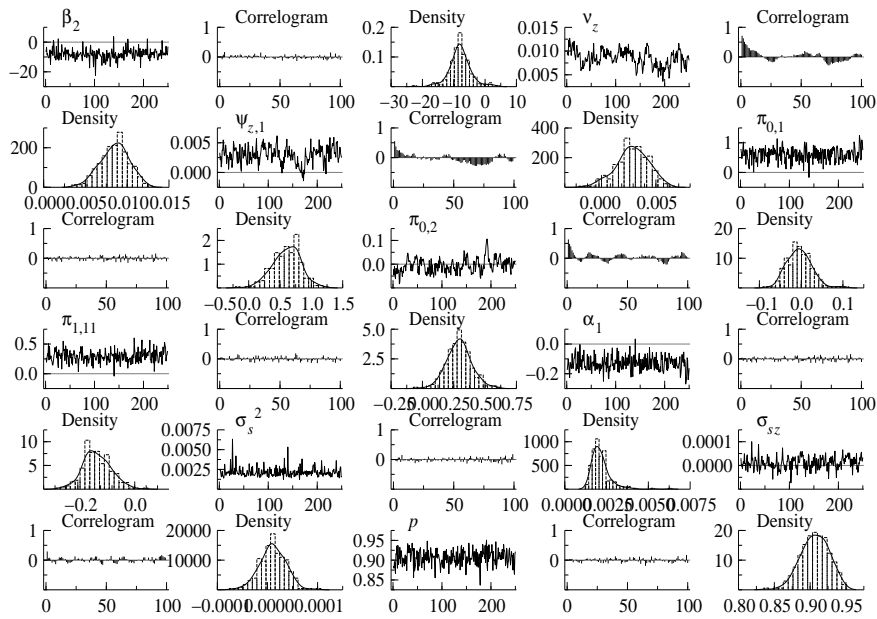
Figure 7: **Gibbs simulations (informative prior).** Posterior distributions of selected parameters, using an informative prior for transition probabilities.

series for the current account and exchange rate from the estimated model. In the second simulation, I set the regime-switching intercept of the current account to its time-invariant average, which I can evaluate on the basis of the estimated values for $\pi_{0,2}$, ν_z, p and q. This exercise is meant to give us an idea of what would happen to the exchange rate once all structural breaks of its fundamental are eliminated. The results, which are shown in figure 9, indicate that without the long swings in the current account, exchange rate fluctuations over the longer term could be much smoother.

6 Conclusions

I have estimated a Markov-switching vector error correction model of the Japanese real effective exchange rate and current account. Two general observations are underlying the model. First, Japan is a country where the exchange rate is predominantly influenced by a single economic fundamental, the current account. Second, the movement of the Japanese current account alters its direction from time to time. Taken together, these two findings imply that the current account exhibits large swings over the years, which translate into similarly large swings of the exchange rate.

In the model, the current account contains a stochastic trend that switches between two regimes, or states, according to an unobservable Markov process. The exchange rate itself is modelled as independent of the regimes. However, since it is cointegrated with
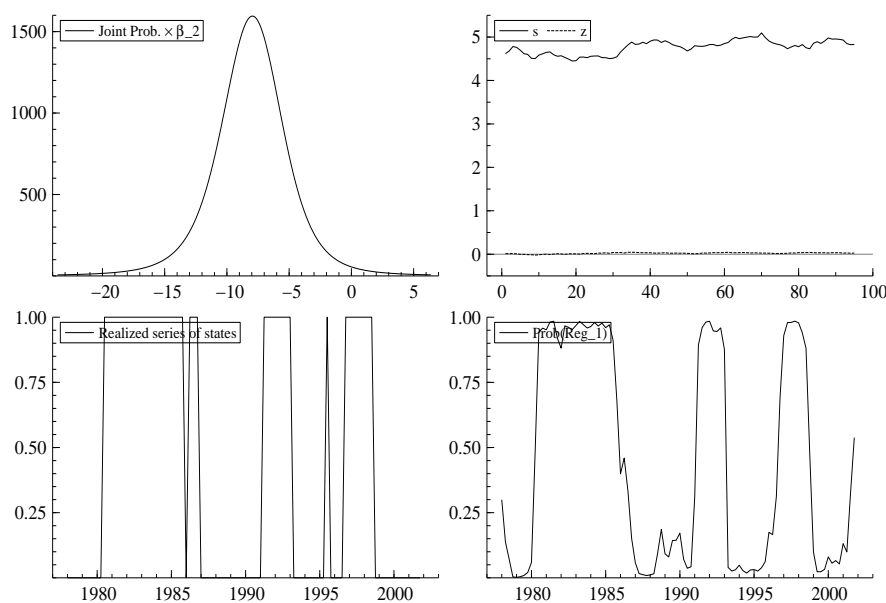
Figure 8: **Regime probabilities (informative prior).** Probability of regime with high current account intercept, using an informative prior for transition probabilities.

the current account, it follows the movements of the current account and adjusts to current account reversals whenever they occur. A simulation exercise demonstrates that the effect of the shifts between the regimes is rather large. The fluctuations of the exchange rate and its fundamental are large when regime switching is allowed for. They become much smaller once the model is simulated without the regime changes.

The paper seeks to explain the strong appreciation of the yen and its large variability, which have had important implications for Japan's economic performance in recent years. But the motivation of this paper goes further. Its more general intention is to highlight the issues on which empirical research on exchanges rates should focus in future.

First, an often raised question in exchange rate economics is whether exchange rates are driven by economic fundamentals (MacDonald 1999). In an influential paper, Meese & Rogoff (1983) demonstrated twenty years ago that a simple random work can outperform many fundamentals-based exchange rate models in terms of out-of-sample forecasting performance. What the present study of Japan seeks to demonstrate is that fundamentals are *strongly* relevant for medium- and long-run exchange rate fluctuations.[7]

---

[7]This is, by the way, just the view of market participants: Cheung & Wong (2000) have recently carried out a survey of practitioners in the interbank foreign exchange markets in Hong Kong, Tokyo, and Singapore. They report that at the medium-run horizon, between 29% (Tokyo) and 35% (Hong Kong) of the replies assert that exchange rate variation is determined by economic fundamentals. The proportion of respondents who hold the same opinion for exchange rates in the long run increases to 76% from Hong Kong and 82% from Tokyo.
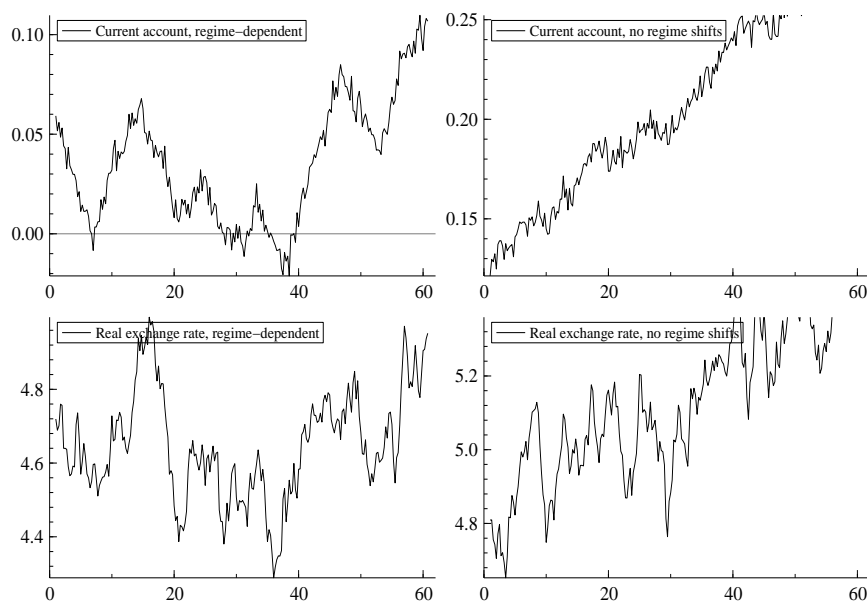
Figure 9: **Simulating exchange rate dynamics.** Simulation of current account (top) and nominal exchange rate (bottom) from the estimated model, over a period of 60 years. Simulations of the original model with switching regimes (panels on the left), and of a single-regime version of the model where the stochastic trend of the current account variable is set to its estimated average (panels on the right).

Second, there has been a lot of controversy over the validity of purchasing power parity (PPP). The PPP puzzle (Rogoff 1996) concerns the question why deviations from PPP are so persistent. The average half-life of a shock to PPP is estimated to be around three to five years, which is difficult to explain with nominal rigidities alone. By studying carefully the experience of an individual country, namely Japan, this paper finds that long-lasting deviations of the real exchange rate from purchasing power parity may be the result of large and persistent movements of the underlying economic fundamentals. Rather than asking why the law of one price does not lead to a faster adjustment towards PPP, the paper suggests that research should focus on the economic forces that produce to deviations from PPP in the first place.

Third, partly as a consequence of the poor empirical performance of PPP models, researchers have recently turned to nonlinear exchange-rate modelling (Michael et al. 1997, Obstfeld & Taylor 1997, Taylor & Peel 2000). Nonlinear time series models of exchange rates are often motivated by the idea that transportation costs reduce arbitrage opportunities in international goods markets, implying that small deviations from PPP can be persistent. This paper shows that what is unsatisfactory with existing nonlinear time series models of exchange rates is that they are generally univariate. Univariate models

28

have difficulties to explain why exchange rates deviate from PPP or why deviations, even large ones, are so persistent.

On the technical side, the paper offers an interesting innovation. The vector error correction model estimated in this paper contains a Markov-switching intercept and seasonal effects only in the equation of the exchange rate fundamental. I show that inference becomes feasible once the model's equations are made independent by making use of the Choleski decomposition of their covariance matrix since univariate inference can then be carried out directly on any equation-specific components of the model. This method is simple to use and it is not difficult to think of applications to other economic problems involving cointegrated systems that are subject to influences affecting variables individually (shifts in regimes, changing volatilities, seasonal effects etc.).

Appendices

A Solution of the theoretical model

In this appendix, I show how the theoretical model in section 2.2 can be solved explicitly using the method of Laplace transforms. The method consists of applying Laplace transforms to each of the equations in the model, solving the system so obtained for the Laplace transforms of all of the variables and applying the inverse Laplace transform to each of the resulting equations.

While the model can be solved explicitly, it must be stressed that the solutions to even rather simple-looking dynamical systems tend to contain unwieldy algebraic expressions that do not allow for an easy interpretation unless strong restrictions are placed on the parameters and on the initial conditions. To avoid excessive complexity, we assume here that the dampening parameter ϕ_1 is zero. The consequences of relaxing this assumptions are explained in the text.

To obtain a solution in terms of the initial conditions for $z(0)$ and $\dot{z}(0)$, first rewrite the model as a system of two equations:

$$\dot{z}(t) = -\phi_2 \xi (z(t) + d(t)), \tag{19a}$$

$$d(t) = -z(t) - \gamma \int_0^t e^{-\gamma(t-\tau)} d(\tau) d\tau. \tag{19b}$$

Now rewrite the model in terms of the Laplace transforms of $z(t)$ and $d(t)$:

$$s^2 Z(s) - \dot{z}(0) - sz(0) = -\phi_2 \xi (Z(s) + D(s)), \tag{20a}$$

$$D(s) = -Z(s) - \frac{\gamma}{s + \gamma} D(s), \tag{20b}$$

where

$$Z(s) := L\{z(t)\} \quad \text{and} \quad D(s) := L\{d(t)\}.$$

Solving these two equations for $Z(s)$ and $D(s)$ yields:

$$Z(s) = \frac{z(0)s^2 + (2\gamma z(0) + \dot{z}(0))s + 2\gamma \dot{z}(0)}{s^3 + 2\gamma s^2 + \gamma \phi_2 \xi}, \tag{21a}$$

$$D(s) = -\frac{z(0)s^2 + (\gamma z(0) + \dot{z}(0))s + \gamma \dot{z}(0)}{s^3 + 2\gamma s^2 + \gamma \phi_2 \xi}. \tag{21b}$$

In this system of equations, there are two initial conditions, $z(0)$ and $\dot{z}(0)$. To keep things simple, we choose to set $z(0) = 0$ from now on. The final solution will therefore depend only on one initial condition, namely $\dot{z}(0)$.

Using the method of partial fractions, the solution can be rewritten in terms of Laplace transforms of the exponential, sine and cosine functions:

$$Z(s) = \frac{A_z}{s + B_z} + \frac{C_z s + D_z}{(s + E_z)^2 + F_z^2}, \tag{22a}$$

$$D(s) = \frac{A_d}{s + B_d} + \frac{C_d s + D_d}{(s + E_d)^2 + F_d^2}, \tag{22b}$$

where

$$A_z = -\frac{\phi_2 \xi \dot{z}(0)}{3B_z \phi_2 \xi + 4\gamma B_z^2 + 2\gamma \phi_2 \xi},$$

$$B_z = \frac{1}{6}\bar{B}_z + \frac{8}{3}\gamma^2 \bar{B}_z^{-1} + \frac{2}{3}\gamma,$$

$$\bar{B}_z = \left(108\gamma \phi_2 \xi + 64\gamma^3 + 12\sqrt{81\gamma^2 \phi_2^2 \xi^2 + 96\gamma^4 \phi_2 \xi}\right)^{\frac{1}{3}},$$

$$C_z = -A_z,$$

$$D_z = \frac{\dot{z}(0)\left(\phi_2^2 \xi^2 + 8\gamma^2 \phi_2 \xi + 4B_z \gamma \phi_2 \xi + 6\xi \phi_2 B_z^2 + 16\gamma^2 B_z^2\right)}{8\xi \phi_2 B_z^2 + 4B_z \gamma \phi_2 \xi + 16\gamma^2 B_z^2 + 3\phi_2^2 \xi^2 + 8\gamma^2 \phi_2 \xi},$$

$$E_z = -\frac{\gamma \phi_2 \xi}{2B_z^2},$$

$$F_z = \frac{1}{2}\sqrt{\frac{4\gamma \phi_2 \xi}{B_z} - \frac{\gamma^2 \phi_2^2 \xi^2}{B_z^4}},$$

$$A_d = \frac{B_d^2 \dot{z}(0)(\gamma^2 + \phi_2 \xi)}{\gamma(5\xi \phi_2 B_d^2 - B_z \gamma \phi_2 \xi + 3\phi_2^2 \xi^2 + 4\gamma^2 B_d^2 + 2\gamma^2 \phi_2 \xi)},$$

$$B_d = B_z,$$

$$C_d = -A_d,$$

$$
\begin{aligned}
D_d = &-\dot{z}(0)(7\gamma^3\phi_2^2\xi^2 + 4\phi_2^3\xi^3\gamma + 4\gamma^5\phi_2\xi + 2\gamma^2 B_d\phi_2^2\xi^2 + \phi_2^3\xi^3 B_d \\
&+2B_d\gamma^4\phi_2\xi + 16\gamma^3 B_d^2\phi_2\xi + 9\phi_2^2\xi^2 B_d^2\gamma + 8B_d^2\gamma^5)/(17\phi_2^2\xi^2 B_d^2\gamma \\
&+16B_d^2\gamma^5 + 32\gamma^3 B_d^2\phi_2\xi + 3\phi_2^3\xi^3 B_d + 6\gamma^2 B_d\phi_2^2\xi^2 + 4B_d\gamma^4\phi_2\xi \\
&+8\gamma^5\phi_2\xi + 15\gamma^3\phi_2^2\xi^2 + 8\phi_2^3\xi^3\gamma),
\end{aligned}
$$

$$E_d = \frac{\phi_2\xi(-B_d^2 + \gamma B_d + \gamma^2)}{2(B_d\phi_2\xi + \gamma B_d^2 + \gamma\phi_2\xi)},$$

$$F_d = \frac{1}{2}\sqrt{\frac{4\gamma\phi_2\xi(B_d^2 - \gamma B_d + \phi_2\xi)}{B_d\phi_2\xi + \gamma B_d^2 + \gamma\phi_2\xi} - \frac{\phi_2^2\xi^2(B_d^2 - \gamma B_d - \gamma^2)^2}{(B_d\phi_2\xi + \gamma B_d^2 + \gamma\phi_2\xi)^2}}.$$

Applying the inverse Laplace transforms to each of the resulting fractions, and making use of the First Shifting Theorem, we obtain the solutions for $z(t)$ and $d(t)$ in equations (8a) and (8b).

B Data

Unless indicated otherwise, all data used for this paper are from the IMF's International Financial Statistics. Data on the Japanese balance-of-payment components has been converted into yen for the estimations, using the IMF's quarterly yen-dollar exchange rate.

C Software

The computations for this paper were carried out using Ox, version 3.0 (see Doornik & Ooms 2001), and PcFiml (see Doornik & Henry 1997). The programs are available from the author upon request. Separate code was written for generating random numbers from the matricvariate Student and inverted Wishart distributions. In order to sample from the matricvariate Student distribution, I used the algorithm in Bauwens et al. (1999, Appendix B.4.5). As regards the inverted Wishart distribution, I translated to Ox a Gauss code that I kindly received from Luc Bauwens.

References

Albert, J. H. & Chib, S. (1993), 'Bayes inference via Gibbs sampling of autoregressive time series subject to markov mean and variance shifts', *Journal of Business and Economic Statistics* **11**(1), 1–15.

Balassa, B. (1964), 'The purchasing power parity doctrine: A reappraisal', *Journal of Political Economy* **72**(6), 584–596.

Bauwens, L., Lubrano, M. & Richard, J.-F. (1999), *Bayesian Inference and Dynamic Econometric Models*, Oxford University Press, New York.

Bauwens, L. & Richard, J.-F. (1985), 'A 1-1 poly-t random variable generator with application to Monte Carlo integration', *Journal of Econometrics* **29**, 19–46.

Brooks, R., Edison, H. J., Kumar, M. & Sløk, T. (2001), 'Exchange rates and capital flows', *International Monetary Fund Working Papers* (190).

Bussiere, M. & Mulder, C. (1999), 'External vulnerability in emerging market economies: How high liquidity can offset weak fundamentals and the effects of contagion', *International Monetary Fund Working Papers* (88).

Cheung, Y.-W. & Wong, C. Y.-P. (2000), 'A survey of market practitioners´ views on exchange rate dynamics', *Journal of International Economics* **51**(2), 401–419.

Doornik, J. A. & Henry, D. F. (1997), *Modelling Dynamic Systems Using PcFiml 9.0 for Windows*, International Thomson Business Press, London, Boston.

Doornik, J. A. & Ooms, M. (2001), 'Introduction to Ox: Ox version 3'.

Eichengreen, B. (1996), *Globalizing Capital*, Princeton University Press, Princeton, New Jersey.

Eichengreen, B. (2003), *Capital Flows and Crises*, The MIT Press, Cambridge, Massachusetts; London.

Engel, C. (1999), 'Accounting for US real exchange rate changes', *Journal of Political Economy* **107**(3), 507–538.

Evans, C. L. & Lyons, R. K. (2002), 'Order flow and exchange rate dynamics', *Journal of Political Economy* **110**(1), 170–180.

Evans, M. D. D. & Lyons, R. K. (2003), 'New micro exchange rate economics'. Unpublished working paper, Haas School of Business, UC Berkeley.

Gelman, A., Carlin, J. B., Stern, H. S. & Rubin, D. B. (1995), *Bayesian Data Analysis*, Chapman and Hall, London.

Girton, L. & Roper, D. (1977), 'A monetary model of exchange market pressure applied to the postwar Canadian experience', *American Economic Review* **67**(4), 537–548.

Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica* **57**(2), 357–384.

Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control* **12**, 231–254.

Kenen, P. B. (2000), *The International Economy*, Cambridge University Press, Cambridge (United Kingdom), New York.

Kim, C.-J. & Nelson, C. R. (1998), 'Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching', *Review of Economics and Statistics* **80**(2), 188–201.

Kim, C.-J. & Nelson, C. R. (1999), *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, MIT Press, Cambridge, Mass.

Lau, E., Baharumshah, A. Z. & Chan, T. H. (2006), 'Current account: Mean-reverting or random walk behaviour', *Japan and the World Economy* **18**, 90–107.

MacDonald, R. (1999), 'Exchange rate behaviour: Are fundamentals important?', *The Economic Journal* **109**, 673–691.

Meese, R. & Rogoff, K. (1983), 'Empirical exchange rate models of the seventies: Do they fit out of sample?', *Journal of International Economics* **14**, 3–24.

Michael, P., Nobay, A. R. & Peel, D. A. (1997), 'Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation', *Journal of Political Economy* **105**(4), 862–879.

Müller-Plantenberg, N. (2006), Japan's imbalance of payments, *in* M. M. Hutchison & F. Westermann, eds, 'Japan's Great Stagnation: Financial and Monetary Policy Lessons for Advanced Economies', MIT Press.

Obstfeld, M. & Taylor, A. M. (1997), 'Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher´s commodity points revisited', *Journal of the Japanese and International Economies* **11**, 441–479.

Ritter, C. & Tanner, M. A. (1992), 'Facilitating the Gibbs sampler: the Gibbs stopper and the griddy-Gibbs sampler', *Journal of the American Statistical Association* **87**(419), 861–868.

Rogoff, K. (1996), 'The purchasing power parity puzzle', *Journal of Economic Literature* **34**, 647–668.

Rosenberg, M. R. (1996), *Currency Forecasting. A Guide to Fundamental and Technical Models of Exchange Rate Determination*, Irwin, London, Chicago, Singapore.

Samuelson, P. A. (1964), 'Theoretical notes on trade problems', *Review of Economics and Statistics* **46**, 145–154.

Taylor, M. P. & Peel, D. A. (2000), 'Nonlinear adjustment, long-run equilibrium and exchange rate fundamentals', *Journal of International Money and Finance* **19**, 33–53.