Modelling regime shifts for foreign exchange market data using hidden Markov models

LIAM PERSSON
Abstract

Financial data is often said to follow different market regimes. These regimes, which not possible to observe directly, are assumed to influence the observable returns. In this thesis such regimes are modeled using hidden Markov models. We will investigate whether the five different currency pairs EUR/NOK, USD/NOK, EUR/USD, EUR/SEK, and USD/SEK exhibit market regimes that can be described using hidden Markov modeling. We will find the most optimal number of states and study the mean, variance, and correlations in each market regime.

Keywords: exchange market data, model selection, hidden Markov model, market regime, correlation
*Modellering av regimskiften för valutamarknadsdata genom dolda Markovkedjor*

**Sammanfattning**


Nyckelord: valuta, modellval, dold Markovkedja, marknadsregimer, korrelation
Acknowledgements

This paper would not have been possible to write without Guillaume Desticourt and Skandinaviska Enskilda Banken (SEB). Desticourt has been helpful with discuss problems and improvements. SEB has also supplied the data that is used in the report.

I will also express my gratitude to my supervisor from KTH Jimmy Olsson who have been very helpful during the project.

Stockholm, May 2021
Liam Persson
Contents

1. Introduction .............................................................................................................................................. 5

2. Data .......................................................................................................................................................... 8

3. Theory ..................................................................................................................................................... 11
   3.1 The foreign exchange market ........................................................................................................... 11
   3.1.1 The foreign exchange rate .......................................................................................................... 11
   3.1.2 Factors that influence the foreign exchange market .................................................................. 11
   3.2 The hidden Markov model ............................................................................................................. 12
   3.2.1 The hidden Markov model .......................................................................................................... 12
   3.2.2 The forward-backward algorithm .............................................................................................. 14
   3.2.3 The Viterbi algorithm ................................................................................................................. 15
   3.2.4 Baum-Welch algorithm .............................................................................................................. 16
   3.2.5 Model selection .......................................................................................................................... 18

4. Results ..................................................................................................................................................... 20
   4.1 Model selection ............................................................................................................................... 20
   4.2 Hidden Markov models with one currency pair ................................................................................. 21
       4.2.1 Returns ..................................................................................................................................... 21
       4.2.2 Transition matrices .................................................................................................................... 23
       4.2.3 Smoothed probabilities ............................................................................................................ 25
       4.2.4 Viterbi path .............................................................................................................................. 26
   4.3 Hidden Markov models with three currency pairs ............................................................................. 28
       4.3.1 EUR currencies ....................................................................................................................... 28
       4.3.2 USD currencies ....................................................................................................................... 33

5. Discussion ............................................................................................................................................... 38
   5.1 Discussion of the results .................................................................................................................. 38
   5.2 Further research .............................................................................................................................. 40

References .................................................................................................................................................... 41
1. Introduction

To determine better methods for risk management is always relevant for the actors on the financial markets. The stock market is often said to be in a bull or a bear market. A bear market is characterized by falling prices and high volatility and a bull market by raising prices during low volatility. One way to look at these two market regimes is as if the returns are affected by some unobserved parameter. The hidden Markov model incorporates latent states that allow such regimes to be modelled efficiently. The hidden Markov model have been shown to be able to detect these different market regimes (Wang et al. 2020).

A well-used method for presenting the risk of a financial portfolio is the Value-at-risk model (VaR) (Guldimann 1996). There are different approaches to such VaR calculations, depending on the estimates of variance and correlation on the different assets in the portfolio (Holton 2003). Research have shown that the correlations between assets vary over time which validates the assumption that they are constant or close to constant (Rickenberg 2019). The risk of a portfolio tends to be underestimated when the correlations are constant since assets tend to have higher correlations during extreme market events meaning different assets often crash at the same time (Rickenberg 2019).

It has been known for a long time that volatility is coming in waves meaning that large changes in prices tend to be followed by more large changes in prices and vice versa (Mandelbrot 1963). Although the fact that both correlations and variance is proven to vary over time some market actors still use constant variance and correlations when determining risk (Hu and Watt 2014). It is therefore highly relevant to investigate how variance and correlations behave in different market regimes to determine a more accurate risk measure of a portfolio.

It has been shown that the foreign exchange market can be described by a hidden Markov model. The model was shown to generate accurate predictions on the mean and variances of the currency pairs (Lee and Ling 2014). The problem of market regimes with different volatility have been researched with various methods for example stochastic volatility models and GARCH models (Alanya and Rodríguez 2018). Still, as far as known to the author of the present thesis, no research on how the correlation matrices look like in the different regimes
has been carried through. Thus, the aim of this project is to examine if it is possible to show that the foreign exchange market seems to have different parameters in the various market regimes. If so, how many different market regimes explain the process most accurate? How the market regimes distinguish from each other in form of expected returns and volatility will also be researched but focus will be on the correlations. The idea is that a better understanding of the behaviours of the foreign exchange market will lead to better insights which will lead to better risk methods and better risk management.

The hidden Markov model (HMM) have been used as a model to identify different market regimes in various financial markets. HMM assume that the process is influenced by some states that are not directly observable and is therefore referred to as hidden states (Westhead et al. 2017). Even if the hidden states are not direct observable, they influence the observations and can therefore be estimated based on observed data (Cappe et al. 2005). Since the model typically tend to stay in the same state, they make it possible to model the returns and consider the serial correlations that market data tend to have (Mandelbrot 1963). The model has also been shown to capture the trending that is often seen in financial data but that is not handled in traditional risk measurement methods like VaR. Studies have shown that HMM are able to construct a portfolio that can create both higher returns and higher risk-adjusted returns at the stock market (Wang et al. 2020). Except for the use in financial data hidden Markov models have also been used in speech recognition for a long time (Rabiner 1989).

An HMM can be contracted with different number of regimes, and it is therefore of great interest to have a method for selecting the most accurate model given the data. The maximum likelihood estimator is often used for variable selection in various mathematical problems. However, this method is not applicable to use directly for HMM since a higher number of states will always result in a higher likelihood for an HMM. There are plenty of different methods that use the maximum likelihood method and then have some penalty for increasing number of estimated parameters. For this paper the two well-known methods Bayesian information criterion (BIC) and Akaike information criteria (AIC) is used. The AIC method by Akaike (1973) was trying to use the log likelihood but to penalties a model with higher number of states. The method BIC method was some years later by Schwarz (1978) and the idea is that a model with fewer number of variables should be chosen in favour of a model with a larger number of variables, just like in the AIC model. The BIC model also takes the
number of observations into account. A model that has more observations can have more parameters without a significant risk of overfitting.

In the light of the previous, the present thesis will explore the following two research questions:

- Can hidden models be used for modelling regimes shifts in foreign exchange data?
- Is the model able to capture correlations between currency pairs?

In section 2 we will get acquainted with the data serving as a basis of the thesis. In section 3 we will review the hidden Markov models and describe the foreign exchange market that will be modelled. In section 4 the results that is obtained from the study will be presented. The results will be discussed and interpreted in section 5.
2. Data

The raw data are supplied by SEB and contains close prices for different foreign exchange pairs. Together with SEB relevant exchange pairs for the study are selected. The daily returns are then calculated by dividing yesterday’s close price by today’s close price. The data contains daily returns from 3928 trading days with the first trading day 2006-03-15 and the last trading day 2021-03-16.

<table>
<thead>
<tr>
<th>Returns for foreign exchange pairs (%)</th>
<th>EUR/NOK</th>
<th>EUR/SEK</th>
<th>EUR/USD</th>
<th>USD/NOK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max</td>
<td>6.663</td>
<td>3.093</td>
<td>3.511</td>
<td>7.525</td>
<td>4.313</td>
</tr>
<tr>
<td>Mean</td>
<td>0.007</td>
<td>0.003</td>
<td>0.001</td>
<td>0.010</td>
<td>0.006</td>
</tr>
<tr>
<td>Min</td>
<td>-4.271</td>
<td>-2.942</td>
<td>-2.404</td>
<td>-4.849</td>
<td>-4.857</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.550</td>
<td>0.453</td>
<td>0.574</td>
<td>0.793</td>
<td>0.751</td>
</tr>
</tbody>
</table>

The returns of the currencies are given in Figures 1-5. When looking at the daily returns it looks like high volatility are coming in waves. Periods of high volatility tend to be followed by periods of high volatility and periods of low volatility tend to be followed by periods of low volatility. In Figure 6 the correlations between the five currency pairs for the whole data period are presented. We can see that especially the currency pairs that contain USD have a high correlation to the other currencies that contain USD.

We can also see that the periods of volatility clusters seem to occur at the same time for all the currency pairs. The clearest examples of this are the great recession in 2008 and the Corona crisis in the beginning of 2020. All five currency pairs have clear volatility spikes during these big market events.
Figure 1 Daily returns for EUR/NOK

Figure 2 Daily returns for USD/NOK

Figure 3 Daily returns for EUR/USD

Figure 4 Daily returns for EUR/SEK

Figure 5 Daily returns for USD/SEK
Figure 6 Correlations between the currency pairs
3. Theory

3.1 The foreign exchange market

3.1.1 The foreign exchange rate

The price of a currency is always relative and priced in relation to another currency. A foreign exchange rate is therefore always given as a pair. A currency pair, say $X = \text{Currency}_1/\text{Currency}_2$ describes how many units of Currency2 you must pay for one unit of Currency1. A higher exchange rate $X$ is equivalent to an increase in the value of Currency1 in relation to Currency2 and vice versa.

If the currencies in a pair change order the exchange rate is inverted:

$$\frac{\text{Currency}_1}{\text{Currency}_2} = X \iff \frac{\text{Currency}_2}{\text{Currency}_1} = \frac{1}{X}$$

The order of the currencies also changes the correlation between the currency pairs. If the currencies in one of the pairs change order the correlation turns negative:

$$\text{Cor}(\frac{\text{Currency}_1}{\text{Currency}_2}, \frac{\text{Currency}_1}{\text{Currency}_3}) = Y \iff \text{Cor}(\frac{\text{Currency}_2}{\text{Currency}_1}, \frac{\text{Currency}_1}{\text{Currency}_3}) = -Y$$

3.1.2 Factors that influence the foreign exchange market

There are plenty of different long-term factors that influence the exchange rates. Some of the most used is the difference in inflation and interest rates, current account deficit and public debt. But there are also factors that is broader like strong economic performance.

There are also short-term factors that can be very influential for the foreign exchange market. Some currencies, as the USD is known as “safe haven” currencies, and they tend to appreciate in relation to most of the other currencies during crises and times of uncertainty (Ranaldo and Söderlind 2010).

It has also been shown that currencies that are well used in the world tend to influence the exchange rate of currencies that is used in the same geographical area. One example of this is that the currencies of European currencies tend to follow the EURO (Yong et al. 2018).
3.2 The hidden Markov model

The model that has been chosen for the research is the hidden Markov model. The model assumes that the output is affected by some state that is not directly observable and is therefore called hidden.

3.2.1 The hidden Markov model

![Figure 7: A hidden Markov model visualized. Inspiration from (Bulla 2006).](image)

The different currency pairs in the model are given by \( C = \{\text{Currency pair}(1), \ldots, \text{Currency pair}(c)\} \) where \( c \) stands for the total number of currency pairs. A hidden model consists of a sequence of observation vectors \( Y_{0:n} = \{y_0, y_1, \ldots, y_n\} \) where every timestep contains one observation for every currency pair.

\[
y_i = \begin{bmatrix} y_{i}^{c(1)} \\ \vdots \\ y_{i}^{c(c)} \end{bmatrix} \text{ for } i = 0, \ldots, n
\]

The matrix \( Y_{0:n} \) is observable and assumed to be dependent upon some hidden state that is not observable directly \( X_{0:n} = \{x_0, x_1, \ldots, x_n\} \) where \( n \) stands for time. The Markov chain \( X \) evolves on some finite set \( X = \{1, 2, \ldots, r\} \) where \( r \) stands for the total amount of states in the model. The set of possible observations is described as \( Y = \mathbb{R}^c \) (Cappe et al. 2005).
The transition probability matrix $Q$ contains the probabilities for transition from the current state to each of the other states $Q = \{q_{ij} = P(X_{k+1} = j|X_k = i)\}$ where $k = 0,\ldots,n$ represents time and $i,j = 1,\ldots,n$ represents states. The probability that the next state will be $j$ given that we know that the current state is $i$ is given by $q_{ij}$ (Cappe et al. 2005).

Probability density function is given by

$$g(x,y;\theta) = \frac{1}{2\pi} \det(\Sigma_x)^{-0.5} e^{-\frac{1}{2}(y-\mu_x)^T \Sigma_x^{-1} (y-\mu_x)}$$

The information in $\theta$ is the transition probabilities $q_{ij}$ for $i,j = 1,\ldots,r$. It also contains the means $\mu_i$ in vector $\mu_i = \begin{pmatrix} \mu_i^{C(1)} \\ \vdots \\ \mu_i^{C(c)} \end{pmatrix}$ for $i = 1,\ldots,r$

The last parameter in $\theta$ is the covariance matrix

$$\Sigma_i = \begin{pmatrix} \text{Cov}(C(1), C(1))_i & \cdots & \text{Cov}(C(1), C(c))_i \\ \vdots & \ddots & \vdots \\ \text{Cov}(C(c), C(1))_i & \cdots & \text{Cov}(C(c), C(c))_i \end{pmatrix}$$

that consists of the covariances between the different currency pairs for every state, $i = 1,\ldots,r$ (Cappe et al. 2005).

Starting probability vector $V = \{v_i = P(X_0 = i)\}$ for $i = 1,\ldots,r$ describes the probability to start in each state.

When operating on a hidden Markov model, one typically faces three problems that are solved with different algorithms

1. To calculate the probability for every state given the observations which is solved by the forward-backward algorithm.
2. The problem to determine the most probable state sequence given the observations, which is solved by the Viterbi algorithm.
3. The last problem is to determine the values for the parameters in the model, and this problem is solved using the Baum Welch algorithm. The algorithm uses the results from the forward-backward algorithm.
3.2.2 The forward-backward algorithm

The forward-backward algorithm is used for calculating the posterior probability for each state for every time observation (Rabiner 1989). These probabilities are also called for the smoothed probabilities. The method makes it possible to determine the posterior probability for each market regime at every time. The method is used in following methods and is also usable if one wants to understand the probability for some state at a specific time. This method can calculate the posterior probability but is not able to calculate the probability for a sequence of states (Westhead et al. 2017).

The filter vectors are given by

$$
\phi_k(i) \triangleq P(X_k = i | Y_{0:k}) \quad \text{for} \quad i \in 1, \ldots, r
$$

and describes the probability to be in each state given the observed observations. Since \( g(x, y) = g(x, Y_k) \) from now on \( g(x, y) \) will be represented by the simplified notation \( g_k(x) \). The forward algorithm uses a recursive formula with the initialization value of

$$
\phi_{0|0}(i) = v(i) \quad \text{for} \quad i \in 1, \ldots, r
$$

The recursive step is given for \( k = 0, \ldots, n \) and \( j \in 1, \ldots, r \)

$$
c_k = \sum_{i=1}^{r} \phi_{k|k-1}(i) g_k(i).
$$

$$
\phi_k(j) = \frac{\phi_{k|k-1}(j) g_k(j)}{c_k}.
$$

$$
\phi_{k+1|k}(j) = \sum_{i=1}^{r} \phi_k(i) q_{ij}.
$$

The backward algorithm uses stored values on \( \phi_k \) for \( k = 0, \ldots, n \) and \( c_k \) for \( k = 0, \ldots, n \).

The Backward algorithm is initialized by

$$
\hat{\beta}_{n|n}(i) = c_k^{-1} \quad \text{for} \quad i = 1, \ldots, r
$$

And the recursive step is given for \( k = n-1, \ldots, 0 \) for \( i = 1, \ldots, r \)

$$
\hat{\beta}_{k|n}(i) = c_k^{-1} \sum_{j=1}^{r} q_{ij} g_{k+1}(j) \hat{\beta}_{k+1|n}(j)
$$

The backward forward algorithm combines the backward and forward algorithm to calculate the marginal smoothing probability for all \( k < n \)

$$
\phi_{k|n}(i) \triangleq P(X_k = i | Y_{0:n}) = \frac{\phi_k(i) \hat{\beta}_{k|n}(i)}{\sum_{j=1}^{r} \phi_k(j) \hat{\beta}_{k|n}(j)}
$$

The bivariate smoothing probabilities is given by

$$
\phi_{k:k+1|n}(i,j) \triangleq P(X_k = i, X_{k+1} = j | Y_{0:n}) = \phi_k(i) q_{ij} g_{k+1}(j) \hat{\beta}_{k+1|n}(j)
$$
The findings from the forward-backward algorithm will be used in the following methods (Cappe et al. 2005).

3.2.3 The Viterbi algorithm

The next step is to determine the optimal path for the state sequence. One possible method, especially for short sequences would be to calculate $P(X|O, \lambda)$ for every state sequence $X$ and then take the most likely sequence. However, this method would be very inefficient and time consuming, especially for longer sequences. An efficient solution to find the optimal state sequence is given by the method that Viterbi (1967) published as A Probabilistic nonsequential decoding algorithm but now is most known as the Viterbi algorithm. This method computes the most likely state sequence over the observed timeframe. The Viterbi Algorithm consider the probabilities for a transition from one state to another and therefore it is possible to determine the most likely state sequence given all data.

To solve the problem with the most likely state sequence $m_k(i)$ is defined as

$$m_k(i) = \max_{(x_0, \ldots, x_{k-1}) \in X^k} \log \phi_{0:k|k}(x_0, \ldots, x_k, i) + \ell_k$$

Where $\ell_k$ stands for the log-likelihood for the observations up to index $k$.

We can use the formula

$$\log \phi_{0:k+1|k+1}(x_0, \ldots, x_{k+1}) = (\ell_k - \ell_{k-1}) + \log \phi_{0:k|k}(x_0, \ldots, x_k) + \log q_{x_k x_{k+1}}(x_0, \ldots, x_k) + \log g_{k+1}(x_{k+1})$$

Where $\phi_{0:k|k}$ is the joint distribution of the states $X_{0:k}$ given the observations $Y_{0:k}$ and the log likelihood up to index $k$ is given by $\ell_k$ to get the Viterbi algorithm to be recursive

$$m_{k+1}(j) = \max_{i \in \{1, \ldots, r\}} [m_k(i) + \log q_{ij}] + \log g_{k+1}(j)$$

The formula is initialized as

$$m_0(i) = \log(\pi(i) g_0(i))$$

Let $\hat{x}_n$ be the state that maximize $m_n(j)$. Then for $k = n-1, n-2, \ldots, 0$ let $\hat{x}_k$ be the state that maximize

$$m_{k+1}(\hat{x}_k) = \max_{i \in \{1, \ldots, r\}} [m_k(i) + \log q_{i\hat{x}_n}] + \log g_{k+1}(\hat{x}_n)$$

The Viterbi algorithm starts with calculating the probability for the most optimal path so far with all possible states at time $k$. Just the highest probabilities leading to each state in time $k$ continue to be used in the following steps. This makes the algorithm much more efficient than it would be to try every possible state sequence (Cappe et al. 2005).
3.2.4 Baum-Welch algorithm

The last problem is to determine the most optimal parameters on the model. Since the parameters of the model is not directly given, they need to be approximated by some method. The Baum-Welch algorithm is a method that is used to estimate the parameters in the transition matrix and the parameters for the distributions in each state and was first published by Baum et al. (1970). The method is starting with some initial guesses of the true parameters that is then updated to find a local optimal solution. It is important to remember that the method does not guarantee a global optimal solution and it is therefore of great importance to run the model with various starting values. The Baum-Welch algorithm is a special case of the Expectation-Maximization algorithm and uses the backward and forward algorithms for updating the parameters. The model starts with some initial guesses for the parameters and then update them in every interaction to get more accurate solutions to the problem.

We introduce $\theta^*$ as the vector with the initial guesses

$$\theta^* = \left( (\mu_i^*), (v_i^*), (q_i^*) \right)_{i = 1,...,r}$$

We want to define the parameters that result in the highest likelihood for the model parameters $\theta$.

$$\theta = \arg \max_{\theta \in \Theta} \ell(\theta)$$

The log likelihood is dependent on the observations and is given as

$$\ell(\theta) \equiv \log \int f(x,y;\theta) dx$$

The joint probability function for the observations given the parameters is given by

$$f_n(x_0:n, y_0:n; \theta) = \pi(x_0; \theta)g(x_0,y_0;\theta)q(x_0,x_1;\theta)g(x_1,y_1;\theta)\ldots q(x_{n-1},x_n;\theta)g(x_n,y_n;\theta)$$

When the logarithm is applied to the formula we get

$$\log f_n(x_0:n, y_0:n; \theta) = \log \pi(x_0; \theta) + \sum_{k=0}^{n-1} \log q(x_k,x_{k+1};\theta) + \sum_{k=0}^{n} \log g(x_k, y_k;\theta)$$

We now introduce the auxiliary function

$$Q(\theta; \theta') \equiv \log \int \log f(x;\theta)p(x;\theta')dx =$$
\[ E_\theta[\log p(X_0; \theta|Y_{0:n})] + \sum_{k=0}^{n-1} E_\theta[\log q(X_k, X_{k+1}; \theta|Y_{0:n})] + \sum_{k=0}^{n} E_\theta[\log g(X_k, Y_k; \theta|Y_{0:n})] \]

That can be simplified to

\[
Q(\theta; \theta') = C - \frac{1}{2} \sum_{k=0}^{n} E_\theta \left[ \sum_{i=1}^{r} \mathbf{1}\{X_k = i\} \left( \log(\text{det}(\Sigma_i)) + (Y_k - \mu_x)\Sigma_x^{-1}(Y_k - \mu_x) \right) | Y_{0:n} \right] \\
+ \sum_{k=1}^{n} E_\theta \left[ \sum_{i=1}^{r} \sum_{j=1}^{r} \mathbf{1}\{(X_{k-1}, X_k) = (i, j)\} \log q_{ij} | Y_{0:n} \right]
\]

where \( C \) is a constant. The formula can be written using the notations for smoothing probabilities that is given in section 3.2.

\[
Q(\theta; \theta') = C - \frac{1}{2} \sum_{k=0}^{n} \sum_{i=1}^{r} \phi_{k|n}(i; \theta') \log(\text{det}(\Sigma_i)) + (Y_k - \mu_x)\Sigma_x^{-1}(Y_k - \mu_x) \\
+ \sum_{k=0}^{n} \sum_{i=1}^{r} \sum_{j=1}^{r} \phi_{k-1|n}(i, j; \theta') \log q_{ij}
\]

The Lagrangian can be used to solve the problem

\[ \mathcal{L}(\theta, \lambda; \theta') = Q(\theta; \theta') + \sum_{i=1}^{r} \lambda_i \left( 1 - \sum_{j=1}^{r} q_{ij} \right) \]

And the derivatives is given by

\[
\frac{\partial}{\partial \mu_i} \mathcal{L}(\theta, \lambda; \theta') = \Sigma_i^{-1} \sum_{k=0}^{n} \phi_{k|n}(i; \theta') (Y_k - \mu_i)
\]

\[
\frac{\partial}{\partial \Sigma_i} \mathcal{L}(\theta, \lambda; \theta') = -\frac{1}{2\text{det}(\Sigma_i)} \sum_{k=0}^{n} \phi_{k|n}(i; \theta') [\Sigma_i - (Y_k - \mu_i)^\top(Y_k - \mu_i)]
\]

\[
\frac{\partial}{\partial q_{ij}} \mathcal{L}(\theta, \lambda; \theta') = \sum_{k=1}^{n} \phi_{k-1|n}(i, j; \theta') \frac{1}{q_{ij}} - \lambda_i
\]

\[
\frac{\partial}{\partial \lambda_i} \mathcal{L}(\theta, \lambda; \theta') = 1 - \sum_{j=1}^{r} q_{ij}
\]

When all derivates are set to the vector of zeros the updating functions can be obtained. The initial guesses are updated with the formulas below.
The algorithm keeps updating the initial guesses until some condition is set. It is important to note that this algorithm finds local optimal solutions, and not always the global optimal solution. It is therefore of great importance to run the model plenty of times with different starting values on $\theta^*$ (Cappe et al. 2005).

### 3.2.5 Model selection

The number of states in a hidden Markov model needs to be decided prior to the calibration. To choose the number of states is an important part of constructing the model. The choosing is a trade-off: fewer states lead to a model that is easier to interpret and minimizes the risk of overfitting. On the other hand, a model with a higher number of states makes the model more accurate. In this paper we will use two different methods for determine the optimal number of states. Both techniques are based on the log likelihood but compensate for the negative aspects of a model with higher number of states. It is also important to remember that the log likelihood is always increasing for higher number of states. It is therefore necessary to have some punishment for a higher number of states when applying log likelihood to a hidden Markov model.

The first model selection method that will be used in this paper is the Bayesian information criterion (BIC) (Schwarz 1978)

$$ BIC = k \log(n) - 2\ell(\theta) $$

Where $k$ stands for the number of parameters that will be determined by the model and $n$ stands for the number of observations. The second method is Akaike information criteria (AIC) (Akaike 1973).

$$ AIC = 2k - 2\ell(\theta) $$

where the parameters have the same notation.
A high value for the log likelihood is preferable over a lower value, but for AIC and BIC the opposite is true. However, it is important to note that this is a value selection model and cannot be used to compare which currency pair that have the most accurate model.
4. Results

4.1 Model selection

For the value selection we are looking at two different methods that both are built upon the log likelihood. Therefore, we start to look at the values of the log likelihood for the different models with different states.

Table 2 Log likelihood

<table>
<thead>
<tr>
<th>Log likelihood</th>
<th>EUR/NOK</th>
<th>USD/NOK</th>
<th>EUR/USD</th>
<th>EUR/SEK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 states</td>
<td>-2626</td>
<td>-4285</td>
<td>-3068</td>
<td>-2033</td>
<td>-4067</td>
</tr>
<tr>
<td>3 states</td>
<td>-2559</td>
<td>-4199</td>
<td>-2992</td>
<td>-1961</td>
<td>-4014</td>
</tr>
<tr>
<td>4 states</td>
<td>-2549</td>
<td>-4184</td>
<td>-2987</td>
<td>-1938</td>
<td>-3999</td>
</tr>
</tbody>
</table>

As already mentioned, it is important to not compare the results from different currency pairs. The results should just compare the different models within each currency to detect the optimal number of states. It is also important to remember that the log likelihood is always increasing for a greater number of states and for model selection the log likelihood by itself is therefore a poor method.

Table 3 Akaike information criteria with optimal value highlighted in grey

<table>
<thead>
<tr>
<th>AIC</th>
<th>EUR/NOK</th>
<th>USD/NOK</th>
<th>EUR/USD</th>
<th>EUR/SEK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 states</td>
<td>5267</td>
<td>8586</td>
<td>6151</td>
<td>4083</td>
<td>8150</td>
</tr>
<tr>
<td>3 states</td>
<td>5149</td>
<td>8429</td>
<td>6015</td>
<td>3952</td>
<td>8058</td>
</tr>
<tr>
<td>4 states</td>
<td>5147</td>
<td>8417</td>
<td>6022</td>
<td>3924</td>
<td>8046</td>
</tr>
</tbody>
</table>

The two methods with different penalty terms typically prefer different models. We can see that the AIC prefer the 4 states model and that BIC prefer the 3 states model for most of the currency pairs. To favour interpretation and understanding the 3-state model will be chosen and in the rest of the report the focus will be on this model.
Table 4 Bayesian information criterion with optimal value highlighted in grey

<table>
<thead>
<tr>
<th>BIC</th>
<th>EUR/NOK</th>
<th>USD/NOK</th>
<th>EUR/USD</th>
<th>EUR/SEK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 states</td>
<td>5317</td>
<td>8636</td>
<td>6219</td>
<td>4133</td>
<td>8200</td>
</tr>
<tr>
<td>3 states</td>
<td>5243</td>
<td>8523</td>
<td>6134</td>
<td>4046</td>
<td>8152</td>
</tr>
<tr>
<td>4 states</td>
<td>5297</td>
<td>8567</td>
<td>6206</td>
<td>4075</td>
<td>8197</td>
</tr>
</tbody>
</table>

4.2 Hidden Markov models with one currency pair

4.2.1 Returns

The states in the hidden Markov model with 3 states are sorted after the volatility in each state from the lowest volatility to the highest. The returns are assumed to follow a normal distribution in each state. This does not implicate that the return in total is following a normal distribution. The distribution \( N(\mu, \sigma^2) \) in each state is given by the mean and standard deviation given in table 5-9.

Table 5 Daily Mean and Standard deviation for the different market regimes for the currency EUR/NOK (converted to yearly)

<table>
<thead>
<tr>
<th>EUR/NOK</th>
<th>Low volatility (State 1)</th>
<th>Medium volatility (State 2)</th>
<th>High volatility (State 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (( \mu ))</td>
<td>-0.02 % (-4.23%)</td>
<td>0.03 % (8.33%)</td>
<td>0.16 % (50.79%)</td>
</tr>
<tr>
<td>Standard deviation (( \sigma ))</td>
<td>0.35 % (5.60%)</td>
<td>0.59 % (9.57%)</td>
<td>1.48 % (23.96%)</td>
</tr>
</tbody>
</table>
Table 6 Daily Mean and Standard deviation for the different market regimes for the currency USD/NOK (converted to yearly)

<table>
<thead>
<tr>
<th>USD/NOK</th>
<th>Low volatility (State 1)</th>
<th>Medium volatility (State 2)</th>
<th>High volatility (State 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (µ)</td>
<td>0.00 % (0.55 %)</td>
<td>0.00 % (-0.51 %)</td>
<td>0.20 % (70.23 %)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.53 % (8.52 %)</td>
<td>0.85 % (13.80 %)</td>
<td>1.85 % (29.88 %)</td>
</tr>
</tbody>
</table>

Table 7 Daily Mean and Standard deviation for the different market regimes for the currency EUR/USD (converted to yearly)

<table>
<thead>
<tr>
<th>EUR/USD</th>
<th>Low volatility (State 1)</th>
<th>Medium volatility (State 2)</th>
<th>High volatility (State 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (µ)</td>
<td>0.01 % (1.42%)</td>
<td>0.01 % (2.39%)</td>
<td>-0.02% (-6.16%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.33% (5.40%)</td>
<td>0.52% (8.28%)</td>
<td>0.89% (14.4%)</td>
</tr>
</tbody>
</table>

Table 8 Daily Mean and Standard deviation for the different market regimes for the currency EUR/SEK (converted to yearly)

<table>
<thead>
<tr>
<th>EUR/SEK</th>
<th>Low volatility (State 1)</th>
<th>Medium volatility (State 2)</th>
<th>High volatility (State 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (µ)</td>
<td>0.01 % (1.20%)</td>
<td>-0.01 % (-1.38%)</td>
<td>0.04 % (12.70%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.28% (4.57%)</td>
<td>0.46% (7.47%)</td>
<td>0.98% (15.83%)</td>
</tr>
</tbody>
</table>
Table 9 Daily Mean and Standard deviation for the different market regimes for the currency USD/SEK (converted to yearly)

<table>
<thead>
<tr>
<th>USD/SEK</th>
<th>Low volatility (State 1)</th>
<th>Medium volatility (State 2)</th>
<th>High volatility (State 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (μ)</td>
<td>0.00% (-0.30%)</td>
<td>-0.01% (-2.58%)</td>
<td>0.10% (29.58%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.53% (8.57%)</td>
<td>0.78% (12.59%)</td>
<td>1.47% (23.79%)</td>
</tr>
</tbody>
</table>

First, we can see that the NOK and SEK tend to depreciate to both EUR and USD during periods of high volatility. We can also see that all currencies depreciate in relation to USD in State 3. This is in line with the theory that says that USD is a “safe haven” that appreciate during times of high uncertainty. The different currencies seem to perform good in different markets. When looking at table 8 we see that the USD appreciate in the dramatic, but rare, high volatility state. We can also see that in the two states with lower volatility the SEK outperform the USD during the time.

When the daily means in State 3 are converted to yearly returns, they are very high. It is important to keep in mind that the market seem to be in the high volatility state rarely and that the periods tend to intensive but short.

In table 1 we could see that USD/NOK had the highest volatility when we looked at the whole period. This is valid even for the state model, the USD/NOK have the highest volatility for each state.

4.2.2 Transition matrices

The transition matrices can be seen below. In the transition matrices the probabilities for the process to be in each state in the next day is shown. The rows are the state in time t and the columns reflect the probability to end up in each state in time t+1.
### Table 10 The transition matrix for the currency EUR/NOK

<table>
<thead>
<tr>
<th>EUR/NOK</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>98.4 %</td>
<td>1.6 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>State 2</td>
<td>3.3 %</td>
<td>94.5 %</td>
<td>2.2 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>13.3 %</td>
<td>86.7 %</td>
</tr>
</tbody>
</table>

### Table 11 The transition matrix for the currency USD/NOK

<table>
<thead>
<tr>
<th>USD/NOK</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>99.8 %</td>
<td>0.1 %</td>
<td>0.1 %</td>
</tr>
<tr>
<td>State 2</td>
<td>0.2 %</td>
<td>99.7 %</td>
<td>0.1 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>2.6 %</td>
<td>97.4 %</td>
</tr>
</tbody>
</table>

### Table 12 The transition matrix for the currency EUR/USD

<table>
<thead>
<tr>
<th>EUR/USD</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>99.5 %</td>
<td>0.5 %</td>
<td>0.1 %</td>
</tr>
<tr>
<td>State 2</td>
<td>0.3 %</td>
<td>99.2 %</td>
<td>0.4 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>1.2 %</td>
<td>98.8 %</td>
</tr>
</tbody>
</table>

### Table 13 The transition matrix for the currency EUR/SEK

<table>
<thead>
<tr>
<th>EUR/SEK</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>98.1 %</td>
<td>1.8 %</td>
<td>0.1 %</td>
</tr>
<tr>
<td>State 2</td>
<td>2.0 %</td>
<td>97.8 %</td>
<td>0.2 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>1.7 %</td>
<td>98.3 %</td>
</tr>
</tbody>
</table>

### Table 14 The transition matrix for the currency USD/SEK

<table>
<thead>
<tr>
<th>USD/SEK</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>99.5 %</td>
<td>0.4 %</td>
<td>0.1 %</td>
</tr>
<tr>
<td>State 2</td>
<td>1.0 %</td>
<td>98.6 %</td>
<td>0.3 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>1.7 %</td>
<td>98.3 %</td>
</tr>
</tbody>
</table>
We can see that the states have a relatively high probability of staying in the same state as the previous trading day. For all processes the probability to be in the same state tomorrow is clearly dominant.

We can also see that the states with a lower volatility tend to be more stable. This is one of the reasons for State 3 to be the rarest state for the models. We can also see that the transitions between State 1 and State 3 seems to be very rare and that the market often walks through State 2 when these shifts are done.

**4.2.3 Smoothed probabilities**

In Figures 8-12 the smoothed probabilities for the different states is given. We can see that the different currency pairs have a clear difference in how stable the smoothed probabilities are. If we compare the smoothed probabilities in Figures 8-9, we can see that the smoothed probabilities are more stable in the latter. The reason for this can be followed back to table 10-11 where we can see that the probabilities to stay in the same state as before is higher for USD/NOK. A reason for these different probabilities for staying in the same state can probably be traced back to the different standard deviation in table 5-6. When we compare the standard deviation, we can see that the standard deviation is higher for USD/NOK. The result of higher standard deviation is that the difference in the probability density functions for the different states are smaller for an observation. A less distinguished difference in the probability density function result in a more stable curve for the probabilities. Even if the different curves are more or less stable, we can still see that they tend to have high probabilities for the same states during the same periods of time.

![Smoothed probabilities using the forward-backward algorithm for EUR/NOK](Figure 8)
**4.2.4 Viterbi path**

The Viterbi path, the state sequence that results in the maximum probability, is given in Figures 13-17 for the different currency pairs. The different currency pairs have different looks but there are also similarities. We can see that the currency pairs tend to be in volatile periods at the same time. The transition matrices influence the Viterbi path and we can see
that the model for EUR/SEK have the lowest probability to stay in the same state. The result of this is a Viterbi path that have a higher number of jumps.

Figure 13 Viterbi path EUR/NOK

Figure 14 Viterbi path USD/NOK

Figure 15 Viterbi path EUR/USD

Figure 16 Viterbi path EUR/SEK

Figure 17 Viterbi path USD/SEK
4.3 Hidden Markov models with three currency pairs

Since we in section 4.2.3 saw that the hidden Markov models for the different currency pairs tend to be in the same state at the same time it may be possible to create one hidden Markov model that include more than one currency pair. A model like this would also be able to investigate how the correlations change during different market regimes. The currency pairs are grouped into two different groups, the currency pair where one of the currencies is USD and the currency pairs where one of the currencies is EUR. Since one of the currency pairs have both USD and EUR the five currency pairs create two groups with three currency pairs in each.

4.3.1 EUR currencies

4.3.1.1 Returns & correlations

The returns and standard deviation in the EUR currencies is very similar to the corresponding values in section 4.2.1. It does not seem to affect the mean and standard deviation if we train hidden Markov model for each of the currency pairs or if we train one big hidden Markov model that is trained with the returns for all the currency pairs with EUR.

Table 15 Daily Mean and Standard deviation for State 1 for the EUR currencies (converted to yearly)

<table>
<thead>
<tr>
<th>EUR currencies</th>
<th>EUR/NOK</th>
<th>EUR/SEK</th>
<th>EUR/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1 Mean (µ)</td>
<td>-0.02 % (⁰⁻⁵.₂₂%)</td>
<td>0.00 % (⁰.₁₆%)</td>
<td>0.01 % (².₀₇%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.34 % (⁵.₅₄%)</td>
<td>0.30 % (⁴.₈₅%)</td>
<td>0.41 % (⁶.₆₀%)</td>
</tr>
</tbody>
</table>
Table 16 Daily Mean and Standard deviation for State 2 for the EUR currencies (converted to yearly)

<table>
<thead>
<tr>
<th>EUR currencies</th>
<th>EUR/NOK</th>
<th>EUR/SEK</th>
<th>EUR/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (μ)</td>
<td>0.03 %</td>
<td>-0.01 %</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>(8.82%)</td>
<td>(-3.13%)</td>
<td>(0.16%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.59 %</td>
<td>0.504 %</td>
<td>0.68%</td>
</tr>
<tr>
<td></td>
<td>(9.48%)</td>
<td>(8.15%)</td>
<td>(10.93%)</td>
</tr>
</tbody>
</table>

Table 17 Daily Mean and Standard deviation for State 3 for the EUR currencies (converted to yearly)

<table>
<thead>
<tr>
<th>EUR currencies</th>
<th>EUR/NOK</th>
<th>EUR/SEK</th>
<th>EUR/USD</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (μ)</td>
<td>0.14 %</td>
<td>0.11 %</td>
<td>-0.06%</td>
</tr>
<tr>
<td></td>
<td>(45.53%)</td>
<td>(32.27%)</td>
<td>(-14.46%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>1.34 %</td>
<td>1.01 %</td>
<td>1.10 %</td>
</tr>
<tr>
<td></td>
<td>(21.75%)</td>
<td>(16.41%)</td>
<td>(17.85%)</td>
</tr>
</tbody>
</table>

The correlations between the currencies are new for the higher dimensional model so we cannot compare the results with the one-dimensional models in section 4.2. However, we can compare the result to the correlations for the total data period in Figure 6.

For the EUR currencies we can see that the only currency pairs that have correlations over 0.2 is the correlation between EUR/SEK and EUR/NOK. The correlation is 0.5 for the whole data period but is lower in State 1 and State 2 and higher in State 3. It seems like the correlation between the currency pair tend to be higher when the volatility is high. One of the reasons for the higher correlations could be that the high volatility state often occurs during global market events like the corona crisis or the great recession.
Figure 18 Correlations for EUR currencies in State 1

Figure 19 Correlations for EUR currencies in State 2
The transition matrix for the EUR currencies differs from the single currency pair models with higher probabilities for transitions to other states. Since there are more parameters, due to the correlations, to train there is a higher risk of overfitting. An overfitted hidden Markov model will lead to higher probability for transitions so there is a risk that these factors are a result of an overfitted model.

Table 18 The transition matrix for the EUR currencies

<table>
<thead>
<tr>
<th>EUR currencies</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>94.0 %</td>
<td>6.0 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>State 2</td>
<td>11.2 %</td>
<td>87.1 %</td>
<td>1.7 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>8.3 %</td>
<td>91.7 %</td>
</tr>
</tbody>
</table>
4.3.1.3 Smoothed probabilities

We can see that the smoothed probability for the EUR currencies is much more unstable than the models that is trained with just one currency pair. Especially State 1 and State 2 is very unstable and a graph like this may be a result of overtraining. The State 3 is relatively stable and look more like the models that was trained with just one currency pair.

![Smoothed probabilities using the forward-backward algorithm for EUR currencies](image1)

4.3.1.4 Viterbi path

The Viterbi path is highly influenced by the smoothed probability and even here we can see that the model tends to jump frequently between the states. Especially the jumps between State 1 and State 2 occurs several times.

![Viterbi path EUR currencies](image2)
4.3.2 USD currencies

4.3.2.1 Returns & correlations

Like pointed out in section 4.3 currency pair EUR/USD is part of both the USD currencies and the EUR currencies. The means and standard deviation for EUR/USD in table 19-21 have small differences to the values obtained in the section 4.3.1.1. The data for the EUR/USD is the same but since the hidden Markov model is trained together with different currency pairs the period in the regimes is not the same.

When compared to the values in section 4.2.1, USD/NOK is the currency that stands out the most. In table 6 the values indicated that the mean was close to 0 for the two states with lowest standard deviation. On the other hand, the combined model in table 19-21 it seems like the higher volatility the market have, the more USD appreciate against NOK.

Table 19 Daily Mean and Standard deviation for State 1 for the USD currencies (converted to yearly)

<table>
<thead>
<tr>
<th>USD currencies State 1</th>
<th>EUR/USD</th>
<th>USD/NOK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (μ)</td>
<td>0.01 %</td>
<td>-0.03 %</td>
<td>-0.01 %</td>
</tr>
<tr>
<td></td>
<td>(2.34%)</td>
<td>(-7.09%)</td>
<td>(-1.79%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.41 %</td>
<td>0.54 %</td>
<td>0.51 %</td>
</tr>
<tr>
<td></td>
<td>(6.56%)</td>
<td>(8.69%)</td>
<td>(8.33%)</td>
</tr>
</tbody>
</table>

Table 20 Daily Mean and Standard deviation for State 2 for the USD currencies (converted to yearly)

<table>
<thead>
<tr>
<th>USD currencies State 2</th>
<th>EUR/USD</th>
<th>USD/NOK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (μ)</td>
<td>0.00 %</td>
<td>0.03 %</td>
<td>-0.01%</td>
</tr>
<tr>
<td></td>
<td>(0.56%)</td>
<td>(8.46%)</td>
<td>(2.60%)</td>
</tr>
<tr>
<td>Standard deviation (σ)</td>
<td>0.66 %</td>
<td>0.88 %</td>
<td>0.84%</td>
</tr>
<tr>
<td></td>
<td>(10.76%)</td>
<td>(14.23%)</td>
<td>(13.54%)</td>
</tr>
</tbody>
</table>
Table 21 Daily Mean and Standard deviation for State 3 for the USD currencies (converted to yearly)

<table>
<thead>
<tr>
<th>USD currencies</th>
<th>EUR/USD</th>
<th>USD/NOK</th>
<th>USD/SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 3</td>
<td>Mean (μ)</td>
<td>0.07 %</td>
<td>0.23 %</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(17.13%)</td>
<td>(84.08%)</td>
</tr>
<tr>
<td></td>
<td>Standard deviation (σ)</td>
<td>1.10 %</td>
<td>1.69 %</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(17.83%)</td>
<td>(27.29%)</td>
</tr>
</tbody>
</table>

The first observation about the correlations is that there are negative correlations for EUR/USD against the other two currency pairs in Figures 23-25. If the EUR/USD would instead be presented as USD/EUR these correlations would, as pointed out in section 3.1.1, turn positive.

In section 2 we saw that the correlations between the USD currencies was much higher than the correlations for the EUR currencies and that hold in each of the market regimes as well. We can also see that the USD currencies, unlike the EUR currencies, have the lowest correlations in State 3.
Figure 24 Correlations for USD currencies in State 2

Figure 25 Correlations for USD currencies in State 3
4.3.2.2 Transition Matrices

We can see that the probability for the model to stay in the same state is much lower for the USD currencies than it was for the single currency models. Since the higher dimensional model have more parameters to train, the risk of overfitting increases. Lower probabilities for staying in the same state is a risk factor for an overfitted model.

Table 22 The transition matrix for the USD currencies

<table>
<thead>
<tr>
<th>USD Currencies</th>
<th>State 1</th>
<th>State 2</th>
<th>State 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 1</td>
<td>93.6 %</td>
<td>6.4 %</td>
<td>0.0 %</td>
</tr>
<tr>
<td>State 2</td>
<td>11.4 %</td>
<td>86.9 %</td>
<td>1.7 %</td>
</tr>
<tr>
<td>State 3</td>
<td>0.0 %</td>
<td>8.6 %</td>
<td>91.4 %</td>
</tr>
</tbody>
</table>

4.3.2.3 Smoothed probabilities

The smoothed probabilities for the USD currencies have very unstable State 1 and State 2. State 3 have similarities with the one-dimensional model, but the other two states do not seem to have clear similarities. Many jumps from close to 0 to close to 1 without clear reason can be an indication of overfitting.

Figure 26 Smoothed probabilities using the forward-backward algorithm for USD currencies

4.3.2.4 Viterbi Path

The Viterbi path is influenced by the smoothed probabilities, and we can see that there are many jumps in the path, especially between State 1 and State 2.
Figure 27 Viterbi path USD currencies
5. Discussion

5.1 Discussion of the results

The intention of the model is to determine how the market is behaving in different market regimes. It is important to remember that the selected time period influences the results. The best example of this can be seen in State 3, which is meant to be a good representation of a high volatility market. If we look at the periods where the models are in State 3, we can see that the great majority of the time is during the great recession of 2008. The result of this is that State 3 seem to describe how the market behaved during a specific market event rather than describe how the market in general behave in a high volatility market. This can also be a reason for the unexpected result that the correlations for the USD currencies decreased during State 3. It is possible that this is a result of the great recession that did hit the European economies later than it hit the American economy. The Euro crisis is also an event that led to high volatility and price movements especially for the EUR currencies. If the model would be trained with data from other crises like the Asian financial crisis of 1997, different states would most likely look different. If the time period that were used for the training would just contain a few months the states would most likely detect another kind of trending. Instead of finding the long-term trends the model would find the short-term trends.

We saw in the findings section that the low and medium volatility states have expected values that is closer to zero. The model suggests that the most dramatic price movements occur during the high volatility state. We can see that the bigger and more well-known currencies like USD and EUR tend to appreciate a lot during these periods and that the returns if annualized is very dramatic. However, it is important to remember that the high volatility state in general just stay for weeks and that the annualized returns is therefore to be comparable to implied volatility that is often priced in annualized numbers. The annualization have used the assumption of independent returns within the state. It is expected from the theory in section 3.1.2 to see that especially USD is appreciating against the Nordic countries during periods of high volatility and uncertainty.

When looking at the smoothed probabilities in the one-dimensional model we can see that most jumps, especially to State 3, can be explained by big market events during the same
time. For example, if we look at the Viterbi path graph for USD/SEK in Figure 17 we can see that the model is in State 3 during a long period of time during the great recession in 2008. We can also see that the model is in this state during the Euro crisis in 2010. The unusual jump from State 1 to State 3 in 2016 is connected to the Brexit referendum that was choking the financial markets. The last jump to State 3 is a result of the corona crisis that have been affecting the financial markets during 2020.

From a mathematical point of view, it is interesting to see that the smoothed probability for USD/NOK in Figure 9 is much more stable than the other currency pairs. One explanation for this is that the standard deviation for the currency pair is higher than the others and that result in a density function that does not punish outliers as strict as the density functions for the other models. The less strict punishment leads to a smoothed probability that does not have the same need for transitions between the states. Since especially the three-dimensional model tend to have a very high number of jumps there is a high risk of overfitting in these models. The number of variables that is trained is increasing because of the correlation matrix and that in addition to the fact that it can be harder to find regimes that are appliable on all three currency pairs is reasons to the unstable smoothed probabilities for the three-dimensional models. An improvement that could have been done is to have a floor that the standard deviation cannot go under. That would result in a density function that behave like the density function for USD/NOK and do not punish outliers and therefore leading a more stable smoothed probabilities graph and lower risk of overfitting.

The different regimes have distinct differences in variance and mean which suggest that the hidden Markov models can capture the regime shifting mechanism for the foreign exchange data. We can also see that the modelled State 3 that is supposed to reflect a high volatile market seem to correspond to the periods that is known on forehand to have been volatile, for example the great recession of 2008 and the Corona crisis. For the higher dimensional model, it is seen that the differences in correlations between the different states are surprisingly small. The differences for the variance and means are minimal between the models that include correlations and the models that does not. The absence of distinct differences suggests that the model is not able to capture correlations between currency pairs in an effectively way. The reason for this can be that the hidden Markov model focuses more on the mean and variance and is unable to capture correlation as an important factor for the market regimes. It is also possible that the reason for the correlations to not influence the model dramatically is
that the difference in correlations between the market regimes is relatively small. In that case the model focus on the differences in means and variances that make a relatively higher impact on the probability density function.

5.2 Further research

In further research a research topic may be to look at how the volatility that is predicted by the hidden Markov model relates to the implied volatility by the prices of options in the market. It would be interesting to investigate if the volatility predictions in the hidden Markov model is priced in for the option prices or if they are not.

It would also be interesting to investigate if a trading model based on the hidden Markov model could perform on historic data and if it is possible to reduce the risk or get higher returns when applying the hidden Markov model.
References


