Manuscript

Forecasting Exchange Rates: The Multi-State Markov-Switching Model with Smoothing

Chunming Yuan*

Department of Economics, University of Maryland, Baltimore County

1000 Hilltop Circle, Baltimore, MD 21250, USA

Abstract

This paper presents an exchange rate forecasting model that outperforms a random walk at short horizons and appears to be robust over different sample spans. Popular smoothing techniques are employed to remove the irregular blips from the noisy data so that the Markov model can capture precisely the trend persistence in exchange rates. Our finding hinges on the fact that exchange rates tend to follow highly persistent trends and the key to beating the random walk is to identify these trends. An attempt to link the trends in exchange rates to the underlying macroeconomic determinants further reveals that fundamentals-based linear models generally fail to capture the persistence in exchange rates and thus are incapable of outforecasting the random walk.

JEL Classification Codes: C32, C53, F31, F47

Keywords: Exchange Rate, Forecasting, Markov-Switching, Smoothing, HP-Filter, Fundamentals-index.

---

* Correspondence to: Chunming Yuan, Department of Economics, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA. Email Address: cmyuan@umbc.edu; Phone: 1-410-455-2314; Fax: 1-410-455-1054.

I am very grateful to Aaron Tornell, Bryan Ellickson, Raffaella Giacomini, and Edward McDevitt for their valuable comments. All remaining errors are my own.

1. Introduction

Explicating the behavior of nominal exchange rates is one of the central themes in international economics. Modeling exchange rates, however, has proved notoriously challenging to economists since the celebrated work of Meese and Rogoff (1983a, b), who found that the fundamentals-based exchange rate models systematically fail to deliver better forecasts than a simple random walk at horizons of up to one year. Subsequent studies, based on longer data sets and employing more sophisticated econometric techniques, attempted to overturn the Meese and Rogoff result but generally turned out to be futile.[1] One prominent exception is the study by Engel and Hamilton (1990) who modeled exchange rates alternating between appreciation and depreciation regimes in a Markovian fashion. They showed that during the floating period of 1973-88, their canonical two-state Markov-switching model captures very well the long-swing feature of major exchange rates and outperforms the random walk both in-sample and out-of-sample at short horizons. Unfortunately, when considering more recent data, their model no longer beats the random walk.

In this paper, we present an exchange rate forecasting model that beats the random walk at short horizons and appears to be robust over different sample spans. The centerpiece of this forecasting model is to exploit the fact that exchange rates tend to follow highly persistent trends and accordingly, the key to beating the random walk is to identify these trends. We extend Engel and Hamilton's (1990) Markov-switching model by allowing for multiple states in which trendless periods are considered in addition to the appreciation and depreciation regimes. In a more important respect, we employ time series filtering techniques to smooth out outliers or transitory blips from the noisy data so as to guarantee that the Markov-switching framework captures more precisely the trend persistence in exchange rates. This practice, to the best of the author's knowledge, presents the first application of combining the Markov-switching model with smoothing techniques to exchange rate forecasting.

Two key observations have motivated us to tailor the standard Markov-switching model. First, imposing only two regimes – appreciation and depreciation – is not consistent with the fact that almost all exchange rates occasionally exhibit range-bound behavior for a sustained period of time (see Fig. 1). This suggests that a third trendless regime might be necessary for better describing the behavior of exchange rates. Second, the standard Markov-switching model is likely to overreact to irregular transitory blips in the data. Since financial time series like exchange rates are often extremely noisy, the oversensitivity of

[1] See Neely and Sarno's (2002) review of the empirical literature on the monetary approach. To mention some, for example, Wolff (1988) used time-varying coefficients, Meese and Rose (1991) allowed for nonlinear formulation of coefficients, Macdonald and Marsh (1997) considered cointegration between fundamentals and exchange rates in an error correction framework, and Killian and Taylor (2001) assumed mean-reverting properties of exchange rates. More recently, Cheung, Chinn, and Pascual (2005) further confirmed Meese and Rogoff's results.

the conventional model tends to induce instability in parameter estimation and misclassification of regime shifts, and in turn undermines its forecastability.[2] Our model corrects these two shortcomings.

Using quarterly data of major exchange rates during 1973-2007, we find that the modified forecasting model achieves considerable forecast accuracy improvement relative to the random walk in terms of forecast mean squared errors (MSE). Specifically, the out-of-sample forecast precision gain, averaging over horizons of up to four quarters, is 3 percent for the Australian dollar, 13 percent for the Canadian dollar, 11 percent for the British pound, and 9 percent for the Japanese yen, respectively. In contrast, the corresponding reduction in MSE from the original model is -2 percent, -7 percent, 9 percent, and -5 percent, respectively. We also experiment our specification using Engel and Hamilton's (1990) dataset. We obtain an average 22 percent increase in forecast accuracy across all currencies, a remarkable rise relative to their 11 percent improvement. When considering different forecast spans, similar results emerge. In general, the standard two-state Markov-switching model fails to offer convincing evidence of outperforming the random walk while our model consistently displays forecast superiority at short horizons across all currencies.

Our study goes further to examine the relationship between exchange rates and the underlying macroeconomic determinants. In particular, given the identified trends in exchange rates and their key role in achieving superior forecastability, it is of great interest to investigate whether these trends are linked in some way to the macroeconomic determinants. To this end, we extract the trend components of the fundamentals-index constructed through a prevailing monetary exchange rate determination model, and then conduct contingency and correlation tests on these two types of trends. Empirical results show that the pattern of the trends in exchange rates is quite different from that of the trends in the fundamentals-index. For example, during the 30 periods of exchange rates' uptrend there are no upward movements in the fundamentals-index while merely 9 periods of downward movements correspond to the 47 periods of exchange rates' downtrend. In addition, the correlation coefficients calculated through a rolling window are generally insignificant and virtually zero since 1995. This suggests that the trends in exchange rates are not related in a linear way to those in the fundamentals-index. Thus, failing to capture the trends in exchange rates can somewhat help explain why the fundamentals-based models do not beat the random walk.

[2] Marsh (2000), for instance, showed that the Markov-switching modeling generally offers sound in-sample fit but fails to deliver superior out-of-sample forecast due to the parameter instability over time. Similarly, Dacco and Satchell (1999) argued that the misclassification of regimes tends to make the Markov-switching models less effective in beating the random walk even if a good in-sample performance has been presented.

The remainder of the paper is structured as follows. Section 2 specifies the Markov-switching model and time series filtering techniques. An unobserved components model is addressed to help choose proper smoothing parameters for the relevant filter. Section 3 describes data, estimation procedure, and parameter estimates. Issues concerning estimation instability and regime misclassification are discussed in details. Section 4 presents forecast performance of the proposed multi-state Markov-switching models in terms of mean squared errors and forecast evaluation from the Diebold-Mariano's test of equal forecast accuracy. Section 5 examines robustness in forecasting superiority of the competing alternative models by varying smoothing parameters and forecasting subsamples. Section 6 analyzes the link between exchange rates and macroeconomic fundamentals in terms of their trend persistence. Section 7 concludes.

2. Model Specification

2.1 The Standard Markov-Switching Model

Following Hamilton (1989), the dynamics of exchange rates are modeled as a state-dependent process where the state is unobserved by the econometrician. Let y_t denote the change of log exchange rate in period t, and suppose that its mean and variance are governed by an unobserved state variable $s_t \in \{1, 2, ..., k\}$, where $s_t = k$ denotes a period of being in state k. The standard k-state Markov switching model can be written as

$$y_t = \mu(s_t) + \sigma(s_t)\varepsilon_t \qquad \text{with } \varepsilon_t \overset{iid}{\sim} N(0,1) \tag{1}$$

such that

$$\begin{cases} y_t = \mu_1 + \sigma_1\varepsilon_t, & \text{if } s_t = 1 \\ y_t = \mu_2 + \sigma_2\varepsilon_t, & \text{if } s_t = 2 \\ \qquad \cdots & \\ y_t = \mu_k + \sigma_k\varepsilon_t, & \text{if } s_t = k \end{cases}$$

Interpretation of the model depends on the value of k. For example, if there are only two states governing the data process, one may view $s_t = 1$ as a period of downtrend of exchange rates associated with a negative mean change while $s_t = 2$ as a period of uptrend of exchange rates corresponding to a positive mean change. In the case of three states, one may include a state of trendless period in which exchange rates fluctuate around a mean zero. Furthermore, a multi-state model is also economically intriguing to nest the flexibility that allows for varying slopes during uptrend/downtrend episodes.

This, nevertheless, raises debates on choosing the optimal value k in modeling the dynamics of exchange rates. Although economic intuition suggests a three-state model may be appropriate for capturing nonlinearity in the data generating process in which exchange rates alternate between sustained periods of appreciation, depreciation, and stagnation, there is virtually no standard distributional theory applicable for evaluating the Markov-switching model against alternatives such as a linear time series model. Some recent studies have proposed unconventional testing procedures attempting to tackle this problem. For example, Cheung and Erlandsson (2005) proposed a simulated likelihood ratio test based on a Monte Carlo method. Nevertheless, their results are fairly sample-specific as they themselves admitted. In this study, we consider $k = 2$ and 3 for the Markov model to compare the forecast performance.

The state variable s_t is assumed to follow an ergodic first-order Markov process and is characterized by the transition matrix:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

where $p_{ij} \equiv \Pr(s_t = j \mid s_{t-1} = i)$ denotes the probability that the process is in state j at time t given that it had been in state i the previous period, and $\sum_j p_{ij} = 1$.

As noted in Hamilton (1989, 1990), one nice feature of this framework is that it accommodates a variety of time series behaviors which are determined endogenously by the estimation procedure instead of imposed exogenously. For example, if estimates show that both μ_i and μ_j are significant with opposite signs and their corresponding transition probabilities, p_{ii} and p_{jj}, are large (say, close to one), it suggests that exchange rates may have sustained periods of uptrend and downtrend states. If the estimates for μ_i and μ_j are both positive (negative) and statistically significant, one can be assured that there may be different mean changes within upward (downward) movements. Alternatively, if estimates show that $p_{i\tau} = p_{is}$, $\forall \tau \neq s$, which means the exchange rate change the current period is completely independent of the state that prevailed last period, it suggests that there is no state-dependent regime shift, that is, the data generating process is simply a random walk. In addition, if one (or more) of the states in the Markov process is absorbing (i.e., once the process enter this state, it remains in the subsequent periods), the time series may be composed of deterministic segments, or structural breaks.

Another merit of this specification is allowing an analyst to form a probabilistic inference about the unobserved s_t based on the estimates of population parameters including state-specific mean, variance, and the transition probabilities p_{ij}. Two types of inferences can be obtained. The so-called smoothed probability, denoted as $\Pr(s_t = j \mid y_1, y_2, ..., y_T)$, which is the probability of being in state j based on the entire observed information, can be derived using an algorithm developed by Kim (1994) or Kim and Nelson (1999). In contrast, the filter probability, denoted as $\Pr(s_t = j \mid y_1, y_2, ..., y_t)$, is the best guess about s_t inferred by the information in the sample data up through time *t*. Algorithms capable of obtaining accurate probabilistic inferences prove to be crucial in forecasting, as this study shows.

2.2 The Markov-Switching Model with Smoothing

To date, a host of empirical attempts of Markov-switching model, including Hamilton (1989) on aggregated output, Filardo(1994) and Birchenhall et al (1999) on business-cycle phases, Cecchetti et al (1990), Abel (1994), and Bauwens et al (2006) on stock returns, and Engel and Hamilton (1990), Engel (1994), and Bollen et al (2000) on exchange rates, among many others, have seen some success in capturing the nonlinearity and regime shifts of the underlying time series, and shown some superiority in forecasting. Nevertheless, the Markov-switching model is not without demerits. Marsh (2000), for instance, showed that Markov models for exchange rates are unstable over time and unsuitable for forecasting. Dacco and Satchel (1999) also argued that the forecast performance of Markov-switching models is very sensitive to misclassification of regimes. Parallelly, Simpson et al (2001) pointed out that financial time series often appear to be quite noisy, particularly due to some extreme observations or outliers, and these irregular observations tend to cause estimation difficulties and sometimes boundary values for the transition probabilities in Markov models.

In light of these findings, we are motivated to incorporate appropriate smoothing techniques into the Markov-switching model, with attempt to alleviate the distortion by irregular components in exchange rates and in turn to enhance its forecastability. The modified model can be written as:

$$y_t = y_t^* + \eta_t \tag{2}$$

$$y_t^* = \mu(s_t) + \sigma(s_t)\varepsilon_t \quad \text{with } \varepsilon_t \overset{iid}{\sim} N(0,1) \tag{3}$$

where y_t^* and η_t are trend component and irregular component respectively. y_t^* is obtained through filtering techniques.[3] In this paper, we employ the Hodrick-Prescott filter (HP-filter, hereafter) to smooth out extreme irregular components from the raw exchange rates.[4]

---

3 In practice, we apply the time series filtering techniques to the raw log exchange rates and y_t^* is calculated as the first difference of the trend component.

4 Another two filtering techniques, the Christiano-Fitzgerald (2003) band pass filter and the simple moving average, are employed in the initial version of this paper. Since no reasonable forecasts are found based on the simple moving average

It is noteworthy that the term "smooth" is loosely defined. Our goal of this practice is to wipe off the outliers that may possibly induce spurious regime shifts but without destroying the intrinsic state-dependence in the dynamics of exchange rates. A solid theoretical derivation for the optimality of "smoothness" to fulfill this purported objective is apparently beyond the scope of this paper. Instead, we first obtain a rough belief on the smoothing parameter through estimating the state-space representation of the HP-filter, and then explore the numerical optimality by examining the effects of varying smoothing parameters according to forecast accuracy improvement.

2.2.1 The HP-Filter

Since originally proposed by Hodrick and Prescott (1980, 1997) who applied this procedure to measure the business-cycle of post-war US quarterly data, the HP-filter has probably become the most popular way of decomposing the economic time series into a growth (trend) component and a cyclical component in the recent years. The HP-filter is implemented by minimizing an objective function that depends on the weighted average of two parts: the squared sum of the cyclical component (the deviation from trend) and the squared sum of the acceleration of the trend component weighted by a parameter, the so-called smoothing λ. If x_t denotes some economic time series, the filter is defined as

$$\min_{\{x_t^*\}_{t=1}^{T}} \left\{ \sum_{t=1}^{T} (x_t - x_t^*)^2 + \lambda \cdot \sum_{t=2}^{T-1} \left[(x_{t+1}^* - x_t^*) - (x_t^* - x_{t-1}^*) \right]^2 \right\} \tag{4}$$

where x_t^* is the trend component and λ is the smoothing parameter. The cyclical component is deviation from the long run path (trend), $x_t - x_t^*$, and smoothness of the trend component is measured by the sum of squares of its second difference:

$$\Delta^2 x_t^* = (1-L)^2 x_t^* = (x_t^* - x_{t-1}^*) - (x_{t-1}^* - x_{t-2}^*) \tag{5}$$

where Δ^2 denote second-order difference, L is the lag operator with $Lx_t = x_{t-1}$.

The HP-filter attempts to minimize the cyclical component, which is equivalent to maximizing the fit of the trend to the series (analogous to OLS), while minimizing the change in the trend's slope. Apparently, these two minimizing efforts contradict each other. In this regard, the smoothing parameter plays a key role in that it determines the trade-off between "goodness of fit" and the smoothness of the trend component. As two extreme cases, when λ=0, the HP-filter returns the original series without

---

(even only two lags used to smooth data) while the band pass filter provides no better results than the HP-filter, these results are not presented here.

smoothing while as $\lambda \to \infty$, the HP-filter returns a linear OLS trend and in turn removes all the cyclical feature associated with the trend component but allows for large fluctuations in the cycle component.

2.2.2 The Smoothing Parameter λ of the HP-Filter

The conventional wisdom on the value of the smoothing parameter suggests to set λ=1600 for quarterly data and λ=100 for annual data.[5] In contrast to the previous efforts which attempt to seek the possibly best value of the smoothing parameter for the HP-filter in order to extract optimal trend component or cyclical component, our practice needs to bear a different kind of burden, in a sense that we shall pick a value for λ such that the regime-switching features in original data remain intact while the extreme noisy components which are likely to distort the estimates from Markov-switching model are removed as thoroughly as possible. Unfortunately, there is no theoretical priori to back up this selection criterion. It is virtually impossible to know to what magnitude a random shock will distort or not the estimation procedure of Markov-switching model given that we are not fully assured that the model is correctly specified as any test for model specification relies on sound estimates free of distortion and inconsistency. In light of the properties of the HP-filter, we can nevertheless maintain a reasonable belief – albeit primitive – that a small value of λ should be taken for the quarterly exchange rates which may guarantee the raw data just "slightly smoothed". For the purpose of this practice, should a value of the smoothing parameter help to enhance the forecastability, one may view it as a good choice, although it may not be an optimal one.

To this end, we first employ an unobserved components model to estimate the smoothing parameter λ. Harvey (1985) explained how to reproduce the HP-filter with the Kalman filter. This is done through the steps as follows: The measurement equation defines the observed variable as the sum of its trend and fluctuations around the trend:

$$x_t = x_t^* + u_t \qquad \text{with } u_t \sim N(0,\sigma_u^2) \tag{6}$$

The state equation defines the change of the growth rate of the trend component which follows a random walk:

$$x_t^* = x_{t-1}^* + g_{t-1} + v_{1,t} \tag{7}$$

$$g_t = g_{t-1} + v_{2,t} \tag{8}$$

5 There are a lot debates on the choice of the value λ in the context of business-cycle measuring. Baxter and King (1999), for example, argued that the smoothing parameter taking value of 10 for annual observations gives better results. Ravn and Uhlig (2002) took an analysis based on frequency domain consideration and show that λ=1600 for quarterly data is inconsistent with λ=100 for annual data, instead a much lower value, λ=100 is preferred. Pedersen (2001) also argued for a value of 1000 for quarterly data and for 3-5 for annual data.

with $v_{1,t} = 0$ and $v_{2,t} \sim N(0,\sigma_v^2)$. And the smoothing parameter, λ, is measured by the ratio σ_u^2 / σ_v^2. A state space model is thus given as:

$$x_t = \Gamma' Z_t + u_t \quad \text{and} \quad Z_t = \Pi Z_{t-1} + V_t \tag{9}$$

with $Z_t = \begin{bmatrix} x_t^* \\ g_t \end{bmatrix}$, $\Gamma = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\Pi = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $V_t = \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix}$, and the distribution of x_t is given by

$$x_t \sim N(\Gamma' \hat{Z}_{t|t-1},\ \Gamma' \Omega_{t|t-1} \Gamma + \sigma_u^2) \tag{10}$$

where $\hat{Z}_{t|t-1}$ is the linear least squares forecasts of the state vector on the basis of data observed through data t, and $\Omega_{t|t-1}$ is the associated mean squared error, represented by the following matrix:

$$\Omega_{t|t-1} = E\left[(Zt - \hat{Z}_{t|t-1})(Zt - \hat{Z}_{t|t-1})'\right] \tag{11}$$

A straightforward estimation procedure for the Kalman filter can be adopted to estimate the unknown variability of the decomposed components. In fact, the smoothing parameter, λ, can be shown to be inversely related to the weight given to current observation in the context of the Kalman filter forecast:

$$\lambda = \frac{1-k^*}{k^*} \beta \tag{12}$$

where k^* is the Kalman filter gain or the weight given to current observation relative to past forecasts, β is some constant value. (See Appendix for a derivation of the link λ to the Kalman filter gain k^*). Preliminary results show that the smoothing parameter varies a lot across different currencies. Typically, a range of values are suggested by the Kalman filter estimation. In this paper, we pick one value which is around the middle point of the interval (an integer, if applicable). We also set λ different values for comparative analysis.

2.2.3 The End-of-Sample Problem of the HP-Filter

Another crucial limitation associated with the HP-filter is the so-called end-of-sample problem, that is, the last a few observations may have disproportionate impacts on the trend at the end of the series. This demerit limits the practical usefulness of the filter, particularly in the context of economic forecast or policy implication. In terms of the Markov-switching model, the last observation of the sample used for estimation, is of particularly importance for out-of-sample forecast since it provides the latest information for updating the predicted conditional probability of being in state j, $\Pr(s_{T+h} = j \mid \Phi_T)$. The common way to reduce this end-point bias problem is to extend the raw series with ARIMA forecasts (see European Commission, 1995). The usefulness of this extension is limited, however, due to the quality of forecasts and uncertainty about how many forecasts are needed. Recently, Bruchez (2003) proposed a simple and natural modification of the HP-filter by assigning different weights of the smoothing parameter. His modified HP-filter is claimed to be robust in the sense that it does not use forecast. It is of computational attraction for us to adopt the similar treatment herein. In contrast to Bruchez (2003) who put more weights on the end-of-sample observations, however, we lower the weights on these observations to avoid the smoothed ones jumping away too far from the original values.

$$\min_{\{x_t^*\}_{t=1}^T} \left\{ \sum_{t=1}^{T} \frac{1}{\lambda_t}(x_t - x_t^*)^2 + \sum_{t=2}^{T-1} \left[(x_{t+1}^* - x_t^*) - (x_t^* - x_{t-1}^*) \right]^2 \right\} \tag{13}$$

with $\lambda_t = \lambda$ for t=3 to T-2; $\lambda_t = \frac{2}{3}\lambda$ for t=2 and T-1; $\lambda_t = \frac{1}{3}\lambda$ for t=1 and T.

3. Estimation

3.1 Data

The data consist of four quarterly spot exchange rates for the Australian Dollar (AUD), the Canadian Dollar (CAD), the British Pound (GBP), and the Japanese Yen (JPY), drawn from the Global Financial Data and the IMF's International Financial Statistics (IFS). All exchange rates are U.S. Dollar (USD) priced, i.e. the amount of USDs per unit of foreign currency. To be comparable in terms of unit measurement, the JPY is scaled by multiplying by 100. The sample contains 137 end-of-quarter observations from the first quarter of 1973 to the first quarter of 2007. These currencies are selected in consideration of data continuity and their importance in the foreign exchange rate markets.[6]

Fig. 1 presents a plot of these exchange rates over the sample period. Some striking features are worth noting. At first sight, most exchange rates indeed seem to be characterized by long swings as documented by Engel and Hamilton (1990), with the CAD and the GBP most remarkable. Whereas the Canadian dollar appeared to be more featured as up-and-down episodes throughout the sample period, the British pound had a "long swing" without apparent trends during 1988-96. Similar trendless periods can also be found in other currencies, as shown in shaded areas. Second, most exchange rates display long run trends. The AUD and the CAD, for instance, experienced a sustained downward movement until 2001 while the JPY moved upward through 1995, notwithstanding relatively shorter courses of opposite movements within these trends. Finally, comovements are found among exchange rates as well.

[6] Germany, France, and Italy are G7 countries and their currencies, historically, were amongst the most important ones in the foreign exchange rate markets while the Euro and the US dollar are now perhaps the best known pair in the world. But due to the cessation of the first three and the short history of the latter, we do not include these currencies in the current study.

Particularly, nearly all exchange rates had a sharp depreciation versus U.S. dollar during 1980-86 while 2001 was another threshold after which they started a sustained appreciation. Among them, the Japanese yen, however, is more peculiar in that it moved differently from the others in most episodes.

3.2 Estimation Procedure

Given that the smoothing parameter of the HP-filter is exogenously chosen, the set of parameters of interest can be summarized as $\theta = \left(\mu_s, \sigma_s^2, p_{ij}\right)$ with s , i , j =1, 2,…, k such that $\sum_j p_{ij} = 1$. Denote $Y_t = (y_1, y_2, ..., y_t)$ a history of observations up to time *t*,[7] and $S_t = (s_1, s_2, ..., s_t)$ historical realizations of state variables up to time *t*. Following the specification described in section 2, it is straightforward to show that the joint probability distribution of the observed data $(y_1, y_2, ..., y_T)$ along with the unobserved states $(s_1, s_2, ..., s_T)$ is given as

$$f(y_1, y_2, ..., y_T, s_1, s_2, ..., s_T; \theta) = \prod_{t=1}^{T} f(y_t \mid s_t; \theta) \cdot p(s_t \mid s_{t-1}; \theta) \tag{14}$$

where

$$f(y_t \mid s_t = j; \theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(y_t - \mu_j)^2}{2\sigma_j^2}\right) \qquad j = 1, 2, ..., k \tag{15}$$

$$p(s_t = j \mid s_{t-1} = i; \theta) = p_{ij} \qquad i, j = 1, 2, ..., k \tag{16}$$

And the log-likelihood function of the observed data can be obtained by summing over all possible values of $(s_1, s_2, ..., s_T)$, a procedure analogous to marginalizing Y_T :

$$\log f(Y_T; \theta) = \log\left(\sum_{s_1=1}^{k} \sum_{s2=1}^{k} ... \sum_{s_T=1}^{k} f(Y_T, S_T; \theta)\right) \tag{17}$$

In practice, construction and numerical maximization of the sample log-likelihood function in this way is computationally intractable, as $(s_1, s_2, ..., s_T)$ may be realized in k^T ways. To this end, a version of the Expectation-Maximization (EM) algorithm proposed by Hamilton (1990) is typically employed to obtain the maximum likelihood estimation. The EM algorithm works well generally in obtaining consistent parameter estimates. However, as Hamilton (1994) pointed out, local maxima may pose a major problem in using EM algorithm to maximize the log likelihood function in (17). In this regard, a

[7] Here we use the unfiltered series $(y_1, y_2, ..., y_t)$ to describe the estimation procedure. The same procedure is used in estimating the filtered model.

grid of stating values are tried and the maximum likelihood estimates are picked at the highest value of the objective function.

3.3 Parameter Estimates

This section reports the parameter estimates based on full-sample data from 1973:Q1 to 2007:Q1 for each currency. As Table 1 shows, the maximum likelihood estimates associated with the 2-state model based on unfiltered log exchange rates predicts a downward trend about 1.6 percent quarterly in the Australian dollar, 0.4 percent in the Canadian dollar, 0.6 percent in the British pound, and 2.6 percent in the Japanese yen while a upward trend about 0.5 percent, 1.2 percent, 0.5 percent, and 2.2 percent quarterly in these corresponding exchange rates. The asymmetry in mean depreciation and appreciation in the Canadian dollar roughly reflects the shape of its plot. In terms of the British pound, however, the mean changes are substantially small in magnitude relative to the historical finding documented by Engel and Hamilton (1990) and Engel (1994). Based on a sample of 1973:Q3 to 1988:Q1, Engel and Hamilton found a 3.8 percent quarterly fall and a 2.7 percent quarterly rise in the pound. This shrinkage can be explained by the fact that since 1988 the pound became more volatile with weak long swings, if there is any. In this regard, more states may be required in order to correctly specify the model for the pound. The model with HP-filter moderately scales up the magnitude of means both for upward and downward trends in these exchange rates. One plausible explanation is that smoothing techniques have filtered out trivial shifts while left relatively sizable shifts in accounting for mean change. Applying the 3-state Markov-switching model to the same exchange rates by imposing a state of mean zero, not surprisingly, magnifies both the upward and downward trends. This is because a considerable amount of trendless shifts, originally classified as uptrend or downtrend, are now excluded in calculating mean change of uptrend or downtrend.

Table 1 also shows that the Canadian dollar and the British pound seem to be well characterized by long swings with sustained appreciation and depreciation regimes. This high persistence of regimes can be represented by the large state-staying probabilities, p_{11} and p_{22}, that is, the probabilities of staying in a state once the process enters it. The large staying probabilities in the unfiltered 2-state model, all exceeding 0.98, imply infrequent switches between states. The expected duration of state j is defined as $1/(1-p_{jj})$. According to this, each state is expected to persist for over 50 quarters or about 12 years on average. This extremely long persistence may be an appropriate depiction of the Canadian dollar's drawn-out depreciation during periods of 1975-1988 and 1992-2003, but apparently contradicts our casual inspection of Fig. 1. As we can see, the appreciation regimes of the Canadian dollar lasted much shorter while the British pound had roughly five years of duration for each state before 1988 and was much more variable afterwards with no distinct upward/downward trends until recent years. This misidentification can be seen being corrected by the model with smoothing. One merit of the filtered model as discussed in section 2 is that it enables the estimation procedure to capture more precisely the signals of genuine regime shifts but to be immune to distracting messages. For example, the estimates from the 2-state HPMS for the Canadian dollar predict a duration of 16 quarters for appreciation while maintains the same prediction on depreciation regime persistence, 50 quarters, as the unfiltered Markov-switching model, which also matches our visual inspection from the plot. In case of the Australian dollar and the Japanese yen, no long swings are predicted by the unfiltered model, according to the state-staying probabilities, which range from 0 to 0.17 for the former and from 0.519 for to 0.757 for the latter. The model with the HP-filter, however, effectively revives these features.

3.4 Regime Classification and Estimation Stability

Fig. 2 presents estimated smoothed probabilities for these currencies. Smoothed probabilities, $\Pr(s_t = j \mid Y_T)$, which are calculated recursively based on the complete sample, generally provide the most informative inference about the state in which the data generating process lies at a particular time. As we can see, in a sharp contrast to the finding by Engel and Hamilton (1990), the two-state Markov-switching specification is no longer capable of capturing the long swings in these exchange rates. For example, it clearly shows that the unfiltered two-state model has totally failed to capture the upward persistence in the Canadian dollar from 1987 to 1992, and missed at least two appreciation episodes in the British pound during 1977-81 and 1985-89. The worst case is in the Australian dollar where it displays messily frequent shifts between two regimes, a pattern in no sense to be viewed as long swings. Introducing an additional state to the Markov-switching model helps slightly in capturing the trend persistence in these exchange rates. Concern about misclassification of regime switches, however, remains. Shifts like the one circled in ellipse in the British pound are much likely to be caused by temporary shocks rather than genuine regime switches. In this vein, both the two-state and three-state Markov-switching models with the HP-filter work much better in classifying trends with different regimes while the latter further precisely distinguishes a trendless regime from uptrends and downtrends.

Fig. 3 examines the stability of smoothed probabilities using different subsamples. It shows clearly that the unfiltered two-state Markov-switching model suffers estimation instability while the filtered model turns out to be temporally consistent. As we can see in the left panel for each currency, the estimated smoothed probabilities from the standard two-state Markov-switching model vary a lot as sample span changes. Taking the British pound as an example, if using the sample composed of data during 1973-95, the unfiltered model produces fairly volatile smoothed probabilities, but when extending the sample up to 1999 or later, the estimated probabilities become "smoothed" but are seriously misclassified as discussed above. Turning to the right panel, one can see that the estimated probabilities from the HP-filtered Markov-switching model are generally unaffected by the change of sample spans, correctly identifying relevant trends.

4. Forecast

The main focus of this exercise is the forecast performance of the modified multi-state Markov-switching model under consideration. It has been well known that forecasting with nonlinear models like Markov-switching is in general much more difficult than forecasting with linear models since the distribution of future shocks on which forecasting is to be grounded may have an arbitrary conditional expectation in a nonlinear scenario rather than a zero mean as in a linear environment.[8] In the same spirit, Clements and Smith (1999) argued that it may not be always possible to exploit nonlinearities to improve forecasts over linear models, even when such nonlinearities are a feature of the data. In regard to Markov-switching models, forecast performance depends on the regime in which the forecast was made. That means misspecification of the type of nonlinearity (regime shifts) can lead to substantial losses in forecast accuracy. One question is thus of particular interest that, given that the filtered model seems to work well in capturing the trend persistence in the dynamics of exchange rates as shown in the preceding section, whether it can outperform some linear alternatives, especially the simple random walk.

In general, the derivation of an optimal predictor is a quite complicated issue in empirical work for nonlinear time series models in that numerical optimization methods like Monte Carlo or bootstrapping are often employed to approximate conditional expectation. There is a great merit associated with Markov-switching models, however, in which an immediate form of the optimal predictor is conveniently available. Following Hamilton (1994), denote $\hat{\xi}_{t|t}$ as a $k \times 1$ vector of conditional probabilities, $P(s_t = j \mid Y_t; \theta)$, for $j = 1, 2, ..., k$, which are inferences about the state at time t. Given the maximum likelihood estimates, θ, the h-period-ahead forecast of y_{t+h}, on the basis of observation of y through time t is given by

$$\hat{y}_{t+h|t} = E\left[y_{t+h} \mid Y_t; \hat{\theta} \right] = \hat{\xi}'_{t|t} \cdot P^h \cdot \hat{\mu} \tag{18}$$

where $\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_k)'$ is the vector of estimates state-dependent mean trends.[9]

---

[8] See Grandger and Terasvirta (1993) and Franses and van Dijk (2000).

[9] Note that given a series of correctly identified smoothed probabilities, $\Pr(s_t = j \mid Y_T)$, in the filtered model, $\hat{\mu}_j$ and $\hat{\sigma}_j^2$ are updated based on the formulae: $\hat{\mu}_j = \sum_t y_t \Pr(s_t = j \mid Y_T) / \sum_t \Pr(s_t = i \mid Y_T)$ and $\hat{\sigma}_j^2 = \sum_t (y_t - \hat{\mu}_j)^2 \Pr(s_t = j \mid Y_T) / \sum_t \Pr(s_t = i \mid Y_T)$, where y_t is unfiltered log exchange rate change.

Given (18), the h-period-ahead forecasts of the level of logarithm of the exchange rate can be calculated as

$$\hat{e}_{t+h|t} = e_t + \sum_{j=1}^{h} \hat{y}_{t+j|t} \tag{19}$$

Finally, the mean squared error (MSE) for forecasting accuracy is given as

$$\frac{1}{T-\tau_0-h+1}\sum_{t=\tau_0}^{T-h}(\hat{e}_{t+h|t} - e_{t+h})^2 \tag{20}$$

where τ_0 is the size of subsample used to estimate parameters.

The standard for measuring forecastability in the context of exchange rates is whether the proposed model can do well in forecasting relative to a random walk. This standard has been widely adopted since Meese and Rogoff's (1983a, b). Despite the consensus on the random walk as the comparison benchmark, there are disputes on whether one should include a drift term with the random walk or not. Although the zero-drift random walk specification is not lacking of proponents, e.g. Meese and Rogoff (1983), Diebold and Nason (1990), and recently Cheung, Chinn, and Pascual (2005), Engel and Hamilton (1990) argued that the random walk with drift is a more reasonable standard of comparison in that the inconsistency in forecasting performance between the driftless random walk and the one with drift is *de facto* in favor of the regime-switching model rather than the random walk, with or without drift. Instead of setting exogenously the drift term to zero, the specification of the random walk with drift allows for the flexibility of updating estimates for the drift term in the context of rolling estimation. In this regard, the random walk with a drift term is used in this paper.[10]

4.1 Forecast Performance

To have clearer vision on the forecast performance, we first apply the multi-state Markov-switching models, both filtered and unfiltered, to Engel and Hamilton's (1990) dataset, which contains quarterly percentage change in dollar exchange rates of the Deutsche mark, the British pound, and the French franc from 1973:Q4 to 1988:Q1.[11] One may note that this dataset is slightly different from the one we are using in that they collected the raw data (level of exchange rates) by taking arithmetic average of the bid and asked prices for the exchange rate for the last day of quarter instead of simply closed prices as in our dataset.

[10] This may be somewhat justified according to the description in section 2 in which almost all exchange rates exhibit apparent upward/downward trends over the whole sample periods.

[11] Downloadable from http://weber.ucsd.edu/pub/jhamilton/markov2.zip.

Table 2 presents the in-sample and out-of-sample mean squared errors of the forecasts. The first two rows of each panel essentially reproduce Engel and Hamilton's results which show that the two-state Markov-switching model achieves superior forecasts relative to the simple random walk.[12] Comparing to their results, remarkably, one can see that the filtered model further enhances the forecast accuracy both in-sample and out-of-sample across all currencies. The average improvement in out-of-sample forecast accuracy from the 3-state filtered Markov-switching model is about 23 percent for the Deutsche mark, 25 percent for the French franc, and 18 percent for the British pound, averaging over forecast horizons up to four quarters, while the corresponding forecast accuracy improvement from the standard two-state model is 9 percent, 11 percent, and 13 percent, respectively. Interestingly, introducing one additional state to model exchange rates generally improves in-sample forecast precision, but fails to escalate forecast reliability out-of-sample in all three exchange rates. This is exactly the case that a good in-sample fit does not necessarily lead to a better out-of-sample forecast, as many other studies documented. One may further notice that, both 2- and 3-state filtered Markov-switching models well outperform the simple random walk using Engel and Hamilton's dataset with the latter slightly more prominent.

The same mechanism is then applied to the new exchange rate sample and the forecast performance is reported in Table 3. The upper panel provides in-sample forecast comparison. A general impression is that the standard Markov-switching model shows virtually no difference from the random walk in terms of the mean squared errors across all currencies with the exception of the Canadian dollar. In a sharp contrast, the model with smoothing works much better than its unfiltered counterpart and can significantly and consistently outforecast the random walk. In particular, the average improvement upon forecast accuracy obtained from the 3-state filtered model is about 8 percent up to four quarters forward forecasts for the Australian dollar, 29 percent for the Canadian dollar, 20 percent for the British pound, and 18 percent for the Japanese yen, respectively. Same as the preceding finding based on Engel and Hamilton's dataset, introducing more states does help improve in-sample forecast performance even when using the new dataset, although the improvement is not as striking as in the previous one.

The out-of-sample forecast performance is of more interest which is shown in the lower panel in Table 3. Similar with the in-sample forecast, the unfiltered Markov-switching model fails to offer convincing evidence of outperforming the random walk. In fact, it works worse than the simple "no change" model across all exchange rates except the British pound. Specifically, the 2-state unfiltered model produces about -2 percent for the Australian dollar, -7 percent for the Canadian dollar, and -5 percent for the Japanese yen, respectively, relative to the random walk in terms of forecast error reduction.

[12] Comparing to EH's original results, the reproduced mean squared errors for the random walk with drift are completely matching, while for MS model they are slightly lower for DEM and FRF and a little higher for GBP due to the estimation error.

Encouraging results are delivered by the filtered model. One can see that the 3-state filtered model is robust in beating the random walk across all currencies at different forecast horizons. Particularly, it achieves average forecast accuracy improvement ranging from a trivial 0.1 percent to a mild 6 percent at the one-quarter horizon, and from a slight 4 percent to a significant 22 percent at the four-quarter horizon. Comparing to the model imposing two regimes, the 3-state model without smoothing obtains no benefit regarding enhancing forecast performance for the Australian dollar, the Canadian dollar, and the British pound, but helps outforecast the random walk for the Japanese yen.

4.2 Forecast Evaluation (The Diebold-Mariano Test)

To further evaluate the forecast accuracy of the standard or tailored Markov-switching model relative to that of the random walk, a test of equal forecast accuracy developed by Diebold-Mariano (1995) is implemented. Let $\hat{e}^{ms}_{t+h|t}$ denote h-step-ahead forecasts for e_{t+h} from Markov models, and $\hat{e}^{rw}_{t+h|t}$ be corresponding forecasts from the random walk. The squared error loss functions are defined as

$$L\ (\hat{e}^{ms}_{t+h|t}) = (\hat{e}^{ms}_{t+h|t} - e_{t+h})^2 \text{ and } L\ (\hat{e}^{rw}_{t+h|t}) = (\hat{e}^{rw}_{t+h|t} - e_{t+h})^2 \qquad (21)$$

And the loss differential is defined as

$$d_t = L\ (\hat{e}^{ms}_{t+h|t}) - L\ (\hat{e}^{rw}_{t+h|t}) \qquad (22)$$

Under the null hypothesis that there is no difference between forecast accuracy of the pair of models being compared, it is equivalent to testing whether the population mean of the loss differentials is zero, that is, $H_0: \ E(d_t) = 0$. The Diebold-Mariano test (DM test, hereafter) is given by

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}} \qquad (23)$$

where $\bar{d} = \frac{1}{T_0}\sum_{t=1}^{T_0} d_t$ is the sample mean of d_t with a total of T_0 forecasts. $\hat{V}(\bar{d})$ is the heteroskedasticity and autocorrelation-consistent (HAC) estimator of the asymptotic variance obtained using the method of Newey and West (1987). The HAC estimator is employed to account for the serial correlation in d_t due to the overlapping data used in calculating h-step-ahead forecasts. The lag-window specification for the HAC estimator of asymptotic variance is based on a Bartlett kernel and a lag-selection procedure by Ng and Perron (1995).

Under the assumptions of covariance stationarity and short-memory for d_t, Diebold and Mariano (1995) show that the DM statistic is distributed under the null hypothesis as standard normal $N(0,1)$. The standard DM test, however, is known to tend to over-reject the null hypothesis in the context of finite samples. Hence, a modified DM test proposed by Harvey et al (1997) is applied here, which is adjusted in a way that takes into account the forecast horizons explicitly:

$$DM^{*} = \left(\frac{T_0 + (1-2h) + h(h-1)/T_0}{T_0} \right)^{1/2} \cdot DM \tag{24}$$

Table 4 reports results of DM test which largely reinforce our findings reported in Table 3. Particularly, the null hypothesis that the 2-state Markov-switching model has the same forecastability as the simple random walk is strongly rejected at 5% significance level. Given values of the MSE ratio higher than one, this implies that the standard model is significantly poorer than the random walk in the context of out-of-sample forecasting. In contrast, the 3-state filtered model is generally significantly better than the random walk according to the DM test.

5. Robustness Checks

5.1 Sample Spans

As Marsh (2000) argued, Markov-switching models for exchange rates may be unstable over time and unsuitable for forecasting. It is of particular interest to investigate whether the forecast performance of the filtered Markov-switching model is temporally consistent. To this end, we recast the exercise using two different sample spans: (1) estimation periods 1973:Q1-1985:Q4, forecast periods 1986:Q1-2007:Q1 and, (2) estimation periods 1973:Q1-2000:Q4, forecast periods 2001:Q1-2007:Q1.

The upper panel of Table 5 presents new forecast results for the Canadian dollar and the British pound.[13] In general, the results are fairly consistent with the prior ones. For example, no forecast superiority of the conventional 2-state Markov-switching model is found when forecasting the Canadian dollar but the 3-state filtered model consistently exhibits strong forecasting power for all currencies across different sample spans. As discussed in section 3, the unfiltered Markov-switching model suffers regime misclassification and estimation instability, which naturally leads to inconsistent forecast performance as sample periods differ. We can see that the conventional 2-state model fails to beat the random walk in forecasting the British pound during the periods of 1986:Q1-2007:Q1, but works well regarding more recent forecast spans, namely 1996:Q1-2007:Q1 and 2001:Q1-2007:Q1. A similar phenomenon emerges when using 3-state unfiltered model to forecast the Canadian dollar. These results are to some extent in line with the finding of Marsh (2000), in that the performance of a particular specification may not be

[13] To conserve space and for convenience, we skip the reports for the other two exchange rates. These results are available upon request.

consistent over different forecasting scenarios. Our model, nevertheless, displays fairly strong robustness in beating the simple random walk.

5.2 Smoothing Effect

Up to now, the smoothing parameter of the HP-filter in our specifications is set exclusively based on the value suggested by the Kalman filter estimation in an unobserved components framework. Specifically, we are using λ=0.3, 5, 300, and 25 for the Australian dollar, the Canadian dollar, the British pound, and the Japanese yen, respectively. To what extend these picks can fulfill our goal of smoothing these price series remains unclear. As discussed above, a rigorous mathematical derivation of the optimal choice of the parameter may be practically impossible due to the indefiniteness of "smoothness". In this regard, we look into the smoothing effect numerically by setting a quite different value against the one suggested for each exchange rate.

The lower panel of Table 5 presents the forecast performance due to smoothing effect. In general, the out-of-sample forecasts are sensitive to differing values of the smoothing parameter. For example, an extremely small or a very large smoothing value makes the filtered Markov-switching model no longer beat the random walk in the case of the Canadian dollar and the Japanese yen. In contrast, the same specification works well for the British pound even using an extremely large smoothing value, say λ=1500. The case of the Australian dollar is also quite different from the other currencies. Since the initially suggested smoothing value for the Australian dollar is rather small, λ=0.3, when setting λ=0.03, the magnitude of this change is trivial and thus no substantial change in forecast performance is entailed. But when setting λ=30, the forecast performance does turn out to be much poorer. In summary, using smoothing values deviating far away from the suggested ones apparently undermines the forecastability of the 3-state filtered model across all currencies. This suggests that a balance between removing irregular components and maintaining state-dependent features in the exchange rates is crucial for achieving forecast superiority.

6. Are Fundamentals Relevant?

In this section, we explore the relationship between trends in exchange rates and trend information in the fundamentals-index which is a linear combination of underlying macroeconomic variables commonly used in monetary exchange rate models. The fundamentals-index can be constructed through an OLS estimation of a first difference specification:[14]

[14] Note that, exchange rates and macroeconomic variables are jointly determined, and it is preferable to apply instrumental variables. Meanwhile, OLS is justified according to Chinn and Meese (1995) in that the gains in consistency are far outweighed by the loss in efficiency, in terms of prediction.

$$\Delta e_t = \Delta X_t \cdot \Pi + u_t \tag{25}$$

where e_t is given as follows:

$$e_t = \beta_0 + \beta_1(m_t - m_t^*) + \beta_2(q_t - q_t^*) + \beta_3(i_t - i_t^*) + \beta_4(\pi_t - \pi_t^*) + v_t \tag{26}$$

where e_t, m_t, q_t, i_t, and π_t are the logarithms at time t of the exchange rate, domestic (U.S.) money supply, output, interest rate, and inflation rate, respectively. Asterisks denote foreign variables. X_t is a vector of relative fundamental variables under consideration. Note that a direct OLS estimation procedure on Eq. (26) may involve spurious regression since these variables are generally nonstationary. Then the fundamentals-index is constructed on the basis of Eq. (26) given the estimated parameters.

The upper left panel of Fig. 4 depicts the fundamentals-index for the Canadian dollar versus the actual log exchange rate data. As one can see that the fundamentals-index follows different movements from the Canadian dollar since the late 1970s, in spite of there being an apparent comovement in the early floating periods. This belief is further reinforced by the upper right panel of Fig. 4, which presents the predicted trends in exchange rates and the counterparts in the fundamental index. Trends are exacted by the modified Markov-switching model and calculated by

$$\hat{y}_t = \Pr(s_t \mid y_1, y_2, ..., y_T)' \cdot \hat{\mu} \tag{27}$$

where $\Pr(s_t \mid y_1, y_2, ..., y_T)$ is the vector of smoothed probabilities of being some particular state for time t, $\hat{\mu} = (\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_k)'$ is the vector of estimates of mean trends. In this figure, a positive value implies that the Canadian dollar moves upward, a negative value is associated with downtrend, while a value around zero implies the currency stagnates at some level without apparent uptrend or downtrend. As one can see, the Canadian dollar generally has different trends from those in the fundamentals-index, except a very short period of overlap during the late 1970s.

The lower left panel of Fig. 4 depicts the correlation coefficients calculated through a rolling window with the first sample window of 1973-83. These correlation coefficients are generally insignificant and are virtually close to zero especially for recent data. The large p-values in the correlation tests have the same conclusion that when considering the entire sample of exchange rates during 1973-2007, these two types of trends exhibit no significant linear linkage. Similarly, in the contingency table, among the total 135 observation periods, there are 30 periods of upward movements, 58 periods of trendless movements, and 47 periods of downward movements in exchange rates, respectively. Given these, there are nonetheless no upward trends in the fundamentals-index corresponding to the uptrends in exchange rates and only 9 out of 47 periods are correctly matched in terms of downward movements in these trends.

This is showing that the trends in exchange rates are generally not related in a linear way to the trends in the fundamentals-index. It thus suggests that, failing to capture the trend information in the exchange rates can help explain why the monetary models are incapable of beating the random walk at short horizons.

7. Conclusion

This paper presents an empirical analysis on forecasting major dollar exchange rates. We show that combining a standard multi-state Markov-switching model with the popular filtering technique significantly enhances both in-sample and out-of-sample forecast performance, and beats a random walk both across currencies and subsamples. The research is inspired by the realization that when using more recent data the standard Markov-switching model popularized by Hamilton (1989) could no longer achieve the same success in escalating forecast accuracy as previously reported in Engel and Hamilton (1990). In addition, preliminary results obtained through the application of the conventional model to the extended time series further increase our awareness that the existence of highly irregular components of the data tends to distort the estimation procedure of the Markov-switching model and thus undermines its forecasting power. Our specification largely eliminates the modeling nuisance and is able to revive the forecast superiority of the Markov-switching model.

Our finding has demonstrated that correctly identifying the persistence in exchange rates plays a key role in achieving superior forecastability to the simple random walk. An attempt to link the trends in exchange rates to those of a fundamentals-index commonly used in the literature reveals that macroeconomic exchange rate models generally fail to capture the persistence in exchange rates in a linear way. This, to some extent, helps explain why fundamentals-based models are incapable of outperforming the random walk. In this spirit, our empirical work corroborates the finding by Kilian and Taylor (2001) that nonlinear dynamics of exchange rates make it difficult for linear structural models to forecast future exchange rate variations.

In the ending of his systematic work, Engel (1994, p.164) suggested that "perhaps the Markov model will perform better in the future, allowing for a third state". The present paper has given a response to this conjecture. A simple extension of the two-state model by including a third trendless regime does improve slightly in-sample forecasts but shows no out-of-sample forecasting superiority in general, as it is now clear that, without removing the irregular components from the noisy data, the multi-state Markov-switching model remains suffering from estimation distortion and regime misclassification. Hence, employing time-series filtering techniques is the more important element in escalating the Markov model’s forecastability.

Several issues nevertheless remain unexplored here. It is of particular interest to understand, for example, why the suggested value of the smoothing parameter for the HP-filter differs so much across different nominal rates. One plausible explanation is that potential structural breaks due to rare events, such as the speculative attack to the British pound in 1992 and the Plaza Accord for the Japanese yen in 1985, may amplify the noise-to-signal ratio. This is partially verified as the suggested value for the British pound is much smaller based on Engel and Hamilton's (1990) dataset than that of extended sample. One may suggest endogeneizing the smoothing parameter in the Markov-switching model instead of setting it exogenously. Our practice, however, indicates that this treatment may not effectively circumvent estimation distortion on the one hand but will surely increase much computational complexity on the other hand. As such, how to find the optimal smoothing parameter remains an open question. Another issue we did not address here is the irregular components removed by the filter. Simply ignoring these extreme parts of the data is vapid from the perspective of modeling. Indeed, they may, notwithstanding the disadvantage to the Markov-switching model, contain some valuable information for forecasting future changes of exchange rates. In this vein, separately modeling the trends and irregular components could possibly offer another promising way to the success of forecasting exchange rates. Finally, as it has been clear that trends in exchange rates provide useful information for explaining the future variation of exchange rates, but the underlying sources driving these trends are still not well understood. Possible explanations may include irrationality of market participants, bubbles, herd behaviors etc. But no universally satisfying answers are found to date. Therefore, it would be economically intriguing in the future research agenda to tackle these issues.

Appendix

The Link of the HP-filter Smoothing Parameter and the Kalman Gain

A general state-space system is given as follows:

$$\xi_{t+1} = f_t \cdot \xi_t + v_{t+1} \quad \text{with } v_t \sim N(0, \sigma_v^2) \tag{28}$$

$$x_t = \xi_t + u_t \qquad \text{with } u_t \sim N(0, \sigma_u^2) \tag{29}$$

The Kalman filter based on Eq. (28) and Eq. 29) is given

$$\hat{\xi}_{t+1} = \hat{\xi}_t \cdot f_t (1 - k_t) + k_t x_t \tag{30}$$

where k_t is the gain of Kalman filter or the weight given to current observation relative to past belief $\hat{\xi}_t$. At steady state, the Kalman gain can be expressed as

$$k^* = \frac{\eta + \sigma_v^2}{\eta + \sigma_u^2 + \sigma_v^2} = \frac{\beta}{\beta + \lambda} \tag{31}$$

where η is the variance of $\hat{\xi}_{t+1}$, $\beta = (\eta + \sigma_v^2) / \sigma_v^2$, and $\lambda = \sigma_u^2 / \sigma_v^2$ is the noise-to-signal ratio. It is easy to get

$$\lambda = \frac{1-k^*}{k^*}\beta \tag{32}$$

That is, the noise-to-signal ratio is the inverse of the weight given to current observation. As k^* goes to zero, that is, forecast pays no attention on current observation while solely relies on the prior belief, the noise-to-signal ratio goes to infinity. In this case, forecast returns a linear trend. In contrast, as k^* goes to unity, forecast solely depends on the current observation which corresponds to a random walk, and the noise-to-signal ratio goes to zero.

Given the state-space representation of the HP-filter in Eqs. (6) - (8), we have the following lemma:

Lemma *Assume the change in trend component, g_t, is proportional to the trend component, x_t^*, the smoothing parameter of the HP-filter is equivalent to the noise-to-signal ratio in Kalman filter, given by Eq. (32).*

Proof Define $g_t = \tilde{f}_t x_t^*$ and $\tilde{f}_t = f_t - 1$, then Eq. (7) can be written as:

$$x_t^* = f_t x_{t-1}^* + v_t \tag{33}$$

which corresponds to Eq. (28). Combine Eq. (33) and Eq. (8), we have shown above, the smoothing parameter, $\lambda = \sigma_u^2 / \sigma_v^2$, take the form of Eq. (32) at the steady state.■

If we further assume the trend component at initial period, x_0^*, is Gaussian, then the distribution of x_t conditional on the history of observations $X_{t-1} = (x_1, x_2, \ldots x_{t-1})$ is Gaussian and is given by

$$x_t \mid X_{t-1} \sim N(\hat{x}_{t|t-1}^*,\ \Omega_{t|t-1} + \sigma_v^2) \tag{34}$$

where $\hat{x}_{t|t-1}^*$ is the forecasts of the state vector on the basis of data observed through data t-1, and $\Omega_{t|t-1}$ is the associated variance of $\hat{x}_{t|t-1}^*$. Given (34), the sample log likelihood is given

$$\sum_{t=1}^{T} \log f_{x_t|X_{t-1}}(x_t \mid X_{t-1}) \tag{35}$$

Expression (35) can then be maximized numerically with respect to unknown parameters $\theta = (f,\ \sigma_u^2,\ \sigma_v^2)$. And thus estimate for the smoothing parameter is given by $\lambda = \sigma_u^2 / \sigma_v^2$.

Reference

Abel, A.B., 1994. Exact solutions for expected rates of return under Markov regime switching: implications for the equity equilibrium puzzle. Journal of Money, Credit, and Banking, Vol. 26, No. 3, 345-361.

Bauwens, L., Preminger, A., Rombouts, J. V. K., 2006. Regime switching GARCH models. Discussion paper, 2006/11, CORE and Department of Economics, University of Catholique de Louvain.

Baxter, M., King, R.G., 1999. Measuring business cycles: approximate band-pass filter for economic time series. Review of Economics and Statistics, Vol. 81, 575-593.

Birchenhall, C.R., Jessen, H., Osborn, D.R., Simpson, P., 1999. Predicting U.S. business cycle regime. Journal of Business and Economic Statistics, 17, 313-323.

Bollen, N.P.B., Gray, S.F., Whaley, R.E., 2000. Regime-switching in foreign exchange rates: evidence from currency option prices. Journal of Econometrics, Volume 94, Issues 1-2, 239-276.

Bruchez, P.A., 2003. A modification of the HP Filter aiming at reducing the end-point bias. Swiss Federal Finance Administration Working Paper.

Cecchetti, S.G., Lam, P.S., Mark, N.C., 1990. Mean reversion in equilibrium asset price. American Economic Review, 80, 398-418.

Cheung, Y., Erlandsson, U.G., 2005. Exchange rates and Markov switching dynamics. Journal of Business & Economic Statistics, Vol. 23, No. 3, 314-320.

Cheung, Y., Chinn, M.D., Pascual, A.G., 2005. Empirical exchange rate models of the nineties: are they fit to survive? Journal of International Money and Finance, 24, 1150-1175.

Chinn, M.D., Meese, R.A., 1995. Banking on currency forecasts: how predictable is change in money? Journal of International Economics, 38, 161-178.

Clements, M.P., Smith, J., 1999. A Monte Carlo study of the forecast performance of empirical SETAR models. Journal of Applied Econometrics, 14, 123-142.

Christiano, L.J., Fitgerald, T.J., 2003. The band pass Filter. International Economic Review, 44(2), 435-466.

Dacco, R., Satchel, S., 1999. Why do regime-switching models forecast so badly? Journal of Forecasting, 18, 1-16.

Diebold, F.X., Nason, J.A., 1990. Nonparametric exchange rate prediction. Journal of International Economics, 28, 315-322.

Diebold F.X., Mariano R., 1995. Comparing predictive accuracy. Journal of Business and Economic Statistics, 13, 253-265.

Engel, C., Hamilton, J.D., 1990. Long swings in the dollar: are they in the data and do markets know it?" American Economic Review, Vol.80, No.4, 689-713.

Engel, C., 1994. Can the Markov-switching model forecast exchange rates? Journal of International Economics, 36, 151-165.

European Commission, 1995. The commission services method for cyclical adjustment of government budget balances. DG II/401/95-EN.

Filardo, A.J., 1994. Business-cycle phases and their transitional dynamics. Journal of Business & Economic Statistics, Vol. 12, No.3, 299-308.

Franses, P. H., van Dijk, D., 2000. Non-linear Time Series Models in Empirical Finance. Cambridge University Press: Cambridge, UK.

Granger, C.W.J., Teräsvirta, T., 1993. Modeling Non-Linear Economic Relationships. Oxford University Press, Oxford.

Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, March 1989, 57, 357-84.

Hamilton, J.D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics, 45, 39-70.

Hamilton, J.D., 1994. Time Series Analysis. Princeton, NJ: Princeton University Press.

Harvey, A.C., 1985. Trends and cycles in macroeconomic time series. Journal of Business and Economic Statistics, 3, 216-227.

Harvey, D., Leybourne, S., and Newbold, P., 1997. Testing the equality of prediction mean squared errors. International Journal of Forecasting,13, 281-291.

Hodrick, R.J., Prescott, E.C., 1980, 1997. Postwar U.S. Business Cycles: an Empirical Investigation. Journal of Money, Credit and Banking, Vol. 29, 1-16.

Killian, L., Taylor, M.P., 2001. Why is it so difficult to beat the random walk forecast of exchange rates? European Central Bank Working paper no. 88.

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

Kim, C., Nelson, C.R., 1999. State-Space Models with Regime Switching, MIT Press, Cambridge, MA

MacDonald, R., Marsh, I., 1997. On fundamentals and exchange rates: a Casselian perspective. Review of Economics and Statistics, 79, 655-664.

Marsh, I.W., 2000. High-frequency Markov switching models in the foreign exchange market. Journal of Forecasting, 123-134.

Meese, R., Rogoff, K., 1983a. Empirical exchange rate models of the seventies: do they fit out of sample? Journal of International Economics, 14, 3-24.

Meese, R., Rogoff, K., 1983b. The out-of-Sample failure of empirical exchange rate models: sampling error or misspecification? in: Jacob Frenkel, ed. Exchange Rates and International Macroeconomics, Chicago: University of Chicago Press, 1983.

Meese, R., Rose, A.K., 1991. An empirical assessment of non-linearities in models of exchange rate determination. Review of Economic Studies 58, 603-619.

Neely, C.J., Sarno, L., 2002. How well do monetary fundamentals forecast exchange rates? Federal Reserve Bank of St. Louis Review 84.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703-708.

Ng, S., Perron, P., 1995. Unit root test in ARMA models with data dependent methods for the selection of the truncation lag. Journal of the American Statistical Association 90: 268-281.

Pedersen, T.M., 2001. The Hodrick-Prescott filter, the Slutzky effect, and the Distortionary effect of filters. Journal of Economic Dynamics and Control, 25, 1081-1101.

Ravn, M.O., Uhlig, H., 2002. On adjusting the HP-filter for the frequency of observations. Review of Economics and Statistics, Vol. 84(2), 371-376.

Simpson, P.W., Osborn, D.R., Sensier, M., 2001. Forecasting UK industrial production over the business cycle. Journal of Forecasting, Vol. 20, 6, 405-424.

Woff, C.P., 1988. Exchange rates, innovations and forecasting. Journal of International Money and Finance, 7(1), 49-61.

Table 1

Maximum Likelihood Estimates of Parameters (Sample 1973:Q1--2007:Q1)

Two-State Model

Parameters	A: Australian Dollar		B: Canadian Dollar		C: British Pound		D: Japanese Yen	
	MS	HPMS	MS	HPMS	MS	HPMS	MS	HPMS
μ_1	-1.631	-2.978	-0.386	-0.612	-0.634	-2.142	-2.591	-1.549
μ_2	0.473	2.196	1.161	1.609	0.475	1.445	2.215	3.349
σ_1	6.355	4.676	2.140	2.134	6.019	5.211	2.607	5.336
σ_2	3.184	3.543	3.576	2.988	3.227	4.329	6.666	5.856
p_{11}	0.000	0.828	0.991	0.980	0.986	0.934	0.519	0.933
p_{22}	0.269	0.824	0.996	0.931	0.996	0.959	0.757	0.915

Three-State Model

Parameters	A: Australian Dollar		B: Canadian Dollar		C: British Pound		D: Japanese Yen	
	MS	HPMS	MS	HPMS	MS	HPMS	MS	HPMS
μ_1	-2.389	-4.568	-0.670	-1.760	-0.631	-3.359	-3.707	-2.840
μ_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
μ_3	3.823	3.685	1.716	1.749	1.474	1.892	2.174	3.516
σ_1	5.454	4.930	2.324	2.099	6.010	4.677	1.907	4.736
σ_2	1.117	3.075	0.476	1.834	2.110	4.932	2.546	5.689
σ_3	1.893	3.244	3.371	3.169	3.949	4.386	6.791	5.779
p_{11}	0.577	0.722	0.761	0.861	0.987	0.931	0.506	0.863
p_{22}	0.715	0.687	0.000	0.891	0.154	0.938	0.920	0.645
p_{33}	0.193	0.735	0.880	0.921	0.000	0.951	0.844	0.909

Note: 1. MS-- standard Markov-switching model;

2. HPMS -- Markov-switching model with HP-filter, λ=0.3, 5, 300,and 25 for AUD, CAD, GBP, and JPY, respectively

3. μ_2 is imposed to be zero for three-state model

Table 2

Forecast Comparison (MSE) based on Engel and Hamilton's Dataset

Forecast Horizons	A: Deutsche Mark				B: French Franc				C: British Pound			
	1	2	3	4	1	2	3	4	1	2	3	4
In-Sample Forecasts (Sample 1973:Q3--1988:Q1)												
RW	38.01	83.79	130.84	199.18	34.37	82.96	143.99	220.88	29.10	76.06	124.45	187.26
MS2†	36.46	75.68	112.29	172.65	30.91	70.71	122.10	191.44	25.08	66.04	108.33	171.45
Improvement†	4.1%	9.7%	14.2%	13.3%	10.1%	14.8%	15.2%	13.3%	13.8%	13.2%	13.0%	8.4%
MS3	36.43	77.07	110.80	169.51	29.33	62.85	103.89	162.28	24.73	63.63	103.48	162.94
Improvement	4.1%	8.0%	15.3%	14.9%	14.7%	24.2%	27.9%	26.5%	15.0%	16.3%	16.9%	13.0%
HPMS2	33.67	69.22	105.12	161.09	28.09	62.22	117.15	195.41	20.42	46.58	78.45	125.84
Improvement	11.4%	17.4%	19.7%	19.1%	18.3%	25.0%	18.6%	11.5%	29.8%	38.8%	37.0%	32.8%
HPMS3*	32.61	70.25	106.60	159.38	27.14	58.84	104.72	175.85	20.21	44.62	67.94	105.37
Improvement*	14.2%	16.2%	18.5%	20.0%	21.0%	29.1%	27.3%	20.4%	30.5%	41.3%	45.4%	43.7%
Out-of-Sample Forecasts (Estimation Sample 1973:Q3--1983:Q4, Forecast Periods 1984:Q1-1988:Q1)												
RW	54.58	141.33	251.62	406.49	52.47	145.59	266.34	426.76	42.35	117.54	186.68	270.43
MS2†	48.36	125.54	230.35	384.59	45.05	127.07	239.52	399.41	35.11	98.44	161.44	253.10
Improvement†	11.4%	11.2%	8.5%	5.4%	14.1%	12.7%	10.1%	6.4%	17.1%	16.2%	13.5%	6.4%
MS3	58.11	154.93	287.17	475.64	59.61	182.73	334.89	521.57	43.37	126.99	215.66	340.89
Improvement	-6.5%	-9.6%	-14.1%	-17.0%	-13.6%	-25.5%	-25.7%	-22.2%	-2.4%	-8.0%	-15.5%	-26.1%
HPMS2	45.50	112.87	204.22	344.29	40.12	104.45	198.48	336.74	33.41	89.72	142.92	225.02
Improvement	16.6%	20.1%	18.8%	15.3%	23.5%	28.3%	25.5%	21.1%	21.1%	23.7%	23.4%	16.8%
HPMS3*	44.14	105.03	187.89	320.05	40.54	106.16	194.26	328.72	34.40	93.29	150.34	238.13
Improvement*	19.1%	25.7%	25.3%	21.3%	22.7%	27.1%	27.1%	23.0%	18.8%	20.6%	19.5%	11.9%

Note: 1. RW -- random walk model; MS2 -- 2-state standard Markov-switching model, and likewise; HPMS2 -- 2-state filtered model with HP-filter, and likewise. Smoothing parameter λ=0.3, 0.3, and 120 for DEM, FRF, and GBP, respectively.

2. MS2† -- specification in Engel and Hamilton (1990)

3. HPMS3* -- preferred model

4. Improvement defined as the percentage of reduction in MSE from the competing model versus the random walk

Table 3

Forecast Comparison (MSE) based on Extended Dataset

Forecast Horizons	A: Australian Dollar				B: Canadian Dollar				C: British Pound				C: Japanese Yen			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
In-Sample Forecasts (Sample 1973:Q1--2007:Q1)																
RW	24.12	47.37	74.66	106.94	6.47	13.48	20.22	30.08	25.76	59.56	86.31	121.62	37.20	82.35	119.24	167.91
MS2†	24.01	47.35	74.33	106.88	6.06	11.99	16.61	24.04	25.59	59.40	86.40	122.42	37.27	82.85	119.02	166.97
Improvement†	0.4%	0.0%	0.4%	0.1%	6.3%	11.1%	17.9%	20.1%	0.7%	0.3%	-0.1%	-0.7%	-0.2%	-0.6%	0.2%	0.6%
MS3	23.94	47.12	74.20	106.42	6.11	12.14	16.45	23.78	25.26	59.30	86.23	122.53	37.14	83.38	120.71	170.24
Improvement	0.7%	0.5%	0.6%	0.5%	5.5%	10.0%	18.7%	21.0%	1.9%	0.4%	0.1%	-0.7%	0.2%	-1.3%	-1.2%	-1.4%
HPMS2	21.69	43.36	70.64	104.19	5.60	10.38	14.14	20.40	23.03	50.03	68.08	94.29	32.48	67.19	93.81	138.85
Improvement	10.0%	8.5%	5.4%	2.6%	13.4%	23.0%	30.1%	32.2%	10.6%	16.0%	21.1%	22.5%	12.7%	18.4%	21.3%	17.3%
HPMS3*	21.23	42.68	70.10	103.53	5.25	9.64	13.16	19.42	22.54	48.93	66.59	91.10	31.60	65.78	93.83	139.96
Improvement*	12.0%	9.9%	6.1%	3.2%	18.7%	28.5%	34.9%	35.4%	12.5%	17.9%	22.8%	25.1%	15.1%	20.1%	21.3%	16.6%
Out-of-Sample Forecasts (Estimation Periods 1973:Q1--1995:Q4, Forecast Periods 1996:Q1-2007:Q1)																
RW	29.66	62.57	113.66	165.46	9.99	21.17	32.61	48.43	12.82	27.42	44.40	61.81	32.97	70.26	109.43	172.09
MS2†	30.03	63.77	116.57	169.40	10.39	22.49	35.05	52.89	11.88	24.69	40.06	57.20	33.61	72.93	116.48	184.84
Improvement†	-1.2%	-1.9%	-2.6%	-2.4%	-4.1%	-6.2%	-7.5%	-9.2%	7.3%	10.0%	9.8%	7.5%	-1.9%	-3.8%	-6.4%	-7.4%
MS3	30.06	63.64	114.03	166.26	10.60	23.17	37.20	56.73	11.72	24.86	40.89	59.32	32.10	66.89	100.91	158.21
Improvement	-1.3%	-1.7%	-0.3%	-0.5%	-6.1%	-9.4%	-14.1%	-17.1%	8.6%	9.3%	7.9%	4.0%	2.6%	4.8%	7.8%	8.1%
HPMS2	28.69	59.41	112.12	164.16	10.00	20.88	31.63	46.09	11.91	24.13	38.09	52.78	32.90	69.44	101.33	171.88
Improvement	3.3%	5.1%	1.4%	0.8%	-0.2%	1.4%	3.0%	4.8%	7.1%	12.0%	14.2%	14.6%	0.2%	1.2%	7.4%	0.1%
HPMS3*	29.64	59.39	110.32	159.19	9.54	19.50	27.15	37.90	12.02	24.35	38.38	52.79	32.23	66.93	94.22	146.85
Improvement*	0.1%	5.1%	2.9%	3.8%	4.5%	7.9%	16.7%	21.7%	6.2%	11.2%	13.5%	14.6%	2.2%	4.7%	13.9%	14.7%

Note: 1. RW -- random walk model; MS2 -- 2-state standard Markov-switching model, and likewise; HPMS2 -- 2-state filtered model with HP-filter, and likewise. Smoothing parameter λ=0.3, 5, 300,and 25 for AUD, CAD, GBP, and JPY, respectively.

2. MS2† -- specification in Engel and Hamilton (1990)

3. HPMS3* -- preferred model

4. Improvement defined as the percentage of reduction in MSE from the competing model versus the random walk

Table 4

Diebold-Mariano Test for Relative Forecastability

Forecast Horizons	A: Australian Dollar				B: Canadian Dollar				C: British Pound				C: Japanese Yen			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
MS2[†] vs. RW																
MSE Ratio	1.012	1.019	1.026	1.024	1.041	1.062	1.075	1.092	0.927	0.900	0.902	0.925	1.019	1.038	1.064	1.074
DM-stat	5.897	5.317	6.931	5.569	5.182	4.298	3.972	4.705	-3.020	-2.542	-2.099	-1.397	7.396	6.155	5.753	5.361
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.011	0.036	0.162	0.000	0.000	0.000	0.000
MS3 vs. RW																
MSE Ratio	1.013	1.017	1.003	1.005	1.061	1.094	1.141	1.171	0.914	0.907	0.921	0.960	0.974	0.952	0.922	0.919
DM-stat	1.845	5.444	1.768	1.652	5.049	3.975	4.487	5.691	-3.194	-2.300	-1.794	-0.791	-3.372	-3.405	-4.072	-4.148
p-value	0.065	0.000	0.077	0.098	0.000	0.000	0.000	0.000	0.001	0.021	0.073	0.429	0.001	0.001	0.000	0.000
HPMS2 vs. RW																
MSE Ratio	0.967	0.949	0.986	0.992	1.002	0.986	0.970	0.952	0.929	0.880	0.858	0.854	0.998	0.988	0.926	0.999
DM-stat	-2.810	-3.620	-0.853	-0.382	0.090	-0.512	-0.753	-1.047	-2.106	-1.987	-1.798	-1.486	-0.093	-0.276	-1.208	-0.022
p-value	0.005	0.000	0.394	0.703	0.928	0.609	0.451	0.295	0.035	0.047	0.072	0.137	0.926	0.783	0.227	0.982
HPMS3[*] vs. RW																
MSE Ratio	0.999	0.949	0.971	0.962	0.955	0.921	0.833	0.783	0.938	0.888	0.865	0.854	0.978	0.953	0.861	0.853
DM-stat	-0.038	-1.949	-1.018	-1.078	-1.331	-1.607	-2.202	-2.424	-2.312	-2.316	-2.136	-1.840	-1.013	-1.142	-2.158	-1.901
p-value	0.969	0.051	0.309	0.281	0.183	0.108	0.028	0.015	0.021	0.021	0.033	0.066	0.311	0.253	0.031	0.057

Note: RW -- random walk model; MS2 -- 2-state standard Markov-switching model, and likewise; HPMS2 -- 2-state Markov-switching model with HP-filter, and likewise. MSE ratios are calculated as the MSE from competing models divided by the MSE from the random walk, and a value less than one means improvement in forecast accuracy. The null hypothesis is H_0: Equal Forecast Accuracy or MSE Ratio=1. The DM statistic is based on the squared error loss function. Asymptotic variance is calculated based on Newey-West's (HAC) estimator with Bartlett kernel and the truncation lag is chosen based on SIC (Schwartz Information Criterion) by Ng and Perron (1995). Forecasts are based on estimation periods 1973:Q1--1995:Q4 and forecast periods 1996Q1-2007Q1.

Table 5

Robustness Checks

Forecast Horizons	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. Different Sample Spans																
	Canadian Dollar								British Pound							
	Forecast 1986:Q1-2007:Q1				Forecast 2001:Q1-2007:Q1				Forecast 1986:Q1-2007:Q1				Forecast 2001:Q1-2007:Q1			
RW	8.16	18.05	29.29	45.50	15.69	32.79	50.50	76.14	25.72	56.60	80.16	117.17	14.76	35.12	56.89	85.78
MS2†	8.16	18.05	29.28	45.48	15.63	32.83	50.45	76.07	26.04	59.64	78.90	117.25	13.68	30.74	47.77	71.09
Improvement†	0.0%	0.0%	0.0%	0.0%	0.4%	-0.1%	0.1%	0.1%	-1.2%	-5.4%	1.6%	-0.1%	7.3%	12.5%	16.0%	17.1%
MS3	8.14	17.96	29.17	45.43	15.09	30.51	42.63	62.33	26.96	62.91	83.03	124.68	13.72	31.14	48.54	72.81
Improvement	0.3%	0.5%	0.4%	0.1%	3.8%	7.0%	15.6%	18.1%	-4.8%	-11.1%	-3.6%	-6.4%	7.1%	11.3%	14.7%	15.1%
HPMS2	8.05	17.43	27.85	44.29	15.49	31.39	46.95	69.09	26.08	58.82	77.81	113.17	13.98	34.01	57.68	89.68
Improvement	1.4%	3.4%	4.9%	2.6%	1.3%	4.3%	7.0%	9.3%	-1.4%	-3.9%	2.9%	3.4%	5.3%	3.2%	-1.4%	-4.6%
HPMS3*	8.14	17.62	27.93	44.49	14.61	28.65	37.77	51.92	25.11	54.16	69.98	98.72	13.12	28.82	44.19	65.81
Improvement*	0.2%	2.4%	4.6%	2.2%	6.9%	12.6%	25.2%	31.8%	2.4%	4.3%	12.7%	15.8%	11.1%	17.9%	22.3%	23.3%
2. Different Smoothing Values																
	A: Australian Dollar				B: Canadian Dollar				C: British Pound				D: Japanese Yen			
RW	29.66	62.57	113.66	165.46	9.99	21.17	32.61	48.43	12.82	27.42	44.40	61.81	32.97	70.26	109.43	172.09
HPMS3(λ_1)	28.92	59.43	109.58	159.58	10.10	21.39	31.04	45.30	12.33	25.54	41.34	57.99	36.12	78.99	116.89	201.85
Improvement	2.5%	5.0%	3.6%	3.6%	-1.1%	-1.0%	4.8%	6.5%	3.8%	6.9%	6.9%	6.2%	-9.6%	-12.4%	-6.8%	-17.3%
HPMS3(λ_2)	29.64	59.39	110.32	159.19	9.54	19.50	27.15	37.90	12.02	24.35	38.38	52.79	32.23	66.93	94.22	146.85
Improvement	0.1%	5.1%	2.9%	3.8%	4.5%	7.9%	16.7%	21.7%	6.2%	11.2%	13.6%	14.6%	2.2%	4.7%	13.9%	14.7%
HPMS3(λ_3)	29.28	63.17	120.00	184.28	10.31	22.12	34.45	52.07	12.30	25.20	40.32	55.80	34.02	74.55	119.03	194.22
Improvement	1.3%	-1.0%	-5.6%	-11.4%	-3.2%	-4.5%	-5.7%	-7.5%	4.1%	8.1%	9.2%	9.7%	-3.2%	-6.1%	-8.8%	-12.9%

Note: 1. RW -- random walk model; MS2 -- 2-state standard Markov-switching model, and likewise; HPMS2 -- 2-state Markov-switching model with HP-filter, and likewise. Smoothing parameter λ= 5 and 300 for CAD and GBP, respectively.

2. MS2† -- specification in Engel and Hamilton (1990)

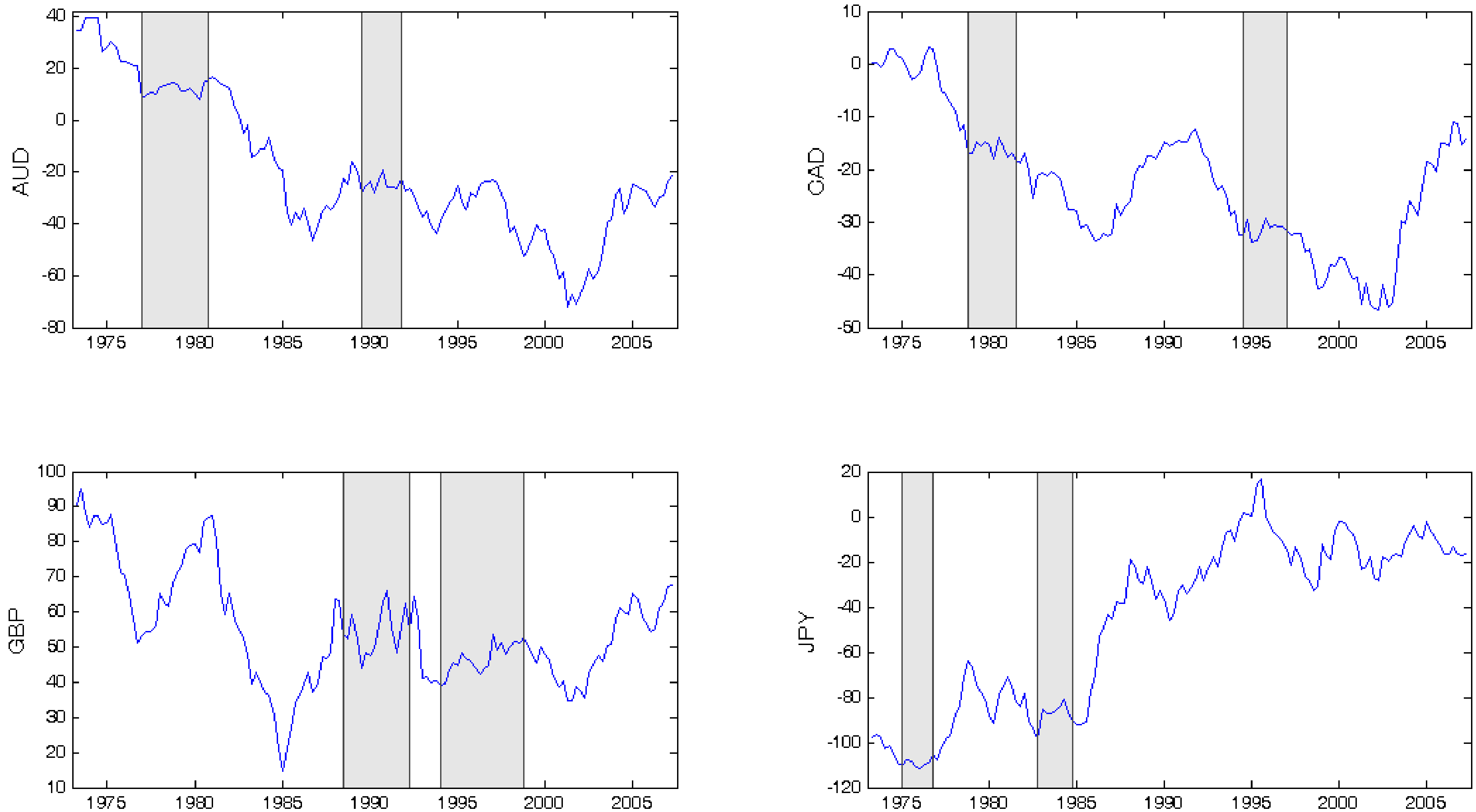
3. HPMS3* -- preferred model

4. Improvement defined as the percentage of reduction in MSE from the competing model versus the random walk

5. λ_1=0.03, λ_2=0.3, λ_3=30 for AUD; λ_1=0.05, λ_2=5, λ_3=100 for CAD; λ_1=10, λ_2=300, , λ_3=1500 for GBP; λ_1=1, λ_2=25, λ_3=500 for JPY, respectively.

Figure 1

Figure 1 Plot of Log Exchange Rates (USD per Unit of Foreign Currency)



Note: shaded areas can be roughly viewed as trendless regimes.

Figure 2

Figure 2 Smoothed Inferences: MS vs. HP-filtered MS Model

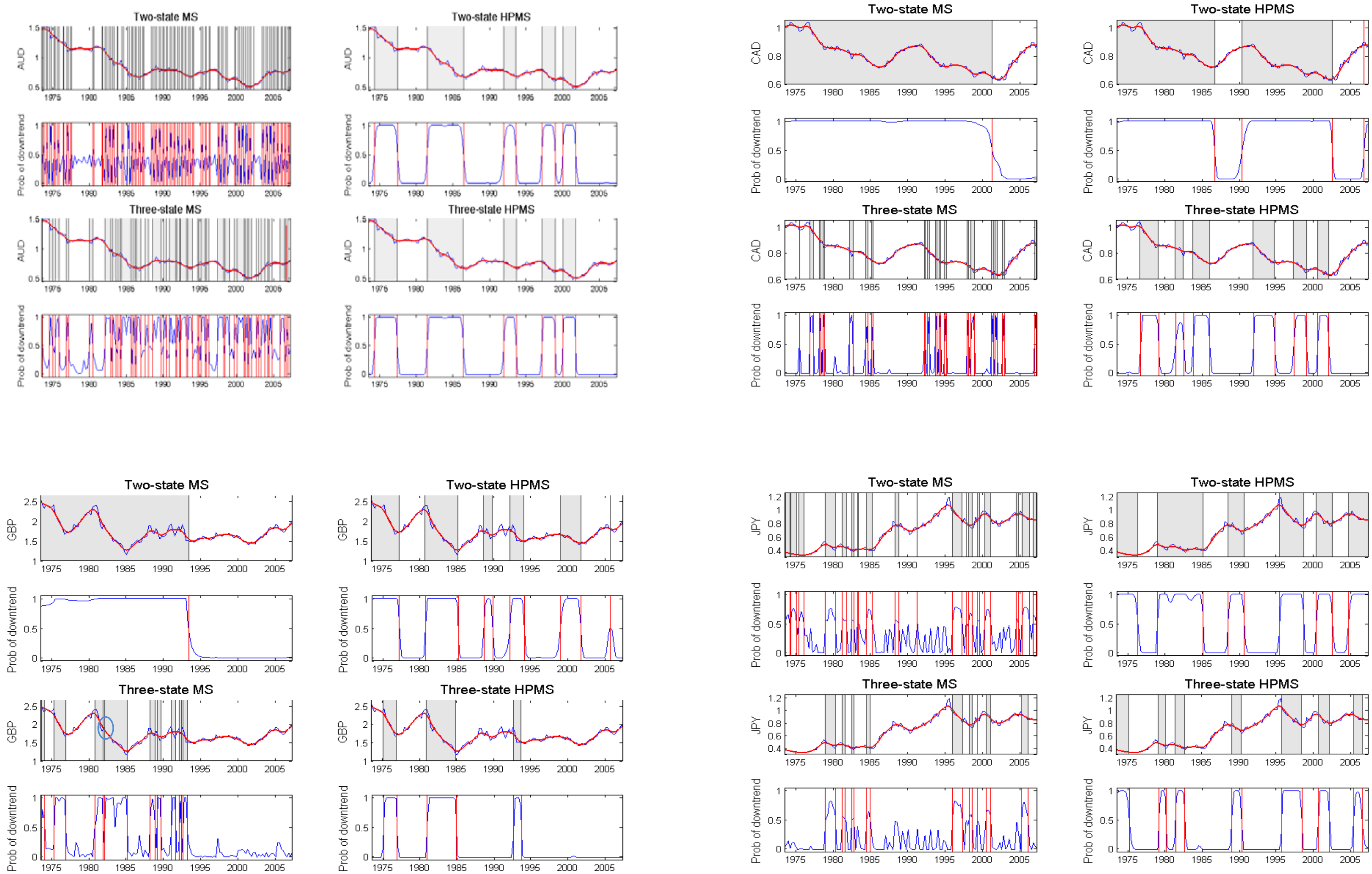


Figure 3

Figure 3 Stability of Smoothed Probabilities: MS vs. HP-filtered MS Model

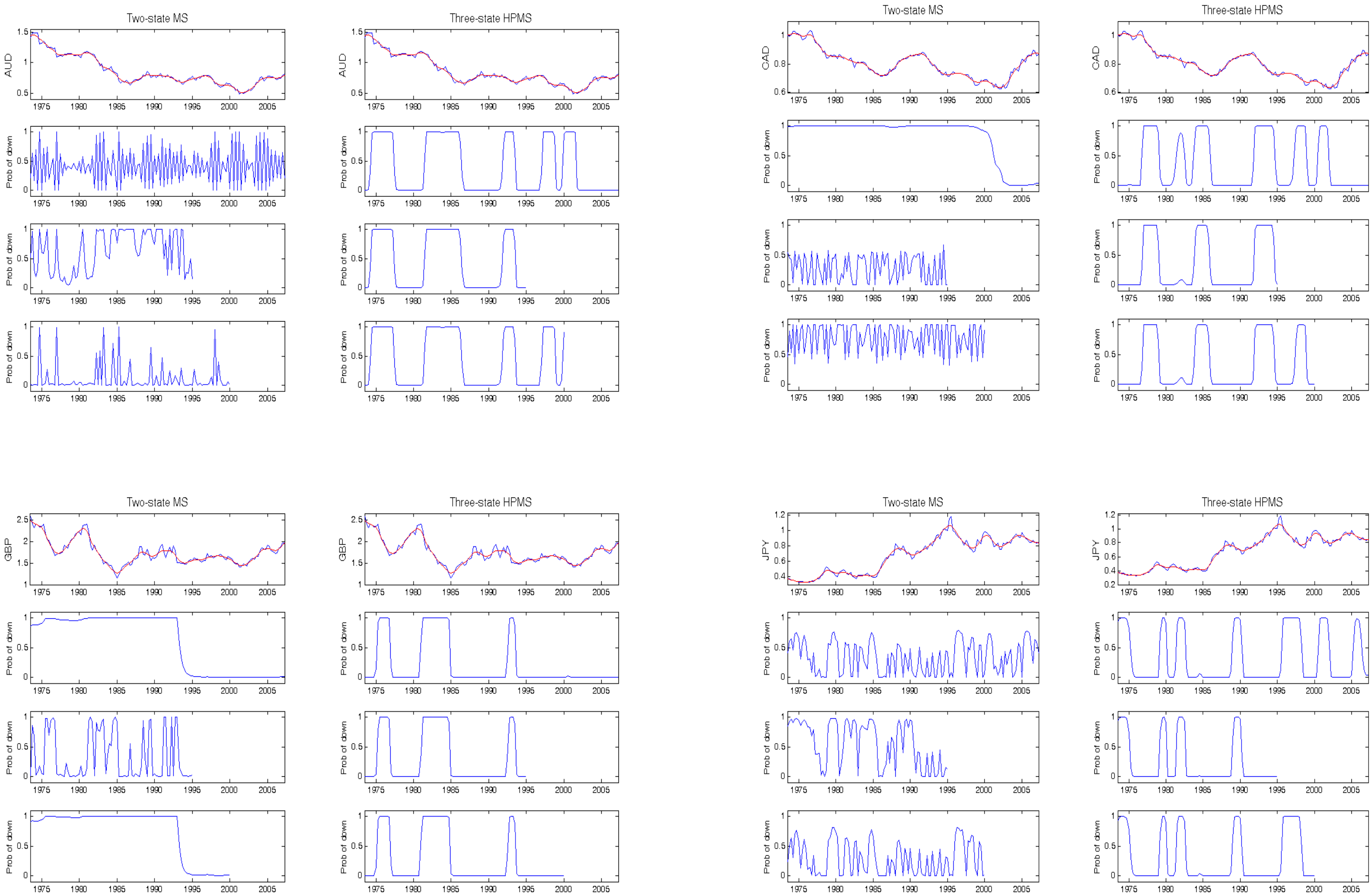
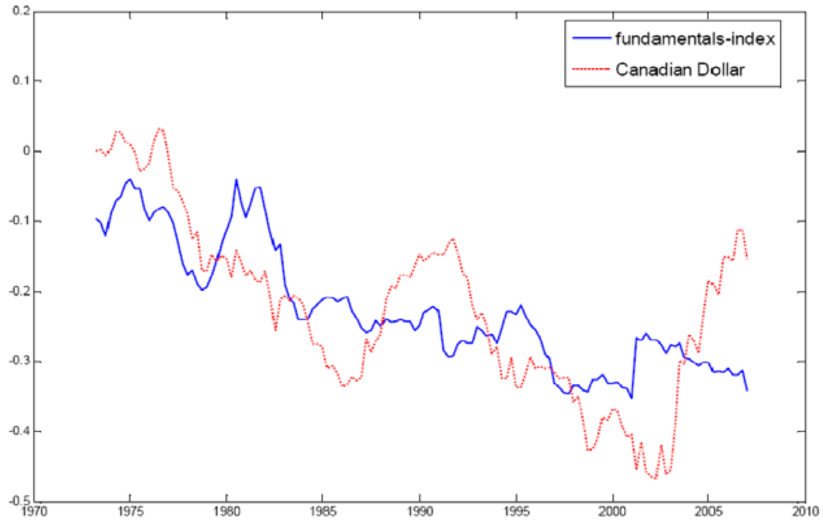
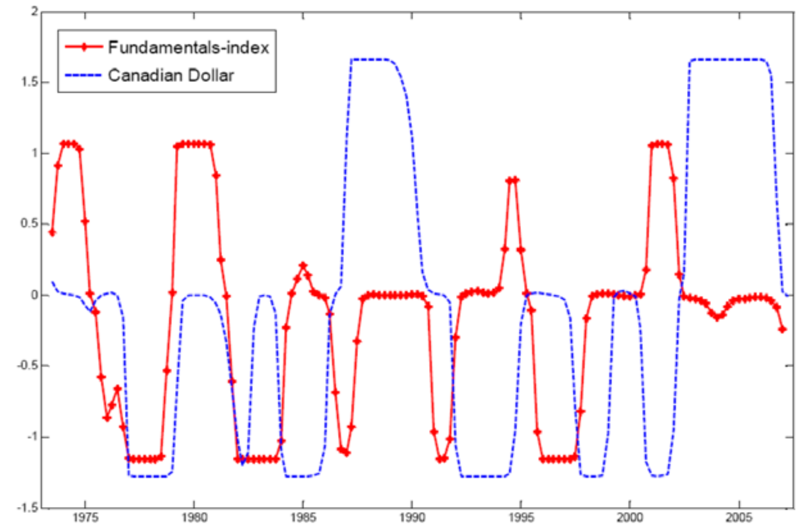


Figure 4

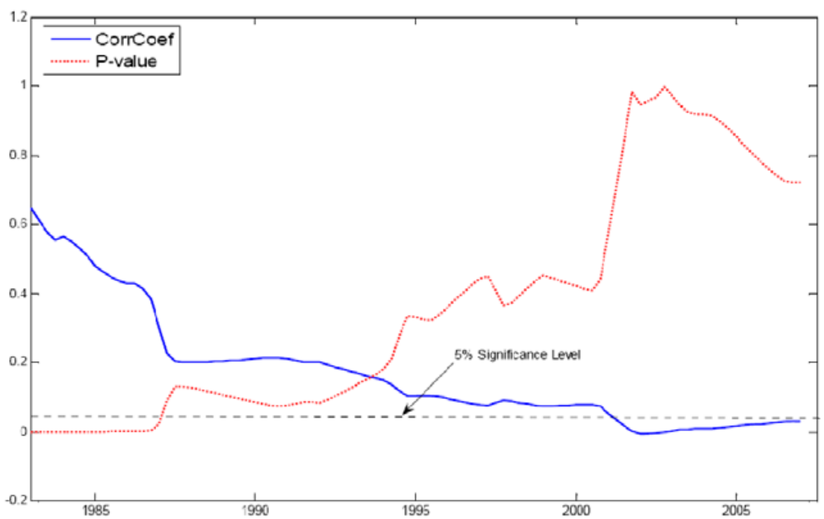
Figure 4 Fundamentals-index vs. Canadian Dollar



A. Fundamentals-index vs. Log Exchange Rate



B. Trend Components



C. Correlations between Trend Components

	$\triangle \hat{f}_t > 0$	$\triangle \hat{f}_t = 0$	$\triangle \hat{f}_t < 0$
$\triangle e_t > 0$	0	15	15
$\triangle e_t = 0$	27	7	24
$\triangle e_t < 0$	16	22	9

Note: $\triangle e_t \in (-2\sigma_{\triangle e},\ 2\sigma_{\triangle e})$ counted as $\triangle e_t = 0$, similarly for $\triangle \hat{f}_t$

D. Contingence Table