Modelling Exchange Rate Volatility with Random Level Shifts

Ye Li∗  Pierre Perron†
Boston University  Boston University

September 11, 2012

Abstract

Recent literature has shown that the volatility of exchange rate returns displays long memory features. It has also been shown that if a short memory process is contaminated by level shifts, the estimate of the long memory parameter tends to be upward biased. In this paper, we directly estimate a random level shift model to the logarithm of absolute returns of the Dollar-Mark and Dollar-Yen exchange rates, in order to assess whether random level shifts can explain this long memory property. The random level shift model is specified as the sum of a short memory and random level shift components. The later is modelled as the cumulative sum of a process which is 0 with probability $1 - \alpha$ and is a random variable with probability $\alpha$. To estimate the model, we transform it to a linear state space form with a mixture of normal innovations, so that we can apply a Kalman filter type algorithm. Our results show that there are few level shifts for the two series, but once they are taken into account, the long memory property of the series disappears. We also provide out-of-sample forecasting comparisons, which show that, most of times, the random level shift model outperforms the popular ARFIMA model in forecasting volatility.

JEL Classification Number: C22.

Keywords: Random Level Shifts, Long-Memory, Forecasting, Volatility.

∗Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (yechina@bu.edu).
†Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (perron@bu.edu).
1 Introduction

A vast literature has documented that various measures of the volatility of asset returns displays features akin to those of a long-memory process. This is also the case for the volatility of exchange rate series; see, e.g., Anderson et al. (2001) and Anderson and Bollerslev (1997), among others. On the other hand, it has been suggested that the long-memory features present in the data could be due to occasional level shifts; see, e.g. Diebold and Inoue (2001). This follows from similar arguments used in Perron (1989, 1990) who showed that changes in level and/or slope of the trend function of a series causes the estimate of the sum of the autoregressive parameters to be biased towards one, suggesting non-stationarity.

Some recent papers have tried to assess whether random level shifts are indeed responsible for this long-memory feature and not simply a theoretical curiosity. Early attempts to that effect include Stărică and Granger (2005) and Granger and Hyung (2003), who argue that for the volatility of stock market indices the evidence for long-memory is weaker when level shifts are taken into account. Stărică and Granger (2005) presented evidence that log-absolute returns of the S&P 500 index is an i.i.d. series affected by occasional shifts in the unconditional variance and show that this specification has better forecasting performance than the more traditional GARCH(1,1) model and its fractionally integrated counterpart. Perron and Qu (2007) analyzed the time and spectral domain properties of a stationary short memory process affected by random level shifts. Perron and Qu (2010) showed that, when applied to daily S&P 500 log-absolute returns over the period 1928-2002, the level shift model explains both the shape of the autocorrelations and the path of log periodogram estimates as a function of the number of frequency ordinates used. Qu and Perron (2011) estimated a stochastic volatility model with level shifts using a Bayesian approach with daily data on returns from the S&P 500 and NASDAQ indices over the period 1980.1-2005.12. They showed that the level shifts account for most of the variation in volatility, that their model provides a better in-sample fit than alternative models and that its forecasting performance is better for the NASDAQ and just as good for the S&P 500 as standard short or long-memory models without level shifts. Lu and Perron (2010) considered a random level shift model for which the series of interest is the sum of a short memory process and a jump or level shift component, modeled as the cumulative sum of a process which is 0 with some probability \(1 - \alpha\) and is a random variable with probability \(\alpha\). They applied it to the logarithm of daily absolute returns for the S&P 500, AMEX, Dow Jones and NASDAQ stock market return indices. The point estimates obtained imply few level shifts for all
series. But once these are taken into account, there is little evidence of serial correlation in the remaining noise and, hence, no evidence of long-memory. Once the estimated shifts are introduced to a standard GARCH model applied to the returns series, any evidence of GARCH effects disappears. They also considered rolling out-of-sample forecasts of squared returns. In most cases, the simple random level shifts model clearly outperforms a standard GARCH(1,1) model and, in many cases, it also provides better forecasts than a fractionally integrated GARCH model. Varneskov and Perron (2011) extended the analysis to introduce random level shifts in a general ARFIMA (autoregressive fractionally integrated moving-average) model. They showed that random level shifts are an essential component to model adequately the volatility of various series, whether from daily data or from realized volatility series constructed using high frequency data. From a forecasting perspective, they showed that the random level shift model is the only one that consistently belong to the 10% Model Confidence Set of Hansen et. al. (2011).

Hence, there is growing evidence that a random level shift model is indeed a genuine contender to explain the long-memory features of volatility. However, most of the results so far pertain to stock market return indices. Little evidence is available concerning the properties of the volatility of exchange rate series. Our goal is to use some of the methodologies recently developed to address this issue. One exception is Morana and Beltratti (2004) who considered structural changes in the realized variance processes of the DM/U.S.$ and Yen/U.S.$ exchange rates. Their results show that the volatility of DM/U.S.$ and Yen/U.S.$ exchange rates show clear evidence of genuine long memory and that the structural changes can only partially explain the long memory features. Their forecasting exercises indicate that for short-term forecasting neglecting the structural changes is not that important, but that accounting for them provides substantial improvements for long-term forecasting. However, as noted by Varneskov and Perron (2011), the results obtained are very different when considering historical spans of daily returns compared to shorter spans of realized volatility series constructed from high frequency data.

In this paper, we follow the approach of Lu and Perron (2010). We consider historical series of daily exchange rates for the Dollar-Yen and Dollar-Mark. We estimate a random level shifts model for log absolute return series, adopting the specification that the series is the sum of a short memory process and level shift component. The level shift component is specified as a mixture model which takes value 0 with probability $\alpha$ and is some random variable with probability $1 - \alpha$. To estimate the model, we cast it in a generalized state space framework with a mixture of normal distributions and use the estimation method developed
by Perron and Wada (2009). We also evaluate the forecasting performance of the random level shifts model relative to the popular ARFIMA model. We show that the former indeed provides improved forecasts in most cases. Also, we document that though few level shifts are present once they are taken into account any evidence of long-memory disappears and what is left is a noise component that is essentially white noise.

The structure of the paper is as follows. Section 2 presents the model and the specifications adopted. Section 3 discusses the method of estimation. The empirical results obtained from estimating the model are presented in Section 4 along with evidence that the level shifts account for all the long-memory features. The forecasting evaluations and comparisons are presented in Section 5. Section 6 offers brief concluding remarks.

2 Model

The random level shift model considered is specified by

\[ y_t = a + \tau_t + c_t \]

where \( a \) is constant term, \( \tau_t \) is the random level shift component and \( c_t \) is a short memory process to model the remaining noise. The level shift component is given by:

\[ \tau_t = \tau_{t-1} + \delta_t \]

where

\[ \delta_t = \pi_t \eta_t \]

with \( \pi_t \) a binomial variable, which takes the value 1 with probability \( \alpha \) and value 0 with probability \( 1 - \alpha \). When \( \pi_t = 1 \), a random level shift \( \eta_t \) happens having a distribution \( \eta_t \sim i.i.d \ N(0, \sigma^2_\eta) \). Furthermore, \( \pi_t, \eta_t \) and \( c_t \) are mutually independent. For the short memory component, in general \( c_t \) can be defined by the process \( c_t = C(L)c_t \) with \( e_t \sim i.i.d \ N(0, \sigma^2_e) \) where \( C(L) = \sum_{i=0}^{\infty} c_i L^i, \sum_{i=0}^{\infty} i|c_i| < \infty \) and \( C(1) \neq 0 \). However, as will be shown for the series considered, once the level shifts are accounted for, barely any serial correlation remains. Accordingly, we shall simply specify \( c_t \) as an AR(1) process. Hence, the model to be used is:

\[ y_t = a + \tau_t + c_t \]
\[ \tau_t = \tau_{t-1} + \delta_t \]
\[ c_t = \phi c_{t-1} + e_t \]
\[ \delta_t = \pi_t \eta_t \]
Note that we can write $\delta_t = \pi_t \eta_1 + (1 - \pi_t) \eta_2$, with $\eta_{it} \sim i.i.d N(0, \sigma_{\eta_1}^2)$ and $\sigma_{\eta_1}^2 = \sigma_{\eta_2}^2$, $\sigma_{\eta_2}^2 = 0$. This allows us to cast the model in a state space framework. More specifically, with the error term being a mixture of two normal distributions, where

$$\Delta y_t = c_t - c_{t-1} + \delta_t$$

and

$$\delta_t = \pi_t \eta_1 + (1 - \pi_t) \eta_2$$

$$c_t = \phi c_{t-1} + e_t$$

In matrix form,

$$\Delta y_t = HX_t + \delta_t$$

$$X_t = FX_{t-1} + U_t$$

In general, when $c_t$ follows an AR(p) process, then

$$X_t = [c_t, c_{t-1}, \ldots, c_{t-p}]'$$

$$F = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \cdots & \phi_p \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

$H = [1, -1, \ldots, 0]$, $U_t$ is a $p$-dimensional normally distributed random vector with zero mean and covariance matrix

$$Q = \begin{pmatrix} \sigma_e^2 & 0_{1 \times (p-1)} \\ 0_{(p-1) \times 1} & 0_{(p-1) \times (p-1)} \end{pmatrix}$$

Comparing this model with the standard state space model, the difference is that the error term is a mixture of two normal distributions.

### 3 Estimation Method

We apply the estimation method proposed by Perron and Wada (2009), see also Wada and Perron (2006). In their paper, they generalized the trend cycle decomposition framework
based on unobserved components with errors that are mixtures of normal distributions, thereby allowing shifts in the slope and level of the trend functions. The main ingredient that underlies the estimation procedure is that the model can be written as a state space model with normal errors occurring in two different possible states. These states can be described by the combined values of the Bernoulli random variables. From this we can generate the log likelihood function from the decomposition of the prediction errors to obtain estimates. Let $Y_t = [\Delta y_1, \Delta y_2, \cdots, \Delta y_t]$ represents the observations available at time $t$ and $\theta = [\sigma_n^2, \alpha, \sigma_e^2, \phi_1, \cdots, \phi_q]$ be the parameter vector to be estimated. The log-likelihood function is

$$\ln(L) = \sum_{t=1}^{T} \ln f(\Delta y_t|Y_{t-1}, \theta),$$

where,

$$f(\Delta y_t|Y_{t-1}, \theta) = \sum_{i=1}^{2} \sum_{j=1}^{2} f(\Delta y_t|s_{t-1} = i, s_t = j, Y_{t-1}, \theta) Pr(s_{t-1} = i, s_t = j|Y_{t-1}, \theta)$$

Here, $s_t$ is an indicator to represent whether or not a random level shift occurs. That is, when $s_t = 1$, then $\pi_t = 1$ and a random level shift happens; when $s_t = 2$, $\pi_t = 0$ and there is no level shift. Let

$$\omega_{ij}^t = f(\Delta y_t|s_{t-1} = i, s_t = j, Y_{t-1}, \theta), \quad i, j \in 1, 2$$

and

$$\tilde{\varepsilon}_{i|t-1}^j = Pr(s_{t-1} = i, s_t = j|Y_{t-1}, \theta)$$

$$= Pr(s_t = j) \sum_{k=1}^{2} Pr(s_{t-2} = k, s_{t-1} = i|Y_{t-1}, \theta)$$

$$= Pr(s_t = j) \tilde{\varepsilon}_{i|t-2}^k, \quad i, j \in 1, 2$$

where,

$$\tilde{\varepsilon}_{i|t-1}^k = Pr(s_{t-2} = k, s_{t-1} = i|Y_{t-1}, \theta)$$

$$= \frac{f(\Delta y_{t-1}|s_{t-2} = k, s_{t-1} = i, Y_{t-2}, \theta) Pr(s_{t-2} = k, s_{t-1} = i, Y_{t-2}, \theta)}{f(\Delta y_{t-1}|Y_{t-2}, \theta)}$$

So we have,

$$\tilde{\varepsilon}_{i|t+1}^k = Pr(s_{t+1} = i, s_t = k|Y_{t}, \theta) = Pr(s_{t+1} = i) \sum_{j=1}^{2} \tilde{\varepsilon}_{j|t}^k.$$
\[
\begin{align*}
\tilde{\epsilon}_{t+1|1}^1 &= \alpha \sum_{j=1}^{2} \tilde{\epsilon}_{t|t}^j = \alpha [\tilde{\epsilon}_{t|t}^1 + \tilde{\epsilon}_{t|t}^2] \\
\tilde{\epsilon}_{t+1|1}^{21} &= \alpha \sum_{j=1}^{2} \tilde{\epsilon}_{t|t}^j = \alpha [\tilde{\epsilon}_{t|t}^1 + \tilde{\epsilon}_{t|t}^2] \\
\tilde{\epsilon}_{t+1|1}^{12} &= (1 - \alpha) \sum_{j=1}^{2} \tilde{\epsilon}_{t|t}^j = (1 - \alpha) [\tilde{\epsilon}_{t|t}^1 + \tilde{\epsilon}_{t|t}^2] \\
\tilde{\epsilon}_{t+1|1}^{22} &= (1 - \alpha) \sum_{j=1}^{2} \tilde{\epsilon}_{t|t}^j = (1 - \alpha) [\tilde{\epsilon}_{t|t}^1 + \tilde{\epsilon}_{t|t}^2]
\end{align*}
\]

In matrix form,
\[
\begin{pmatrix}
\tilde{\epsilon}_{t+1|1}^1 \\
\tilde{\epsilon}_{t+1|1}^{21} \\
\tilde{\epsilon}_{t+1|1}^{12} \\
\tilde{\epsilon}_{t+1|1}^{22}
\end{pmatrix} =
\begin{pmatrix}
\alpha & \alpha & 0 & 0 \\
0 & 0 & \alpha & \alpha \\
1 - \alpha & 1 - \alpha & 0 & 0 \\
0 & 0 & 1 - \alpha & 1 - \alpha
\end{pmatrix}
\begin{pmatrix}
\tilde{\epsilon}_{t|t}^1 \\
\tilde{\epsilon}_{t|t}^{21} \\
\tilde{\epsilon}_{t|t}^{12} \\
\tilde{\epsilon}_{t|t}^{22}
\end{pmatrix}
\]

The conditional likelihood function for \( \Delta y_t \) is:
\[
f(\Delta y_t|s_{t-1} = i, s_t = j, Y_{t-1}, \theta) = \frac{1}{\sqrt{2\pi}|f_t^{ij}|^{-1/2}} \exp\left( -\frac{v_t^{ij} (f_t^{ij})^{-1/2} v_t^{ij}}{2} \right)
\]
where \( v_t^{ij} \) is the prediction error,
\[
v_t^{ij} = \Delta y_t - \Delta y_{t|t-1}^{ij} = \Delta y_t - E[\Delta y_t|s_t = i, s_{t-1} = j, Y_{t-1}, \theta]
\]
and \( f_t^{ij} = E(v_t^{ij} v_t^{ij}) \) is the prediction error variance. The prediction \( \Delta y_{t|t-1} \) based on past information does not depend on the state of time \( t \), but \( \Delta y_t \) does. The basic inputs are predictions for the state variables and their variances, which are
\[
X_{t|t-1}^i = F X_{t|t-1} \\
P_{t|t-1}^i = F P_{t|t-1} F' + Q
\]
The prediction error is \( v_t^{ij} = \Delta y_t - H X_{t|t-1}^{ij} \), so that \( f_t^{ij} = H P_{t|t-1}^{ij} H' + R_j \) where \( R_j \) is the variance of the error term, which takes two possible values: \( R_j = \sigma_\eta^2 \) with probability \( \alpha \) when
\( \pi_t = 1, R_j = 0 \) with probability \((1 - \alpha)\) when \( \pi_t = 0 \). Applying the updating formula, given \( s_t = j, s_{t-1} = i \), we obtain:

\[
X_{ij}^t = X_{ij}^{t-1} - P_{ij}^t H'(HP_{ij}^{t-1}H' + R_j)^{-1}(\Delta y_t - HX_{ij}^{t-1})
\]

\[P_{ij}^t = P_{ij}^{t-1} - P_{ij}^{t-1} H'(HP_{ij}^{t-1}H' + R_j)^{-1} HP_{ij}^{t-1}
\]

As in Perron and Wada (2009), to reduce the dimension of the estimation problem, we adopt the re-collapsing procedure of Harrison and Stevens (1976), given by:

\[
X_{jt}^t = \frac{\sum_{i=1}^{2} Pr(s_{t-1} = i, s_t = j|Y_t, \theta) X_{ij}^{t-1}}{Pr(s_t = j|Y_t, \theta)}
\]

and

\[
P_{jt}^t = \frac{\sum_{i=1}^{2} Pr(s_{t-1} = i, s_t = j|Y_t, \theta) [P_{ij}^{t-1} + (X_{jt}^t - X_{ij}^{t-1})(X_{jt}^t - X_{ij}^{t-1})]}{Pr(s_t = j|Y_t, \theta)}
\]

4 Empirical Results for Exchange Rate Returns

We consider the random level shift model for two daily exchange rate returns series, the Dollar-Yen and Dollar-Mark (both daily series from 10/11/1983 to 7/30/2010; 6,994 observations; obtained from the CRSP database). We apply our level shift model to log-absolute returns since they do not suffer from a non-negativity constraint as do, say, absolute or squared returns. There is also no loss relative to using squared returns in identifying level shift since log-absolute returns are a monotonic transformation. Since we wish to identify the probability of shifts and their locations, the fact that log-absolute returns are quite noisy is not problematic since our methods are robust to the presence of noise. Another reason is the fact that for many asset returns, a log-absolute transformation yields a series that is closer to being normally distributed (see, e.g., Andersen, Bollerslev, Diebold and Labys, 2001). When returns are zero or close to it, the log absolute value transformation implies extreme negative values. Using our method, these outliers would be attributed to the level shift component and thus bias the probability of shifts upward. To avoid this problem, we bound absolute returns away from zero by adding a small constant, i.e., we use \( \log(|r_t| + 0.001) \), a technique introduced to the stochastic volatility literature by Fuller (1996).

We first discuss some features of the series. Figure 1 presents the autocorrelation functions up to 100 lags. They show the autocorrelations to decay slowly, a feature typical of a long-memory process. To provide further evidence about this long-memory feature, we estimate the long-memory parameter \( d \) using the log-periodogram estimator of Geweke and
Porter-Hudak (1983) with a trimming value of \( m = T^{1/2} \). The estimates obtained were 0.32 (s.e.=0.03) for the Dollar-Yen and 0.34 (s.e.=0.03) for the Dollar-Mark, strongly suggesting the presence of long-memory. The results are qualitatively the same using other values for the bandwidth.

Our aim is to assess whether this long-memory feature is genuine or caused by the presence of level shifts. To that effect, we estimate the RLS model. For the specification of the short memory component, we consider two cases: 1) \( c_t = e_t \), so that the parameters to be estimated are \((\sigma_e^2, \sigma_\eta^2, \alpha)\); 2) \( c_t = \phi_1 c_{t-1} + e_t \), so that \( \phi_1 \) is an additional parameter to be estimated. The initial value for the state vector is \( X_{0|0} = (0, 0)' \) and the initial value for the covariance matrix is set to

\[
P_{0|0} = \begin{pmatrix}
\sigma_e^2 & 0 \\
0 & 0 \\
\end{pmatrix}.
\]

The estimates are presented in Table 1. The first feature of interest is that in the specification with an AR(1) component, the estimate of \( \phi_1 \) is very close to zero, indicating the near absence of serial correlation in the noise component once level shifts are accounted for. Hence, in what follows, we concentrate on the case with i.i.d. errors. The probability of shifts is small and imprecisely estimated. However, even if it is not statistically significant, as we shall see, it is practically very significant. One way to see this is that the standard deviation of the shifts component is larger than the standard deviation of the noise. Hence, it plays an important role. Given the point estimate of the probability of shifts, one can deduce an implied estimate of the number of shifts occurring in the sample. For the Dollar-Mark it is 8 and for the Dollar-Yen it is 5. As we shall see, even when such few shifts are taken into account the properties of the remaining noise is dramatically altered; a feature we discuss next.

4.1 The Effect of Level Shifts on the Long Memory Property

Given the estimation results, we seek to assess whether or not the random level shift component can explain the long memory property of the exchange rate returns. The strategy we adopt is the following. Given the estimated number of shifts, we estimate the break dates and regime specific means using the method of Bai and Perron (2003). Once these are obtained, we estimate the noise component as the difference between the original series and the fitted level shift process.

To be more specific, let \( m \) be the number of breaks (8 for the Dollar-Mark and 5 for the Dollar-Yen), \( T_i \) \((i=1, \cdots, m)\) be the break dates (with the convention that \( T_0 = 0, T_{m+1} = \))
and \( \{ u_i; i = 1, \ldots, m+1 \} \) be the means within each regime. The method of Bai and Perron (2003) allows obtaining estimates of the break dates \( \{ \hat{T}_i; i = 1, \ldots, m \} \) and regime-specific means \( \{ \hat{u}_i; i = 1, \ldots, m+1 \} \) as global minimizers of the objective function:

\[
\sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (y_t - u_i)^2
\]

The noise component, say \( \hat{c}_t \), is then obtained as \( \hat{c}_t = y_t - \sum_{i=1}^{m+1} \hat{u}_i DU_{i,t} \), where \( DU_{i,t} = 1 \) if \( \hat{T}_{i-1} < t \leq \hat{T}_i \) and 0, otherwise. To get a better view of the implied level shift process and its relation to the volatility of the exchange rate series, Figure 2 presents a graphs of the fitted level shift process in conjunction with a smoothed estimate of the log-absolute returns, obtained using a standard Gaussian kernel.

The autocorrelation function of \( \hat{c}_t \) is presented in Figure 3, which shows that basically no serial correlation is left once the level shifts are taken into account. We estimated the long-memory parameter \( d \) using the log-periodogram estimator of Geweke and Porter-Hudak (1983) with the same trimming value of \( m = T^{1/2} \), applied to the noise component \( \hat{c}_t \). The estimates obtained were \(-0.05\) (s.e.=0.03) for the Dollar-Yen and \(-0.02\) (s.e.=0.04) for the Dollar-Mark, reinforcing the conclusion that the long-memory feature in the data disappears once the level shifts are taken into account. Even if the level shifts are few in number, they can fully explain and account for the long-memory features of the exchange rate series. To get a better view of the implied level shift process and its relation to the volatility of the exchange rate series, Figure 3 presents a graphs of the fitted level shift process in conjunction with a smoothed estimate of the log-absolute returns, obtained using a standard Gaussian kernel.

5 Forecasting

We now consider the performance of the random level shift model with white noise errors in forecasting volatility proxied by squared returns relative to the ARFIMA model. The reason to make the comparisons with the ARFIMA model is that it is generally perceived as the best forecasting model for asset volatility. In order to assess the robustness of the results, we adopt two different designs; one follows Stáricá and Granger (2005) and Lu and Perron (2010), the other follows the framework of Varneskov and Perron (2010).

For the first experiment, we start forecasting at observations 2,000, and re-estimate the models every 20 days, at which point forecasts of up to 200 days are constructed. Since the proxy of realized squared returns are quite noisy, to reduce the effect of sampling variability
we follow Stărică and Granger (2005) in the construction of a metric to gauge the relative performance. Let $\hat{\sigma}^2_{t+p}$ be a $p$-step ahead forecast of $\sigma^2_{t+p}$, the variance of returns $r_t$ at time $t + p$, proxied by the squared demeaned returns. Let $n$ be the number of forecasts produced, then the estimated MSE is constructed as:

$$MSE(p) = \frac{1}{n} \sum_{t=1}^{n} (\hat{r}^2_{t,p} - \sigma^2_{t,p})$$

where $\sigma^2_{t,p} = \sum_{k=1}^{p} \hat{\sigma}^2_{t+k}$ and $\hat{r}^2_{t,p} = \sum_{k=1}^{p} r^2_{t+k}$ is the realized volatility over the interval $[t+1, t+p]$. The relative forecasting performance of two models is evaluated by the ratio of their MSEs.

For the random level shift model, since the noise component is serially uncorrelated and the level shifts are unforecastable, the best predictor is the mean of the last regime. The issue then becomes how to obtain a good estimate of the current mean of a regime at a given date at which the forecasts are made, without using information after that date. For reasons discussed in Lu and Perron (2010), we resort to using a backwards CUSUM procedure, as in Pesaran and Timmerman (1999). At each forecasting period, we use the CUSUM test of Brown, Durbin and Evans (1975). We determine the cutoff point to get the mean to forecast as the first time the CUSUM statistic crosses one of the critical lines, determined by the criterion that the probability of at least one of the last 1,000 cumulative sums of standardized recursive residuals crossing a line is 10%. The CUSUM is a procedure that effectively indicates the date at which a forecast failure occurs and is, accordingly, the best suited from a forecasting perspective (see, e.g., Pesaran and Timmerman, 1999). The results are presents in Figure 4. They show that the RLS model provides better forecasts than the ARFIMA model for a wide range of forecast horizons.

To assess whether the increased forecast accuracy is statistically significant we consider tests based on a Mincer and Zarnowitz (1969) type of regression given by

$$r^2_{t+p} - r^2_t = b_1(f^{LS}_{t,p} - r^2_t) + b_2(f^{ARFIMA}_{t,p} - r^2_t) + u_t$$

where $f^{LS}_{t,p}$ denotes the $p$-step ahead forecast of $\sigma^2_{t+p}$ from the random level shift model and $f^{ARFIMA}_{t,p}$ denotes the $p$-step ahead forecast from the ARFIMA model. The goal of this regression is to see if the forecasts from the ARFIMA model are uncorrelated with the forecast errors from the random level shift model. This is done by testing the null hypothesis $H_0^3 : (b_1, b_2) = (0, 1)$ using a standard Wald test with an asymptotic chi-square distribution. We can also test if the forecast errors from the random level shift model are uncorrelated
with the forecasts from ARFIMA models. This is done by testing $H_0^B: (b_1, b_2) = (1, 0)$. The p-values of such tests are presented in Table 2. For the Dollar-Mark series, the results are clear and informative. For most horizons, the null hypothesis $H_0^A$ cannot be rejected at the 5% significance level, while the null hypothesis $H_0^B$ can be rejected. In all cases, the p-values of the test for $H_0^A$ are higher than the p-values for the test $H_0^B$. The results therefore provide evidence that the RLS model provides statistically significant improvements in forecasting relative to the ARFIMA model. For the Dollar-Yen series, the results are more ambiguous. Again, the p-values of the test for $H_0^A$ are higher than the p-values for the test $H_0^B$ for all forecasting horizons. However, in most cases both tests lead to a rejection at the 5% level, indicating that both models fail to yield accurate forecasts. Nevertheless, for very long horizons (160 and 200 steps), the RLS model performs significantly better.

The second design to assess the relative forecasting performance follows the method adopted by Varneskov and Perron (2011). The first main difference is that for random level shift model, we obtain $\tau$-step ahead forecasts directly from the filtered estimates obtained when estimating the state-space model. The $\tau$-step ahead forecast is then given by:

$$\hat{y}_{t|t+\tau} = y_t + F \sum_{i=0}^{1} \sum_{j=0}^{1} Pr(s_{t+1} = j) Pr(s_t = i|Y_t))H_{ij}^{\tau|t}$$

The second difference pertains to the method to compare the forecasts. We consider out-of-sample forecasting of the last 900 ($T^{out} \in [1, 900]$) days of the two exchange rates series, which are also divided equally into three subperiods to assess the robustness of the results, $T^{out} \in [1, 300], T^{out} \in [301, 600]$ and $T^{out} \in [601, 901]$ . We compare three models, the Random Level Shift, ARFIMA(1, d, 1) and ARFIMA(0, d, 0) and consider direct $\tau$-step ahead forecasting for three different horizons $\tau = (1, 5, 10)$. The $\tau$-step ahead forecasts are defined as $\bar{y}_{t+\tau|t} = \sum_{s=1}^{\tau} \hat{y}_{t+s|t}$. Similarly the cumulative volatility proxy is defined by $\bar{\sigma}^2_{t,\tau} = \sum_{s=1}^{\tau} y_{t+s}$. We use the mean square forecast error (MSFE) criterion defined as:

$$MSFE_{\tau} = \frac{1}{T_f} \sum_{t=1}^{T_f} (\bar{\sigma}^2_{t,\tau} - \bar{y}_{t+\tau|t})^2$$

where $T_f$ is the total number of forecasts produced. The forecasts are evaluated and compared using the 10% Model Confidence Set (MCS) of Hansen et al. (2011). The MSFE’s and accompanying MCS p-values (in parenthesis) for multiple and pairwise comparisons of the three models are shown in Table 3. In 19 out of 24 cases, the RLS model belong to the 10% Model Confidence Set for all and pairwise comparisons as well as all pairwise compar-
isons. Again, the evidence about the superiority of the RLS model is strongest in the case of the Dollar-Mark exchange rate. Hence, in general, the evidence points to the fact that in many cases the RLS model provides statistically better forecasts, while for some others the difference is not statistically significant.

6 Conclusion

We considered historical series of daily exchange rates for the Dollar-Yen and Dollar-Mark. We estimated a random level shifts model for the log absolute return series, adopting the specification that the series is the sum of a short memory process and level shift component. We documented that though few level shifts are present once they are taken into account any evidence of long-memory disappears and what is left is a noise component that is essentially white noise. We also evaluated the forecasting performance of the random level shift model relative to the popular ARFIMA model. We showed that the former indeed provides improved forecasts in most cases. Our paper therefore adds to the recent literature that considered the volatility of stock market indices, by showing that a random level shift model is indeed a serious contender to explaining the long-memory features of the volatility of exchange rate series.
References


Table 1: Estimates of the RLS model with $c_t = e_t$ and $c_t = \phi_1 c_{t-1} + e_t$.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_\eta$</th>
<th>$\alpha$</th>
<th>$\sigma_c$</th>
<th>$\phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dollar-Yen</td>
<td>1.043 (0.046)</td>
<td>0.0007</td>
<td>0.7522</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.132 (0.247)</td>
<td>0.0005</td>
<td>0.7533</td>
<td>0.0309</td>
</tr>
<tr>
<td>Dollar-Mark</td>
<td>0.678 (0.2418)</td>
<td>0.0012</td>
<td>0.748</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2289 (4.377)</td>
<td>0.0089</td>
<td>0.7488</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Table 2: Comparison of forecasting performance between the RLS and ARFIMA models.

<table>
<thead>
<tr>
<th>Horizon (days)</th>
<th>p-values for the Wald Statistic</th>
<th>Dollar-Mark</th>
<th>Dollar-Yen</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H^A_0$</td>
<td>$H^B_0$</td>
<td>$H^A_0$</td>
<td>$H^B_0$</td>
</tr>
<tr>
<td>20</td>
<td>0.0726</td>
<td>0.0081</td>
<td>0.0024</td>
</tr>
<tr>
<td>40</td>
<td>0.1524</td>
<td>0.0214</td>
<td>0.0016</td>
</tr>
<tr>
<td>60</td>
<td>0.0514</td>
<td>0.0030</td>
<td>0.0000</td>
</tr>
<tr>
<td>80</td>
<td>0.3888</td>
<td>0.0604</td>
<td>0.0000</td>
</tr>
<tr>
<td>100</td>
<td>0.0198</td>
<td>0.0007</td>
<td>0.0254</td>
</tr>
<tr>
<td>120</td>
<td>0.2739</td>
<td>0.0484</td>
<td>0.0216</td>
</tr>
<tr>
<td>140</td>
<td>0.0262</td>
<td>0.0023</td>
<td>0.0000</td>
</tr>
<tr>
<td>160</td>
<td>0.0020</td>
<td>0.0005</td>
<td>0.0809</td>
</tr>
<tr>
<td>180</td>
<td>0.6402</td>
<td>0.1731</td>
<td>0.0286</td>
</tr>
<tr>
<td>200</td>
<td>0.9325</td>
<td>0.3196</td>
<td>0.2084</td>
</tr>
</tbody>
</table>

Table 2: Comparison of forecasting performance between the RLS and ARFIMA models.
### a) Dollar-Mark exchange rates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RLS</td>
<td>0.442(0.001)</td>
<td>3.881(0.002)</td>
<td>5.299(0.002)</td>
<td>0.699(0.003)</td>
<td>3.878(0.003)</td>
<td>8.978(0.003)</td>
<td>0.470(0.001)</td>
</tr>
<tr>
<td>ARFIMA(0,0)</td>
<td>0.445(0.059)</td>
<td>3.695(0.141)</td>
<td>5.805(0.199)</td>
<td>0.712(0.076)</td>
<td>4.177(0.0169)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARFIMA(1,1)</td>
<td>0.445(0.741)</td>
<td>2.891(1.000)</td>
<td>7.341(0.626)</td>
<td>0.575(0.657)</td>
<td>4.401(0.0856)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### b) Dollar-Yen exchange rates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RLS</td>
<td>0.538(0.245)</td>
<td>3.881(0.002)</td>
<td>10.460(0.006)</td>
<td>0.691(0.716)</td>
<td>5.675(0.863)</td>
<td>14.624(1.000)</td>
<td>0.534(0.533)</td>
</tr>
<tr>
<td>ARFIMA(0,0)</td>
<td>0.530(1.000)</td>
<td>3.695(1.000)</td>
<td>9.752(1.000)</td>
<td>0.686(1.000)</td>
<td>5.620(1.000)</td>
<td>14.784(1.000)</td>
<td></td>
</tr>
<tr>
<td>ARFIMA(1,1)</td>
<td>0.632(1.000)</td>
<td>3.868(1.000)</td>
<td>9.584(1.000)</td>
<td>0.622(0.255)</td>
<td>4.659(0.0001)</td>
<td>12.585(0.000)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Forecast evaluations of the RLS, ARFIMA(0,0,d) and ARFIMA(1,d,1) models. 

(0) indicates that model belongs to 10% MCS using all comparisions, (c) indicates that the RLS model belongs to the 10% MCS using all pairwise comparisons.
Figure 1: Autocorrelations of log absolute returns
Figure 2: Level shift component and smoothed volatility series
Figure 3: Autocorrelations of the residuals from the fitted level shift process
Figure 4: MSE ratio