Continuous Time Regime Switching Model Applied to Foreign Exchange Rate *

Stéphane GOUTTE‡ AND Benteng ZOU †

October 10, 2012

Abstract

Continuous time modified Cox-Ingersoll-Ross (1985) stochastic model is employed, combining with Hamilton (1989) type Markov regime switching framework, to study daily foreign exchange rates, where all parameter values depend on the value of a continuous time Markov chain. The Expectation-Maximization algorithm is extended, generalized, applied to a more general class of regime switching models and used to study some exchange rate data. We compare the obtained results with non regime switching models and notice that the regime switching outcomes match much better the reality than the others without Markov switching; and two regimes in most of the cases are better than more regimes.

Keywords Foreign exchange rate; Regime switching model; Expectation-Maximization algorithm; financial crisis.

MSC Classification (2010): 91G70 60J05 91G30

JEL Classification (2010): F31 C58 C51 C01

*We appreciate enormously the valuable discussion in the early stage of this work with Michel Beine, Yin-Wong Cheung, Gautam Tripathi and Rafal Weron. But of course, all eventual mistakes and errors are ours.

†Centre National de la Recherche Scientifique (CNRS), Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, Universités Paris 7 Diderot. Supported by the FUI project $R = MC^2$. Mail: goutte@math.univ-paris-diderot.fr

‡CREA, Université du Luxembourg. Mail: benteng.zou@uni.lu
1 Introduction

Markov Switching model defines two or more states or regimes, and hence, it can present the dynamic process of variables of concern vividly and provide researchers and policy makers with a clear clue of how these variables have evolved in the past and how they may change in the future.

Engle and Hamilton (1990) first study the exchange rate behavior using non mean-reverting Markov switching model basing on quarterly data in exchange rate, during the period of 1973-1988, and find that Markov switching model is a good approximation to the series. Engle (1994) extends this work and studies whether Markov switching model is a useful tool for describing the behavior of 18 exchange rates and he concludes that the Markov switching model fits well in-sample for many exchange rates, but the Markov model does not generate superior forecasts to a random walk or the forward rate. Engle and Hakkio (1996) examine the behavior of European Monetary System exchange rates using Markov switching model and find that the changes in exchange rate match the periodic extreme volatility. Marsh (2000) goes one step further and study the daily exchange rates of three countries against the US dollar by applying Markov switching model and concludes that the data are well estimated by Markov switching model though the out-of-sample forecasting are very poor due to parameter instability. And Bollen et al. (2000) examine the ability of regime switching model to capture the dynamics of foreign exchange rates and their test shows that a regime-switching model with independent shifts in mean and variance exhibits a closer fit and more accurate variance forecasts than a range of other models, though the observed option prices do not fully reflect regime switching information.

Recently, Bergman and Hansson (2005) notice that the Markov switching model is good to describe the exchange rates of six industrialized countries against the US dollar. Cheung and Erlandsson (2005) test three dollar-based exchange rates by quarterly and monthly data, respectively, and notice that monthly data “...unambiguous evidence of the presence of Markov switching dynamics”. Their findings suggest that “data frequency, in addition to sample size, is crucial for determining the number of regimes”. More recently, Ismail and Isa (2007) employ Markov switching model to capture regime shifts behavior in Malaysia ringgit exchange rates against four other countries between 1990 and 2005. They conclude that Markov Shifting model is found to successfully capture the timing of regime shifts in the four series.

Except the above mentioned work, Lopez (1996) studies the exchange rate market in the long run (and short run) by specially taking into account the central bank regime shift and claims that the central bank activity do have long term effects on exchange rates (except the short term impacts).
These analysis confirms that the Hamilton’s Markov switching model is a good way to study the exchange rate behavior given the fact that the real world economies are changing from regime to regime due to different crisis and/or policies. Exchange rate regime changes and their effects on some key macroeconomic variables are studied by Caporale and Pittis (1995), who offer some insight of the effects of some regime changes on the real world.

In this present paper, we would rather see from a different direction than them. That is, we would like to see what if the macroeconomic regime switches, how the exchange rate would follow? To capture the stochastic nature of foreign exchange rate, we modify the Cox, Ingersoll and Ross (1985) stochastic interest rate model to measure exchange rate, particularly with a mean reverting part. Furthermore, we also calibrate the model basing on real daily exchange rate data from Jan. 2000 until May 2012. To understand and catch the complete information of switching regimes, we present an extended and generalized Expectation-Maximization algorithm and followed by some comparison with respect to some other non-regime switching models. By doing so, we are convinced that stochastic exchange rates, under regime switching model, can sharply catch the regime switching time and period. In addition, we also observe that two type of regimes: good and bad economic performance (or normal and crisis periods), is better for most of the exchange rates studies than more regimes. And that confirms again the data frequency argument of Cheung and Erlandsson (2005). The very recent paper of Naszodi (2011) is close to our idea of switching regime effects on exchange rate. However, our regime shifting is more general than Naszodi (2011), in which the switching is only about the exchange rate regime changes “from free floating to a completely fixed one”, such as “the adoption of the Euro”.

Nonetheless, in doing so, we employ time series filtering and smoothing technique to smooth out the noisy data. This technique promises that our results capture more precisely the trend of exchange rates than the standard Hamilton’s Markov switching model, which could be over-affected due to noise in the data, and hence misleading the regime switching results.

To our knowledge this is the first time that combination of Cox-Ingersoll-Ross (hereafter in short CIR) framework with Markov regime switching model is employed to study foreign exchange rates, though similar idea is used recently by Driffill and Kenc (2009) in studying bonds prices, however, there is no data smoothing process in their work. In this paper, with refined and modified filtering and smoothing algorithm, we show that the regime switching Cox-Ingersoll-Ross model fits better foreign exchange rate data than, firstly, non regime switching CIR and, secondly, other non regime switching models.

For the above findings of maximum likelihood, we employ an extension of the well known

\[1\] As to this points, see more clearly in Dacco and Satchel (1999).
Expectation-Maximization (EM) algorithm, which was named and explained first by Dempster, Laird and Rubin (1977). Our setting is developed in Hamilton (1989a, b) and generalized in Choi (2009) or more recently in Janczura and Weron (2011). In short, this algorithm works in two steps: firstly the Expectation step (E-step) where all the probabilities of the model are calculated basing on the current estimation of parameters; then secondly, using these probabilities, the Maximization step (M-step) is done by maximizing the expected log-likelihood found in the E-step, and update estimated parameters will be used in the next E-step. Nonetheless, this is a classical likelihood maximization steps but the likelihood function is weighted with the probabilities calculated in E-step. Indeed, we work in a regime switching model. Hence, we will use a generalization of the (EM)-algorithm in order to apply to our general regime switching model.

This paper is arranged as following: Section 2 presents the exchange rate Cox-Ingersoll-Ross model with regime switching. Section 3 documents real data, refined and modified filtering and smoothing algorithm. Then, Section 4 presents the main findings including simulation analysis, different types of comparison results. Section 5 studies some forecasting and Section 6 concludes.

2 The model

In this section, we first introduce the notion of continuous time Markov chain on finite space, then general continuous time regime switching model is presented. Finally, some examples of real word regime switching are followed. Let $T (> 0)$ be a fixed maturity time and denote by $(\Omega, \mathbb{F} := (\mathcal{F}_t), 0 \leq t \leq T, \mathbb{P})$ an underlying probability space.

**Definition 2.1** Let $(X_t)_{t \in [0, T]}$ be a continuous time Markov chain on finite space $S := \{1, 2, \ldots, K\}$. Denote $\mathcal{F}_t := \sigma(X_s); 0 \leq s \leq t$, the natural filtration generated by the continuous time Markov chain $X$. The generator matrix of $X$ is then denoted by $\Pi^X$ and it is given by

$$\Pi^X_{ij} \geq 0 \quad \text{if } i \neq j \text{ for all } i, j \in S \quad \text{and} \quad \Pi^X_{ii} = -\sum_{j \neq i} \Pi^X_{ij} \quad \text{otherwise.} \quad (2.1)$$

**Remark 2.1** The quantity $\Pi^X_{ij}$ represents the intensity of the jump from state $i$ to state $j$.

We can now give the definition of global regime switching model.
\textbf{Definition 2.2} For all $t \in [0, T]$, let $X_t$ be a continuous time Markov chain on finite space $S := \{1, \ldots, K\}$ defined as in Definition 2.1. A continuous time Regime switching model (RS-M) is a stochastic process $(r_t)$ which is solution of the stochastic differential equation given by
\begin{equation}
    dr_t = (\alpha(X_t) - \beta(X_t) r_t) \, dt + \sigma(X_t) \, (r_t)^{\delta(X_t)} \, dW_t \quad \text{with} \quad r_0 = r \in \mathbb{R}^+
\end{equation}
where $\alpha(X_t)$, $\beta(X_t)$, $\sigma(X_t)$ and $\delta(X_t)$ are functions of the Markov chain $X$. Hence, they are constants which take values in $\alpha(S)$, $\beta(S)$, $\sigma(S)$ and $\delta(S)$
\begin{align*}
    \alpha(S) := \{\alpha(1), \ldots, \alpha(K)\} &\in \mathbb{R}^{K^*}, & \beta(S) := \{\beta(1), \ldots, \beta(K)\}, \\
    \sigma(S) := \{\sigma(1), \ldots, \sigma(K)\} &\in \mathbb{R}^{K^+} & \delta(S) := \{\delta(1), \ldots, \delta(K)\} &\in \mathbb{R}^K
\end{align*}
For all $j \in \{1, \ldots, K\}$, we impose the condition $2\alpha(j) \geq \sigma(j)^2$ to ensure the positivity of the process $r$.

\textbf{Remark 2.2} – The above (RS-M) model is a continuous time regime switching diffusion with drift $\mu(r_t, X_t) = (\alpha(X_t) - \beta(X_t) r_t)$ and volatility $\sigma(r_t, X_t) = \sigma(X_t) (r_t)^{\delta(X_t)}$. For simplicity, we will denote the values $\alpha(X_t)$, $\beta(X_t)$, $\sigma(X_t)$ and $\delta(X_t)$ by $\alpha_t$, $\beta_t$, $\sigma_t$ and $\delta_t$.

– The drift factor, $(\alpha_t - \beta_t r_t)$, ensures mean reversion of the process towards the long run value $\frac{\alpha_t}{\beta_t}$, with speed of adjustment governed by the strictly positive parameter $\beta_t$. From economic point of view, if the value of $\beta_t$ is large then the dynamic of the process $r$ is almost near the value of the mean, even if there is a spike at time $t \in [0, T]$. Then, for a small time period $\epsilon$, the value of $r_{t+\epsilon}$ will be again close to the value of the mean.

– $\alpha(S) \in \mathbb{R}^{K^*}$ means that all values of the parameter $\alpha$ in each state have to be strictly positive, since we have the condition $2\alpha(j) \geq \sigma(j)^2$ for every $j \in S$. Moreover, $\sigma$ is the volatility so it has to be positive. Hence, $\sigma(S) \in \mathbb{R}^{K^+}$.

\textbf{Remark 2.3} – It is obvious that in this model there are two sources of randomness: the Brownian motion $W$ appearing in the dynamic of $r$ and the Markov chain $X$. Hence, there is a randomness due to the information market which is the initial continuous filtration $\mathcal{F}$ generated by the Brownian motion $W$; and another randomness due to the Markov chain $X$, $\mathcal{F}^X$. We assume that $W$ and $X$ are mutually independent.

– This independence implies that the Markov chain is an exogenous factor of the market information. Thus, it can be seen as an exogenous factor such as an economic impact factor. An economic interpretation of this is that the Markov chain can represent a credit rating of a firm $A$. Indeed, assume that (RS-M) models the spread of a firm $A$ then the Markov chain can be the credit rating of this firm given by an exogenous rating company as “Standard and Poors”. Then it is natural to think that the dynamic of the spread of the firm $A$ depends on the value of this notation $X$ (for more detail, see Goutte and Ngoupeyou [24]).
Hence, model \((\text{RS-M})\) (2.2) is a mean reverting model with local volatility. Moreover, this model is construct to encompass most of the financial models stated in the literature. Indeed, we obtain:

- **Regime switching Cox Ingersoll Ross (RS-CIR).**
  Taking \(\delta(X_t) \equiv \frac{1}{2}\), we obtain a standard Cox-Ingersoll-Ross (CIR) model with regime switching parameters:
  \[
  dr_t = (\alpha(X_t) - \beta(X_t)r_t) dt + \sigma(X_t)\sqrt{r_t}dW_t. \tag{2.3}
  \]

- **Regime switching Vasicek (RS-V).**
  Taking \(\delta(X_t) \equiv 0\), it yields a Vasicek model with regime switching parameters:
  \[
  dr_t = (\alpha(X_t) - \beta(X_t)r_t) dt + \sigma(X_t)dW_t. \tag{2.4}
  \]

- **Regime switching mean reverting Geometric Brownian motion (RS-R-GM).**
  Taking \(\delta(X_t) \equiv 1\), we have a geometric Brownian motion with mean reverting and regime switching parameters:
  \[
  dr_t = (\alpha(X_t) - \beta(X_t)r_t) dt + \sigma(X_t)r_tdW_t. \tag{2.5}
  \]

- **Regime switching Constant of elasticity Variance (RS-CEV)**
  Taking \(\alpha(X_t) \equiv 0\), it follows a constant of elasticity variance (CEV) with regime switching parameters:
  \[
  dr_t = -\beta(X_t) r_t dt + \sigma(X_t)(r_t)^{\delta(X_t)} dW_t. \tag{2.6}
  \]

The use of Hamilton’s (1989) Markov-switching models to study business cycle, economic growth and unemployment rate et al is not new. Here, we just mention a few. In his seminal paper, Hamilton (1989) already notices that Markov-switching models are able to reproduce the different phase of the business cycles and captures the cyclical behavior of the U.S. GDP growth data. McConnell and Perez-Quiros (2000) use an augmented model, compare with Hamilton’s original model and test the data up to the late 1990s. They notice the recessions clearer in their series by the augmented power. Kontolemis (2001) applies multivariate version of the model used by Engle and Hamilton (1990) to time series composing the composite coincident indicator in the U.S. data in order to identify the turning points for the U.S. business cycle. More recently, Bai and Wang (2010) go one step further by allowing changes in variance and show that their restricted model well identifies both short-run regime switches and long-run structure changes in the U.S. macroeconomic data.

In Europe, Ferrara (2008) employs Markov-switching model to construct probabilistic indicators and serves as useful tools for providing original qualitative information for economic
analysis especially “to monitor on a monthly basis turning points in the business cycle in French industry and those in the acceleration cycles in the French economy as a whole” and “Indicators of this nature are currently being developed for the euro area as a whole”. Billio and Casarin’s (2010) recent working paper study the Euro area by considering monthly observation from January 1970 until May 2009 of the industrial production index. They find that their new class of Markov switching latent factor model (with stochastic transition probability) “implies a better description of the dynamics of the Euro-zone business cycle”.

Basing on the above facts that Markov switching models capture the economic cycles and regime switching, therefore, we would like to see how the exchange rates would behave, do the exchange rates follow the economic regime switching and how large (or small) are the effects?

In order to answer these questions, in the following sections, we first introduce some real data followed by some calibration, estimation, comparison with some other models, and some forecasting for exchange rates.

3 Data and Estimations Methods

In this section, we first state real foreign exchange rate data for different currencies. In order to study and get as much information as possible, we then present the Estimation Method: the extended and generalized Expectation-Maximization algorithm.

3.1 Data

Our samples of exchange rate include Euros VS Dollars, Yuan VS Dollars, Euro VS Yen, Euro VS Livre (GB) and Euro VS Yuan. Figure 1, 2 and 3 show the historical value of the corresponding daily foreign exchange rates over the period of January 1st, 2000 until May 28, 2012, except for the foreign exchange rate Yuan VS Dollars which begins in January 2006 since before it is a fixed constant.

---

2The idea of using Markov switching model was first proposed by Baron and Baron (2002).
3The data are taken on the web site [http://fxtop.com/fr/](http://fxtop.com/fr/)
Figure 1: On left: Price of 1 Euro in Dollars between Jan. 2000 and May 2012. On right: Price of 1 Yuan in Dollars between Jan. 2006 and May 2012.


Figure 3: Price of 1 Euro in Yuan between Jan. 2006 and May 2012.
We begin by giving in Table 1 some general descriptive statistics for all foreign exchange rate data.

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro/Dollars</td>
<td>0.8324</td>
<td>1.5849</td>
<td>1.2190</td>
<td>0.1965</td>
<td>-0.0033</td>
<td>0.0032</td>
</tr>
<tr>
<td>Yuan/Dollars</td>
<td>6.3809</td>
<td>8.0702</td>
<td>7.1376</td>
<td>0.5164</td>
<td>0.0771</td>
<td>0.1298</td>
</tr>
<tr>
<td>Euro/Yen</td>
<td>90.5300</td>
<td>168.77</td>
<td>127.9484</td>
<td>19.2980</td>
<td>2222.8422</td>
<td>312563.5673</td>
</tr>
<tr>
<td>Euro/Livre</td>
<td>0.5794</td>
<td>0.9610</td>
<td>0.72803</td>
<td>0.0994</td>
<td>0.0005</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics

It is easy to see that this historical data of exchange rates have significant different economic fluctuation during this time interval. To study the above presented data, we rely on an estimation method: Expectation-Maximization algorithm.

### 3.2 Estimation Method

As mentioned in the introduction, the Expectation-Maximization EM-algorithm we are applying here is a generalization and extension of the EM-algorithm developed in Hamilton (1989a, b) and generalized in Choi (2009) or more recently in Janczura and Weron (2011). We applied this estimation method to the general regime switching model.

Suppose that the size of historical data is $M + 1$. Let $\Gamma$ denote the corresponding increasing sequence of time where this data value are taken:

$$\Gamma = \{t_j; 0 = t_0 \leq t_1 \leq \ldots t_{M-1} \leq t_M = T\}, \quad \text{with} \quad \Delta t = t_j - t_{j-1}.$$  

Then, the discretized approximation model of (2.2) is given by

$$r_{tk} - r_{tk-1} = (\alpha_{tk} - \beta_{tk} r_{tk-1}) \Delta t + \sigma_{tk} (r_{tk-1})^{\delta_{tk}} \Delta W_{tk}, \quad \text{for all} \quad k \in \{1, \ldots, M\}. \quad (3.7)$$

Here, the time step $\Delta t$ is equal to one since we have uniform equidistant time data values. Then, $\Delta W_{tk} \sim \sqrt{\Delta t} \epsilon_{tk} = \epsilon_{tk}$ where $\epsilon_{tk} \sim \mathcal{N}(0, 1)$. Hence, it yields

$$r_{tk} - r_{tk-1} = (\alpha_{tk} - \beta_{tk} r_{tk-1}) + \sigma_{tk} (r_{tk-1})^{\delta_{tk}} \epsilon_{tk},$$

$$r_{tk} = \alpha_{tk} + (1 - \beta_{tk}) r_{tk-1} + \sigma_{tk} (r_{tk-1})^{\delta_{tk}} \epsilon_{tk}. \quad (3.8)$$

We will denote in the sequel by $F_{tk}^r$ the vector of historical value of the process $r$ until time $t_k \in \Gamma$. Thus, $F_{tk}^r$ is the vector of the $k + 1$ last value of the discretized model defined in (3.8) and therefore, $F_{tk}^r = (r_{t_0}, r_{t_1}, \ldots, r_{tk})$.  

9
To estimate the optimal set of parameters \( \hat{\Theta} := (\hat{\alpha}_i, \hat{\beta}_i, \hat{\sigma}_i^2, \hat{\delta}_i, \hat{\Pi}) \), for \( i \in S \), we use the EM-algorithm where the set of parameter \( \Theta \) is estimated by an iterative two-step procedure. First, the Expectation procedure or E-step: We evaluate the smoothed and filtered probability. In fact the filtered probability is given by the probability such that the Markov chain \( X \) is in regime \( i \in S \) at time \( t \) with respect to \( \mathcal{F}_t^r \). And the smoothed probability is given by the probability such that the Markov chain \( X \) is in regime \( i \in S \) at time \( t \) with respect to all the historical data \( \mathcal{F}_t^r \).

Second, the Maximization step, or M-step: We estimate all the parameters of the vector \( \Theta \) using maximum likelihood estimation and the probability obtained in the E-step. More precisely, the process is giving as follow:

**Proposition 3.1** The estimation method is given by the following procedure.

1. Starting with an initial vector set \( \Theta^{(0)} := (\alpha_i^{(0)}, \beta_i^{(0)}, \sigma_i^2(0), \delta_i(0), \Pi^{(0)}) \), for all \( i \in S \). Fixed \( N \in \mathbb{N} \), the maximum number of iteration we authorize for this method (for the step 2 and 3 of EM-algorithm). And fixed a positive constant \( \varepsilon \) as a convergence constant for the estimated log likelihood function.

2. Assume that we are at the \( n + 1 \leq N \) steps, calculation in the previous iteration of the algorithm yields vector set \( \Theta^{(n)} := (\alpha_i^{(n)}, \beta_i^{(n)}, \sigma_i^2(0), \delta_i(0), \Pi^{(n)}) \).

**E-Step :** Filtered probability: For all \( i \in S \) and \( k = \{1, 2, \ldots, M\} \), evaluate the quantity

\[
P \left( X_{tk} = i | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right) = \frac{P \left( X_{tk}, r_{tk} | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)}{f \left( r_{tk} | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)}
\]

\[
= \frac{P \left( X_{tk} = i | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right) f \left( r_{tk} | X_{tk} = i ; \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)}{\sum_{j \in S} P \left( X_{tk} = j | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right) f \left( r_{tk} | X_{tk} = j ; \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)}
\]  (3.9)

with

\[
P \left( X_{tk} = i | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right) = \sum_{j \in S} P \left( X_{tk} = i, X_{tk-1} = j | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)
\]

\[
= \sum_{j \in S} P \left( X_{tk} = i, X_{tk-1} = j | \Theta^{(n)} \right) P \left( X_{tk-1} = j | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)
\]

\[
= \sum_{j \in S} \Pi_j^{(n)} \left( X_{tk-1} = j | \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right)
\]  (3.10)

where \( f \left( r_{tk} | X_{tk} = i ; \mathcal{F}_{tk-1}^r ; \Theta^{(n)} \right) \) is the density of the process \( r \) at time \( t_k \) conditional that the process is in regime \( i \in S \). Observed by (3.8), that given \( \mathcal{F}_{tk-1}^r \), the process \( r_{tk} \).
has a conditional Gaussian distribution with mean
\[ \alpha_i^{(n)} + (1 - \beta_i^{(n)}) r_{tk-1} \]
and standard deviation \( \sigma_i^{(n)} (r_{tk-1}) \delta_i^{(n)} \), whose density function is given by
\[
f \left( r_{tk} | X_{tk} = i; F_{tk-1}^r; \Theta^{(n)} \right) = \frac{1}{\sqrt{2\pi} \sigma_i^{(n)} |r_{tk-1}|^{\delta_i^{(n)}}} \exp \left\{ - \frac{\left( r_{tk} - (1 - \beta_i^{(n)}) r_{tk-1} - \alpha_i^{(n)} \right)^2}{2\sigma_i^{(n)} |r_{tk-1}|^{2\delta_i^{(n)}}} \right\}. \tag{3.11} \]

Smoothed probability: For all \( i \in S \) and \( k = \{M - 1, M - 2, \ldots, 1\}, \)
\[
P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) = \sum_{j \in S} \left( P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) P \left( X_{tk+1} = j | F_{tk}^r; \Theta^{(n)} \right) \Pi_{ij}^{(n)} \right) / P \left( X_{tk+1} = j | F_{tk}^r; \Theta^{(n)} \right). \tag{3.12} \]

**M-Step:** The maximum likelihood estimates \( \Theta^{(n+1)} \) for all model parameters is given, for all \( i \in S \), by
\[
\alpha_i^{(n+1)} = \frac{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}} \left( r_{tk} - (1 - \beta_i^{(n+1)}) r_{tk-1} \right)}{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}}},
\]
\[
\beta_i^{(n+1)} = \frac{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}} r_{tk-1} - B_1}{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}}},
\]
\[
\sigma_i^{2(n+1)} = \frac{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}} \left( r_{tk} - \alpha_i^{(n+1)} - (1 - \beta_i^{(n+1)}) r_{tk-1} \right)^2}{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right)},
\]
where
\[
B_1 = r_{tk} - r_{tk-1} - \frac{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}} (r_{tk} - r_{tk-1})}{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}}},
\]
\[
B_2 = \frac{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}} r_{tk-1}}{\sum_{k=2}^{M} P \left( X_{tk} = i | F_{tk}^r; \Theta^{(n)} \right) |r_{tk-1}|^{-2\delta_i^{(n)}}} - r_{tk-1}.
\]
The fourth parameter \( \delta_i^{(n+1)} \) is obtained by a numerical maximization of the likelihood function. Finally, the transition probabilities are estimated according to the following formula
\[
\Pi_{ij}^{(n+1)} = \frac{\sum_{k=2}^{M} P \left( X_{tk} = j | F_{tk}^r; \Theta^{(n)} \right) \Pi_{ij}^{(n)} P \left( X_{tk-1} = i | F_{tk-1}^r; \Theta^{(n)} \right)}{\sum_{k=2}^{M} P \left( X_{tk-1} = i | F_{tk-1}^r; \Theta^{(n)} \right)} \tag{3.13}.
\]
3. Denote by \( \Theta^{(n+1)} := \left( \alpha^{(n+1)}_i, \beta^{(n+1)}_i, \sigma^{2(n+1)}_i, \delta^{(n+1)}_i, \Pi^{(n+1)} \right) \), the new parameters of the algorithm and use it in step 2 until the convergence of the EM-algorithm. In fact, we stop the procedure if one of the following conditions are verified:

(a) We have done \( N \) times the procedure.
(b) The difference between the log likelihood at step \( n+1 \) and at step \( n \), satisfied the relation

\[
\log L(n + 1) - \log L(n) < \varepsilon.
\]

(3.14)

Remark 3.4  
(1) Proof of obtaining estimators \( \alpha^{(n+1)}_i, \beta^{(n+1)}_i \) and \( \sigma^{2(n+1)}_i \) are demonstrated in Lemma 3.1 of Janczura and Weron (2011). Formula to obtain all \( \Pi^{(n+1)}_{ij} \) are deduced from Kim (1994).

(2) Since the log likelihood function is increasing in each iteration of the procedure, we don’t need to take absolute value of the left hand side of inequality (3.14).

(3) In our case (i.e. regime switching), the standard log-likelihood function without regime switching \( \sum_{k=1}^M \log \left( f \left( r_{tk} | \mathcal{F}^{r}_{tk-1}; \Theta^{(n)} \right) \right) \) has to be weighted with the corresponding smoothed inference. Each observation \( r_{tk} \) belongs to the \( i \)th state with probability \( P \left( X_{tk} = i | \mathcal{F}^r_{tk-1}; \Theta^{(n)} \right) \). Hence, the regime switching log-likelihood function is:

\[
L(\Theta) = \sum_{k=1}^M \log \left( f \left( r_{tk} | \mathcal{F}^{r}_{tk-1}; \Theta^{(n)} \right) \right) P \left( X_{tk} = i | \mathcal{F}^r_{tk-1}; \Theta^{(n)} \right).
\]

(3.15)

(4) The standard error of each estimator is obtained by taking the square root of the \( (i, i) \)-th entry of the inverse of \( -\mathbb{H}(\hat{\Theta}) \) where \( \mathbb{H}(\hat{\Theta}) \) is the Hessian matrix defined by

\[
\mathbb{H}(\hat{\Theta})_{ij} = \left( \frac{\partial^2 L(\hat{\Theta})}{\partial \Theta_i \partial \Theta_j} \right), \quad i, j \in S.
\]

4  Main Findings and Models Comparison

Starting from a model with two regimes \( S = \{1, 2\} \), which represent two states of the economy: good and bad economic performance or a "normal" and crisis economy. We would like to discover:

i. Among regime switching models, which is the best regime switching model regarding foreign exchange rate data.
ii. Regime switching against non regime switching models, does the regime switching model offers better results?

Based on general model (2.2), we have 5 regime switching models: the more general model (RS-M) given by (2.2), the models (RS-CIR), (RS-V), (RS-R-GM) and (RS-CEV). We will estimate the parameters of each model on each foreign exchange rate data. To run the estimation procedure, we need to take an initial parameters $\Theta(0)$. For this, we take an initial regime distribution equals to $(\frac{1}{2}, \frac{1}{2})$. That is, beginning with the same probability in each regime, we take an initial transition matrix $\Pi(0)$ of the Markov chain $X$, such that, $\Pi_{11}^{(0)} = \Pi_{22}^{(0)} = \frac{1}{2}$. And finally, initial parameters values $(\alpha(0), \beta(0), \sigma(0), \delta(0))$ are given by the results of a global maximum likelihood estimation on the corresponding models without regime shift.

4.1 Parameters estimation

The Tables (2), (3), (4), (5) and (6) present values of the parameters estimation for each models.

<table>
<thead>
<tr>
<th></th>
<th>RS-M</th>
<th>RS-CIR</th>
<th>RS-V</th>
<th>RS-R-GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_1$</td>
<td>0.140761</td>
<td>0.500000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>1.093691</td>
<td>0.500000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
<tr>
<td>$\hat{\alpha}_1$</td>
<td>0.000340 (0.0054)</td>
<td>0.004508 (0.0060)</td>
<td>0.000099 (0.0055)</td>
<td>0.006285 (0.0056)</td>
</tr>
<tr>
<td>$\hat{\alpha}_2$</td>
<td>0.180753 (0.0548)</td>
<td>0.010250 (0.0118)</td>
<td>0.184747 (0.0542)</td>
<td>0.008511 (0.0111)</td>
</tr>
<tr>
<td>$\hat{\beta}_1$</td>
<td>-0.001088 (0.0046)</td>
<td>0.001781 (0.0050)</td>
<td>-0.001513 (0.0046)</td>
<td>0.003397 (0.0047)</td>
</tr>
<tr>
<td>$\hat{\beta}_2$</td>
<td>0.135189 (0.0405)</td>
<td>0.010276 (0.0096)</td>
<td>0.138021 (0.0398)</td>
<td>0.009030 (0.0097)</td>
</tr>
<tr>
<td>$\hat{\sigma}_1$</td>
<td>0.017021 (0.0000)</td>
<td>0.015023 (0.0000)</td>
<td>0.017343 (0.0000)</td>
<td>0.013763 (0.0000)</td>
</tr>
<tr>
<td>$\hat{\sigma}_2$</td>
<td>0.021777 (0.0001)</td>
<td>0.025140 (0.0001)</td>
<td>0.030345 (0.0001)</td>
<td>0.023305 (0.0001)</td>
</tr>
<tr>
<td>$\Pi_{11}^X$</td>
<td>0.993486</td>
<td>0.991849</td>
<td>0.993597</td>
<td>0.990263</td>
</tr>
<tr>
<td>$\Pi_{22}^X$</td>
<td>0.982696</td>
<td>0.983720</td>
<td>0.983147</td>
<td>0.976474</td>
</tr>
<tr>
<td>$\pi_1$</td>
<td>0.726504</td>
<td>0.666359</td>
<td>0.724667</td>
<td>0.707275</td>
</tr>
<tr>
<td>$\pi_2$</td>
<td>0.273496</td>
<td>0.333641</td>
<td>0.275333</td>
<td>0.292725</td>
</tr>
</tbody>
</table>

Table 2: Maximum Likelihood estimation results for the Euro-Dollars data with standard errors.
\begin{tabular}{|c|c|c|c|c|}
\hline
& RS-M & RS-CIR & RS-V & RS-R-GM \\
\hline
\hat{\delta}_1 & 8.688204 & 0.500000 & 0.000000 & 1.000000 \\
\hat{\delta}_2 & -1.957911 & 0.500000 & 0.000000 & 1.000000 \\
\hat{\alpha}_1 & 0.021557 (0.0067) & 0.002560 (0.0177) & 0.019259 (0.0060) & 0.005650 (0.0182) \\
\hat{\alpha}_2 & 0.005200 (0.0195) & 0.002820 (0.0009) & 0.001121 (0.0024) & 0.002973 (0.0009) \\
\hat{\beta}_1 & 0.001989 (0.0027) & 0.001503 (0.0025) & 0.002873 (0.0009) & 0.002873 (0.0009) \\
\hat{\beta}_2 & 0.001989 (0.0027) & 0.001503 (0.0025) & 0.002873 (0.0009) & 0.002873 (0.0009) \\
\hat{\sigma}_1 & 0.000000 (0.0000) & 0.006980 (0.0000) & 0.003307 (0.0000) & 0.002652 (0.0000) \\
\hat{\sigma}_2 & 0.871350 (0.0803) & 0.001249 (0.0000) & 0.018447 (0.0000) & 0.000476 (0.0000) \\
\hline
\hat{\Pi}_{11}^X & 0.881806 & 0.984915 & 0.957572 & 0.983178 \\
\hat{\Pi}_{22}^X & 0.933623 & 0.956565 & 0.985652 & 0.954240 \\
\pi_1 & 0.359628 & 0.742231 & 0.252713 & 0.731195 \\
\pi_2 & 0.640372 & 0.257769 & 0.747287 & 0.268805 \\
\hline
\end{tabular}

Table 3: Maximum Likelihood estimation results for the Dollars-Yuan data with standard errors.

\begin{tabular}{|c|c|c|c|c|}
\hline
& RS-M & RS-CIR & RS-V & RS-R-GM \\
\hline
\hat{\delta}_1 & 1.797463 & 0.500000 & 0.000000 & 1.000000 \\
\hat{\delta}_2 & -0.083423 & 0.500000 & 0.000000 & 1.000000 \\
\hat{\alpha}_1 & 14.918395 (4.6358) & 7.525180 (3.0508) & 17.774657 (5.9303) & 6.889107 (3.0148) \\
\hat{\alpha}_2 & -0.363333 (0.7003) & 0.004430 (0.7695) & 0.199365 (0.6992) & 0.088266 (0.7828) \\
\hat{\beta}_1 & 0.122275 (0.0390) & 0.068257 (0.0266) & 0.143222 (0.0455) & 0.062993 (0.0271) \\
\hat{\beta}_2 & -0.003321 (0.0054) & -0.00544 (0.0060) & -0.00105 (0.0054) & 0.00099 (0.0062) \\
\hat{\sigma}_1 & 0.000813 (0.0000) & 0.397770 (0.0254) & 5.199643 (5.8142) & 0.035694 (0.0002) \\
\hat{\sigma}_2 & 3.143649 (0.6962) & 0.183135 (0.0024) & 2.129099 (0.3167) & 0.016173 (0.0000) \\
\hline
\hat{\Pi}_{11}^X & 0.890571 & 0.966836 & 0.899831 & 0.970122 \\
\hat{\Pi}_{22}^X & 0.988872 & 0.995772 & 0.991836 & 0.996123 \\
\pi_1 & 0.092303 & 0.113062 & 0.075357 & 0.114846 \\
\pi_2 & 0.907697 & 0.886938 & 0.924643 & 0.885154 \\
\hline
\end{tabular}

Table 4: Maximum Likelihood estimation results for the Euro-Yen data with standard errors.

Where the quantities \( \pi_1 \) and \( \pi_2 \) represent the stationary distribution of the Markov chain given by

\[
(\pi_1, \pi_2) = \left( \frac{1 - \hat{\Pi}_{22}^X}{2 - \hat{\Pi}_{11}^X - \hat{\Pi}_{22}^X}, \frac{1 - \hat{\Pi}_{11}^X}{2 - \hat{\Pi}_{11}^X - \hat{\Pi}_{22}^X} \right).
\]
Furthermore, it is clear that the presence of mean reverting effect (i.e. the parameter $\hat{\beta}_i$) in each data can’t be rejected.

For each data, taking the model which obtain the higher log likelihood value (see Table 5 and Table 6).

<table>
<thead>
<tr>
<th></th>
<th>RS-M</th>
<th>RS-CIR</th>
<th>RS-V</th>
<th>RS-R-GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_1$</td>
<td>-4.137988</td>
<td>0.500000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>3.049220</td>
<td>0.500000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.018711 (0.0066)</td>
<td>0.008197 (0.0045)</td>
<td>0.008158 (0.0046)</td>
<td>0.006397 (0.0032)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.004255 (0.0043)</td>
<td>0.004782 (0.0074)</td>
<td>0.007253 (0.0090)</td>
<td>0.009330 (0.0137)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.286561 (0.0098)</td>
<td>0.011400 (0.0066)</td>
<td>0.011340 (0.0067)</td>
<td>0.008588 (0.0045)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.004260 (0.0061)</td>
<td>0.005297 (0.0094)</td>
<td>0.008402 (0.0110)</td>
<td>-0.016901 (0.0202)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.001188 (0.0000)</td>
<td>0.008090 (0.0000)</td>
<td>0.006728 (0.0000)</td>
<td>0.010927 (0.0000)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.022192 (0.0000)</td>
<td>0.016107 (0.0000)</td>
<td>0.014717 (0.0000)</td>
<td>0.023550 (0.0001)</td>
</tr>
</tbody>
</table>

Table 5: Maximum Likelihood estimation results for the Euro-Livres data with standard errors.

<table>
<thead>
<tr>
<th></th>
<th>RS-M</th>
<th>RS-CIR</th>
<th>RS-V</th>
<th>RS-R-GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_1$</td>
<td>4.265033</td>
<td>0.500000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>-0.878781</td>
<td>0.500000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.281880 (0.1386)</td>
<td>0.327537 (0.1549)</td>
<td>0.116381 (0.0981)</td>
<td>0.313203 (0.1560)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.128251 (0.0825)</td>
<td>0.082623 (0.0575)</td>
<td>0.048449 (0.0591)</td>
<td>0.081319 (0.0543)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.036466 (0.0180)</td>
<td>0.040929 (0.0180)</td>
<td>0.013811 (0.0106)</td>
<td>0.039361 (0.0184)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.012602 (0.0084)</td>
<td>0.007405 (0.0061)</td>
<td>0.003647 (0.0064)</td>
<td>0.007305 (0.0059)</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.000018 (0.0000)</td>
<td>0.065897 (0.0005)</td>
<td>0.179799 (0.0028)</td>
<td>0.022532 (0.0001)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>1.049041 (0.0851)</td>
<td>0.041546 (0.0001)</td>
<td>0.107939 (0.0012)</td>
<td>0.013496 (0.0001)</td>
</tr>
</tbody>
</table>

Table 6: Maximum Likelihood estimation results for the Euro-Yuan data with standard errors.
(13), we can plot Figures 10, 11 and 12 which give the evolution of the values of parameters in each regime during all the estimation procedure. In all the figures in this subsection, the regime 1 will be in color blue and regime 2 in color red. It is obvious that, in all the cases, convergence of the EM-Algorithm happens in less than 40 steps.

Figures 13, 14 and 15 give the smoothed and the filtered probabilities basing on real data. Usually, “smoothed probabilities allow for the most information ex-post analysis of the data, while filtered probabilities are useful for forecasting” as stated by Calvet and Fisher (2008).

4.2 Value of the Regime duration time.

Basing on and using the above estimations and classifications, we can reproduce the exchange rates, whose origin real data are presented at the beginning of last section. Figures 4, 5 and 6 give trajectories of foreign exchange rate with respect to the value of the current regime. Again, we take the model which obtain the higher log likelihood value (see Table (13)).

The left graph in Figure 4 indicates clearly two significantly different time periods. The first one, in blue, corresponds to an increasing time period where the value of the change is better for Euro zone. This can be seen from the value of estimating parameters. Indeed, in this regime the speed of adjustment parameter \( \hat{\beta} \) is close to zero (\( \hat{\beta}_1 = 0.001294 \)) which means that the Euro-dollar exchange rate dynamic has a mean reversion close to zero. The second one, in red, corresponds to a more volatile time period where the volatility in this regime equals \( \hat{\sigma} = 0.025852 \) against \( \hat{\sigma} = 0.014986 \) as in regime 1. This shows an increasing of the volatility which equals to [4]

---

4For more detail, see for example, Calvet and Fisher (2008).
72.51%. Hence, all the crisis periods fall into this regime which are the periods (1) between January 2000 and March 2001 and (2) from the autumn 2008 global financial crisis afterward.\footnote{The first Euro crisis, as addressed by BusinessWeek on October 2, 2000, that “The euro is in crisis, and as it goes, so may go the future of the New Europe. After a flawless and much-acclaimed debut just 20 months ago, Europe’s new single currency has lost more than 25% of its value against the dollar—and there is still no bottom in sight.” And this down-move of Euro to Dollar ended at the starting of recession in the U.S. economy from March 2001. In deed, the NBER’s Business Cycle Dating Committee has determined that a peak in business activity occurred in the U.S. economy in March 2001. That is the end of an expansion and the beginning of a recession. As this committee also announced later on March 17, 2003, that this recession finished in 8 months, that is, the beginning of 2002. However, this expansion did not last too long, global financial crisis which started in the U.S. in December 2007, resulted in the collapse of large financial institutions, the bailout of banks by national governments and downturns in stock markets around the world. It contributed to a significant decline in economic activity, leading to a severe global economic recession in 2008-2009. The financial crisis was triggered by a complex interplay of valuation and liquidity problems in the United States banking system in 2008.}

Similar finding is also presented in the right graph of Figure\cite{4} which reads that there are two different time periods: regime 1 corresponding to a time period where the value of the change is better for Dollar zone; and a second regime which corresponds to stable or constant period as the crisis-mode policy taken by the People’s Bank of China.\footnote{It is worth noticing that the financial crisis which broke out in the United States in 2008 shot the global financial markets and dented investment confidence. The People’s Bank of China then took a crisis-mode policy by stopping the gradual appreciation of the RMB against dollar: The yuan/dollar rate has been stable at about 6.86± 0.3 percent since July 2008. Zhou Xiaochuan, governor of the central bank, said in March 2010 that the exchange rate policy China took amid the crisis was part of the government’s stimulus packages, and would exit “sooner or later” along with other crisis-measures. China’s June 20, 2010, announcement that it would allow more flexibility in its yuan exchange rate regime meant an end to the crisis-mode policy the government took to cushion the blow from the global financial crisis. Zhao said when the RMB exchange rate regime becomes more market-oriented, China’s export businesses should take more responsibilities and become more self-reliant.}

Furthermore, we can remark that the volatility of this foreign exchange rate is very close to zero: 0.006980 in regime 1 and 0.001249 in regime 2.

For the Euro/Yen calibration, we can see on the left graph of Figure\cite{5} that is the case where one regime corresponds to standard dynamic and the other one catches the spikes of the dynamics. The regime 1 (blue color) documents the two crisis time periods mentioned above.

This crisis regime has a very high value for the speed of adjustment parameter, \(\hat{\beta}_1 = 0.068287\). This is typically a spike regime where the value of the foreign exchange rates change brutally, then returns quickly to the mean value. And of course the volatility in the crisis regime is bigger than the volatility in the standard economy regime. \(\hat{\sigma}_1 = 0.396827\) against \(\hat{\sigma}_2 = 0.182396\), this corresponds to an increasing of 117.56%.

For the Euro/Livre calibration shows on the right graph of Figure\cite{5} that regime 2, in red, corresponds to a crisis time period. Thus, the autumn 2008 crisis and the time period between
January 2000 and March 2001 fall in this regime.

The foreign exchange rate dynamic in this crisis time period has, again, a higher estimated volatility than in the standard regime (in blue). Indeed, $\hat{\sigma}_2 = 0.016095$ and $\hat{\sigma}_1 = 0.007954$, this is an increasing of 102.35% of the volatility. We observe again that the speed of adjustment parameter is bigger in the crisis regime, $\hat{\beta}_2 = 0.005374$ against $\hat{\beta}_1 = 0.002115$ (+154.09%).

Finally, the Euro/Yuan calibration presented in Figure 6 states the same regime cut as Euro/Livre foreign exchange rate. But here the impact of the crisis is less pronounced in term of volatility, only +60.70% than in term of the speed of adjustment +334.57%. 

Figure 5: Foreign exchange rate with respect to the regime state for: on left: Euro/Yen and on right: Euro/Livres.

Figure 6: Foreign exchange rate with respect to the regime state for Euro/Yuan.
4.3 Good Classification Measures

An ideal model is that classifying regimes sharply and having smoothed probabilities which are either close to zero or one. In order to measure the quality of regime classification, we propose two measures:

(1) The regime classification measure (RCM) introduced by Ang and Bekaert (2002) and generalized for multiple state by Baele (2005).

(2) The smoothed probability indicator.

The detail definition and formulation are the following:

**Regime classification measure:** Let $K (> 0)$ be the number of regimes, the RCM statistics is then given by

$$RCM(K) = 100 \left( 1 - \frac{K}{K-1} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{K} \left( P \left( X_t = i \mid \mathcal{F}_t ; \hat{\Theta} \right) - \frac{1}{K} \right)^2 \right)$$  \hspace{1cm} (4.16)

where the quantity $P \left( X_{tk} = i \mid \mathcal{F}_{tk} ; \Theta^{(n)} \right)$ is the smoothed probability given in (3.12) and $\hat{\Theta}$ is the vector parameter estimation results. The constant serves to normalize the statistic to be between 0 and 100. Good regime classification is associated with low RCM statistic value: a value of 0 means perfect regime classification and a value of 100 implies that no information about regimes is revealed.

**Smoothed probability indicator:** A good classification for data can be also seen when the smoothed probability is less than 0.1 or greater than 0.9. Then this means that the data at time $t \in [0, T]$ is with a probability higher than 0.9 in one of regimes for the 10% error and higher than 0.95 for the 5% error. We will call this percentage as the smoothed probability indicator with $p\%$ error and we will denote here by $P_{p\%}$.

In the following, we evaluate the RCM statistics and the smoothed probability indicators for all foreign exchange rates data and all models. The results are stated in Table 7.

Table 7 clearly documents that for all foreign exchange rate data the regime classification measure (RCM) is close to zero and the smoothed probability indicator is in most cases close to 90%. This indicates that the two regimes obtained via the EM-algorithm classify the data in a very good way. And hence, there exists different regimes in the dynamics of foreign exchange rate. Therefore, it is better to take into account the existence of this regime switching in modeling foreign exchange rate dynamics.
<table>
<thead>
<tr>
<th></th>
<th>RS-M</th>
<th>RS-CIR</th>
<th>RS-V</th>
<th>RS-R-GM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euro/Dollars</td>
<td>11.03</td>
<td>18.53</td>
<td>10.51</td>
<td>19.35</td>
</tr>
<tr>
<td>Euro/Dollars</td>
<td>Perc (10%)</td>
<td>89.43</td>
<td>82.60</td>
<td>88.99</td>
</tr>
<tr>
<td>Yuan/Dollars</td>
<td>20.46</td>
<td>6.25</td>
<td>5.72</td>
<td>7.70</td>
</tr>
<tr>
<td>Yuan/Dollars</td>
<td>Perc (10%)</td>
<td>75.69</td>
<td>94.80</td>
<td>93.06</td>
</tr>
<tr>
<td>Euro/Yen</td>
<td>10.71</td>
<td>6.61</td>
<td>7.72</td>
<td>6.43</td>
</tr>
<tr>
<td>Euro/Yen</td>
<td>Perc (10%)</td>
<td>89.65</td>
<td>95.81</td>
<td>93.17</td>
</tr>
<tr>
<td>Euro/Livre</td>
<td>25.95</td>
<td>8.77</td>
<td>8.60</td>
<td>6.85</td>
</tr>
<tr>
<td>Euro/Livre</td>
<td>Perc (10%)</td>
<td>70.92</td>
<td>92.29</td>
<td>90.97</td>
</tr>
<tr>
<td>Euro/Yuan</td>
<td>22.18</td>
<td>22.32</td>
<td>28.46</td>
<td>20.24</td>
</tr>
<tr>
<td>Euro/Yuan</td>
<td>Perc (10%)</td>
<td>71.14</td>
<td>77.09</td>
<td>71.14</td>
</tr>
</tbody>
</table>

Table 7: RCM statistics and percentage given by the smoothed probability indicator for 10%: $P^{10\%}$.

### 4.4 Model fitting

In this subsection, some interesting tests are done to show which is the best regime switching model in the sense that it fits better foreign exchange rate data. For this aim, we evaluate the log likelihood values of each models obtained in the calibration. Thus, a likelihood maximization procedure is used. Furthermore, we calculate the Akaike information criterion (AIC)\(^7\) and the Bayesian information criterion (BIC)\(^8\) which are given by

\[
AIC = -2 \ln(L(\hat{\Theta})) + 2 \times k,
\]

and

\[
BIC = -2 \ln(L(\hat{\Theta})) + k \ln(n)
\]

where \(L(\hat{\Theta})\) is the log-likelihood value obtained with the estimated parameters, \(k\) is the degree of freedom of each models and \(n\) the number of data for each historical foreign exchange rates.

All the results are stated in Tables 13, 14 and 15 at the end of the paper. The decision rule is the following: Given a set of candidate models for the data, the preferred model is the one

---

\(^7\)The Akaike information criterion is a measure of the relative goodness of fit of a statistical model. It was developed by Hirotugu Akaike, under the name of “an information criterion” (AIC), and was first published by Akaike (1974) in \[1\].

\(^8\)The Bayesian information criterion (BIC) or Schwarz criterion is a criterion for model selection among a finite set of models. It is based, in part, on the likelihood function, and it is closely related to Akaike information criterion (AIC). It was developed by Gideon E. Schwarz (1978) in \[25\].
with the minimum AIC or BIC value. Hence, we can see from Table [14 and 15] that in all the cases, regime switching models give a less AIC and BIC values than corresponding non regime switching models. This demonstrates that regime switching models fit better the data than non regime switching model. Moreover, Table [13] gives the log likelihood values of each models and it is clear that this value is always higher for regime switching models than non regime switching models. This confirms the previous statement.

Let us now check which is the better regime switching model. To make a choice, we have to take into account two things. Firstly, the log likelihood value given by the model. Indeed, higher in this value and better the fit of the data, better is the model. But, secondly, we have to weight these values with the values given by the (RCM) in Table [7] which measures the good classification of the data. Indeed, even if a model has a higher log likelihood value it is important that its RCM to be close to zero. As an example, if we took the Euro/Yen results, the higher log likelihood value is obtained by the (RS-M) model (-1012.42) but this value is closed to the one given by the (RS-V) model (-1015.83). So if comparing their (RCM), we see that (RS-M) model has a (RCM) equals to 10.71 but the (RS-V) has a value equals to 7.72. In conclusion, the choice of the (RS-V) model is better to fit this data since is obtained a log likelihood value close to the best model but obtained a better classification of the regime. Regarding all the results, the choice of the (RS-CIR) or (RS-R-GM) seems to be the best model to fit well data and obtain a good classification of the data with significant regime periods.

4.5 Impact of regime switching in each parameters

In this section, we would like to see what would happen if one of the three parameters of the model (2.2) do not depend on the regime switching process. To answer this question, we run another simulation and show in Table [8] for the Euro/Dollars foreign exchange rate data and the RS-CIR model, by assuming that one of the parameters don’t depend on the regime switching. This exercise gives a log likelihood value less than the RS-CIR model. This means that the CIR model, where all parameters depend on the regime switching process, fits better to the real data. Therefore, assuming that the speed of adjustment process $\beta$ or the volatility parameter $\sigma$ are equal in each regime give a worse fit on data than in the RS-CIR model.

Furthermore, similar conclusion can be obtained if we take another foreign exchange rate data or another regime switching model.
### 4.6 Three Regimes case

One step further from the previous subsections, we would like to see what would be the outcomes if there exist three Markov switching regimes. One could capture “normal” economic dynamics, a second presents for “crisis” and the last one states “good” economic performance. Can more regimes capture more precisely the economic and financial dynamics, what would be the gain and what could be the lost if more regimes are introduced?

We estimate models in the special case of 3 regimes (i.e. \( S = \{1, 2, 3\} \)) and evaluate the corresponding Regime classification measure given in (4.16) for \( K = 3 \), the smoothed probability indicator with 10% errors. The results are stated in Table 9.

It is clear from the Table 9 that the regime classification measure (RCM) is bigger in three regimes cases than in the two regimes case for almost all the data and regime switching models. Moreover, we notice that in many cases the three regimes cases gives very bad classification with respect to the two regimes cases. Indeed, in the case of Euro/Livre, using the (RS-CIR) model, only 34.36% of the data are good classified for 10% error under the three regimes cases against 92.29% with the two regimes cases.

Nevertheless, we also observe that in the case of Euro/Dollars, the RCM value obtained by the (RS-CIR) model with 3 regimes is smaller than that with 2 regimes: 10.83 against 18.53. This results is confirmed by the value of the smoothed probability indicator. Indeed, for 10% error, 90.60% of the data are good classified in the 3 regimes model while only 82.60% in the 2 regimes case.

As to the Euro/Dollar exchange rate, Figure 7 displays that the three regime case separates better the second regime (in red) than in the two regimes case. The three regimes cases differentiate the two level of the more volatile time periods (the red regime obtained in the two regimes cases): the first one, in green, correspond to the lower value time period and the second one, in red, the higher value time period. Hence, this two periods are differentiate by long mean level value \( \frac{\alpha_2}{\beta_2} = \frac{0.1833}{0.1371} = 1.3370 \) for the regime 2, in red, and \( \frac{\alpha_3}{\beta_3} = \frac{0.1153}{0.1280} = 0.9008 \) for the regime 3.
Table 9: RCM statistics in the case of two and three regimes and percentage given by the smoothed probability indicator for 10% in the case of 2 and 3 regimes.

The truth that three regime case gives better calibration results than those of the two regimes is only due to this special form of the data’s plot. If we do the same calibration for the Euro/Yuan exchange rate models with the (RS-CIR) too, we don’t obtain better results with three regimes. Actually, it’s even worse, as we saw in Table 9, we find a RCM value of 63.48 against 8.77 and the good classify only 34.36% of the data against 92.29% in the two regimes case. These results can be see in the figure calibration results given in Figure [8] where we clearly observe that it seems to be very difficult to find significant economic or financial interpretations of this classification.
Figure 7: Euro/Dollar foreign exchange rate models with (RS-CIR) with respect to the regime state for: on left: two regimes and on right: three regimes.

Figure 8: Euro/Livres foreign exchange rate models with (RS-CIR) with respect to the regime state for: on left: two regimes and on right: three regimes.
As a conclusion, two regimes seems to be the best choice because it gives significant better results in most cases. In the cases where results are better with three regimes, the gain in term of classification is very small regarding the lost in the cases where three regimes give worst results. Hence, it’s better to always take two regimes rather than three.

5 Forecasting

Given observations \((r_{t_0}, r_{t_1}, \ldots, r_{t_k})\) taken at time \(t_k, k \in \{0, \ldots, M\}\), we can define the mean-squared error (MSE) for a given model as

\[
MSE = \frac{1}{M+1} \sum_{i=1}^{M+1} (\hat{r}_{ti} - r_{ti})^2
\]  

(5.17)

where \(\hat{r}_{ti}\) is the predicted value of \(r_{ti}\) given all information available up to time \(t_{i-1}\) (i.e., \(\mathcal{F}_{t_{i-1}}^r\)). Moreover, we recall first that \(\Delta_t = t_i - t_{i-1}\) is a constant and

\[
\hat{r}_{ti+\Delta_t} = \mathbb{E}[r_{ti+\Delta_t}|\mathcal{F}_{ti}^r], \quad \forall i \in \{0, \ldots, M - 1\}
\]

is a function of \(r_{ti}\) and the current estimates of all the parameters. By (3.7), the distribution of \(r_{ti+\Delta_t}\) is normal with mean

\[
r_{ti} + \left(\hat{\alpha}_{ti+1} - \hat{\beta}_1 r_{ti}\right) \Delta_t.
\]

Hence, for two state regime switching models, we get

\[
\mathbb{E}[r_{ti+\Delta_t}|\mathcal{F}_{ti}^r] = \pi^1 \cdot \left(\mathbb{E}[r_{ti+\Delta_t}|X_{ti+\Delta_t} = 1; \mathcal{F}_{ti}^r] + \mathbb{E}[r_{ti+\Delta_t}|X_{ti+\Delta_t} = 0; \mathcal{F}_{ti}^r] \right).
\]

So, it yields

\[
\mathbb{E}[r_{ti+\Delta_t}|\mathcal{F}_{ti}^r] = \pi^1 \cdot r_{ti} + \left(\hat{\alpha}_1 - \hat{\beta}_1 r_{ti}\right) \Delta_t) + \pi^2 \cdot \left(\mathbb{E}[r_{ti+\Delta_t}|X_{ti+\Delta_t} = 1; \mathcal{F}_{ti}^r] + \mathbb{E}[r_{ti+\Delta_t}|X_{ti+\Delta_t} = 0; \mathcal{F}_{ti}^r] \right).
\]

In Table 10, 11 and 12 we display the MSE statistics for ours different models and for each foreign exchange rates data. Keeping the last 60 values of \(r\) for the forecast estimation, we estimate all the model parameters using the first \(M - 60\) values of \(r\). In other words, this \(M - 60\) first data serves as a training data set in order to learn the parameters of each model. Then, starting from the \(M - 60\) values, perform the one step ahead, the two step ahead and the three step ahead predictions via using these parameters’ estimation.

These results are remarkable close. Nevertheless, we can see that the (RS-R-GM) or (RS-CIR) are the two models which give often the best results. However, we can remark that taking a foreign exchange rate data, the model which minimizes the forecasting error is the same for
Table 10: MSE comparison of each model and data for One Ahead Forecast.

<table>
<thead>
<tr>
<th></th>
<th>Euro/Dollars</th>
<th>Yuan/Dollars</th>
<th>Euro/Yen</th>
<th>Euro/Livre</th>
<th>Euro/Yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RS-M)</td>
<td>0.6698 × 10^{-3}</td>
<td>0.3328 × 10^{-3}</td>
<td>4.2696</td>
<td>0.1183 × 10^{-3}</td>
<td>0.0278</td>
</tr>
<tr>
<td>(RS-CIR)</td>
<td>0.6839 × 10^{-3}</td>
<td>0.3131 × 10^{-3}</td>
<td>4.2648</td>
<td>0.1182 × 10^{-3}</td>
<td>0.0275</td>
</tr>
<tr>
<td>(RS-V)</td>
<td>0.6698 × 10^{-3}</td>
<td>0.6079 × 10^{-3}</td>
<td>4.3282</td>
<td>0.1181 × 10^{-3}</td>
<td>0.0276</td>
</tr>
<tr>
<td>(RS-R-GM)</td>
<td>0.6835 × 10^{-3}</td>
<td>0.3130 × 10^{-3}</td>
<td>4.3499</td>
<td>0.1180 × 10^{-3}</td>
<td>0.0275</td>
</tr>
</tbody>
</table>

Table 11: MSE comparison of each model and data for Two Ahead Forecast.

<table>
<thead>
<tr>
<th></th>
<th>Euro/Dollars</th>
<th>Yuan/Dollars</th>
<th>Euro/Yen</th>
<th>Euro/Livre</th>
<th>Euro/Yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RS-M)</td>
<td>0.0012</td>
<td>0.0007</td>
<td>9.6947</td>
<td>0.1992 × 10^{-3}</td>
<td>0.05</td>
</tr>
<tr>
<td>(RS-CIR)</td>
<td>0.0012</td>
<td>0.0006</td>
<td>9.6535</td>
<td>0.1986 × 10^{-3}</td>
<td>0.049</td>
</tr>
<tr>
<td>(RS-V)</td>
<td>0.0012</td>
<td>0.0018</td>
<td>9.9092</td>
<td>0.1982 × 10^{-3}</td>
<td>0.0489</td>
</tr>
<tr>
<td>(RS-R-GM)</td>
<td>0.0012</td>
<td>0.0006</td>
<td>9.9624</td>
<td>0.1978 × 10^{-3}</td>
<td>0.0489</td>
</tr>
</tbody>
</table>

Table 12: MSE comparison of each model and data for Three Ahead Forecast.

<table>
<thead>
<tr>
<th></th>
<th>Euro/Dollars</th>
<th>Yuan/Dollars</th>
<th>Euro/Yen</th>
<th>Euro/Livre</th>
<th>Euro/Yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RS-M)</td>
<td>0.0015</td>
<td>0.0009</td>
<td>14.9629</td>
<td>0.2685 × 10^{-3}</td>
<td>0.0684</td>
</tr>
<tr>
<td>(RS-CIR)</td>
<td>0.0016</td>
<td>0.0008</td>
<td>14.8482</td>
<td>0.2669 × 10^{-3}</td>
<td>0.0663</td>
</tr>
<tr>
<td>(RS-V)</td>
<td>0.0015</td>
<td>0.0034</td>
<td>15.4086</td>
<td>0.2662 × 10^{-3}</td>
<td>0.0669</td>
</tr>
<tr>
<td>(RS-R-GM)</td>
<td>0.0016</td>
<td>0.0008</td>
<td>15.4871</td>
<td>0.2651 × 10^{-3}</td>
<td>0.0660</td>
</tr>
</tbody>
</table>

all forecasting delay. Moreover, this best forecasting model corresponds to the best estimated model (see subsection 4.4) which gives at the same time a higher log likelihood value and a smaller (RCM) statistic.

6 Conclusion

Theoretically, empirically, politically and academically, there have been enormous studies and analysis about currency exchange rates, the corresponding effects, the courses of volatilities, and so on. The previous findings are mixed and each has its own focus. In this present framework, we initially introduce a continuous time mean reverting Cox-Ingersoll-Ross model with regime switching parameters to model foreign exchange rates in order to understand whether
it is possible that stochastic exchange rates catch the real world regimes switching, such as financial crisis and economy taking off, and how fast the exchange rates react to these kind of changes. We have clearly documented that mean reverting regime switching Cox-Ingersoll-Ross model fits much better foreign exchange rate data than non regime switching models. Moreover, the regime switching process (i.e. a homogeneous continuous time Markov chain on a finite state space \( S \)) allows us to highlight some economic and financial time period where dynamics of foreign exchange rates are significantly different.

Furthermore, we extend the expectation-maximization algorithm, with filtering and smoothing technique to smooth out noisy data, to calibrate regime switching model and show that only a few number of step is needed to obtain a very good calibration. Thus, this refined and modified filtering and smoothing algorithm could be used for other studies and tests of time series related topics, such as the macroeconomic effects of tax changes and some more detail can also be found in Romer and Romer (2010).
References


<table>
<thead>
<tr>
<th>Model</th>
<th>Euro/Dollars</th>
<th>Yuan/Dollars</th>
<th>Euro/Yen</th>
<th>Euro/Livre</th>
<th>Euro/Yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RS-M)</td>
<td>1135.51</td>
<td>-</td>
<td>-1012.42</td>
<td>1517.18</td>
<td>236.13</td>
</tr>
<tr>
<td>(RS-CIR) and (CIR)</td>
<td>1138.89</td>
<td>1098.88</td>
<td>-1023.79</td>
<td>1513.04</td>
<td>231.94</td>
</tr>
<tr>
<td>(RS-V) and (V)</td>
<td>1135.16</td>
<td>1092.37</td>
<td>-1015.83</td>
<td>1509.58</td>
<td>230.84</td>
</tr>
<tr>
<td>(RS-R-GM) and (R-GM)</td>
<td>1138.26</td>
<td>1085.60</td>
<td>-1026.06</td>
<td>1514.88</td>
<td>232.43</td>
</tr>
</tbody>
</table>

Table 13: Log likelihood values of each model and data.

<table>
<thead>
<tr>
<th>Model</th>
<th>Euro/Dollars</th>
<th>Yuan/Dollars</th>
<th>Euro/Yen</th>
<th>Euro/Livre</th>
<th>Euro/Yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RS-M)</td>
<td>-2251.02</td>
<td>-1874.00</td>
<td>2044.84</td>
<td>-3014.35</td>
<td>-452.27</td>
</tr>
<tr>
<td>(RS-CIR) and (CIR)</td>
<td>-2251.79</td>
<td>-2191.75</td>
<td>-1778.70</td>
<td>-3010.08</td>
<td>-447.89</td>
</tr>
<tr>
<td>(RS-V) and (V)</td>
<td>-2254.31</td>
<td>-2178.75</td>
<td>-1779.94</td>
<td>-3003.17</td>
<td>-445.67</td>
</tr>
<tr>
<td>(RS-R-GM) and (R-GM)</td>
<td>-2260.52</td>
<td>-2165.21</td>
<td>-1778.94</td>
<td>-3013.75</td>
<td>-448.87</td>
</tr>
</tbody>
</table>

Table 14: Akaike information criterion (AIC) of each model and data.

<table>
<thead>
<tr>
<th>Model</th>
<th>Euro/Dollars</th>
<th>Yuan/Dollars</th>
<th>Euro/Yen</th>
<th>Euro/Livre</th>
<th>Euro/Yuan</th>
</tr>
</thead>
<tbody>
<tr>
<td>(RS-M)</td>
<td>-2209.84</td>
<td>-1837.37</td>
<td>2086.02</td>
<td>-2973.17</td>
<td>-411.09</td>
</tr>
<tr>
<td>(RS-CIR) and (CIR)</td>
<td>-2218.84</td>
<td>-2179.40</td>
<td>-1749.40</td>
<td>-2977.14</td>
<td>-414.94</td>
</tr>
<tr>
<td>(RS-V) and (V)</td>
<td>-2221.37</td>
<td>-2166.40</td>
<td>-1750.64</td>
<td>-2970.23</td>
<td>-412.73</td>
</tr>
<tr>
<td>(RS-R-GM) and (R-GM)</td>
<td>-2227.58</td>
<td>-2152.85</td>
<td>-1749.63</td>
<td>-2980.80</td>
<td>-415.92</td>
</tr>
</tbody>
</table>

Table 15: Bayesian information criterion (BIC) of each model and data.
Figure 10: Evolution of the calibrated parameters values: on left: Euro/Dollars and on right: Yuan/Dollars.

Figure 11: Evolution of the calibrated parameters values: on left: Euro/Yen and on right: Euro/Livres.
Figure 12: Evolution of the calibrated parameters values for Euro/Yuan.

Figure 13: Smoothed and Filtered probabilities for: on left: Euro/Dollars and on right: Yuan/Dollars.

Figure 14: Smoothed and Filtered probabilities for: on left: Euro/Yen and on right: Euro/Livres.
Figure 15: Smoothed and Filtered probabilities for Euro/Yuan.