HEC MONTRÉAL

Regime-Switching Correlations with Exogenous Economic Variables

par
Paul Kelendji

Directeurs de recherche: David Ardia et Geneviève Gauthier

Sciences de la gestion
(Ingénierie Financière)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M. Sc.)

Décembre 2023

Contents

1 Introduction

Modelling dynamic correlations in financial markets is crucial for risk management and asset allocation. A simplistic approach would be to assume constancy over time, represented by a fixed correlation matrix. However, this assumption falls short during periods of market stress or rapid economic change, when correlations can change abruptly. Such a static perspective risks underestimating the probability and impact of extreme events, potentially leading to inadequate risk management strategies and portfolio allocations.
To better capture the evolving nature of market conditions, we can employ regime-switching models, which allow the parameters governing the behavior of financial time series to change when the regime, or the underlying state of the market, changes. The Markov-Switching model, a particular type of regime-switching model, employs a Markov chain to represent the stochastic process of moving between different regimes. In financial applications, these regimes might correspond to periods of high or low market volatility, growth or recession in the economy, or bull or bear market conditions.
This thesis delves into the dynamics of daily exchange rates among major world currencies—the Japanese yen (JPY), Euro (EUR), and Canadian dollar (CAD) against the U.S. dollar (USD). By employing and extending the Regime-Switching Dynamic Correlation (RSDC) model by Pelletier (2006), which incorporates the principles of Markov-Switching to capture the dynamic nature of correlations, we aim to provide a more granular and precise representation of the intricate financial interdependencies at play.

1.1 Survey of Literature

Regime-switching models have evolved substantially since their inception, serving as vital tools in economic and financial analysis. This section traces the lineage of these models, highlighting significant contributions and innovations that have shaped current methodologies.

1.1.1 Early Developments and Applications (1989-2000)

The early literature on Markov-switching models, notably beginning with Hamilton's seminal work (Hamilton, 1989; Hamilton, 1990), marked an era where these models were primarily used for analyzing business cycles and financial data series. These models conceptualized the notion of regime shifts, elucidating the dynamic nature of various economic variables through states or regimes that capture the underlying processes and potential future trends. Key applications during this period extended across business cycles, interest rates, and asset pricing, with notable studies by Sola and Driffill (1994), Asea and Blomberg (1998), and Town (1992) deepening the understanding of regime shifts in their respective domains.

1.1.2 Advancements and Diversification (2000-Present)

Post-2000, the literature on Markov-switching models diversified into addressing more complex models and technical issues. Studies like Akintug and Rasmussen (2005) expanded the application to hydrologic time series, while others like Pei Yin (2007) applied Hidden Markov

Models (HMM) to volatility estimation and price prediction, indicating the model's superior performance over GARCH models. The innovation of time-varying transition probabilities, as proposed by Bazzi et al. (2014), furthered the evolution of these models, allowing for more nuanced analysis of economic indicators like U.S. Industrial Production growth.
Recent contributions, such as those from Hwuy et al. (2016) and Nystrup et al. (2017), have introduced models that adaptively estimate parameters, capturing long memory in return volatility and improving forecasting performance. Chang et al. (2017) presented a novel approach to regime-switching that introduces an autoregressive latent factor, and Augustyniak, Bauwens et al. (2019) proposed the factorial hidden Markov volatility process, enhancing both in-sample fit and out-of-sample forecasting.
The evolution of Markov-switching methodologies has seen a significant shift from being predominantly used in economic and financial time series analysis to a broader array of applications and technical refinements post-2000. This thesis aims to build on this rich literature by applying advanced regime-switching models to the dynamic correlations of major world currencies, thereby offering novel insights and contributing to the continuous development of this influential analytical framework.

1.2 Research Contribution

Building upon the vast array of literature, this research positions itself at the intersection of Markov-switching models and real-world economic indicators. This study contributes to the existing literature by integrating the methodology of Pelletier (2006) with contemporary economic indicators as exogenous variables, an approach that has not been extensively explored in the context of foreign exchange rates. The primary innovation of this thesis is the use of these exogenous economic indicators—particularly GDP growth rates and the S&P market index—to enhance the predictive power and inferential accuracy of the Regime-Switching Dynamic Correlation (RSDC) models for foreign exchange rates.
Specifically, the thesis investigates the daily exchange rates of major world currencies, focusing on understanding the dynamics of the Japanese yen (JPY), Euro (EUR), and Canadian dollar (CAD) in relation to the U.S. dollar (USD). Through empirical analysis, it has been demonstrated that economic indicators such as interest rates, inflation differentials, and trade balances exert a significant impact on exchange rate fluctuations. The inclusion of these variables in the RSDC model has shown to improve the detection of regime shifts, which is particularly evident during periods of economic uncertainty and market turbulence.
The practical implications of this research are far-reaching, offering valuable insights for policymakers, investors, and economists who grapple with the complexities of foreign exchange markets. By augmenting the Regime-Switching Dynamic Correlation (RSDC) model with relevant economic indicators as exogenous variables, this work not only enhances the model's analytical precision but also provides a more comprehensive understanding of how external economic factors interact with currency dynamics.
The findings of this research underscore the increased inferential power of the RSDC model when it is informed by exogenous economic indicators. The model's ability to predict regime switches and correlation dynamics has been improved by incorporating GDP and S&P index data. This suggests that economic performance and market sentiment, as captured by these indicators, are integral to the behavior of currency exchange rates.

The contribution of this thesis lies in its innovative approach to enriching the regime-switching framework with exogenous economic variables, thus offering a new perspective on the understanding of dynamic correlations in exchange rates. This endeavor not only aligns with the ongoing evolution of Markov-switching methodologies but also paves the way for more sophisticated tools in economic forecasting and financial analysis, which may include the exploration of additional indicators or the application of high-frequency data to enhance model accuracy.

1.3 Overview of this Thesis

This thesis is organized as follows:

- **Chapter 2: Methodology** describes the statistical models and computational techniques employed in this study, with an emphasis on the extensions made to the Regime-Switching Dynamic Correlation (RSDC) model.
- **Chapter 3: Data and Preliminary Analysis** presents the data sets used, their sources, and some preliminary analyses to show the need for regime-switching models in capturing the dynamic of exchange rates.
- **Chapter 4: Empirical Results** discusses the findings from the application of the RSDC model to the daily exchange rates between the Japanese yen (JPY), Euro (EUR), and Canadian dollar (CAD) against the U.S. dollar (USD).
- **Chapter 5: Conclusions and Future Work** summarizes the key contributions of this thesis and proposes avenues for future research.

2 Methodology

2.1 Pelletier's (2006) RSDC model

We examine a multivariate time series with K components, $\{Y_t\}_{t=1}^{T}$, where K represents the number of time series. The variability in the dynamic relationship between time series is decomposed into standard deviations and correlations, where the correlation values transit between different regimes via a Markov chain:

$$Y_t = H_{t,S_t}^{1/2} Z_t \quad \text{with} \quad \{Z_t\}_{t\in\mathbb{N}} \sim \text{i.i.d}\,(0, I_K)^1. \tag{1}$$

where $\{S_t\}$ is a Markov chain representing to the state at time t, and I_K is the identity matrix of dimension K, which is a $K \times K$ square matrix with ones on the main diagonal and zeros elsewhere.. We decompose the time-varying covariance matrix H_{t,S_t} as follows

$$H_{t,S_t} = D_t \Gamma_{S_t} D_t, \tag{2}$$

where D_t is a diagonal matrix composed of the standard deviations, $\sigma_{k,t}$, with $k = 1, \ldots, K$, and Γ_{S_t} is the correlation matrix.

The ARMACH model is employed to model the standard deviations $\sigma_{k,t}$ of each component k in the diagonal matrix D_t. This process, as proposed by Taylor (1986), directly models the time-varying standard deviations. The representation is given as

$$\sigma_{k,t} = \omega_k + \alpha_k |y_{k,t-1}| + \gamma_k \sigma_{k,t-1}, \tag{3}$$

where $y_{k,t-1}$ is the component k at time $t-1$ of the multivariate time series Y_{t-1}.

The correlation matrix Γ_{s_t} varies between different regimes, with distinct correlation values assigned to each regime. The transition between these regimes is governed by a Markov chain, represented by the transition probability matrix Π. Each element $p_{i,j}$ in this matrix represents the probability of moving from regime i in period t to regime j in period $t+1$.

For the scope of this thesis, we will be considering only two regimes, therefore Π, the transition probability matrix, is a 2×2 matrix. It is represented as

$$\Pi = \begin{bmatrix} p_{1,1} & 1 - p_{1,1} \\ 1 - p_{2,2} & p_{2,2} \end{bmatrix}, \tag{4}$$

where $p_{i,j}$ denotes the probability of transitioning from state i at time $t-1$ to state j at time t, for i, j in $\{1, 2\}$.

$$p_{i,j} = P[S_t = j \mid S_{t-1} = i]. \tag{5}$$

The Markov chain used in our model adheres to the standard assumptions of aperiodicity, irreducibility, and ergodicity, which are crucial for its statistical properties. These concepts are elaborated in Chapter 4 of Ross (1993).

[1]The square root of the covariance matrix H_{t,S_t}, denoted $H_{t,S_t}^{1/2}$, in the context of this model, is used to transform a vector of i.i.d standard normal random variables Z_t into a vector Y_t with the desired covariance structure. The Cholesky decomposition is one common method to compute this matrix square root, ensuring that the transformed variables Y_t have the specified covariance matrix H_{t,S_t}.

2.2 Constrained Specification

In this section, we aim to maintain numerical stability and simplify computational efforts by considering a constant standard deviation matrix. This is in contrast to the dynamic standard deviation matrix typically employed. Instead of directly modeling the time-varying volatility for each series, we normalize the observations using a series of estimated standard deviations derived from the ARMACH(1,1) model. This normalization process allows us to apply a model with constant volatility parameters to the standardized series, focusing our attention on the dynamics of the correlation matrix rather than the complexities introduced by time-varying volatility.

More precisely, let each component of the original data be a time series of length T. For $k = 1, \ldots, K$, we have

$$y_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,T}].$$

To model the time-varying volatility of each series, we fit an ARMACH(1,1) model for each one. This process involves estimating the model parameters ω_k, α_k, and γ_k that best capture the conditional variance dynamics of each series. We employ the Maximum Likelihood Estimation (MLE) method, which optimizes the likelihood function given the observed data. The MLE approach is particularly effective for time series data, as it takes into account the temporal dependencies within the data. The estimated standard deviations from the ARMACH model are thus given by

$$\hat{\sigma}_k = [\hat{\sigma}_{k,1}, \hat{\sigma}_{k,2}, \ldots, \hat{\sigma}_{k,T}].$$

The initial conditional variance $\hat{\sigma}_{k,1}^2$ can be estimated using the sample variance of the returns, which provides a straightforward starting point. Alternatively, one may consider the long-term variance of the process given by $\frac{\omega_k}{1-(\alpha_k+\gamma_k)}$, under the assumption that $\alpha_k + \gamma_k < 1$ for the process to be stationary. This long-term variance reflects the expected level to which the process will revert over time in the absence of new shocks.
We then transform the original series by dividing each observation by its corresponding estimated standard deviation to obtain the filtered series:

$$u_k = \frac{y_k}{\hat{\sigma}_k} = \left[\frac{y_{k,1}}{\hat{\sigma}_{k,1}}, \frac{y_{k,2}}{\hat{\sigma}_{k,2}}, \ldots, \frac{y_{k,T}}{\hat{\sigma}_{k,T}}\right] = [u_{k,1}, u_{k,2}, \ldots, u_{k,T}].$$

Dividing the original data by the estimated conditional standard deviations allows for normalization with respect to the volatility of the series. This process reduces the effect of varying volatility levels over time, making the data more consistent for analysis. It helps in focusing on the core movements of the series, independent of their volatility magnitude.

Through a filtering process, we refine our model to focus on the variability represented by the standard deviations, effectively simplifying D_t to the identity matrix I under the assumption of filtered series. This assumption posits that the filtered components u_t have unit variance. Furthermore, by assuming the innovations Z_t are normally distributed, we establish a filtered framework for our time series analysis.
The filtered series U_t is therefore represented as

$$U_t|s_t = \Gamma_{s_t}^{1/2} Z_t \sim N(0_K, \Gamma_{s_t}) \tag{6}$$

The vector 0_K in Equation (6) is a zero vector of dimension K.
Equation (6) provides the distribution of the filtered series U_t, a result of the standardization procedure and assumptions. A comprehensive validation of this approach, as well as the detailed outcomes of the filtering process within the variance model, is provided Appendix A.

2.3 Extension With Time-Varying Transition Probabilities

A method proposed by Diebold et al. (1994) is employed to incorporate time-varying transition probabilities within regimes and to introduce state dependence into the dynamic correlations model.
The state dependent probability matrix updates its elements at each time step, conditioned on the accumulated set of information Ω_t, which comprises past filtered series:

$$\Omega_t = (U_t, U_{t-1}, \ldots, U_1). \tag{7}$$

In this thesis, the accumulated information set, Ω_t, is primarily derived from the filtered exchange rates of the Japanese yen (JPY), Euro (EUR), and Canadian dollar (CAD) against the US dollar (USD), as presented in Section(3.1).

The transition probability of moving from state i to state j at time $t+1$ is denoted as:

$$p_{i,j}(t+1) = P[S_{t+1} = j | S_t = i, \Omega_t]. \tag{8}$$

This probability is modeled as a logistic function of a set of exogenous variables and parameters:

$$p_{i,j}(t+1) = \frac{1}{1 + \exp[-x_t^\top \beta_{s_t}]}. \tag{9}$$

Here, x_t represents the column vector of exogenous variables at time t, including a constant term, and is specified as

$$x_t = \begin{pmatrix} 1 \\ x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{M,t} \end{pmatrix}, \tag{10}$$

and β_{s_t} is the column vector of parameters associated with the current regime s_t, expressed as:

$$\beta_{s_t} = \begin{pmatrix} \beta_{0,s_t} \\ \beta_{1,s_t} \\ \beta_{2,s_t} \\ \vdots \\ \beta_{M,s_t} \end{pmatrix}. \tag{11}$$

Each $x_{i,t}$ is the value of the i^{th} exogenous variable at time t, and each β_{i,s_t} represents the parameter for the i^{th} exogenous variable in state s_t. When no exogenous variables are considered, these vectors contain only the constant term, reducing the model to one with constant transition probabilities represented by a static matrix Π.

2.4 Estimation Algorithm

2.4.1 Concepts

Definition 2.1 (Forecast, Filtered, and Smoothing Probabilities)**.** Forecast probability refers to the probability that the state of a Markov chain S_t will take on a certain value s_t, given information that was known prior to time t. Filtering probability, on the other hand, is the probability that S_t will take on the value s_t given information from the past and current state, while smoothing probability considers the full sample information to determine the probability that S_t will take on the value s_t:

$$\begin{aligned} P_{\text{forecast}} &= P[S_t = s_t \mid \Omega_{t-1}, \theta], \\ P_{\text{filtered}} &= P[S_t = s_t \mid \Omega_t, \theta], \\ P_{\text{smoothing}} &= P[S_t = s_t \mid \Omega_T, \theta]. \end{aligned}$$

The parameter vector, denoted as θ, includes two sets of coefficients, β_1 and β_2, and two sets of correlation coefficients, Γ_1 and Γ_2. Specifically,

$$\theta = [\beta_1, \beta_2, \Gamma_1, \Gamma_2]$$

where each of these parameters is a vector.

In this thesis, we focus on the estimation method based on the maximum likelihood function (see Section 2.4.2 for details). The accuracy of estimating the state variable S_t depends on evaluating its conditional forecasts (conditional expectations) of $S_t = s_t$ based on various information sets. These forecasts encompass the predictions of conditional probability, filtered probability, and smoothing probability.

The parameters of the Markov-switching model are estimated by considering the joint conditional probability of future states as a function of the joint conditional probabilities of current states and the transition probabilities. This is described by the filtering process (Definition 2.2).

Definition 2.2 (Filtering Process)**.** The process of filtering involves transforming the conditional probabilities of the current states through a dynamic system represented by the transition probability matrix to obtain the conditional probabilities of future states, that is

$$P[S_t = s_t \mid \Omega_{t-1}, \theta] = \sum_{s_{t-1}} P[S_t = s_t \mid S_{t-1} = s_{t-1}, \Omega_{t-1}, \theta] \times P[S_{t-1} = s_{t-1} \mid \Omega_{t-1}, \theta]$$

2.4.2 Estimating Parameters

The parameter vector θ for the proposed model is estimated using the maximum likelihood method. The log-likelihood function, denoted by $l(u_1, u_2, \ldots, u_T|\theta)$, is a function of the vector of parameters θ. It is defined as the sum of the natural logarithm of the conditional density function of the observed variables, given the information available up to time t. In mathematical terms, the log-likelihood function can be written as:

$$l(u_1, u_2, \ldots, u_T|\theta) = \sum_{t=1}^{T} \log\left(f\left(U_t = u_t|\Omega_{t-1}, \theta\right)\right), \tag{12}$$

where T is the sample size, and f represents the conditional density function of the observed variables. This function f describes the probability distribution of the observed variables U_t given the information set Ω_{t-1} and the parameter vector θ. Specifically, conditioned on the regime state, the distribution of U_t follows a multivariate normal distribution as indicated by Equation (6). The explicit form of the density function f is presented in this section.

Hamilton's 1989 paper proposed an algorithm to estimate parameters of a switching process when the true state of the system at any given time is unobservable. Basically, we want to jointly estimate (i) the parameters of the model conditional on being in either state 1 or state 2 and (ii) the probability that we are in state 1 or state 2 at a particular time.

Let $\Omega_t = \{u_1, u_2, \ldots, u_{t-1}, u_t\}$ denote the information available up to time t, where u_t is a vector of size K and corresponds to the observation of the K components at time t.

1. **Time $t-1$ state (previous state):**

$$\xi_{i,t-1} = P[S_{t-1} = i|\Omega_{t-1}, \theta]$$

2. **State transition from i to j (state propagation):**

$$p_{ij}(t) = P[S_t = j|S_{t-1} = i, \Omega_{t-1}, \theta]$$

   which represents the probability of state i at time $t-1$ transitioning to state j at time t.

3. **Densities under the regime j at time t (data observations and state dependent errors):**

$$\eta_{j,t} = f(u_t|S_t = j, \Omega_{t-1}, \theta) = \frac{1}{(2\pi)^{\frac{K}{2}}|\Gamma_j|^{\frac{1}{2}}} e^{-\frac{1}{2}u_t'\Gamma_j^{-1}u_t}$$

4. **Conditional density of the time t observation (combined likelihood with state being collapsed):**

$$f(u_t|\Omega_{t-1},\theta) = \sum_{i=1}^{2}\sum_{j=1}^{2} \xi_{i,t-1} p_{ij}(t) \eta_{j,t}$$

- Start with the definition of conditional probability of observing u_t given Ω_{t-1} and θ,

$$f(u_t|\Omega_{t-1},\theta).$$

- Expand this probability by considering all possible transitions between states i and j, where $i, j \in \{1, 2\}$,

$$f(u_t|\Omega_{t-1},\theta) = \sum_{i=1}^{2}\sum_{j=1}^{2} P(S_{t-1}=i, S_t=j|\Omega_{t-1},\theta) \cdot f(u_t|S_t=j,\Omega_{t-1},\theta).$$

- Recognize that $P(S_{t-1}=i, S_t=j|\Omega_{t-1},\theta)$ can be decomposed into the product of the probability of being in state i at time $t-1$ and the transition probability to state j at time t,

$$P(S_{t-1}=i, S_t=j|\Omega_{t-1},\theta) = \xi_{i,t-1} \cdot p_{ij}(t).$$

- Note that $f(u_t|S_t=j,\Omega_{t-1},\theta)$ is the likelihood of observing u_t given the state j, which is represented by $\eta_{j,t}$,

$$f(u_t|S_t=j,\Omega_{t-1},\theta) = \eta_{j,t}.$$

- Substitute these expressions back into the expanded probability,

$$f(u_t|\Omega_{t-1},\theta) = \sum_{i=1}^{2}\sum_{j=1}^{2} \xi_{i,t-1} \cdot p_{ij}(t) \cdot \eta_{j,t}.$$

5. **Time t state (corrected from previous state):**

$$\xi_{j,t} = P[S_t=j|\Omega_t,\theta] = \frac{\sum_{i=1}^{2} \xi_{i,t-1} p_{ij}(t) \eta_{j,t}}{f(u_t|\Omega_{t-1},\theta)}$$

- The probability of being in state j at time t, given all information up to t and parameters θ, is formulated as the ratio of the joint probability of observing u_t under state j and being in state j at t, to the total probability of observing u_t under any state. This is expressed as:

$$P[S_t=j|\Omega_t,\theta] = \frac{P[S_t=j, u_t|\Omega_{t-1},\theta]}{P[u_t|\Omega_{t-1},\theta]}.$$

- The numerator represents the joint probability of transitioning to state j at time t and observing u_t, which is obtained by summing over all possible previous states i, considering the transition probabilities $p_{ij}(t)$ and the likelihood of u_t given state j, $\eta_{j,t}$:

$$P[S_t = j, u_t|\Omega_{t-1}, \theta] = \sum_{i=1}^{2} \xi_{i,t-1} \cdot p_{ij}(t) \cdot \eta_{j,t}.$$

- The denominator, $P[u_t|\Omega_{t-1}, \theta]$, normalizes the probability by considering all possible states at time t, ensuring the sum of probabilities over all states equals 1. It is calculated by summing the joint probabilities over all state transitions and observations of u_t:

$$P[u_t|\Omega_{t-1}, \theta] = \sum_{j=1}^{2} \sum_{i=1}^{2} \xi_{i,t-1} \cdot p_{ij}(t) \cdot \eta_{j,t}.$$

- Combining these elements yields the updated probability for state j at time t, reflecting the latest observed data u_t and the transition dynamics encoded in θ:

$$P[S_t = j|\Omega_t, \theta] = \frac{\sum_{i=1}^{2} \xi_{i,t-1} \cdot p_{ij}(t) \cdot \eta_{j,t}}{\sum_{j=1}^{2} \sum_{i=1}^{2} \xi_{i,t-1} \cdot p_{ij}(t) \cdot \eta_{j,t}} = \frac{\sum_{i=1}^{2} \xi_{i,t-1} \cdot p_{ij}(t) \cdot \eta_{j,t}}{f(u_t|\Omega_{t-1}, \theta)}$$

Iterate steps 1) through 5) from $t = 1$ to T.

As a result of executing this iteration, the sample conditional log-likelihood of the observed data is calculated, and is optimized numerically[2] to find the best fitting set $\hat{\theta}$.

Regime-Switching estimation involves the numerical optimization of a high-dimension parameter vector, θ. Given the nature of this task, it is beneficial to employ vectorized operations. Many programming languages, including Python and MATLAB, are designed to handle vectorized operations more efficiently, leading to faster computation times. Therefore, we use vectorization in our algorithm implementation to optimize performance:
First, we have

$$\Pi(t) = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix} (t)$$

which is the transition probability matrix.
Let

$$P[S_t = s_t|\Omega_t, \theta] = \xi_{t|t} = \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \end{bmatrix}$$

[2]Differential evolution is a type of numerical optimization algorithm that is particularly useful for estimating the parameters of complex nonlinear functions. It iteratively explores the parameter space to find an optimum of the objective function, leveraging a population-based approach that does not require derivative information. Differential evolution is known for its robustness and efficiency and has been successfully applied to a variety of optimization problems, including those in econometrics and finance. Therefore, it can be a good choice for estimating the parameters of the regime switching model in this thesis, which involves a complex nonlinear log-likelihood function.

$$f_t = \begin{bmatrix} f(U_t = u_t | s_t = 1, \Omega_{t-1}, \theta) \\ f(U_t = u_t | s_t = 2, \Omega_{t-1}, \theta) \end{bmatrix}$$

and

$$\xi_{t|t-1} = \begin{bmatrix} P[S_t = 1 | \Omega_{t-1}, \theta] \\ P[S_t = 2 | \Omega_{t-1}, \theta] \end{bmatrix}$$

Now the series of inference and forecast probabilities regarding the unobserved state (regime) that we need can be expressed as:

$$\xi_{t|t} = \frac{1}{\xi'_{t|t-1} f_t} \xi_{t|t-1} \odot f_t$$

where $\odot$ corresponds to element-wise multiplication, and

$$\xi_{t+1|t} = \Pi(t+1)\xi_{t|t}$$

from the definition of the Filtering Process (2.2).

The vectorized algorithm is recursively performed the following way:

1. **Initialize a guess (expectation) for the probabilities of each state at time zero ($\xi_{1|0}$)**
   One effective approach to initialize $\xi_{1|0}$ is to use the limiting probabilities (π_1 and π_2) of the Markov process. These limiting probabilities represent the long-term stable state probabilities to which the system eventually converges. Mathematically, these probabilities are found by solving:

$$\begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \Pi(1) \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix},$$
$$\pi_1 + \pi_2 = 1$$
$$\pi_1, \pi_2 > 0.$$

   For a two-state system, the solution is:

$$\pi_1 = \frac{1 - p_{22}(1)}{1 - p_{11}(1) + 1 - p_{22}(1)},$$
$$\pi_2 = \frac{1 - p_{11}(1)}{1 - p_{11}(1) + 1 - p_{22}(1)}$$

   These probabilities provide a stable and informed initial guess for $\xi_{1|0}$.

2. **Use data available at time 1 to update this expectation:**

$$\xi_{1|1} = \frac{1}{\xi'_{1|0} f_1} \xi_{1|0} \odot f_1.$$

3. **Use the transition probabilities to form an expectation about the state probabilities for time 2:**

$$\xi_{2|1} = \Pi(2)\xi_{1|1}.$$

4. **Continue this process until $t = T$.**

2.5 Smoothing

In a Markov-Switching model, the regime is considered unobservable and likelihood maximizing parameter values are obtained using the observed data. There are two main objectives of this estimation:

1. Finding the likelihood maximizing regime-dependent parameters and
2. Inferring the prevailing regime at specific points in time.

To make inferences about the regime at a particular time, a smoothing algorithm is employed, which uses all available data and optimized model parameters to determine the most likely regime at each time period. The main difference between the smoothing algorithm and the filter is that the former considers all available data up to time T, while the latter considers all data up to time $t-1$ to predict the regime at time t. The smoothing algorithm, developed by Kim (1994), involves the following steps:

Consider the joint probability that the regime is in state j at time t and state k at time $t+1$ based on full information:

$$\begin{aligned} P[S_t = j, S_{t+1} = k|\Omega_T] &= P[S_{t+1} = k|\Omega_T] \times P[S_t = j|S_{t+1} = k, \Omega_T] \\ &\approx P[S_{t+1} = k|\Omega_T] \times P[S_t = j|S_{t+1} = k, \Omega_t] \\ &= \frac{P[S_{t+1} = k|\Omega_T] \times P[S_t = j, S_{t+1} = k|\Omega_t]}{P[S_{t+1} = k|\Omega_t]} \\ &= \frac{P[S_{t+1} = k|\Omega_T] \times P[S_t = j|\Omega_t] \times P[S_{t+1} = k|S_t = j, \Omega_t]}{P[S_{t+1} = k|\Omega_t]}. \end{aligned}$$

The smoothing algorithm employed in this study follows the approach developed by Kim (1994), which necessitates an approximation to compute the joint probability of regime states. While this approximation facilitates the computation process, it is acknowledged that a more precise calculation could be achieved through the application of a forward-backward algorithm. The forward-backward method is a dynamic programming approach that avoids approximations by systematically processing the sequential data in both forward and backward passes, thus ensuring that the computation of smoothed probabilities leverages all available information without the need for simplifying assumptions. Rabiner's (1989) tutorial on hidden Markov models provides an in-depth explanation of this algorithm and its applications, underscoring its utility in accurately estimating the states of a system based on observed sequences.

The mechanics of the smoother involve using the updated regime probability from the last iteration of the smoother, denoted as $P[S_t = s_t|\Omega_t, \theta]$ (previously represented by $\xi_{t|t}$), along with the updated probability from the filter loop at $T-1$, denoted as $P[S_{T-1}|\Omega_{T-1}]$ (previously represented by $\xi_{T-1|T-1}$), and the transition probabilities $P[S_T = 1|S_{T-1} = 1]$ and $P[S_T = 2|S_{T-1} = 2]$, which are the model parameters p_{11} and p_{22} (or β_1 and β_2 for the time varying model), respectively.

Given these quantities, we can calculate the smoothed regime probabilities:

$$P[S_{T-1}=1, S_T=1|\Omega_T] = \frac{P[S_T=1|\Omega_T] \times P[S_{T-1}=1|\Omega_{T-1}] \times P[S_T=1|S_{T-1}=1, \Omega_{T-1}]}{P[S_T=1|\Omega_{T-1}]}$$
$$P[S_{T-1}=1, S_T=2|\Omega_T] = \frac{P[S_T=2|\Omega_T] \times P[S_{T-1}=1|\Omega_{T-1}] \times P[S_T=2|S_{T-1}=1, \Omega_{T-1}]}{P[S_T=2|\Omega_{T-1}]}$$

The smoothed probability that the prevailing regime at time $t = T-1$ was in state $s_t = 1$ is then:

$$P(S_{T-1}=1|\Omega_T) = P(S_{T-1}=1, S_T=1) + P(S_{T-1}=1, S_T=2)$$

therefore,

$$P(S_{T-1}=2|\Omega_T) = P(S_{T-1}=2, S_T=2) + P(S_{T-1}=2, S_T=1)$$

and stepping backwards in time,

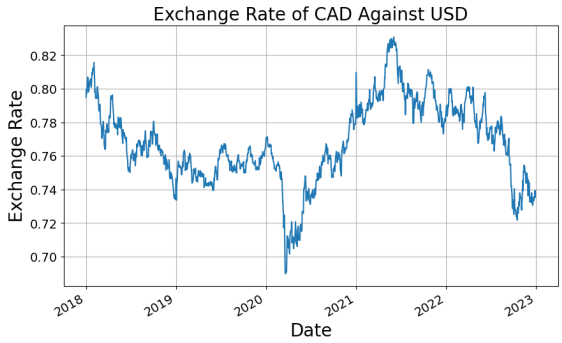
$$P(S_1=2|\Omega_T) = P(S_1=2, S_2=2) + P(S_1=2, S_2=1)$$

We compute the smoothed regime probabilities for each time point. These probabilities can be used to infer the start and end dates of different regimes by assigning to each date the regime that has the highest smoothed probability, which can provide valuable insights into the dynamics of the underlying process.

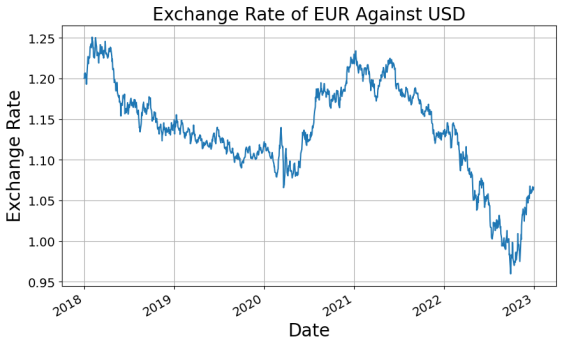
3 Data

3.1 Dependant Variables

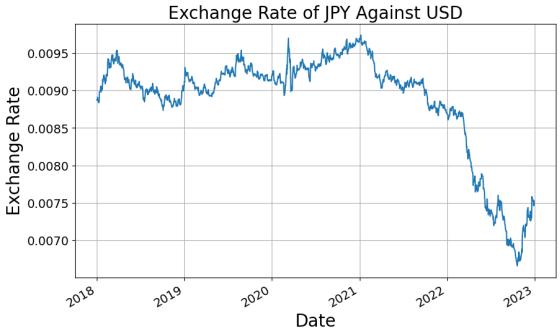
In this section, we present the data used in our analysis. The dependant variables consist of daily exchange rates of the Japanese yen (JPY), Euro (EUR), and Canadian dollar (CAD) against the US dollar (USD). These exchange rates are studied over a period from January 1st, 2018, to April 31st, 2022, depicted in Figure 1, and a test set extending from May 1st, 2022, to January 1st, 2023. The choice of these currencies is motivated by their significant role in the global foreign exchange market. The time period was chosen to provide a recent and relevant context for our analysis.



(a) CAD/USD



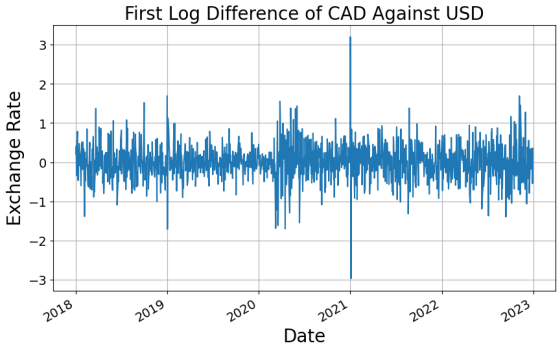
(b) EUR/USD



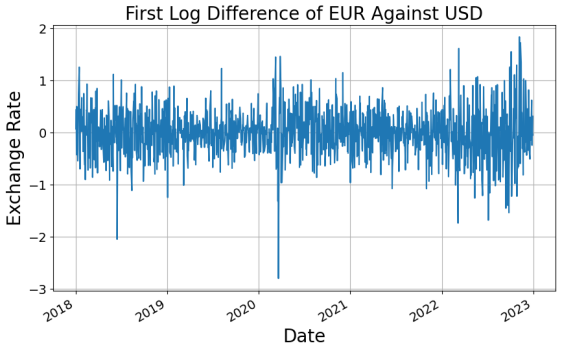
(c) JPY/USD

Figure 1: Exchange rate series. Time series plots of daily exchange rates of JPY, EUR, and CAD against USD from January 1st, 2018, to April 31st, 2022. The plots reveal the dynamic nature of exchange rates, with periods of relative stability interrupted by episodes of volatility.

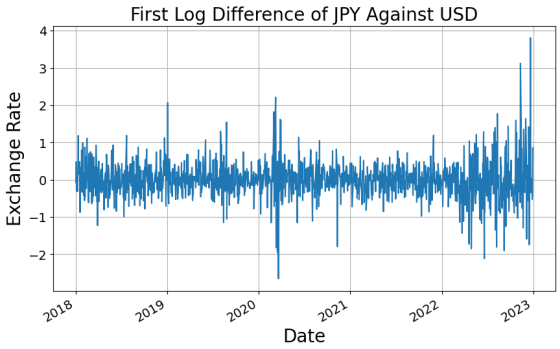
Next, we focus on the first log differences of these exchange rates, centered and multiplied by 100, as it is the correlation of these differences that we are primarily interested in for our analysis. The log differences of each series are presented in Figure 2.
It is worth noting that the date range of our dataset differs from that of Pelletier (2006). The reason for this is that we incorporate exogenous variables into our analysis that were



(a) Log Difference CAD/USD



(b) Log Difference EUR/USD



(c) Log Difference JPY/USD

Figure 2: First Log Differences of Exchange Rates. These plots display the first log differences of the JPY/USD, EUR/USD, and CAD/USD exchange rates, centered and multiplied by 100. This transformation highlights the daily changes in rates and is crucial for our analysis of dynamic correlations.

not available during the time of those datasets.

To prepare our data for analysis, we apply the ARMACH(1,1) model to the first difference of the logarithm of each series, which is then filtered. This process helps us in modeling and understanding the volatility patterns inherent to the exchange rate series. Figure 3 illustrates the time-varying conditional standard deviations for each exchange rate series derived from this model, showcasing the periods of volatility and stability for the JPY/USD, EUR/USD, and CAD/USD pairs.

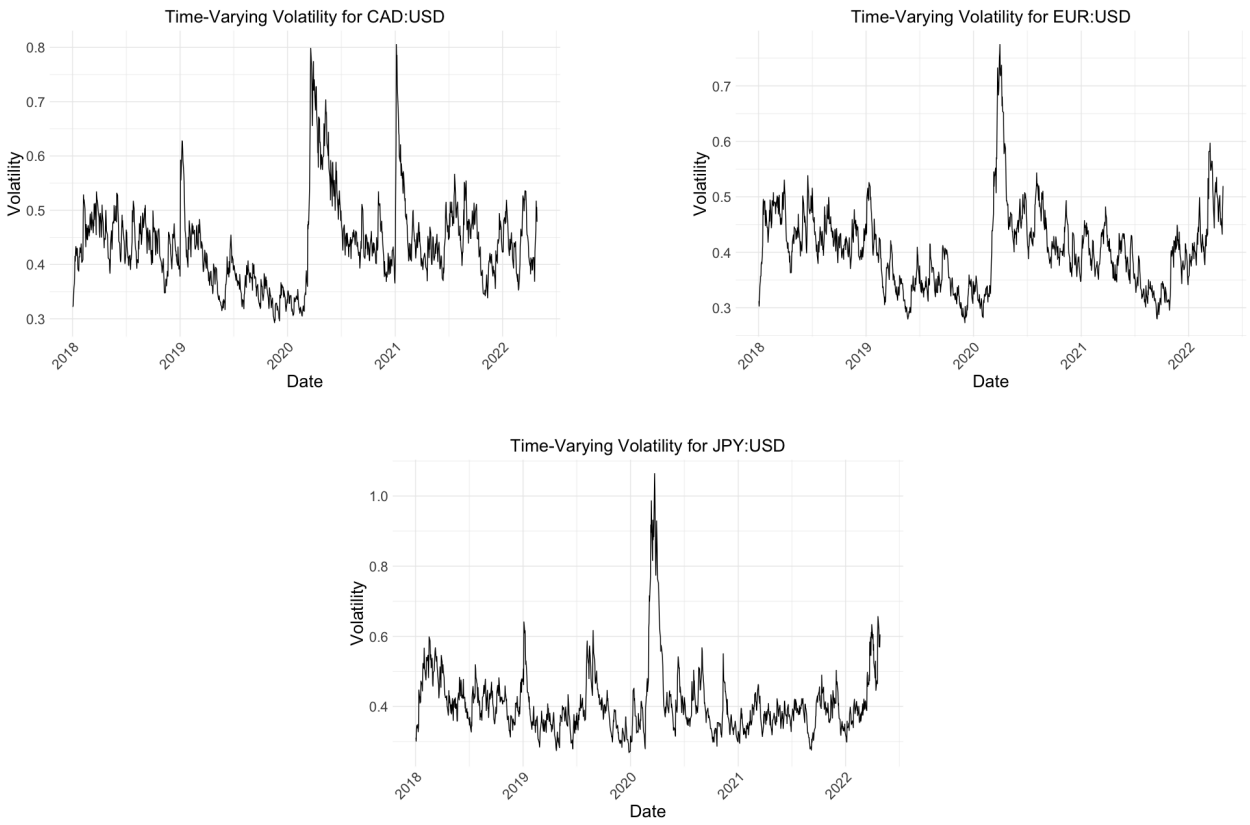


Figure 3: Each subplot represents the time series of the fitted standard deviations for the transformed exchange rate between the specified currencies.

Following the application of the ARMACH(1,1) model, we calculate the parameters for each currency series. Table 1 presents the estimated parameters for the ARMACH(1,1) models, detailing the specific values of ω, α, and γ for the JPY/USD, EUR/USD, and CAD/USD exchange rates.

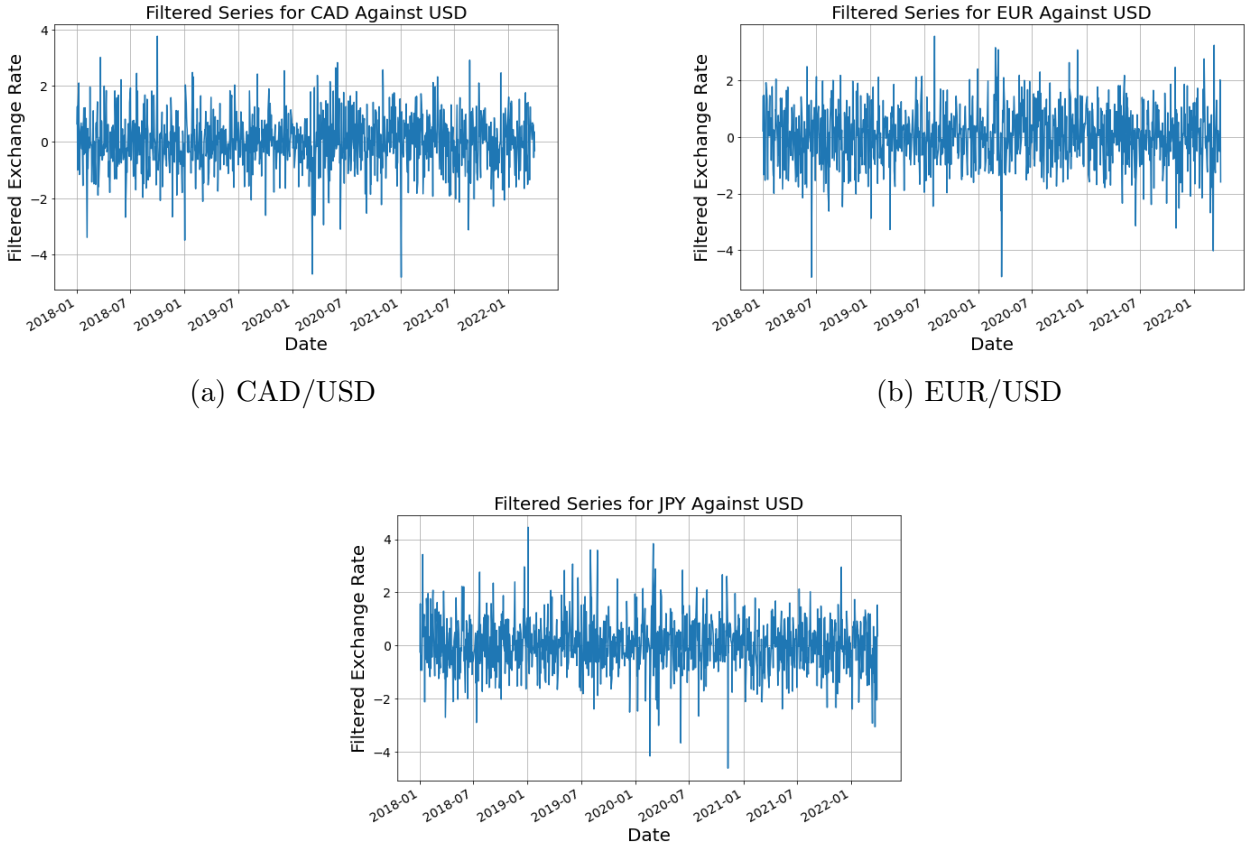
Table 1.
Estimated Parameters for the ARMACH(1,1) Model
The parameters are estimated on each of the series obtained by taking the first difference of the logarithm of each series, multiplying by 100, and subtracting the sample mean. These estimations are according to the ARMACH(1,1) model as described in Equation (3).

Currency	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\gamma}$
JPY/USD	0.0272	0.1063	0.8535
EUR/USD	0.0137	0.0716	0.9116
CAD/USD	0.0190	0.0773	0.9002

The estimated parameters of the ARMACH(1,1) model, as presented in Table 1, offer insights into the volatility dynamics of the respective currency exchange rates. The estimated parameter $\hat{\omega}$ represents the long-term average volatility, indicating the baseline level of fluctuations in the exchange rates. A higher value of $\hat{\omega}$ suggests a greater inherent volatility

in the currency's value. The estimated parameter $\hat{\alpha}$ captures the response of volatility to market shocks from the previous day, reflecting how quickly the market reacts to new information. A larger $\hat{\alpha}$ indicates a more sensitive response, leading to higher volatility following market movements. Lastly, the estimated parameter $\hat{\gamma}$ measures the persistence of volatility over time. Values close to 1 suggest that the effects of shocks on volatility are long-lasting, indicating a high degree of volatility clustering in the currency exchange rate series. This analysis provides a quantitative foundation for understanding the behavior of volatility in foreign exchange markets, highlighting differences in the stability and reaction to new information among the currencies studied.

The final step in our data preparation is the creation of filtered series. These series, depicted in Figure 4, are the result of multiplying the first difference of the logarithms by 100, subtracting the sample mean, and applying the variance model parameters from Table 1. The series in Figure 4 are referred to as 'filtered series' and are used for further analysis of dynamic correlations between exchange rates.



(a) CAD/USD

(b) EUR/USD

(c) JPY/USD

Figure 4: Filtered exchange rate series. The series represent the filtered rates for CAD, EUR, and JPY against USD, obtained by applying the ARMACH(1,1) model to the first difference of the logarithm of each series, multiplied by 100, and subtracting the sample mean.

An essential aspect of our analysis is the investigation of dynamic correlations among the

exchange rates. To provide a visual representation of these correlations, we calculate rolling correlations using a 45-day window. Figure 5 illustrates the time series of the rolling correlations for each currency pair. This visualization helps in understanding how correlations between the EUR/JPY, EUR/CAD, and CAD/JPY currency pairs evolve over time.

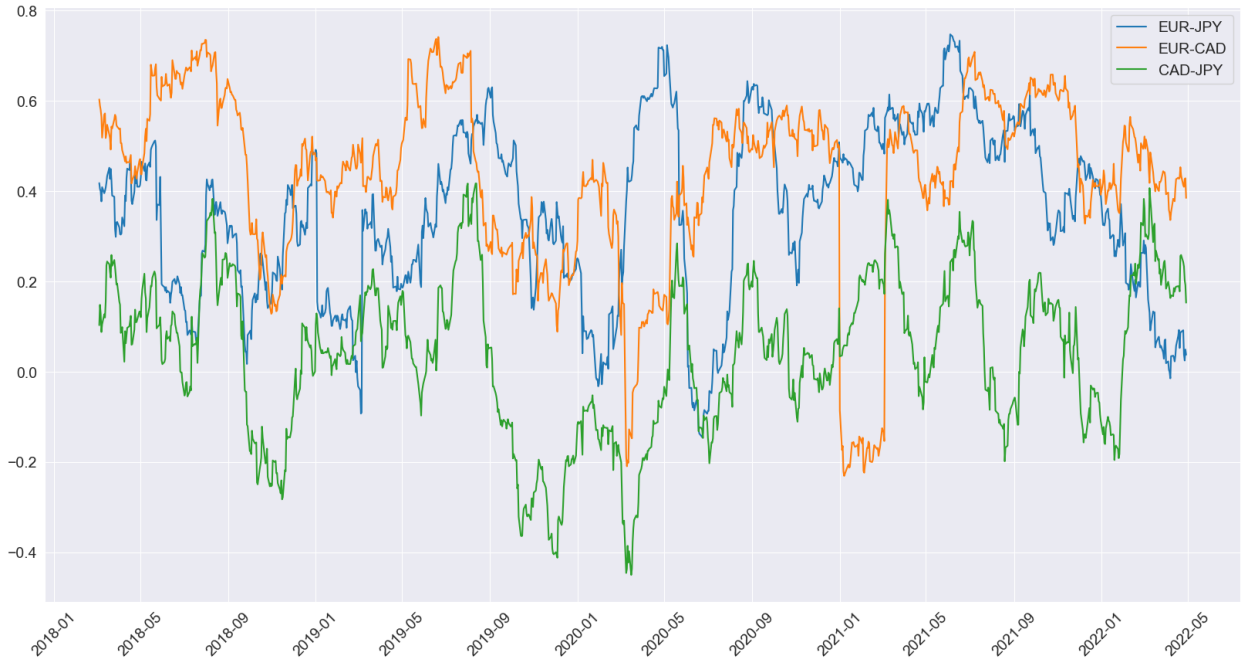


Figure 5: Rolling correlations between the EUR/JPY, EUR/CAD, and CAD/JPY currency pairs over time with a 45-day rolling window.

The rolling correlations provide insight into the fluctuating relationship between the currency pairs over the observed period, highlighting times of convergence and divergence that may correspond to economic events or market sentiment shifts. The average correlation offers a summary measure of the overall market correlation structure and serves as a precursor to more complex analyses within this thesis.

3.2 Explanatory Variables

The set of explanatory variables selected for this study is based on their established impact on exchange rates, as documented in economic literature. The variables encompass a range of economic indicators and market data, each serving to elucidate different aspects influencing currency valuation. Data for these variables was acquired through the Federal Reserve Economic Data (FRED) database, ensuring comprehensive coverage and robustness of information. Below is an outline of the chosen variables and the rationale for their inclusion:

1. **Gross Domestic Product (GDP):** Quarterly GDP growth rates are interpolated to a daily frequency using the last available value method to represent economic performance without introducing forward-looking bias. GDP is a primary indicator of

economic strength, and its growth rate directly correlates with the health of a country's currency.

2. **Consumer Price Index (CPI):** Monthly CPI data, converted to a daily measure, accounts for inflation rate effects. A lower inflation rate typically indicates a stronger currency value, thus affecting exchange rates.

3. **Market Indices:** Indicators like the S&P 500, FTSE 100, and the Nikkei 225 reflect overall economic performance and investor sentiment, both of which have significant bearings on currency values.

4. **Volatility Index (VIX):** Known as the "fear index," the VIX gauges market risk and sentiment. Its inclusion helps assess how market volatility influences currency valuation.

5. **Oil Prices (CL=F):** As a crucial commodity, oil prices affect the economic stability of exporting and importing countries, thereby influencing their exchange rates.

6. **Gold Prices (GC=F):** Gold prices serve as a barometer for economic uncertainty and can affect currency stability and, consequently, exchange rates.

7. **10-Year Treasury Note Yield (TNX):** The yield on this note is a benchmark for global finances, and shifts in the yield suggest changes in U.S. economic expectations, which are closely tied to currency fluctuations.

For each of these explanatory variables, both the level (price) and the first moment (return) were considered to capture different aspects of their influence on exchange rates. Additionally, all variables were normalized to enable a direct comparison of coefficient parameters across different datasets. This normalization facilitates an equitable assessment and interpretation of the influence exerted by each variable within our models, ensuring a balanced and comprehensive analytical approach.
Each variable thus provides a distinct and quantifiable perspective on the dynamic forces shaping exchange rates, ensuring our model comprehensively captures the multifaceted nature of currency fluctuations.

4 Empirical Results

In this section, we present and discuss the results of our model estimation. We consider various combinations of exogenous variables, each representing a different possible set of influences on the exchange rates under study. The performance of each model is evaluated using several metrics, including the log-likelihood, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the likelihood ratio (LR) test statistic and its p-value, and the out-of-sample log-likelihood.
The formulas for AIC and BIC, based on the log-likelihood function l defined in Equation (12) and the number of parameters q, are as follows:

$$\text{AIC} = 2q - 2l$$

$$\text{BIC} = q\log(T) - 2l$$

Additionally, we evaluate the out-of-sample likelihood, calculated using Equation (13). This likelihood is calculated using the parameters estimated from the training set and is evaluated on the data from the test set, covering observations from $T+1$ to $T+n$, where n is the number of observations in the test set.

$$l(u_{T+1}, u_{T+2}, \ldots, u_{T+n}|\hat{\theta}) = \sum_{t=T+1}^{T+n} \log\left(f\left(U_t = u_t|\Omega_{t-1}, \hat{\theta}\right)\right) \tag{13}$$

Subsequently, we conduct a comprehensive analysis to compare the best-fitting extended model against the restricted model.

4.1 Results Overview

In model selection, the objective is to optimize the trade-off between the fit of the model to the observed data and its complexity, which is defined by the number of parameters. To this end, evaluation criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were employed. Both AIC and BIC integrate the log-likelihood of the model while penalizing for the number of parameters, albeit to varying degrees. AIC is generally more permissive regarding model complexity as compared to BIC.
Additionally, out-of-sample log-likelihood was introduced as another evaluation criterion, serving to assess the generalizability of each model to unseen data.
Initial analyses involved the exploration of various configurations of exogenous variables. Subsequently, models were ranked based on their AIC, BIC, and out-of-sample log-likelihood values (LL OOS). Detailed results of this rigorous model selection process are presented in Table 2.
The Likelihood Ratio test p-value (LR) is also provided in these tables. Although not a direct measure of goodness-of-fit, the LR test is instrumental in comparing the adequacy of two nested models. Specifically, the restricted model is considered a special case of the unrestricted model. The test statistic is derived from the ratio of the likelihoods of these two models, and the corresponding p-value quantifies the statistical significance of the differences in fit between them. For an elaborate discussion of the LR test, readers are directed to Section C in the appendix.

In summary, lower values of AIC or BIC, in conjunction with higher out-of-sample log-likelihood, are indicative of a more favorable balance between model fit and complexity.
The correlation estimates are denoted by the symbol Γ, followed by a subscript that includes two asset numbers and a state number. The asset numbers correspond to the assets being compared, and the state number represents the state.
Specifically, the first number in the subscript corresponds to the asset number: 1 for CAD/USD, 2 for EUR/USD, and 3 for JPY/USD. The second number in the subscript also corresponds to an asset number, following the same numbering system. The third number in the subscript represents the state, with 1 indicating state 1 and 2 indicating state 2.
These correlation estimates provide insight into the relationships between different currency pairs in each state, which can be useful for understanding the dynamics of the exchange rates under different conditions.

4.2 Discussion of Results

It's noteworthy to mention that in our analysis, despite considering both the level (price) and the first moment (return) for each variable, the models incorporating returns consistently outperformed those with level variables. This was reflected in the selection of models for Table 2, where all variables are in their return form. This observation underscores the importance of focusing on the rate of change or relative movement in the variables, rather than their absolute levels, in capturing the nuances of exchange rate dynamics.
From Table 2, it's evident that the combination of Japan GDP, S&P Index, Yield Index as exogenous variables yields the lowest AIC score, making it the top-performing model based on this criterion. This could suggest that the dynamics of the Japanese GDP, the performance of the stock market (as measured by the S&P Index), and the yield index play a significant role in influencing the correlation dynamics between the exchange rates.

Another interesting observation is that models solely relying on stock market data, like the S&P Index, also seem to perform exceptionally well, indicating the prominent role of equity market movements in shaping currency relationships.
However, we highlight the relatively close AIC scores between the top models. This suggests that while there are differences in how well each model fits the data, the distinctions might not be pronounced. Consequently, the choice between these models may depend on the specific goals of the analysis or the ease of acquiring and processing the required data for the exogenous variables.

The Bayesian Information Criterion (BIC) is another tool used to judge the quality of a model in light of its complexity. It penalizes complex models more harshly than AIC. Therefore, BIC can be especially useful when preferring simpler models or when there's a risk of overfitting. The model with just the S&P Index as an exogenous variable outperforms others based on the BIC, suggesting that, when penalizing for complexity, a model focusing on the stock market provides a satisfactory fit. This underlines the potential influence and significance of stock markets in determining exchange rate dynamics.

A notable mention is the restricted model (with no exogenous variables) which also ranks relatively high based on BIC, pointing to the efficiency of a simpler model.

Table 2. Top 3 Models for Each Metric

This table presents the top three models ranked according to various metrics: AIC, BIC, and out-of-sample log-likelihood (LL OOS). Rows 1 to 3 list the models with the lowest AIC values, indicating the best fit with the least complexity. Rows 4 to 6 display the models according to BIC, which considers a larger penalty for complexity, and rows 7 to 9 are ranked based on LL OOS, reflecting the model's predictive performance on new data.

Exogenous Variables	**q**	**AIC**	**BIC**	**LR**	**LL OOS**	$\Gamma_{12,1}$	$\Gamma_{12,2}$	$\Gamma_{13,1}$	$\Gamma_{13,2}$	$\Gamma_{23,1}$	$\Gamma_{23,2}$
[Japan GDP, S&P Index, Yield Index]	14	8542	8612	0.00%	-973	0.613	0.087	0.202	-0.245	0.502	0.075
[Japan CPI, S&P Index, Yield Index]	14	8547	8617	0.00%	-983	0.627	0.108	0.199	-0.202	0.493	0.117
[USA CPI, S&P Index, Yield Index]	14	8549	8619	0.01%	-981	0.615	0.088	0.203	-0.246	0.498	0.078
[S&P Index]	10	8549	8599	0.00%	-988	0.709	0.227	0.300	-0.139	0.602	0.163
None	8	8565	8605	NA	-983	0.661	0.228	0.216	-0.110	0.589	0.138
[Japan GDP, S&P Index, Yield Index]	14	8542	8612	0.00%	-973	0.613	0.087	0.202	-0.245	0.502	0.075
[Japan GDP, S&P Index, Yield Index]	14	8542	8612	0.00%	-973	0.613	0.087	0.202	-0.245	0.502	0.075
[Germany GDP, Germany CPI, S&P Index]	8	8565	8605	98.02%	-975	0.645	0.140	0.242	-0.193	0.523	0.137
[Japan GDP, USA GDP, Oil Index, Yield Index]	14	8542	8612	100.00%	-976	0.553	0.181	0.088	-0.145	0.455	0.157

Interestingly, the Japan GDP, S&P Index, Yield Index combination, which was the top performer for AIC, still retains a prominent position in the BIC rankings. This overlap between AIC and BIC hints at the robustness of this particular model's performance.

Both AIC and BIC, while approaching model evaluation from slightly different perspectives, highlight the substantial influence of stock market indices (like the S&P) on exchange rate correlations. Moreover, while some combinations of exogenous variables provide optimal fits, the distinctions between the top models' performances are relatively nuanced, suggesting that model selection should be driven by the specific goals of the analysis, data availability, and computational efficiency.

The Out-of-Sample Log-Likelihood (OOS LL) is a key metric for evaluating the model's predictive accuracy on new data that the model has not seen during training. A higher OOS LL value suggests a model that generalizes better to new data points, which is essential for predictive tasks.

From Table 2, the model with Japan GDP, S&P Index, and Yield Index not only performs well according to AIC and BIC but also shows a superior OOS LL score of -973. This further solidifies the argument that this model is both well-fitted and generalizable.

Conversely, the restricted model, despite its relative simplicity, managed to achieve a relatively good OOS LL score of -983, as seen in Table 2. This aligns with its high BIC rank, emphasizing that a less complex model can still offer good predictive power.

Interestingly, the S&P Index model, while dominating in BIC, shows a slightly lower OOS LL score of -988. This might suggest that while the model is parsimonious and fits the in-sample data well, its predictive accuracy could be slightly compromised.

This multi-criterion analysis—comprising AIC, BIC, and OOS LL—provides a more holistic view of model performance. It reiterates the substantial role of the S&P Index and the Japan GDP, among others, in determining exchange rate dynamics. It also implies that depending on the objective—be it explanatory power or predictive accuracy—different models may be more appropriate.

Overall, the OOS LL metric complements the AIC and BIC metrics, offering an additional layer of validation in model selection.

Ultimately, these findings emphasize the interconnectedness of global financial markets. The intricate web of relationships between stock markets, economic indicators, and exchange rates showcases the complexities inherent in predicting currency movements.

4.3 Best-Fitting Model comparison with Restricted Model

In this subsection, we delve deeper into the results by comparing the restricted model, which includes no exogenous variables, with two unrestricted models: the best-fitting model with one exogenous variable and the best-fitting model with multiple exogenous variables. This

Table 3. Transition Probabilities Parameters for Restricted and Unrestricted Models
This table presents the transition probability estimates (beta parameters) and the average transition probabilities for each state ($\hat{p}_{11}$ and $\hat{p}_{22}$) for three distinct models. Model 1 is the restricted model with no exogenous variables. Model 2 incorporates the S&P Index as an exogenous variable. Model 3 includes the S&P Index, Japan's GDP, and the Yield Index as exogenous variables. The beta parameters represent the logistic probabilities parameters for transitioning from one state to another, while $\hat{p}_{11}$ and $\hat{p}_{22}$ represent the average transition probabilities for each state.

Models	$\hat{\beta}_{0,1}$	$\hat{\beta}_{0,2}$	$\hat{\beta}_{1,1}$	$\hat{\beta}_{1,2}$	$\hat{\beta}_{2,1}$	$\hat{\beta}_{2,2}$	$\hat{\beta}_{3,1}$	$\hat{\beta}_{3,2}$	$\hat{p}_{11}$	$\hat{p}_{22}$
Model 1	1.888	1.502	–	–	–	–	–	–	0.869	0.818
Model 2	2.074	1.484	1.483	-0.180	–	–	–	–	0.877	0.812
Model 3	2.927	2.191	0.378	-2.363	1.320	5.011	0.437	8.966	0.921	0.512

comparison allows us to better understand the impact of including exogenous variables on the performance of the model. These models are presented in Table 3.
To complement the analysis presented in Table 3, it is beneficial to visualize the behavior of the exogenous variables included in the models. This visualization helps in understanding the characteristics and fluctuations of these variables over time, which may have influenced the modells performance.
The plot in Figure 6 showcases the individual time series of the three key exogenous variables: Japan's GDP, the S&P Index, and the Yield Index.

4.3.1 Comparison of Models

First, we compare the transition probability estimates of the models. Transition probabilities represent the likelihood of moving from one state to another in the next time period, given the current state. By comparing these estimates across models, we can gain insights into how the inclusion of exogenous variables affects the dynamics of the exchange rates. For instance, if the transition probabilities of the unrestricted models are significantly different from those of the restricted model, this would suggest that the exogenous variables have a substantial impact on the state transitions.

Next, we compare the forecasted probabilities and the smoothed probabilities of each model. The forecasted probabilities represent the model's predictions of the future state, given the current state and the exogenous variables, while the smoothed probabilities represent the model's estimates of the current state, given all past and future observations. By comparing these probabilities, we can assess the accuracy of the model's predictions and its ability to capture the underlying dynamics of the exchange rates. If the forecasted probabilities are close to the smoothed probabilities, this would indicate that the model is accurately capturing the state transitions and is therefore a good fit for the data.

Through these comparisons, we aim to identify the model that provides the best balance of accuracy and complexity, and thus the most reliable insights into the dynamics of the exchange rates.

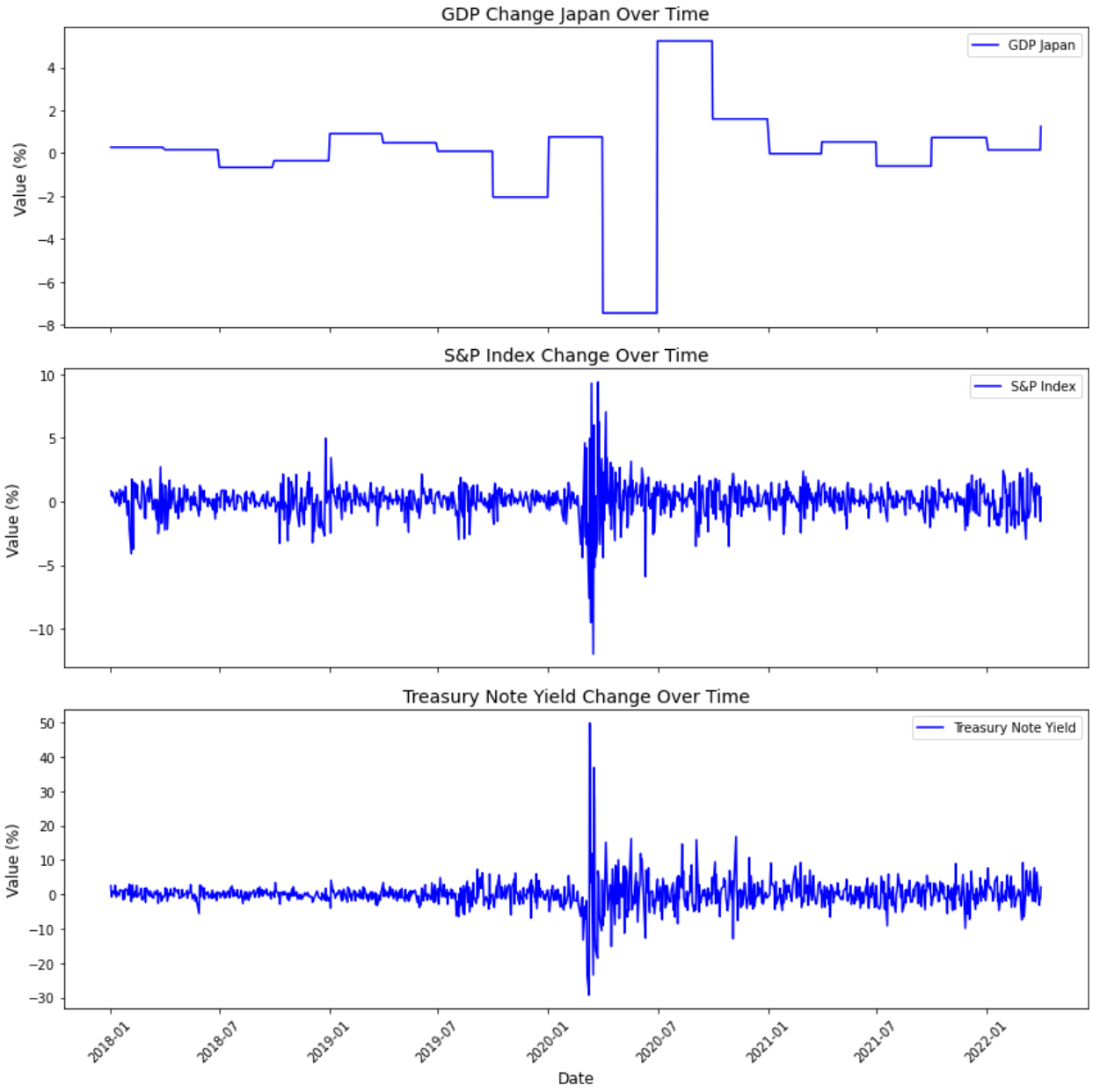


Figure 6: Time Series of Exogenous Variables Employed in Model 3

The figure illustrates the time series of Japan's GDP, the S&P Index, and the Yield Index. Observing these variables' trends and volatility offers insight into the economic and market conditions that may affect the dynamic correlations between exchange rates.

4.3.2 Transition Probabilities

We defined in section 2 the transition probabilities to correspond to

$$p_{i,j}(t+1) = \frac{1}{1+\exp[-x_t\beta_{s_t}]}$$

For the restricted model (Model 1: None), we can directly evaluate the recurrent probabilities using the β parameters from Table 3:

$$\hat{p}_{1,1} = \frac{1}{1+\exp[-\hat{\beta}_{0,1}]} = \frac{1}{1+\exp[-1.888]} = 0.869$$

$$\hat{p}_{2,2} = \frac{1}{1+\exp[-\hat{\beta}_{0,2}]} = \frac{1}{1+\exp[-1.502]} = 0.818$$

For the unrestricted models (Model 2 and Model 3), the transition probabilities can be calculated using their respective β parameters, as listed in Table 3. Figures 7 and 8 specifically present the transition probabilities p_{11} and p_{22} for Models 2 and 3, respectively.

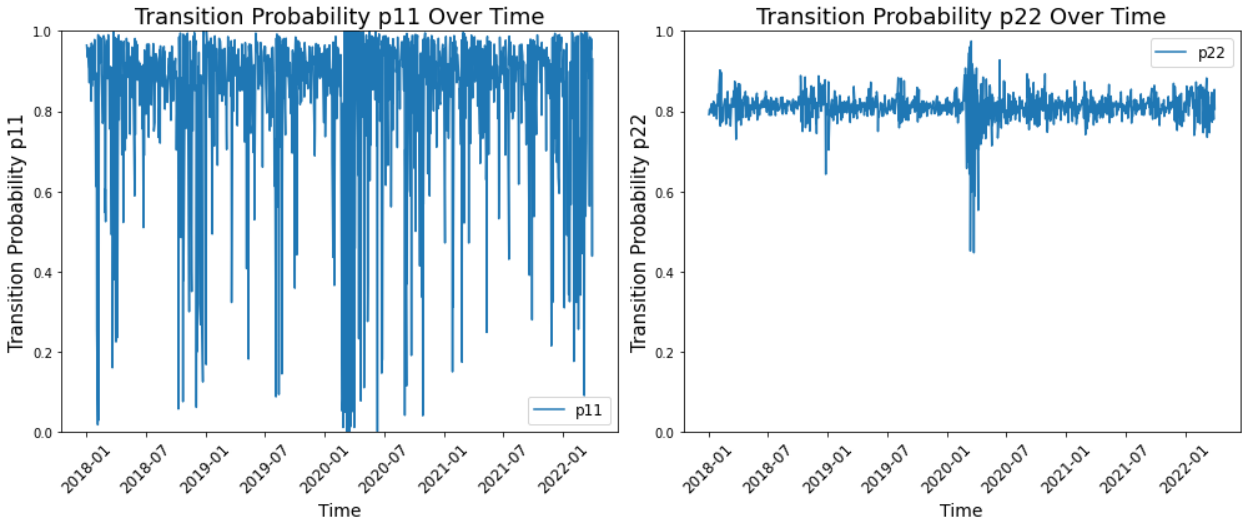


Figure 7: Temporal evolution of transition probabilities for Model 2, which incorporates the S&P Index as an exogenous variable. The figure contains two plots: the first plot represents $p_{1,1,t}$ and the second plot represents $p_{2,2,t}$.

The size of the beta coefficients serves as an excellent metric to gauge the model's responsiveness to the exogenous variable, the S&P Index. A higher absolute value of the beta coefficients indicates that even minor variations in the S&P Index can bring about significant changes in the transition probabilities. The transition probability p_{22} hovering around the 0.80 mark suggests a moderate sensitivity to the S&P Index. However, the sharp volatility rise at time step 580, aligning with the onset of the COVID-19 pandemic, reveals that unprecedented events can amplify this sensitivity dramatically.
The transition probability p_{11} displays pronounced volatility, fluctuating extensively between 0 and 1. Yet, the fact that it is commonly observed around 0.95 signifies a prevailing tendency

for the system to stay in state 1 once entered. The extensive volatility range of p_{11} underscores the system's susceptibility to external factors, indicating that the persistence in state 1 can be easily disrupted by minor variations in the S&P Index. This high volatility can be ascribed to the beta corresponding to the S&P 500 for state 1, which at 1.484, is significantly larger in magnitude than the corresponding beta for state 2, which stands at -0.180. This means that the transition probabilities for state 1 are substantially more reactive to shifts in the S&P Index compared to state 2.

Recurrent probabilities from the restricted model pointed towards one state's heightened recurrence over the other. In the context of Model 2, the more stable nature of p_{22} versus the high volatility of p_{11} underlines that while state 2 is relatively consistent and less influenced by external factors, state 1 is more capricious, primarily due to the pronounced influence of the S&P Index on its transition probability.

To sum up, the behavior of transition probabilities underscores the profound effects that external financial variables can exert on exchange rate dynamics. The disruptive influence of significant exogenous events, such as the COVID-19 pandemic, further emphasizes the sensitivity of these transition probabilities to worldwide financial occurrences. With the pronounced influence of the S&P Index on transition probabilities, especially on state 1, financial analysts and policymakers should remain alert to potential abrupt shifts in exchange rate behaviors governed by global financial market movements.

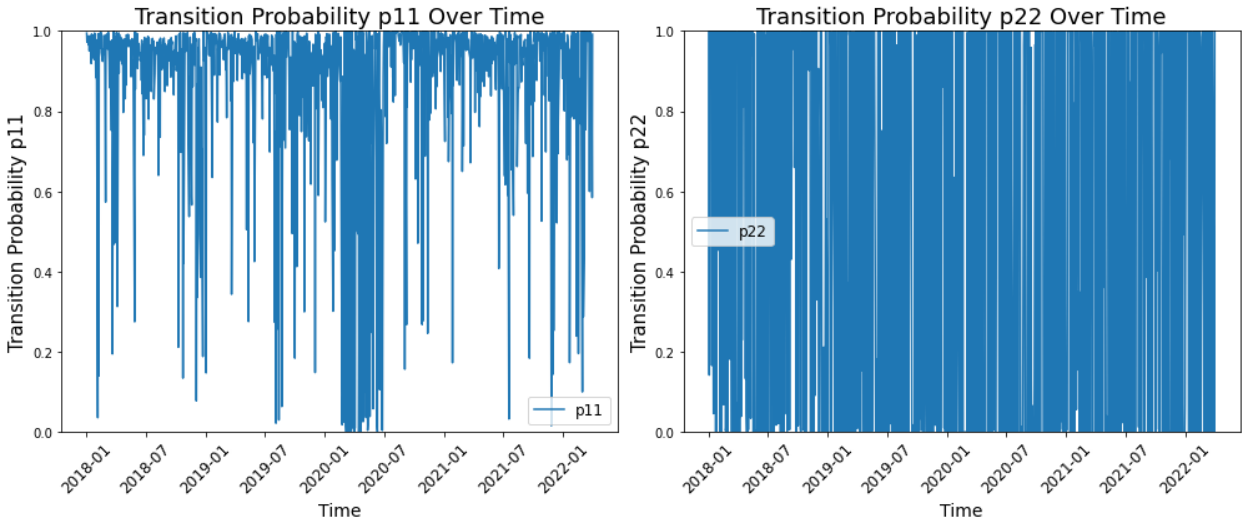


Figure 8: Temporal evolution of transition probabilities for Model 3, which includes the S&P Index, Japan's GDP, and the Yield as exogenous variables. The figure contains two plots: the first plot represents p_{11} and the second plot represents p_{22}.

The incorporation of multiple exogenous variables, namely the S&P Index, Japan's GDP, and the Yield, brings forth a richer, albeit more complex, narrative in the temporal evolution of transition probabilities in Model 3, as depicted in Figure 8.

The transition probability p_{22} in Model 3 showcases an exceptionally high degree of volatility, as can be observed in the second plot of Figure 8. So much so that its fluctuations are almost indiscernible on a visual representation. Such extreme volatility suggests that this state's transitions are highly sensitive to the incorporated exogenous variables, with their effects

being compounded when integrated into the model. The extreme sensitivity and resulting erratic behavior could potentially make predictions challenging for this state, highlighting the risks of introducing too much complexity without additional refinement.
In stark contrast, p_{11} demonstrates a recurrent behavior, centering primarily around the 0.98 mark. Nonetheless, its occasional plunges between the extremities of 0 and 1 reveal that, while this state is predominantly stable, it isn't entirely immune to external influences. Such oscillations emphasize that even states with a recurrent nature can experience instances of pronounced susceptibility to exogenous variables.
The vastly disparate behaviors of p_{11} and p_{22} can be effectively elucidated by examining their corresponding betas. The betas that are larger in absolute value naturally render the model more sensitive to variations in the corresponding exogenous variables. In the context of Model 3, the overwhelming sensitivity and volatility of p_{22} is a direct outcome of its betas being significantly larger in absolute value than those governing p_{11}. This highlights that while large betas can offer keen insights into the influence of exogenous factors, they can also introduce pronounced volatility, potentially rendering certain states too erratic for practical application.
In short, Model 3 tries to take a closer look by considering more factors, but this also means it's a bit trickier to balance. For people using this model, like analysts or decision-makers, it's important to remember that it can be a bit unpredictable and to be cautious when using it to make decisions. The lesson from Model 3 is that we need to keep tweaking and improving our models to make sure they're helpful and not just complicated for the sake of being complicated.

4.3.3 Smoothed and Forecast Probabilities

Notably, while there is a general alignment between the two probabilities, discrepancies emerge, especially at the peaks. The mismatch at these critical junctures highlights moments where the model's predictions were not in tandem with the subsequent realized states. This divergence underscores the challenges of forecasting even in the absence of exogenous influences and emphasizes the need for refining the model to better align predictions with realized transitions.

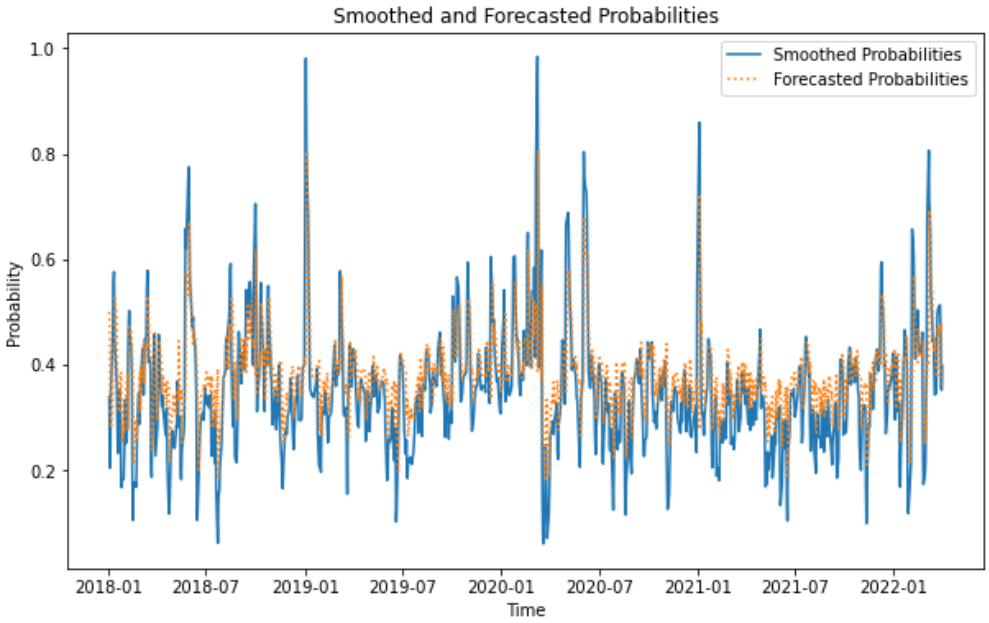


Figure 9: Smoothed and Forecast Probabilities Model 1 of state 2

Figure 9 illustrates the temporal interplay between the smoothed and forecast probabilities of state 2 for Model 1, which operates without any exogenous variables. The blue line represents the smoothed probabilities, which are calculated based on the past and current states, while the orange line represents the forecasted probabilities, which are predictions of future states.

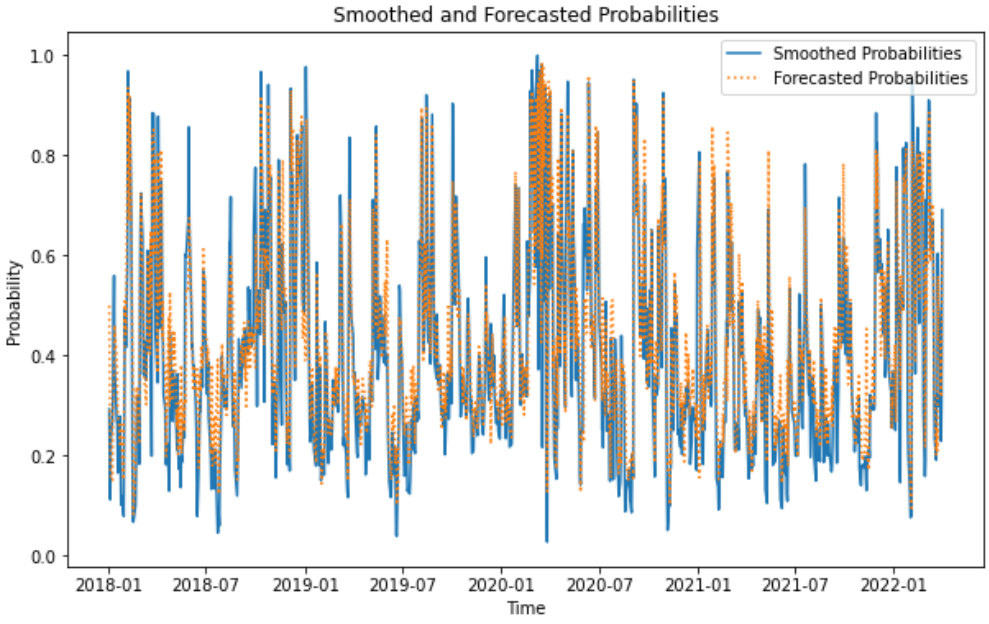


Figure 10: Smoothed and Forecast Probabilities Model 2 of state 2

Figure 10 depicts the smoothed probabilities of state 2 over time for Model 2, which includes the S&P index as an exogenous variable.

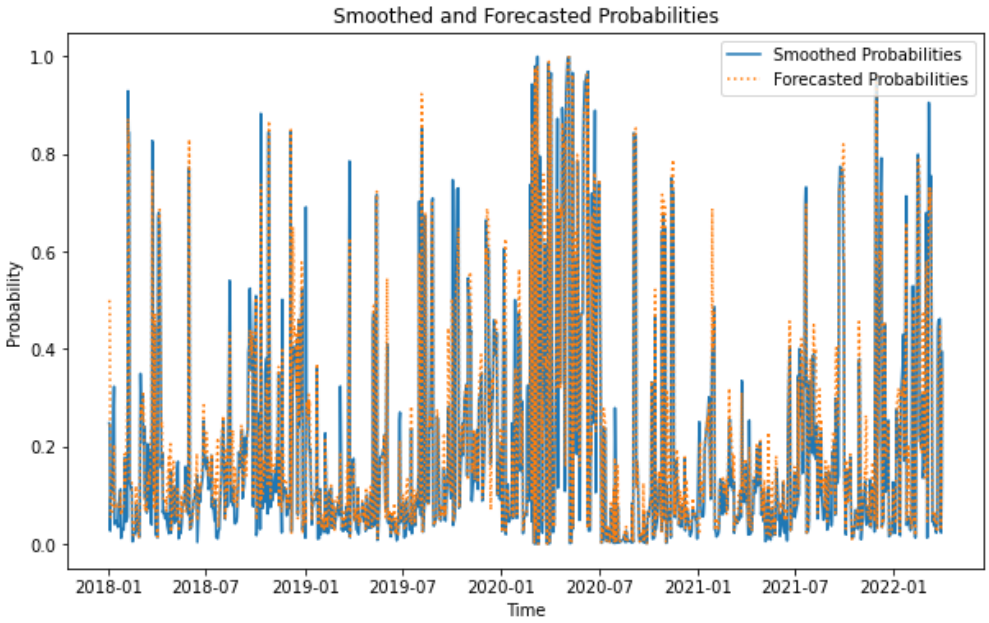


Figure 11: Smoothed and Forecast Probabilities Model 3 of state 2

Figure 11 illustrates the smoothed probabilities of state 2 over time for Model 3, which includes the S&P Index, Japan GDP, and the Yield as exogenous variables. The smoothed probabilities are indicative of the model's ability to capture the transition dynamics between states.

Noticeable improvement is observed in the alignment between the two probabilities of Model 3 compared to Model 1, underscoring the influence of the S&P Index in honing the model's predictions. However, this enhanced accuracy comes at the expense of increased volatility in the forecast probabilities, suggesting that while the inclusion of the S&P Index aids in prediction accuracy, it also introduces an element of sensitivity to the model's anticipations. A profound amplification of the alignment between forecast and smoothed probabilities is evident when compared to both Models 1 and 2, reinforcing the potency of incorporating a wider set of exogenous variables. Yet, this heightened congruence is coupled with a substantial surge in volatility, emphasizing that while the integrative approach of Model 3 sharpens predictive accuracy, it simultaneously escalates its susceptibility to rapid fluctuations. As we progressed from Model 1 to Model 3, a clear trend emerged in the behavior of both smoothed and forecast probabilities. This trend offers a compelling narrative about the role of exogenous variables in modeling and the trade-offs inherent in their inclusion.
The most evident observation is the increased alignment between the forecast and smoothed probabilities as we move from the restricted Model 1 to the more complex Model 3. This can be attributed to the richer informational content brought about by the inclusion of exogenous variables. Specifically, these variables contribute supplementary data that help the model anticipate state transitions more accurately. The S&P Index in Model 2 and the additional incorporation of Japan's GDP and the Yield in Model 3 provide the model with a broader context, thereby allowing it to better predict the likelihood of transitioning to state 2 based on more diverse and potentially interconnected economic indicators.

While the enhanced alignment of probabilities is undeniably a positive outcome, it comes with the added cost of increased volatility. The reason for this trade-off is embedded in the nature of exogenous variables themselves. By including more variables, the model becomes more sensitive to fluctuations in any one of them. Hence, while it can more accurately predict transitions based on current data, it is also more reactive to changes in that data, leading to greater volatility. The larger absolute values of the betas for the incorporated exogenous variables, especially in Model 3, make the model hyper-responsive to shifts in these indicators, amplifying the volatility observed in the transition probabilities.

The observed trends between the three models underscore a fundamental modeling consideration. There exists an inherent tension between accuracy and stability. While the inclusion of more exogenous variables can offer better predictive prowess, it can simultaneously introduce unpredictability in the model's outputs. This necessitates a careful evaluation of the trade-offs involved. Practitioners must decide on the optimal balance, ensuring that they aren't overly sacrificing stability for short-term accuracy.

In conclusion, as we delve into increasingly sophisticated models by incorporating more exogenous variables, it becomes imperative to understand and anticipate the dynamic effects these variables can have. Recognizing the interplay between accuracy and volatility can guide future modeling efforts, ensuring that chosen models are both reflective of the data's complexity and robust against its inherent unpredictabilities.

5 Critique and Future Research

5.1 Critical Review of the Current Work

The work presented in this thesis represents a significant step in modeling dynamic correlations in currency markets, notably through the enhancement of the Regime-Switching Dynamic Correlation (RSDC) model with key economic indicators. By integrating measures like the VIX index, the thesis offers a nuanced approach to understanding how macroeconomic factors influence regime shifts in correlation structures. However, the use of daily exchange rates, while insightful for day-to-day market movements, may overlook the finer details that only high-frequency data can capture. Additionally, while the VIX index provides a broad measure of market volatility, it doesn't encapsulate the full depth of information available in the options market, such as the volatility spread across different strike prices.

5.2 Avenues for Future Research

The present thesis has shed light on the multifaceted domain of dynamic correlation within financial markets, yet it represents merely the commencement of a broader scholarly endeavor. Subsequent inquiries may seek to augment this foundational work by incorporating high-frequency intraday data. Such an inclusion promises to elucidate the transient dynamics of market behavior in response to economic events and news, offering a window into the market's immediate and granular responses, which elude the scope of daily data.

Additionally, the options market harbors a wealth of nuanced data in the form of advanced implied volatility metrics. A meticulous examination of volatility skewness and the volatility surface could afford researchers with a more refined understanding of market sentiment and the anticipatory stances of market actors regarding currency trajectories.

The deployment of sophisticated machine learning algorithms presents a promising avenue for modeling the intricate, nonlinear interdependencies endemic to financial markets. These innovative methodologies are poised to detect latent patterns within correlation dynamics, thus providing robust tools for risk management and augmenting the efficacy of predictive models.

Furthermore, the exploration of a broader array of macroeconomic and financial indicators could serve as a pivotal step towards a more comprehensive understanding of market dynamics. These indicators may offer prognostic value, heralding impending shifts in market regimes or oscillations in market volatility.

This thesis serves as an academic precursor, inviting future researchers to delve deeper into the complex web of currency correlations. It is a call to enrich our collective intellect and to enhance our strategic interactions with the ever-evolving dynamics of global financial markets.

References

[1] Akintug, B. and Rasmussen, T.N. (2005). A Markov-Switching Model of Annual Hydrologic Time Series. Water Resources Research, 41(5), W05006.

[2] Asea, P.K. and Blomberg, S.B. (1998). Lending Cycles. Journal of Econometrics, 83(1-2), 89-128.

[3] Augustyniak, M., Bauwens, L., et al. (2019). Factorial Hidden Markov Volatility Models. Journal of Financial Econometrics, 17(1), 1-29.

[4] Bazzi, M., Blasques, F., Koopman, S.J., and Lucas, A. (2014). Time-Varying Transition Probabilities for Markov Regime Switching Models. Journal of Time Series Analysis, 35(5), 451-468.

[5] Chang, Y., Tsiaplias, S., and Harris, D. (2017). Modelling Regime Switches with an Autoregressive Latent Factor Approach. Journal of Econometrics, 196(2), 291-306.

[6] Filardo, A.J. and Gordon, S.F. (1998). Business Cycle Durations. Journal of Econometrics, 85, 99-123.

[7] Hamilton, J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica, 57(2), 357-384.

[8] Hamilton, J.D. (1990). Analysis of Time Series Subject to Changes in Regime. Journal of Econometrics, 45(1-2), 39-70.

[9] Hamilton, J.D. (2005). Regime-Switching Models. Prepared for: Palgrave Dictionary of Economics, Department of Economics, University of California, San Diego.

[10] Hwuy, L., Hsu, C., and Lin, T. (2016). Adaptive Estimation of Regime-Switching Models: Application to the Stock Market. Journal of Business & Economic Statistics, 34(2), 270-285.

[11] Kim, C.-J. (1994). Dynamic Linear Models with Markov-Switching. Journal of Econometrics, 60, 1-22. North-Holland.

[12] Kuan, C.-M. (2002). Lecture on the Markov Switching Model. Institute of Economics, Academia Sinica.

[13] Nystrup, P., Madsen, H., and Lindström, E. (2017). Long Memory and Regime Switching: A Simulation Study on the Markov-Switching ARFIMA Model. Computational Statistics & Data Analysis, 114, 40-57.

[14] Pelletier, D. (2006). Regime Switching for Dynamic Correlations. Journal of Econometrics, 131, 445–473.

[15] Sola, M. and Driffill, J. (1994). Understanding the Regime-Switching Behavior of Interest Rates. Journal of Monetary Economics, 34(2), 163-183.

[16] Town, R. (1992). Regime-Switching Stochastic Processes: Applications in the Bond Market. Finance and Stochastics, 6(3), 265-284.

[17] Yin, P. (2007). Estimating Volatility Using Hidden Markov Models. Journal of Financial Econometrics, 5(4), 504-532.

[18] Rabiner, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257-286.

A Monte-Carlo Study to Validate our Implementation

To validate our implementation, we set the parameter vector,

$$\theta = [\Gamma_{1,2,1}, \Gamma_{1,3,1}, \ldots, \Gamma_{1,K,1}, \Gamma_{12,2}, \Gamma_{1,3,2}, \ldots, \Gamma_{1,K,2}, p_{1,1}, p_{2,2}],$$

and we proceed as follows :

1. We attribute a value for each of the parameters to estimate, let $\Gamma_{i,j,1} = 0.8754$, $\Gamma_{i,j,2} = 0.4739$ for all i and j, and $p_{1,1} = 0.93$, $p_{2,2} = 0.85$.

2. For $l = 1, \ldots, 100$, we simulate a $T = 1000$ steps Markov-Chain governed by

$$\Pi = \begin{bmatrix} p_{1,1} & 1 - p_{1,1} \\ 1 - p_{2,2} & p_{2,2} \end{bmatrix}.$$

   We obtain a vector $\mathbf{S_l} = [s_{l,1}, s_{l,2}, \ldots, s_{l,1000}]$ where $s_{l,t}$ corresponds to the state at time step t for the simulation l.

   We also simulate a vector

$$y_l = [y_{l,1}, y_{l,2}, \ldots, y_{l,1000}]$$

   where $y_{l,t}$ is a simulated from distribution $N(0_K, \Gamma_{s_t})$.

3. For each of our $\mathbf{y}_l$ vector, we apply the Hamilton Filter (see Subsection 2.4) to estimate the parameters. We get

$$\hat{\theta} = [\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{100}].$$

4. For each of the parameters in θ, using $\hat{\theta}$, we construct a boxplot to visualize the distribution of the estimates around the true parameter values. This allows us to assess the accuracy of the Hamilton Filter in retrieving the true parameter values and to evaluate the precision and variability of the estimates.

We apply the preceding methodology for $K = 3$, which is the number of assets studied in this thesis. The results are presented in the following tables and figures.

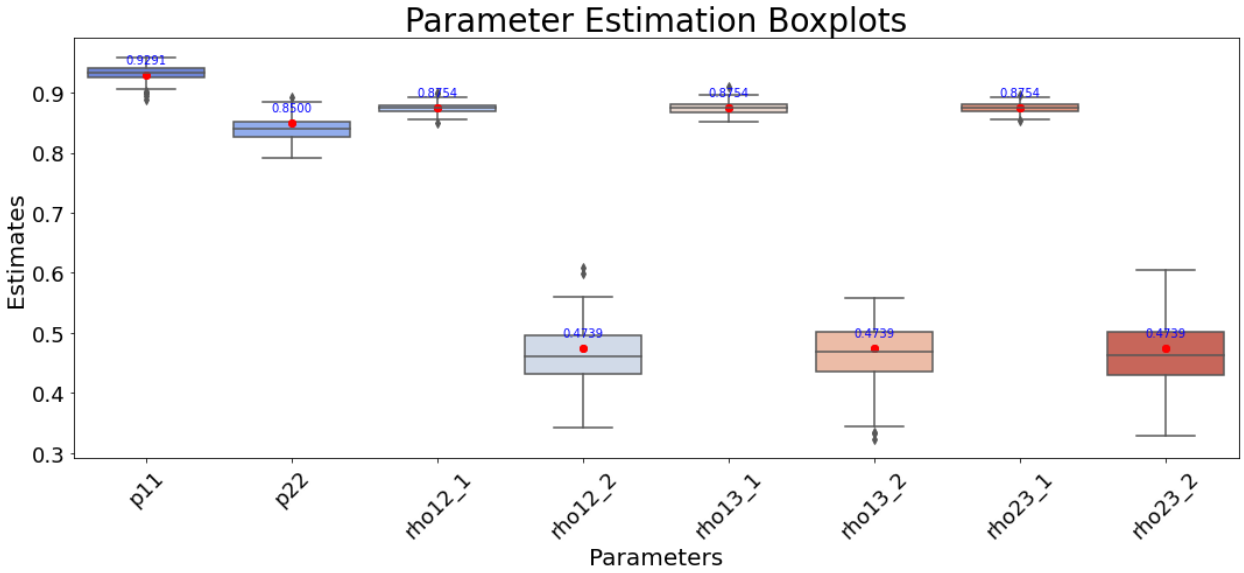


Figure 12: This figure showcases the distribution of estimated parameters for a system comprising three assets, derived from 100 Monte-Carlo simulations. The first two boxplots correspond to the parameters $p11$ and $p22$, respectively. The subsequent boxplots represent correlations, alternating between states: $\Gamma_{1,2,1}$, $\Gamma_{1,2,2}$, $\Gamma_{1,3,1}$, $\Gamma_{1,3,2}$, $\Gamma_{2,3,1}$, and $\Gamma_{2,3,2}$. Each boxplot represents the spread of estimated values for a specific parameter in the vector. The true parameter values, used for simulation, are indicated by a bold point. The results provide insights into the accuracy and consistency of the Hamilton Filter in estimating the model parameters across multiple simulations.

Table 4. T-test Results for Model Parameters
This table displays t-test analyses for estimated model parameters, comparing each estimated parameter ($\hat{\theta}$) with its true value (θ_{True}). The tests are grounded in 100 Monte-Carlo simulations, with the t-statistic calculated by $t = \frac{\bar{\theta} - \theta_{\text{True}}}{s/\sqrt{n}}$, where $\bar{\theta}$ is the sample average of the estimated parameters, s is the sample standard deviation, and n is the number of simulations, which is 100 is this experiment. P-values lower than 5% (i.e., $p < 0.05$) suggest that the estimations significantly differ from the true values. The parameters under examination include transition probabilities (p_{11}, p_{22}) and the correlation parameters (Γ).

Parameter	**T-statistic**	**P-value (%)**
$p11$	-0.063	95.02
$p22$	-4.518	0.00
$\Gamma_{1,2,1}$	-1.359	17.77
$\Gamma_{1,2,2}$	-2.349	2.12
$\Gamma_{1,3,1}$	-0.639	52.48
$\Gamma_{1,3,2}$	-2.016	4.71
$\Gamma_{2,3,1}$	-0.591	55.62
$\Gamma_{2,3,2}$	-2.256	2.67

B Quasi Replication of Pelletier(2006) results

In this replication demonstration, we employ the same exchange rate dataset that was used by Pelletier (2006), which originally comes from Harvey et al. (1994) and Kim et al. (1998). This dataset comprises the closing exchange rates of four currencies—Pound, Deutschmark, Yen, and Swiss Franc—all against the US dollar. The data spans from October 1, 1981, to June 28, 1985, encompassing a total of 946 observations for each currency. The time-series for these exchange rates are visually represented in Figure 14.

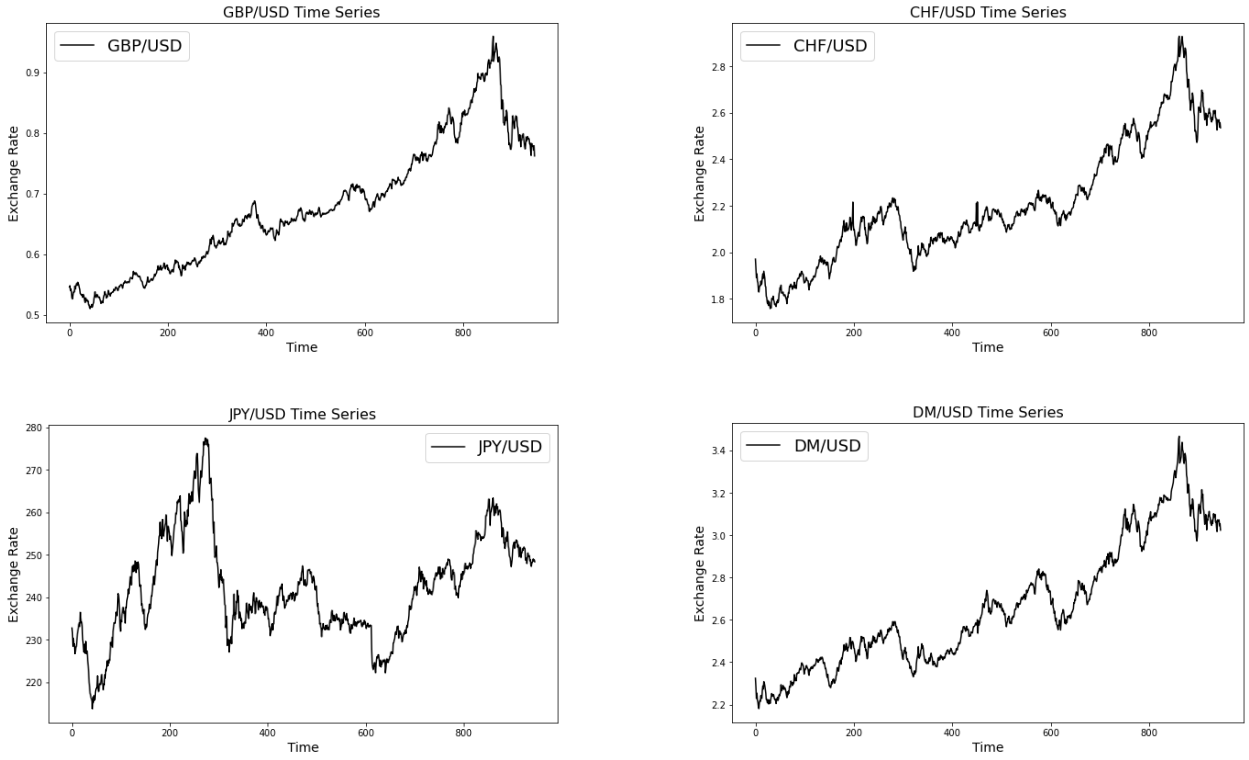


Figure 13: Exchange rate time-series for four currencies against the USD.

Following the preprocessing steps delineated by Pelletier, we calculate 100 times the first difference of the logarithm for each of the currency series and then subtract the sample mean. These are the transformed series on which the volatility model is applied, and they are represented as Y_t in Equation 1. The time-series for these transformed data are visually represented in Figure 14.

This dataset serves as an invaluable resource for our analysis. It was originally used by Harvey et al. (1994) to demonstrate a multivariate stochastic volatility model, assuming constant correlations over time. In our study, we replicate and extend Pelletier's RSDC model with the goal of scrutinizing this assumption and shedding further light on the dynamics of these exchange rates.

Our model serves as a sub-model of the one proposed by Pelletier (2006). Specifically, before applying the filter for regime-switching detection, we standardize the dataset using a variance model. In this regard, the time-varying variance-covariance matrix is decomposed, and Pelletier's ARMACH(1,1) process is used to model the conditional standard deviations:

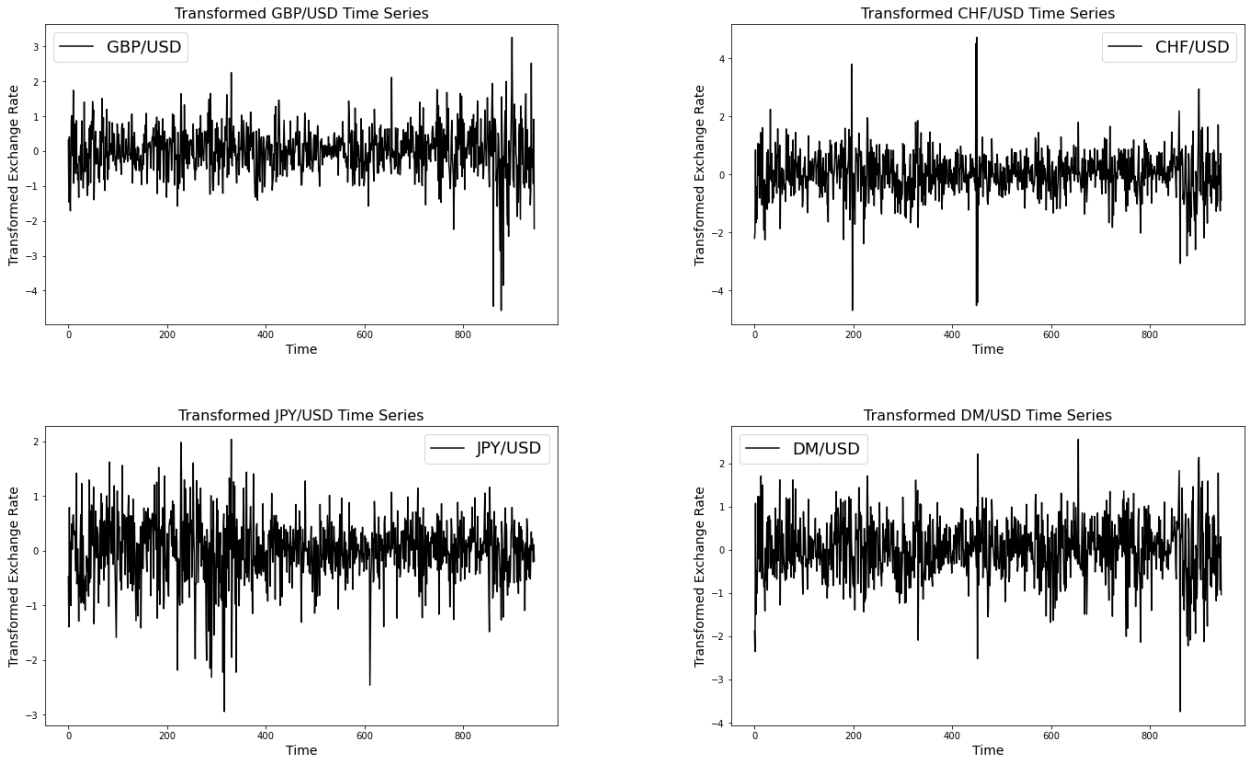


Figure 14: Exchange rate time-series for four currencies against the USD.

$$s_t = \omega + \tilde{\alpha}|y_{t-1}| + \gamma s_{t-1} \tag{14}$$

The transformation we apply aligns with Equation 14, where each univariate exchange rate time-series is filtered by their respective conditional standard deviations. To elaborate, consider the four series of exchange rates as shown in Equation (14):

$$\text{For } k = 1, \ldots, 4, y_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,946}]$$

We fit an ARMACH(1,1) model to each series, obtaining conditional standard deviations s_k for each series as follows:

$$\text{For } i = 1, \ldots, 4, s_k = [s_{k,1}, s_{k,2}, \ldots, s_{k,946}]$$

The final transformation leads to a filtered series u_k, given by:

$$\text{For } k = 1, \ldots, 4, u_k = \frac{y_k}{s_k} = \left[\frac{y_{k,1}}{s_{k,1}}, \frac{y_{k,2}}{s_{k,2}}, \ldots, \frac{y_{k,946}}{s_{k,946}}\right] = [u_{k,1}, u_{k,2}, \ldots, u_{k,946}], U_t = \begin{pmatrix} u_{1,t} \\ u_{2,t} \\ u_{3,t} \\ u_{4,t} \end{pmatrix}$$

The estimated parameters for the ARMACH(1,1) models applied to each series are summarized in Table 5. The time-varying conditional standard deviations are plotted in Figure 15. Finally, the filtered series U_t are depicted in Figure 16.

Table 5. Estimated Parameters for the ARMACH(1,1) Models
This table displays the estimated parameters (ω, α_1, and γ_1) for the ARMACH(1,1) models as defined by Equation 14. These parameters are essential for understanding the volatility and persistence in the exchange rates for various currencies including GBP/USD, DM/USD, JPY/USD, and CHF/USD.

Currency	ω	α_1	γ_1
GBP/USD	0.0147	0.1157	0.8902
DM/USD	0.0436	0.1302	0.8335
JPY/USD	0.0059	0.0287	0.9676
CHF/USD	0.1149	0.1656	0.7251

We apply the Hamilton Filter (see Subsection 2.4) on the transformed datasets. Our estimation results, including both the correlation parameters and transition probabilities, are showcased in Table 6. The standard deviations of these parameter estimates, indicative of their precision, are presented below each estimate in parentheses. These were calculated using the Fisher Information Matrix, details of which are provided in Appendix Section D. For comparative purposes, we juxtapose our findings with those reported by Pelletier (2006), enabling a direct assessment of the generalizability and effectiveness of our generalized model.

Table 6: Comparison of Estimated Correlation Parameters and Transition Probabilities

	Our Model		Pelletier (2006)	
Estimates	Regime 1	Regime 2	Regime 1	Regime 2
Γ_{12}	0.8315	0.3284	0.8754	0.4011
	(0.0416)	(0.0164)	(0.0292)	(0.0098)
Γ_{13}	0.7007	0.0269	0.7656	0.1859
	(0.0350)	(0.0013)	(0.0363)	(0.0996)
Γ_{14}	0.8084	0.2280	0.8569	0.3255
	(0.0404)	(0.0114)	(0.0283)	(0.1275)
Γ_{23}	0.8111	0.3729	0.8471	0.4739
	(0.0406)	(0.0186)	(0.0181)	(0.0843)
Γ_{24}	0.9378	0.3924	0.9510	0.5626
	(0.0469)	(0.0196)	(0.0061)	(0.1871)
Γ_{34}	0.8271	0.1935	0.8617	0.3250
	(0.0414)	(0.0097)	(0.0184)	(0.0166)
p_{11}	0.9496	-	0.9291	-
	(0.0475)		(0.0356)	
p_{22}	0.6791	-	0.6666	-
	(0.0339)		(0.0605)	

This table presents a comparative analysis of the estimated correlation parameters (Γ) and transition probabilities (p_{11} and p_{22}) between our generalized model and the RSDC model reported in Pelletier (2006). Standard deviations are provided below each estimate in parentheses to assess the precision of the estimates.

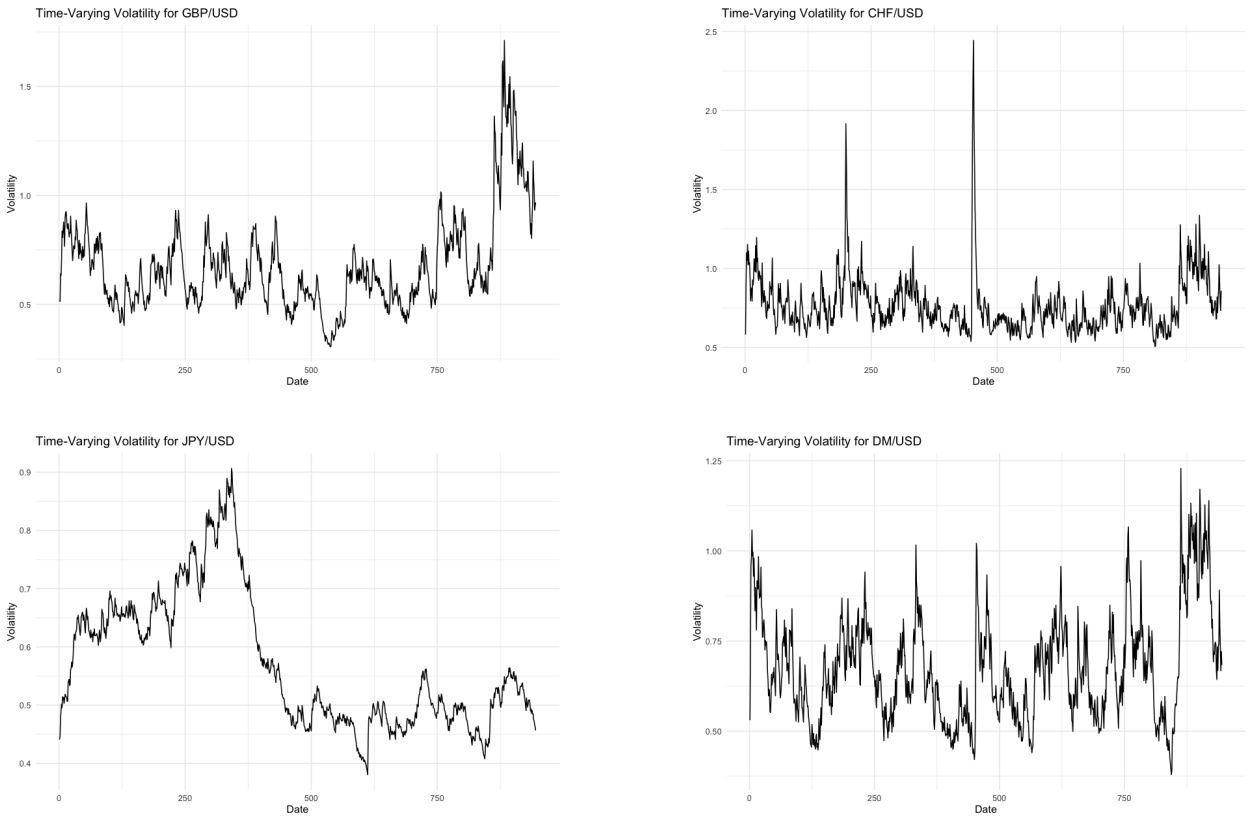


Figure 15: Time-varying conditional standard deviations for each series

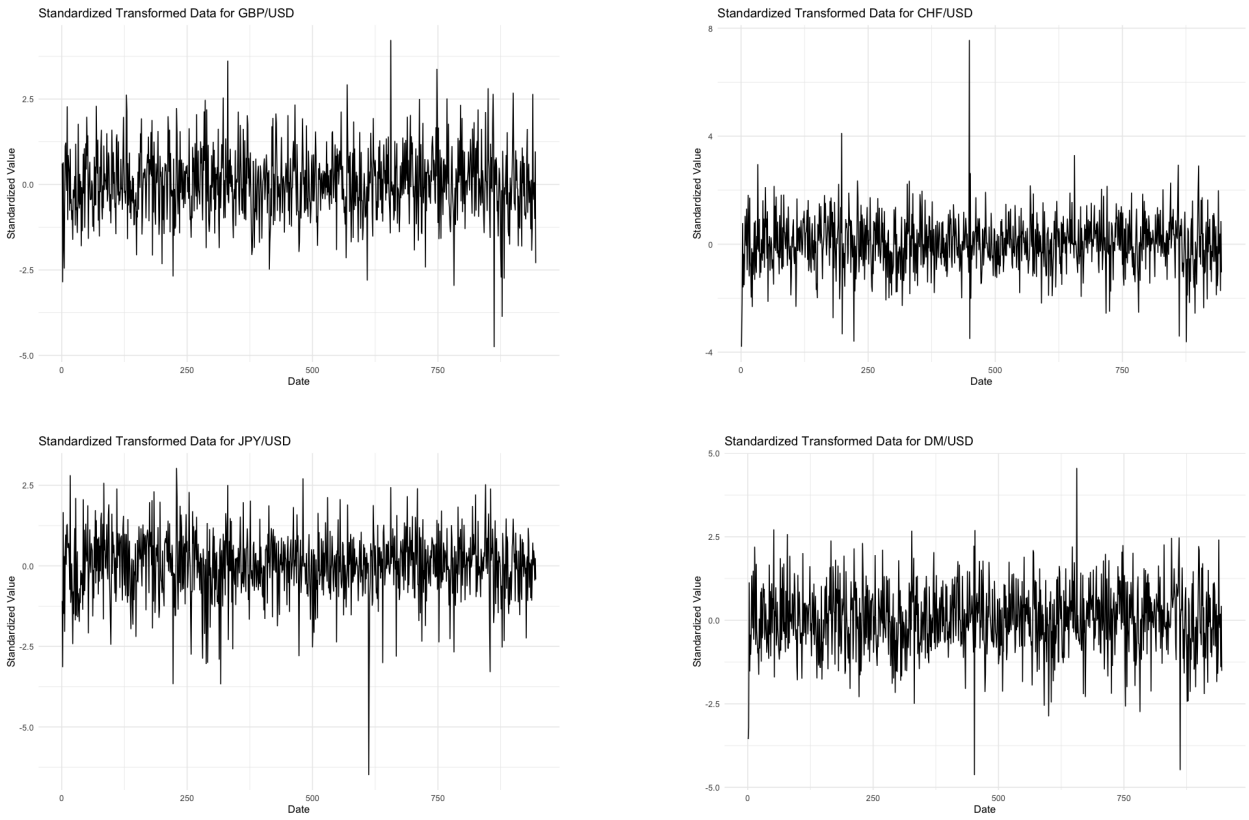


Figure 16: Filtered exchange rate series

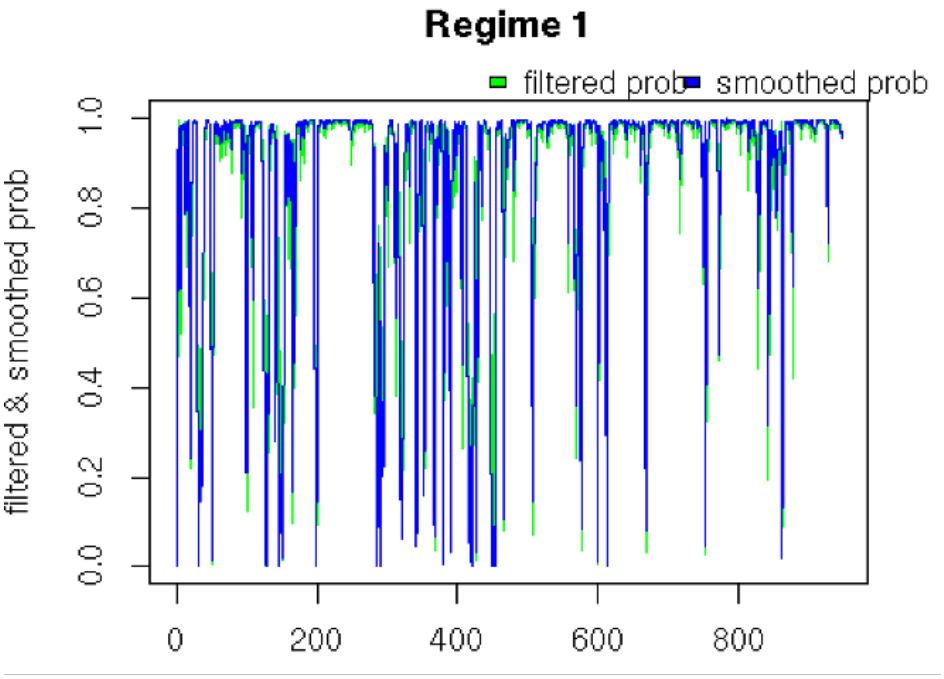


Figure 17: Smoothed and Filtered probabilities

(a) This figure illustrates the smoothed and filtered probabilities obtained from our quasi replication of Pelletier's methodology. It provides a visual representation of the forecasted state probabilities over time, allowing for a direct comparison with the original results presented in Pelletier (2006).

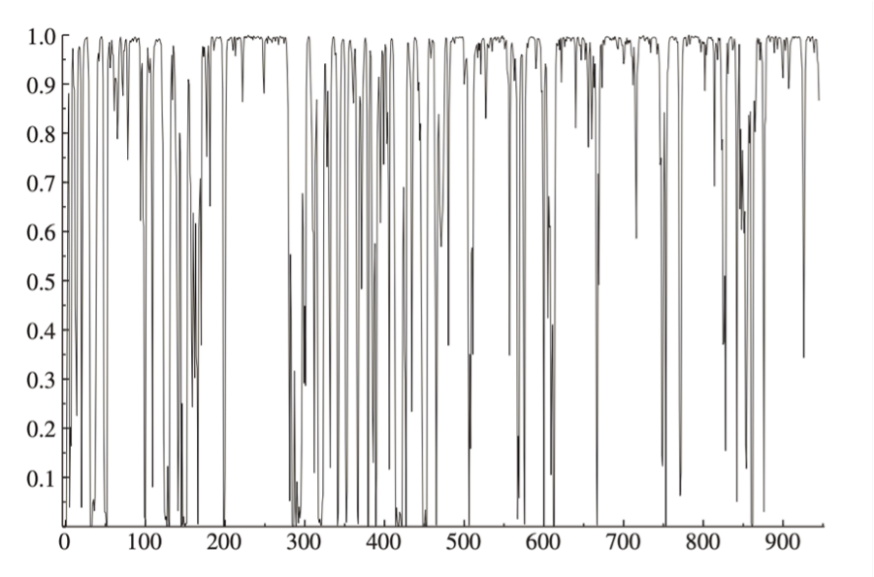


Figure 18: Smooth probabilities of the two-regime RSDC model.

(a) This figure, extracted directly from Pelletier (2006), showcases the smoothed probabilities derived from the two-regime RSDC model. It serves as the benchmark against which our quasi-replication's results are evaluated, highlighting the similarities and potential differences between the two sets of findings.

C Likelihood Ratio Test

The Likelihood Ratio (LR) test is a statistical method used to compare two nested models to ascertain which one provides a better fit to the observed data. Specifically, it tests the null hypothesis H_0 that a simpler (restricted) model is true against an alternative hypothesis H_1 that a more complex (extended) model is true. The LR test statistic is calculated as follows:

$$LR = -2\left(\log \text{Likelihood of Restricted Model} - \log \text{Likelihood of Extended Model}\right)$$

Under the null hypothesis, the LR test statistic follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters between the two models. The steps to conduct an LR test are as follows:

1. Fit the restricted and extended models to the data and calculate their log-likelihoods.
2. Compute the LR test statistic using the formula above.
3. Determine the degrees of freedom, which is the difference in the number of parameters between the extended and restricted models.
4. Use the chi-square distribution to find the p-value corresponding to the computed LR test statistic and the determined degrees of freedom.
5. Reject or fail to reject the null hypothesis based on the p-value. Typically, a p-value below 0.05 is considered significant, indicating that the extended model provides a significantly better fit to the data than the restricted model.

By performing the LR test, researchers can rigorously assess whether the addition of extra parameters in the extended model meaningfully improves the fit, or if the simpler restricted model is sufficient.

D Calculation of Standard Deviations using Fisher Information Matrix

The Fisher Information Matrix is an essential tool in statistical estimation for measuring the amount of information that an observable random variable carries about an unknown parameter upon which the probability of the random variable depends. In the context of our work, we employ it to calculate the standard deviations of our parameter estimates.
Recall the log-likelihood function $l(u_1, u_2, \ldots, u_T|\theta)$ as defined in Equation 12. After finding the set of parameters $\hat{\theta}$ that maximize this function, we evaluate the Hessian matrix H at these estimated parameters:

$$H = -\nabla^2 l(\hat{\theta}),$$

where $\nabla^2 l(\hat{\theta})$ is the second derivative of the log-likelihood function evaluated at $\hat{\theta}$.
The Fisher Information Matrix $I(\hat{\theta})$ is then obtained as the inverse of the Hessian:

$$I(\hat{\theta}) = H^{-1}.$$

The diagonal elements of $I(\hat{\theta})$ provide the variances of the parameter estimates, and taking the square root of these diagonal elements gives us their standard deviations. These standard deviations are reported in Table 6 in the Appendix section B.

```python
import numpy as np
import pandas as pd

def hamilton_filter(theta, y, x):
    """
    Compute the log-likelihood of the Hamiltonian Markov Switching Model with two regimes.

    Parameters:
    theta (np.array): Parameter vector including betas and rhos.
    y (np.array): Response variable, assumed to be a two-dimensional numpy array.
    x (np.array): Explanatory variables, assumed to be a numpy array.

    Returns:
    float: Negative log-likelihood of the model.
    """

    # Reshape y and x if they are Pandas DataFrames
    if not isinstance(y, np.ndarray):
        y = y.values
    if not isinstance(x, np.ndarray):
        x = x.values

    n_var, n_exog = y.shape[1], x.shape[1] if len(x.shape) > 1 else 1
    n_rho = n_var * (n_var - 1) // 2

    # Split theta into betas and rhos
    beta_1, beta_2 = theta[:n_exog], theta[n_exog:2*n_exog]
    rho_1 = [2 / (1 + np.exp(-rho_val)) - 1 for rho_val in theta[2*n_exog:2*n_exog + n_rho]]
    rho_2 = [2 / (1 + np.exp(-rho_val)) - 1 for rho_val in theta[2*n_exog + n_rho:2*n_exog + 2*n_rho]]

    # Create and populate correlation matrices
    corr_mtx_1, corr_mtx_2 = np.ones((n_var, n_var)), np.ones((n_var, n_var))
    corr_mtx_1[np.triu_indices(n_var, 1)] = rho_1
    corr_mtx_2[np.triu_indices(n_var, 1)] = rho_2
    corr_mtx_1 = np.triu(corr_mtx_1).T + np.triu(corr_mtx_1, 1)
    corr_mtx_2 = np.triu(corr_mtx_2).T + np.triu(corr_mtx_2, 1)

    # Compute inverse covariance matrices and determinants
    inv_cov_mtx_1 = np.linalg.pinv(corr_mtx_1)
    inv_cov_mtx_2 = np.linalg.pinv(corr_mtx_2)
    det_inv_cov_mtx_1 = np.linalg.det(2 * np.pi * corr_mtx_1)**(-1/2)
    det_inv_cov_mtx_2 = np.linalg.det(2 * np.pi * corr_mtx_2)**(-1/2)

    # Conditional densities
    f1 = det_inv_cov_mtx_1 * np.exp(-0.5 * np.einsum('ij,ij->i', np.dot(y, inv_cov_mtx_1), y))
    f2 = det_inv_cov_mtx_2 * np.exp(-0.5 * np.einsum('ij,ij->i', np.dot(y, inv_cov_mtx_2), y))
    f = np.column_stack((f1, f2))
```

```
48
49      # State probabilities and model likelihood
50      S_inf = np.zeros_like(f)
51      S_forecast = np.zeros_like(f)
52      model_lik = np.zeros(y.shape[0])
53      p11 = 1 / (1 + np.exp(-np.dot(beta_1, x[0])))
54      p22 = 1 / (1 + np.exp(-np.dot(beta_2, x[0])))
55      S_inf[0, :] = np.array([p11, p22]) * f[0, :] / np.dot([1, 1], np.array
    ([p11, p22]) * f[0, :])
56      S_forecast[0] = [0.5, 0.5]
57
58      for t in range(y.shape[0] - 1):
59          p11 = 1 / (1 + np.exp(-np.dot(beta_1, x[t])))
60          p22 = 1 / (1 + np.exp(-np.dot(beta_2, x[t])))
61          P = np.array([[p11, 1 - p11], [1 - p22, p22]])
62          S_forecast[t + 1, :] = P.T @ S_inf[t, :]
63          S_inf[t + 1, :] = S_forecast[t + 1, :] * f[t + 1, :] / (S_forecast
    [t + 1, :] @ f[t + 1, :])
64          model_lik[t + 1] = np.dot([1, 1], S_forecast[t + 1, :] * f[t + 1,
    :])
65
66      # Compute log-likelihood
67      logl = np.sum(np.log(model_lik[1:]))
68      return -logl if np.isfinite(logl) else -1e10
69
70
71
72
73  def Smooth_hamilton_filter(theta, y, x, reg='Smth'):
74      """
75      Compute smoothed and forecasted probabilities for a Hamiltonian Markov
        Switching Model.
76
77      Parameters:
78      theta (np.array): Array of parameters, including betas and rhos.
79      y (np.array): Response variable, assumed to be a two-dimensional numpy
        array.
80      x (np.array): Explanatory variables, assumed to be a numpy array.
81      reg (str): Specifies the type of probabilities to use for regime
       determination.
82
83      Returns:
84      pd.DataFrame: DataFrame with smoothed probabilities, forecast
       probabilities,
85                    filter probabilities, regime and switch indicators.
86      """
87
88      # Ensure y and x are numpy arrays
89      if not isinstance(y, np.ndarray):
90          y = y.values
91      if not isinstance(x, np.ndarray):
92          x = x.values
93
94      n_var, n_exog = y.shape[1], x.shape[1] if len(x.shape) > 1 else 1
```

```
    n_rho = n_var * (n_var - 1) // 2

    # Extract betas and rhos from theta
    beta_1, beta_2 = theta[:n_exog], theta[n_exog:2*n_exog]
    rho_1 = [2 / (1 + np.exp(-rho_val)) - 1 for rho_val in theta[2*n_exog
: 2*n_exog + n_rho]]
    rho_2 = [2 / (1 + np.exp(-rho_val)) - 1 for rho_val in theta[2*n_exog
+ n_rho : 2*n_exog + 2*n_rho]]

    # Create correlation matrices
    corr_mtx_1, corr_mtx_2 = np.ones((n_var, n_var)), np.ones((n_var,
n_var))
    corr_mtx_1[np.triu_indices(n_var, 1)] = rho_1
    corr_mtx_2[np.triu_indices(n_var, 1)] = rho_2
    corr_mtx_1 = np.triu(corr_mtx_1).T + np.triu(corr_mtx_1, 1)
    corr_mtx_2 = np.triu(corr_mtx_2).T + np.triu(corr_mtx_2, 1)

    # Calculate inverse covariance matrices and determinants
    inv_cov_mtx_1 = np.linalg.inv(corr_mtx_1)
    inv_cov_mtx_2 = np.linalg.inv(corr_mtx_2)
    det_inv_cov_mtx_1 = np.linalg.det(2 * np.pi * corr_mtx_1)**(-1/2)
    det_inv_cov_mtx_2 = np.linalg.det(2 * np.pi * corr_mtx_2)**(-1/2)

    # Conditional densities
    f1 = det_inv_cov_mtx_1 * np.exp(-0.5 * np.einsum('ij,ij->i', np.dot(y,
    inv_cov_mtx_1), y))
    f2 = det_inv_cov_mtx_2 * np.exp(-0.5 * np.einsum('ij,ij->i', np.dot(y,
    inv_cov_mtx_2), y))
    f = np.column_stack((f1, f2))

    # Initialize forecast and filtered state probabilities
    S_forecast = np.zeros((y.shape[0], 2))
    S_inf = np.zeros_like(S_forecast)

    # Calculate initial state probabilities
    p11 = 1 / (1 + np.exp(-np.dot(beta_1, x[0])))
    p22 = 1 / (1 + np.exp(-np.dot(beta_2, x[0])))
    S_inf[0, :] = (np.array([p11, p22]) * f[0, :]) / (np.array([1, 1]) @ (
np.array([p11, p22]) * f[0, :]))
    S_forecast[0] = [0.5, 0.5]

    # Iterate over time series
    for t in range(y.shape[0] - 1):
        p11 = 1 / (1 + np.exp(-np.dot(beta_1, x[t])))
        p22 = 1 / (1 + np.exp(-np.dot(beta_2, x[t])))
        P = np.array([[p11, 1 - p11], [1 - p22, p22]])
        S_forecast[t + 1, :] = P.T @ S_inf[t, :]
        S_inf[t + 1, :] = (S_forecast[t + 1, :] * f[t + 1, :]) / (
S_forecast[t + 1, :] @ f[t + 1, :])

    # Smooth the probabilities
    T = y.shape[0]
    P_smooth = pd.DataFrame({'s1': np.zeros(T), 's2': np.zeros(T)})
    P_smooth.iloc[T - 1] = S_inf[T - 1]
```

```
    for is_ in range(T - 2, -1, -1):
        p11 = 1 / (1 + np.exp(-np.dot(beta_1, x[is_])))
        p22 = 1 / (1 + np.exp(-np.dot(beta_2, x[is_])))

        # Calculate smoothed probabilities
        p1 = (S_inf[is_ + 1, 0] * S_inf[is_, 0] * p11) / S_forecast[is_ +
    1, 0]
        p2 = (S_inf[is_ + 1, 1] * S_inf[is_, 0] * (1 - p11)) / S_forecast[
    is_ + 1, 1]
        p3 = (S_inf[is_ + 1, 0] * S_inf[is_, 1] * (1 - p22)) / S_forecast[
    is_ + 1, 0]
        p4 = (S_inf[is_ + 1, 1] * S_inf[is_, 1] * p22) / S_forecast[is_ +
    1, 1]
        P_smooth.iloc[is_, 0] = p1 + p2
        P_smooth.iloc[is_, 1] = p3 + p4

    # Compile results into DataFrame
    S_forecast_df = pd.DataFrame(S_forecast, columns=['ForeP1', 'ForeP2'])
    S_inf_df = pd.DataFrame(S_inf, columns=['FiltP1', 'FiltP2'])
    result_df = pd.concat([P_smooth, S_forecast_df, S_inf_df], axis=1)
    result_df.columns = ['SmthP1', 'SmthP2', 'ForeP1', 'ForeP2', 'FiltP1',
    'FiltP2']

    # Determine regime and switching points
    reg_ind = reg + 'P1'
    result_df["regime"] = np.where(result_df[reg_ind] > 0.5, 1, 2)
    result_df["switch"] = np.where(result_df["regime"] != result_df["
    regime"].shift(1), 1, 0)

    return result_df
```

Listing 1: Hamiltonian Markov Switching Model Code