UNIVERSITY OF CAPE TOWN

MASTERS DISSERTATION

The Impact of the Carry Trade on Global Currency Markets

Author:
Steven SMIT

Supervisor:
Dr. Etienne PIENAAR
Co-supervisor:
Dr. Daniel POLAKOW

A dissertation submitted in fulfillment of the requirements for the degree of Masters in Mathematical Statistics

in the

Department of Statistical Sciences

October 18, 2019

Declaration of Authorship

I, Steven SMIT, declare that this thesis titled, "The Impact of the Carry Trade on Global Currency Markets" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Signed by candidate

Date:

UNIVERSITY OF CAPE TOWN

Abstract

Department of Statistical Sciences

Masters in Mathematical Statistics

The Impact of the Carry Trade on Global Currency Markets

by Steven SMIT

This work analyses the effect of the carry trade factor, statistically derived from a comprehensive basket of currencies, on currencies in various heuristically defined global risk appetite regimes. Findings of a heightened (lessened) impact of this factor for Emerging/Commodity (Developed/European) currencies in the presence of high risk are presented. The risk appetite process is additionally analysed by modelling it as a Markov-switching model, providing evidence of three inherent regimes, with properties roughly consistent with findings in the literature.

Acknowledgements

Firstly, I'd like to thank Daniel Polakow for his help and support throughout this dissertation, in particular for his help in getting the content of Chapter 2 published, for which he is second author. Secondly, Etienne Pienaar for his support and very helpful comments and amendments, particularly for Chapter 3.

Thirdly, Prescient Securities as well as my employer, Kantar, for financial assistance throughout the past two years.

Finally, I'd like to thank my wife for her tremendous and unwavering support throughout the process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The importance of the carry trade and its effect on currency markets is well known to researchers. Lustig, Roussanov and Verdelhan (2011:3732) analysed the principal components of interest-rate ordered currency portfolios, providing evidence that the second most important dimension to returns, after general Dollar risk, is a so-called "carry trade risk factor". Brunnermeier, Nagel and Pedersen (2008) document that currencies with similar interest rates move together, suggesting that the carry trade has an effect on currency markets.

The currency "carry trade" is the act of borrowing capital in a low-interest currency and investing the proceeds into a higher interest rate currency (e.g. Burnside (2012)). For example (and ignoring frictional costs to transacting), borrowing 1 Japanese Yen at 0.075% for a maturity of 3 months, converting into South African Rand at the ratio of 1 Yen to 0.12 Rand, and purchasing a 3 month South African treasury bond/note yields 7.33% (Organisation for Economic Co-operation and Development [OECD], 2017). The "carry" is then the interest differential of $(7.33 - 0.075)\% = 7.255\%$ (annualised). The return to the trade, at expiry, comprises the carry and any additional profit or loss imposed by the fluctuation of the Rand/Yen exchange rate.

This dissertation seeks to further examine and document the relationship between global risk appetite (using the CBOE VIX Index as a proxy), the carry trade and currency markets. First, the hypothesis that the relationship between the global carry trade factor and currencies differs in three constructed "risk" states of the world, based on the level of the VIX Index, is tested.

This work also seeks to examine the concept and existence of these risk states. Following Hamilton's (1989) paper detailing the estimation of a Markov-switching autoregressive model, Chen and Tsay (2007) introduced an algorithm to estimate a Markov-switching autoregressive moving-average model in the frequentist framework. This dissertation estimates such a model, compares to the nested mean-only and AR Markov-switching models for two and three states, and estimates the probability of the number of states (from none to three) using a Bayesian approach - to provide evidence of intrinsic regimes in the risk-appetite process (proxied for by the VIX Index).

The structure of the dissertation proceeds as follows: Chapter 2 documents the relationship between global risk appetite, the carry trade and global currency markets. Chapter 3 analyses the existence of intrinsic regimes in the global risk appetite process. Chapter 4 provides a brief validation of the model (as identified in Chapter 3) and Chapter 5 concludes.

Chapter 2

The Relationship Between the Carry Trade, Emerging/Commodity Currencies and Risk

2.1 Introduction

This chapter examines the relationship between the currency carry trade and the global currency market against the backdrop of global risk appetite. This is done by assessing whether the relationship between the carry trade risk factor and global currencies differs in three constructed states of the world, "low risk", "normal risk" and "high risk".

Evidence supporting the findings of two common risk factors in currency markets, namely a currency market risk factor ("Dollar risk") and a carry trade factor, are documented in Lustig, Roussanov and Verdelhan (2011:3732). A third "Asian markets" factor was identified and noted by observing the loadings of the currencies on the principal components of the total sample, and by observing the correlation with an equally weighted portfolio of currencies clustered using a k-means algorithm. This algorithm identified three clusters of currencies according to loadings on the principal components, being: "Developed/European Currencies", "Emerging/Commodity Currencies" and "Asian Currencies".

Global risk appetite states were then constructed by observing the level of the CBOE VIX Index. By conducting a partial F-test of a full model (including all interaction terms between the first three principal components and the risk states) and a restricted model (excluding interaction terms), evidence was found that, for a large majority of currencies, there is a significant interaction between principal components coefficients and risk state dummy variables, suggesting that the risk state alters the impact of the common currency factors on currencies.

Furthermore, it was observed that there was an increase in the proportion of total variation explained in the "High" risk state for Emerging/Commodity cluster currencies, while the opposite was observed for Developed/European cluster currencies. When analysing the regressions of currencies against the carry trade factor, the striking increase in R^2 for Emerging/Commodity currencies when moving into the "High" risk state was noted, as well as the statistical significance of the coefficients. Analysing the scatterplot of currency returns against the derived carry trade factor, it was noted that the plots tended to span the origin, implying that positive (negative) changes in carry trade risk are associated with negative (positive) changes in

the currency. The interpretation of these results is that a heightened global risk state is associated with a strengthening of the relationship between the carry trade and Emerging/Commodity currencies, and is associated with the carry trade driving a considerable portion of currency variation.

The significance of a joint interaction between the three principal components across many currencies, in addition to that between the principal components and the risk states, was noted. These findings provide a fertile avenue for future research to elaborate upon.

2.2 Review of the Literature

The persistence and popularity of the carry trade is closely related to the theory of Uncovered Interest Parity (UIP), with the existence of positive excess returns to the carry trade being linked to the empirical failure of the UIP (Curcuru, Vega & Hoek, 2010:437). The theory of UIP broadly states that in risk-neutral settings, bilateral exchange rates ought to depreciate according to the interest rate differential between the higher interest rate currency and the lower interest rate currency.

In their survey of the literature, Froot and Thaler (1990) documented the empirical failure of the UIP and the findings which suggest currencies *appreciate* according to interest rate differentials. They go on to note that academics typically argue that the findings imply the existence of time-varying risk premia or errors in expectations.

Burnside et al. (2011) found that common sources of risk failed to explain carry returns but found evidence that the returns to the carry trade were compensation for rare-event risk. Brunnermeier et al. (2008) found that controlling for the VIX Volatility Index (used as a proxy for risk appetite) reduced the predictiveness of interest rates on currencies, partially explaining the failure of UIP. They also note that carry losses correlate positively with positive changes in the VIX Index, suggesting the increased VIX is a sign of decline in the risk appetite of traders and induces an unwinding of carry positions. Christiansen, Ranaldo and Söderlind (2011) modelled carry trade returns by allowing for time-varying risk premia; regime variables were modelled using a variety of standard risk proxies, including the CBOE VIX Index. They provide evidence that factors explaining carry trade returns depend on the regime and that during high risk regimes, exposure to systematic (stock market) risk is increased.

Lustig, Roussanov and Verdelhan (2011) found that low interest rate currencies appreciated in times of high carry trade risk, while high interest rate currencies depreciated. They noted that this was not due to variation in interest rates, but due to variations in spot currency returns. Furthermore, they noted that in times of high equity volatility, low interest rate currencies offer higher returns while high interest rate currencies offer lower returns. The opposite was found to be true in times of lower equity market volatility.

Also of interest, Menkhoff et al. (2012:712) found a similar result for global FX volatility, also showing that it accounted for more than 90% of the cross-sectional variation in excess returns of their constructed carry trade portfolios.

What is not clear is the effect that the carry trade factor has on currency markets when examined against the backdrop of global risk appetite (using the CBOE Implied Volatility (VIX) Index as a proxy for risk appetite). Brunnermeier et al. (2008) documented an aspect of this relationship, namely that between the VIX and the carry trade. Lustig, Roussanov and Verdelhan (2011:3732) derived a Dollar risk and carry trade factor (a "slope" and "level" factor) and analysed the effect these had on interest-rate sorted currency portfolios. This chapter builds upon these studies by focusing on the effect of the carry trade factor on currencies, when examined against the backdrop of global risk appetite "state".

The suspicion amongst investment professionals and currency traders, is that the appetite foreigners have for investing in carry currencies (such as the ZAR) is contingent on the global risk backdrop, since carry as an asset class is intrinsically risky and a function of both liquidity and yield of the targeted currency, with liquidity often being the determinant of eventual profit (or loss) when the global investment backdrop changes from a perception of low- (risk-on) to high-risk (risk-off). Within the context of the South African marketplace, Polakow and Flint (2015:4) noted in their analysis of the factors affecting the South African equity market the dual impact of the global "risk-on/risk-off" signal as well as interest rates on the South African equity market, and noted the interaction required further investigation. This dynamic seems to be established therefore in the equity market, and the analysis could simply be expanded to any emerging market currency (i.e. ones with relatively high interest rates).

2.3 Data

Historical daily currency exchange rate data was obtained for 115 currencies against the US Dollar. The currency pairs were denominated in USD per 1 unit Foreign Currency. This was reduced to 29 currencies (see Table 2.1) in all (this was done owing to poor resolution in the dataset, further details provided in Section 2.4.1). Currencies were chosen by starting with the currencies of the constituent countries of the G20 (24 in all). The currencies of Bosnia and Herzegovina, New Zealand, Norway, Singapore and Switzerland were further included, based on a selection methodology further detailed in Section 2.4.1. The time period of the sample ran from 2 January 1990 to 13 April 2017 ($n = 7118$ days).

TABLE 2.1: Currencies in the analysis

G20 Currencies*	ARS, AUD, BRL, CAD, CNY, EUR, INR, IDR, JPY, KRW, MXN, RUB, SAR, ZAR, TRY, GBP, SEK, HRK, CZK, HUF, PLN, BGN, DKK, RON
Additional Selected Currencies**	NOK, CHF, NZD, SGD, BAM

*Currencies of constituent members of the European Union (excluding the USD, as the sample is denominated against USD) as at May 2017
**Methodology detailed in Section 2.4.1

This paper follows the approach of Brunnermeier et al. (2008) in selecting the CBOE VIX Index as a proxy for risk-appetite globally. Daily VIX Index values for the period 2 January 1990 until 13 April 2017 ($n = 7118$ days) were obtained. All data were

acquired from Bloomberg (2017). The 2nd of January 1990 was selected as it was the first date that CBOE VIX Index values were available.

Weekend dates were removed from the dataset and all dates with missing values which occurred after the currency began trading were imputed by filling forward to avoid a look-ahead bias. Dates before trading were either imputed with 0 for the principal components analyses (PCA) or left as missing and dropped for the regressions. Daily percentage returns were calculated for all currencies in the analysis.

2.4 Methods & Results

2.4.1 Principal Components and Clustering Analysis for Factor Detection

The sample of currencies was reduced from an original 115, owing to the poor resolution observed in the dataset. That is, the first 10 components in the PCA of the original 115 currency dataset only explained roughly 27% of the total variation. The panel of 29 currencies were selected from the original data initially by selecting the 24 constituent currencies of the G20. A further five currencies, the Bosnia-Herzegovina convertible Mark (BGM), the New Zealand Dollar (NZD), the Norwegian Krone (NOK), the Swiss Franc (CHF) and the Singaporean Dollar (SGD) were added as they individually added to the total variance explained by the first three principal components of the scaled and centred data (details on the cut-off for principal components, as well as the full procedure, detailed below). A fourth currency, the Bhutanese Ngultrum, also added to total variance, however was excluded as it is pegged to, and has perfect sample correlation with, the Indian Rupee. This panel was examined by analysing the principal components of the scaled and centred data (missing values were replaced with zeroes after the scaling and centring process).

A PCA is a statistical procedure for dimensionality reduction which reveals orthogonal components, ordered by the proportion of the total variance in the dataset that they explain. The covariance matrix A of the scaled and centred currency data is decomposed into its constituent eigenvalues (λ) and eigenvectors (v); these are obtained by solving the following matrix equation:

$$Av = \lambda v.$$

The eigenvalues are numerically sorted in descending order. The eigenvector corresponding to the first eigenvalue is the first principal component, the second eigenvector corresponds to the second principal component, etc. To obtain the variation explained by each principal component, the proportion attributable to each eigenvalue of the sum of the eigenvalues is calculated.

The Kaiser-Guttman criterion (Guttman, 1954) was used to pick a lower bound of the acceptable principal components to retain. In addition, a scree plot, another heuristic used to decide how many principal components to keep, was analysed and a number of principal components were selected and kept for further analysis.

The loadings of the panel of currencies against these principal components were calculated and the k-means clustering algorithm (MacQueen, 1967) was applied to the result. An elbow plot was used to identify a reasonable number of clusters to use. The "elbow" represents the point after which there are relatively small decreases in

within-cluster sum of squares (that come from adding more clusters), suggesting that this point represents the appropriate number of clusters to use.

PCA is known for rendering components that are not necessarily easily interpretable. Further analysis is then required and an inference made as to the nature of the dimensions. With that in mind, the loadings of the currencies on principal components, as well as the correlation between the principal components and the clusters were analysed, and identifiers (or labels) for each of the principal components were inferred.

2.4.2 PCA Results

Table 2.2 contains the proportion of total variance explained by the first 5 principal components in the dataset. The Kaiser-Guttman criterion suggests that all principal components from the sixth onwards ought to be discarded, as their eigenvalues are less than 1. The scree plot is presented in Figure 2.1. Analysis of the scree plot suggests that a reasonable number of principal components to retain is three.

TABLE 2.2: Proportion of total variance explained by the first five principal components

Principal Component	Total Variance Explained
1	30.24%
2	7.73%
3	4.86%
4	4.02%
5	3.75%

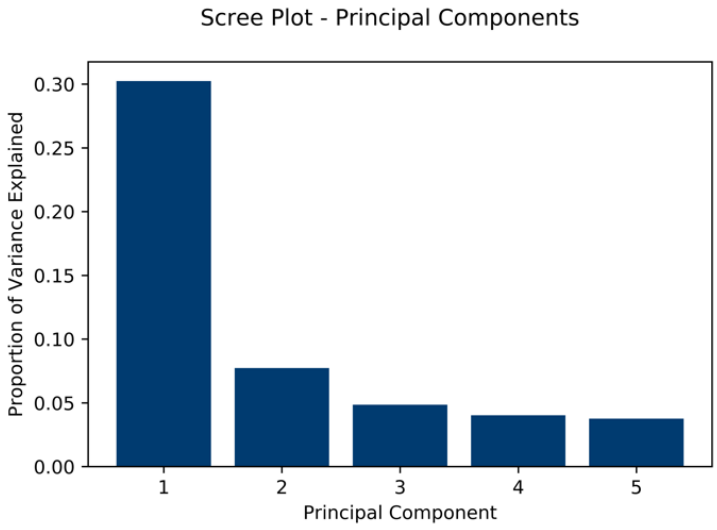

FIGURE 2.1: Scree plot of the variance explained by the first five principal components

Table 2.3 displays the loadings of the currencies in the sample against the first three principal components. The currencies all load positively along the first principal

component (with the exception of the Saudi Arabian Riyal), while the currencies load positively and negatively along the second principal component.

Analysis of the elbow plot obtained from the k-means clustering algorithm (Figure A.1 in Appendix A) suggested that three or four clusters were a reasonable number to select. Three were chosen for parsimony, but the groupings when identifying four clusters are detailed in Table A.1 in Appendix A.

The clusters were formed by applying the k-means algorithm to the loadings of these currencies on the first three principal components. The clusters therefore represent groupings of currencies which behave similarly in the presence of the three most important (from a variance perspective) market factors. Labels were attached to each group based on their general composition. The chosen labels were "Developed/European Currencies", "Emerging/Commodity Currencies" and "Asian Currencies" and this is shown in Table 2.3, with each currency indexed by its currency cluster.

The Developed/European Currencies cluster contains a mix of European and Developed market currencies, with exceptions like the CNY (Chinese Renminbi), ARS (Argentine Peso) and SAR (Saudi Arabian Riyal). A possible explanation for their inclusion in this cluster is that all of these currencies had or have had, for a significant period of the sample, a pegging of some form to the US Dollar. The Emerging/Commodity Currencies cluster and Asian Currencies cluster are so named for the mix of emerging market & commodity currencies and Asian currencies respectively.

The similarity of loadings of a majority of currencies on the first principal component suggests that it represents a general Dollar market risk factor, as originally noted in Lustig, Roussanov and Verdelhan (2011). The correlation between this component and an equally weighted portfolio invested in all 29 currencies in the sample is 0.95, lending support to this idea. Furthermore, the correlation between this component and the earliest available values of the Bloomberg Dollar Spot Index (a measure of Dollar strength) is -0.93. The sign of this correlation is to be expected as the currencies in the sample were assessed as "Dollars per foreign currency" - Dollar strength would result in declining values for any of the currencies in the sample.

The general trend of positive and negative loadings on the second principal component for the Developed and Emerging Markets clusters respectively, suggests that the factor affects these two clusters of currencies in opposite ways. Indeed, the correlation between this factor and the difference between two equally-weighted portfolios invested in the Developed Market and Emerging Market clusters respectively, is 0.85. Furthermore, the correlation between this factor and the earliest available daily changes in the Deutsche Bank Currency Harvest Global (USD) index (a proxy for global carry appetite) was -0.61. Katzke and Polakow (2017) observed a similar correlation and documented the oscillatory nature of the loadings of this index on the second principal component. This finding lends support to the idea that the second principal component is a carry trade factor; a similar finding to Lustig, Roussanov and Verdelhan (2011).

The third principal component appears to load ambiguously on both the Developed and Emerging Markets clusters, but consistently has high loadings on the Asian Market cluster, which suggests that it is a factor unique to Asian currencies. The correlation between the returns of an equally weighted portfolio invested in the Asian currencies cluster, and the third principal component, is 0.84. While this is

TABLE 2.3: Loadings of each currency on the first three principal components

		Principal Component		
Cluster	**Currency**	**1**	**2**	**3**
Developed/European Currencies	ARS	0.010	-0.031	0.018
	CNY	0.011	-0.009	0.062
	EUR	0.294	0.234	-0.038
	JPY	0.096	0.324	0.297
	SAR	-0.009	0.025	0.133
	GBP	0.230	0.132	0.051
	SEK	0.281	0.140	-0.022
	HRK	0.206	0.056	-0.086
	CZK	0.249	0.037	-0.092
	BGN	0.126	0.052	-0.060
	DKK	0.296	0.270	-0.017
	RON	0.096	-0.006	-0.081
	NOK	0.283	0.129	-0.004
	CHF	0.234	0.343	0.056
	BAM	0.174	0.043	-0.042
Emerging/Commodity Currencies	AUD	0.228	-0.284	0.047
	BRL	0.099	-0.242	-0.061
	CAD	0.203	-0.289	-0.067
	MXN	0.103	-0.339	-0.098
	RUB	0.067	-0.112	-0.021
	ZAR	0.199	-0.251	-0.078
	TRY	0.116	-0.191	-0.128
	HUF	0.251	-0.062	-0.128
	PLN	0.232	-0.114	-0.117
	NZD	0.228	-0.238	0.067
Asian Currencies	INR	0.068	-0.147	0.232
	IDR	0.028	-0.088	0.545
	KRW	0.046	-0.145	0.550
	SGD	0.207	-0.110	0.350

by no means conclusive evidence, it is suggestive of this factor being an "Asian market" factor. It is also worth noting that while the clustering algorithm placed JPY in the Developed/European currencies cluster, it is geographically Asian and this is reflected in its somewhat large and positive loading on this "Asian market" factor.

2.4.3 Construction of VIX States

Three risk states were constructed by selecting thresholds of the VIX Index which split the states into "low", "normal" and "high". The selections were informed by analysing the distribution of historical VIX Index values. The first and second thresholds were selected to be the 15^{th} and 85^{th} percentiles of the sample distribution of VIX Index values, in order to capture "low" and "high" risk states. The idea is to capture the upper and lower tails of the distribution and use them as proxies for the "low" and "high" risk states.

2.4.4 The VIX Index States Thresholds

State 1 comprised VIX Index observations below (or equal to) the 15^{th} percentile ($n = 1070$). State 2 comprised observations which were greater than the 15^{th} percentile and below (or equal to) the 85^{th} percentile ($n = 4980$). Finally, state 3 comprised observations greater than the 85^{th} percentile ($n = 1068$). The 15^{th} percentile for the VIX Index observations was 12.68 and the 85^{th} percentile was 26.09.

2.4.5 The Model

The linear model (Equation 2.1) regressing each currency's returns against the selected principal components, dummy variables representing the risk state and all interactions between the VIX state dummies and principal components, was estimated using Ordinary Least Squares (OLS) regression. Observations with missing values were dropped from the regression. The principal components were obtained from a PCA that included the currency being regressed, to facilitate a common interpretation. It was found that the ordination of principal components shifted when excluding certain currencies.

$$\begin{aligned} Y_i = \alpha + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_4 S_2 + \beta_5 S_3 \\ + \beta_6 PC_1 * S_2 + \beta_7 PC_2 * S_2 + \beta_8 PC_3 * S_2 + \beta_9 PC_1 * S_3 + \beta_{10} PC_2 * S_3 + \beta_{11} PC_3 * S_3. \end{aligned} \tag{2.1}$$

where:

Y_i = the returns of the i^{th} currency in the sample
PC_1 = the first principal component
PC_2 = the second principal component
PC_3 = the third principal component
S_2 = the indicator/dummy variable denoting risk state 2 ("Normal")
S_3 = the indicator/dummy variable denoting risk state 3 ("High")

The first hypothesis of interest is that there is no interaction between risk states and the principal components. Formally, the null hypothesis is:

$$\beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = \beta_{11} = 0.$$

against the alternative that at least one of the β coefficients is not equal to 0.

The hypothesis was tested using the partial F-test by comparing the full model of Equation 2.1 to the restricted model of Equation 2.2:

$$Y_i = \alpha + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_4 S_2 + \beta_5 S_3. \tag{2.2}$$

On the basis of these results for all currencies in the sample, model fit was then assessed across states, and regressions (stratified by risk state) of currencies against the chosen principal components were run. These regressions are referred to as the stratified regressions of the full model.

Finally, the same regressions were run for each currency, restricted to the second principal component (or the derived carry trade factor) to analyse the coefficients, as well as the effect on model fit. This model is described by Equation 2.3.

$$Y_i = \alpha + \beta_1 PC_2. \tag{2.3}$$

2.4.6 Model Estimation Results

Table A.2 (noted in Appendix A) displays p-values for the partial F-test of the hypothesis that there is no interaction between risk state and the principal component regressors. The p-values suggest that one can reject the null hypothesis and proceed with the conclusion that there is an interaction effect, with the exception of the Russian Rouble (RUB) (p = 0.2). This is not a trivial finding, and documents what the author believes is, for the first time, a statistically significant interaction between currency drivers and risk state.

Appendix B contains the details and results of a hypothesis test of the joint interaction between the principal components; the findings of which provide evidence of a joint interaction between the currency market factors for most currencies.

Figure 2.2 displays the average R^2 per currency cluster (with the exclusion of the pegged currencies, ARS, CNY and SAR from the Developed/European Currencies cluster*) for the full model (Equation 2.1). There is a clear decline in R^2 as one moves from the Developed/European cluster to the remaining clusters, suggesting that the currencies of the Developed/European cluster tend to be less idiosyncratic. The full results from the stratified regressions of the full model, for each currency in the sample, are presented in Table A.3 in Appendix A.

Figure 2.3 displays, for the model described in Equation 2.3, the average R^2 per currency cluster, split by risk state. One should note the increase in R^2 when moving into the "High" (State 3) risk state for the Emerging/Commodity currencies cluster. The opposite is observed with the Developed/European currencies cluster, where there is a drop-off in R^2 when moving from the "Low" (State 1) risk state to the "High" risk state. The Asian currencies cluster clearly follows the same pattern as the Emerging/Commodity cluster, albeit with a lower average model fit for this carry

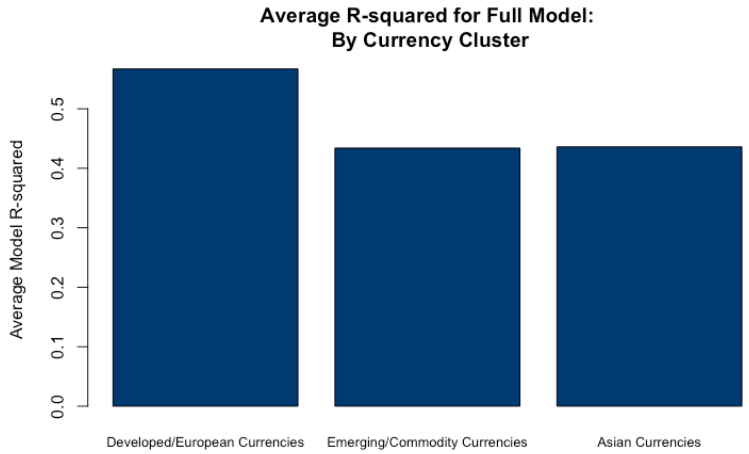

FIGURE 2.2: Average R^2 for the full model per currency cluster*

*The R^2 for the full model of the currencies excluded from this cluster is 0.01 (ARS), 0.05 (CNY) and 0.03 (SAR). When including these currencies, the average R^2 per cluster is roughly the same: Developed/European Currencies (0.46), Emerging/Commodity Currencies (0.43) and Asian Currencies (0.44)

trade factor-only model. The full results from the stratified regressions of this carry trade factor-only model, for each currency in the sample, are presented in Table A.4 in Appendix A.

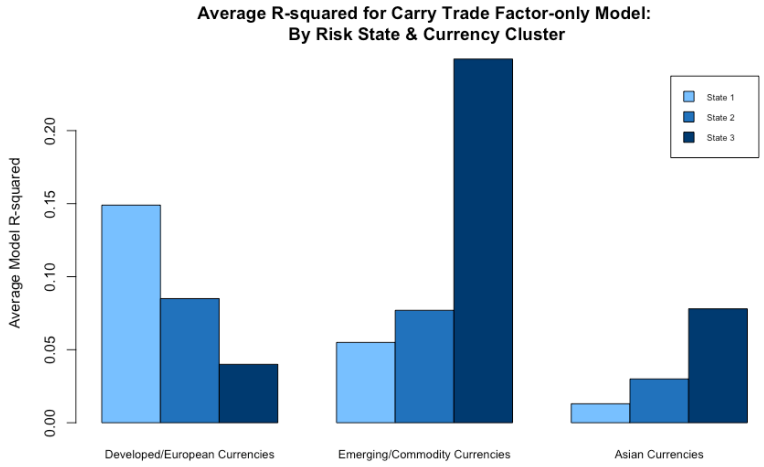

FIGURE 2.3: Average R^2 for the carry-trade factor only model, split by currency cluster and risk state

Figure 2.4 displays the scatterplots of the daily currency (in the emerging markets cluster) percentage changes (y-axis) against the second principal component/carry trade factor (x-axis). With the possible exception of the Russian Rouble, one notes a general negative slope spanning the origin.

Figure A.2 (found in Appendix A) displays the same scatterplot, considering only those observations falling in the "High" risk state. Again, one notes the general negative slope spanning the origin, with the exception of the Russian Rouble.

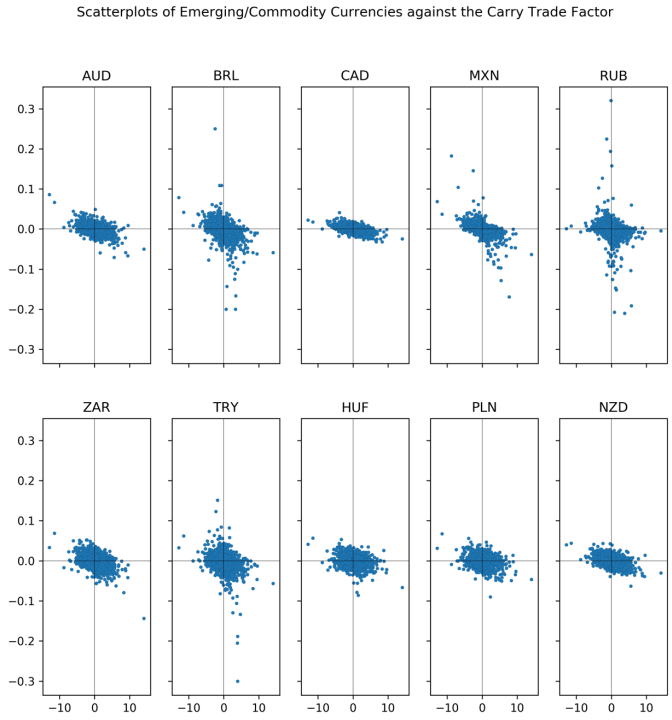

FIGURE 2.4: Scatterplots of Emerging/Commodity currencies against the carry trade factor

Table 2.4 displays average daily returns in percent (with standard deviations in parentheses) for each identified cluster within the three risk states. There tends to be a decline in average daily returns as one moves from State 1 to State 3. Emerging/-Commodity Markets appear to suffer most on average in State 3, while the Asian currencies cluster performs best in State 1. Standard deviation of returns also increases appreciably in all clusters when moving from State 1 to State 3. The Developed/European Currencies cluster tends to have the lowest standard deviation in all states, except in State 1 where the Asian Currencies cluster has the lowest. The Emerging/Commodity Currencies cluster has the highest standard deviation across states, while simultaneously possessing the worst return profile.

TABLE 2.4: Average daily returns per cluster by risk state. Standard deviation is noted in parenthesis

State	Developed/European Currencies	Emerging/Commodity Currencies	Asian Currencies
1	0.00% (0.37%)	-0.03% (0.38%)	0.01% (0.20%)
2	-0.02% (0.44%)	-0.03% (0.46%)	-0.01% (0.46%)
3	-0.00% (0.52%)	-0.09% (0.86%)	-0.04% (0.73%)

2.5 Discussion

The findings suggest that most of the variation in global currency markets can be accounted for by three key factors. Furthermore, the findings corroborate those in the literature. Namely, that the first principal component represents a Dollar risk factor while the second principal component is a carry trade factor. It is suggested that the third principal component represents an Asian markets factor.

An interesting result worth noting is the low resolution (observed in the total variance explained by the principal components) in the full 115 currency sample. This finding suggests a high degree of idiosyncratic noise across the broader currency market: the total variance explained by the first 10 dimensions, constructed to be orthogonal and maximise the remaining variance not captured by the dimension prior, is roughly 27%. If there was not a high degree of idiosyncratic noise, then currencies would move according to a handful of dimensions/factors and the total variance explained by these dimensions would be higher.

The total panel of currencies was reduced to 29 currencies, representing currencies from constituents of the G20, as well as Bosnia-Herzegovina, New Zealand, Norway, Singapore and Switzerland, which were chosen according to the improvement they brought to the variance explained in a PCA focusing on the first 3 components. The resolution in the reduced currency sample improved to 67.68% variance explained by the first 10 principal components. The drop-off in variance explained, from the general market factor (30.24%) to the carry trade factor (7.73%), is apparent and suggests that in general, the carry trade factor may not be a very important factor explaining broader returns on a linear basis.

Interestingly, the results presented in Section 2.4.6 provide evidence of an interaction effect between risk state and the principal components of currency markets. The results held for the majority of currencies, with the exception being the Russian Rouble with a p-value of 0.2. The implication for the remaining currencies is evidence of a joint interaction between the key currency market factors and the risk states. Appendix B details the test of the hypothesis of a joint interaction between the principal components, providing evidence to suggest a joint interaction for the South African Rand (ZAR), Australian Dollar (AUD), Chinese Renminbi (CNY), Euro (EUR) and Japanese Yen (JPY) among others.

It is worth noting that there are two possibilities, either this evidence suggests that equity implied volatility has a significant impact on the relationship between the currencies and the related market risk. Or, the VIX Index is not solely representing equity implied volatility and is indeed, as suggested by this work, a proxy for global risk appetite.

Table A.3 in Appendix A displays results of estimating the full model of each currency against the first three principal components, stratified by risk state. Currencies in the Emerging/Commodity cluster tend to have relatively large increases in R^2 when moving into the "High" risk state. The same pattern was observed for the Asian cluster currencies, but with lower overall R^2. Currencies in the Developed/European cluster tend to have the highest R^2 in the "Low" risk state. The regression coefficients for the carry trade factor tend to be significant across states and negative in the Emerging/Commodity cluster, but positive in the Developed/European cluster, as one would expect. It was also observed that average R^2 is higher for the

Developed/European cluster, when compared to the other two clusters, suggesting the Developed/European cluster currencies are less idiosyncratic.

To understand the strength of the relationship between the carry trade factor and currencies in different states, Table A.4 in Appendix A displays the regressions of each currency in the sample on the carry trade factor only. These results clearly display the jump in model fit when moving into the "High" risk state for the majority of Emerging/Commodity cluster currencies. The regression coefficient for the carry trade factor is significant across states for this cluster, save for the Hungarian Forint (HUF, State 1, p=0.25, State 2, p=0.39) and the Polish Złoty (PLN, State 1, p=0.55). Furthermore, the coefficient for the carry trade factor is the most negative in the "High" state for all Emerging/Commodity cluster currencies, except the Mexican Peso.

These results provide strong evidence that in "High" risk states, the carry trade factor specifically impacts Emerging/Commodity cluster currencies and drives a solid proportion of their variation. The opposite is true for Developed/European cluster currencies; in "High" risk states, the R^2 tends to drop considerably, with the notable exception of the Japanese Yen (which conventional wisdom suggests is a key currency in the carry trade), where it is observed to increase for the carry trade factor-only model.

Chapter 3

Regime-Switching Models of Risk Appetite

3.1 Introduction

Results of the analysis in Chapter 2 showed that there was a statistically significant interaction between the risk state and the three currency market factors that were extracted using the PCA. The risk states in question were chosen according to a simple heuristic: values above the 85^{th} percentile were deemed "high risk", values less than or equal to the 15^{th} percentile were deemed "low risk" and those falling between were considered "normal risk". Such an approach does not constitute a parametric "model" of the risk appetite and the choice of percentiles are somewhat arbitrary. The results of the previous section indicate there is value in analysing the returns to currencies, contingent on these risk states. This chapter will further examine the concept of the risk "state" and attempt to infer the true data generating process for risk appetite, assuming that it is governed by a Markov-switching process.

Three different Markov-switching model classes are considered: mean-only, autoregressive, and autoregressive moving-average. The estimation of these models, with a focus on the autoregressive moving-average model, is discussed and the results of simulations to calculate the bias of the estimator are conducted, with the results presented in Appendix C. The results of estimating a battery of models varying by number of states, model class, parameter switching scheme and error distribution are presented to inform model selection. The choice of number of states (also called "regimes") is an important one as the true model could be linear in nature. The parallel sampling method as described in Congdon (2006:349) is used to estimate the posterior probability of the best model being either linear (one effective state), or containing two or three states.

The layout of the chapter is as follows: a review of the Markov-switching time-series model literature, a discussion of the data used and a review of previous applications of such models to the data in question, the methodology employed and finally the results obtained. The chapter concludes by discussing applications to the methodology of Chapter 2.

3.2 Review of the Regime-Switching Literature

Regime-switching models are attractive because they permit the existence of multiple states of the world in which a process evolves. At least one parameter of the process will differ across states, allowing for richer modelling of the underlying process. This is important as economic (Hamilton, 1989) and financial data (Ang & Timmermann, 2012) have a tendency to display nonlinear or abrupt shifts in behaviour. This chapter will discuss the particular case where the state or regime variable is unobserved, and inference must be done to determine the state that the process is in. Specifically, this chapter will look at the Markov-switching time-series model, a popular non-linear approach to time-series modelling (Alvarez-Plata & Schrooten, 2006).

Hamilton (1989), building on Goldfeld and Quandt's (1973) Markov-switching regression model, developed the Hamilton filter to estimate a Markov-switching autoregressive model to describe the evolution of the post-war US real GNP. The filter is an alternative approach to that explored by Sclove (1983), who, treating the unobserved states as observed, picked those values for the states which maximised a joint likelihood function of the states and GNP. Hamilton's approach involves numerically maximising the marginal likelihood function and then inferring the states using the estimated parameters of the model. The model specification was an AR(4) model with a deterministic trend, switching according to two states of the world - "normal" and "recessionary". The states which the model identified as recessionary and normal were found to be similar to those identified by the National Bureau of Economic Research's dating of business cycles, suggesting the model could be used as an objective method to infer the business cycle.

The Markov-switching model has been applied to many more time-series in the literature, with Jeanne and Masson (2000) modelling estimates of devaluation probabilities for the French franc as a two-state Markov-switching model, where only the deterministic trend is allowed to vary by state. They compared this model to the linear case, noting a substantially better fit in the two-state model. Discussing the likelihoods of the two models, they noted the complication presented by comparing a model with nuisance parameters using the likelihood ratio test but went on to suggest that even using a "conservative" critical value, the finding was significant. Alvarez-Plata and Schrooten (2006) modelled a constructed index of speculative pressure on the Argentinean Peso as a Markov-switching model and inferred state probabilities for being in either the "tranquil" (probability of devaluation is low) or the "high economic tensions" state (probability of devaluation is high). Furthermore, they used the upper bound for the likelihood ratio test significance level, derived in Davies (1977), to compare their model to a no-state one, rejecting the null hypothesis of no switching.

The models discussed did not include a moving-average (MA) term into the time-series specification and the estimation of such models is a natural extension. Chen and Tsay (2007) explored the inclusion of the MA term in the Markov-switching model equation using two algorithms to overcome the challenges presented by the path-dependence of the unobservable error term. This path-dependence results in an exponential increase in the possible paths of the model, NT where N represents the number of states of the model and T represents the number of observations, making it computationally challenging to estimate. They adopted the approach of Gray (1996) in replacing the error at time t with its conditional expectation, calling this the

Hamilton-Gray algorithm. They include an extension to the model (the Extended Hamilton-Gray algorithm) which makes more efficient use of previous information in generating the conditional expectations of the lagged error, than the Hamilton-Grey algorithm.

An alternative strand of literature concerns using Bayesian methods for the estimation of these models. For example, Kim and Kim (2015) discuss approaches in the literature for implementing Markov Chain Monte-Carlo estimation for Markov-switching ARMA models.

Of interest is the appropriate choice of model and selection among candidate models. Francq and Zakoian (2001) derive the autocovariance function for Markov-switching ARMA models, as well as stationarity conditions. Smith et al. (2006) derived a "Markov switching criterion", an estimate of Kullback-Leibler divergence, the motivation for which was owing to the finding that Akaike's (Akaike, 1998) Information Criterion (AIC) leads one to over-select parameters (particularly the true number of states) in the Markov-switching model.

A natural question is whether the estimated model is statistically significantly different from the equivalent, linear model (that is, one with a single state). Such a question cannot be answered using standard likelihood ratio tests, as some model parameters are not identified under the null (Di Sanzo, 2009:154). Hansen (1992) developed a bound for the standardized likelihood ratio which allows the calculation of p-values. Carrasco et al. (2014) developed a test statistic for testing the constancy of parameters in not just Markov-switching models, but more generally, models with random coefficients. Di Sanzo (2009) developed a residual bootstrap procedure to test the null hypothesis of no switching for the Markov-switching AR model, finding better performance than the Hansen and Carrasco et al. tests.

It is, however, not clear how to use the tests described above to select between two competing models of multiple states. Congdon (2006:349) (also discussed in Zucchini, MacDonald and Langrock (2016:116)) take a Bayesian approach to attempt to answer the question of the true number of states in the model, by describing a method for estimating the posterior probability of the number of states by parallel sampling of posterior distributions of the models.

3.3 Data

This chapter uses the CBOE VIX Index (first panel of Figure 3.1) as a proxy for global risk appetite. The time period considered is from 3 January 2000 to 13 April 2017 ($n = 4509$), to ensure a suitably large sample while lightening computational load. All data were obtained from Bloomberg (2017).

Romo (2011) modelled the VIX Index using Markov-switching specifications, analysing a two-state AR(1) model with three different specifications for the variance of the error term: error variance depending upon the state z, error variance with an ARCH specification, and error variance with a GARCH specification. Errors were assumed to be distributed according to the student t-distribution, with the latter two models found to have better out-of-sample performance.

Baba and Sakurai (2011:1415) modelled the VIX Index as an n-state Markov-switching process, using the Markov-switching Criterion (MSC) of Smith et al. (2006) to select

three as the optimal number of states ("tranquil", "turmoil", "crisis"). It is worth noting that they constrained the probability matrix to rule out transitions from tranquil to crisis-states, for computational efficiency (p. 1416). Performing a likelihood ratio test, they found that, at the 5% level, the restricted model did not perform significantly differently to a model without this restriction. They introduced time-varying probabilities for the regimes as well, using a set of macro-economic variables as predictors, finding 5 year and 10-year Treasury spreads (spreads calculated relative to the 3-month Treasury bill) had a statistically significant effect on the transition probability from the "tranquil" to the "turmoil" state. The tranquil, turmoil and crisis states were found to have a persistence of 29, 9.1 and 2 months, which agrees with common sense, if arguably too persistent for the crisis state.

This chapter contributes to the literature by modelling the CBOE VIX Volatility Index as a Markov-switching ARMA model, that is, including the moving average term, estimated using the Hamilton-Gray algorithm of Chen and Tsay (2007). The model is compared to the nested mean-only and AR Markov-switching models, of varying switching combinations, for Normal and t-distributed errors and for two, three-state specifications (with four-state results noted). A model selection procedure is conducted on the basis of AIC (Akaike, 1998) & BIC (Schwarz, 1978), as well as model fit diagnostics, and the correct number of states is validated using a parallel sampling approach in the Bayesian framework.

An initial modelling concern is the stationarity of this time-series. The sample autocorrelation (ACF) and partial autocorrelation (PACF) functions are presented in the second and third rows of Figure 3.1. Visual inspection of the series suggests that it may be non-stationary and may require differencing to obtain a stationary series. One can note the slow decay of the sample ACF, which also suggests the data require differencing and are non-stationary.

Panel 1 (left) of Figure 3.2 displays the data after first-differencing, as well as the associated sample ACF & PACF. The lack of autocorrelation suggests that further differencing is not appropriate for this series, however the sample ACF and PACF for the squared differences of the VIX Index in panel 2 (right) of Figure 3.2 suggest a variance which varies over time.

As the time series plots suggest that the time series considered doesn't require further differencing, this chapter will therefore focus on modelling the changes in the VIX Index.

3.4 Methodology

The general form of the Markov-switching model considered is described by the following equation:

$$y_t = \alpha_{S_t} + \sum_{i=1}^{p} \beta^i_{S_t} y_{t-i} + \sum_{j=1}^{q} \theta^j_{S_t} \epsilon_{t-j} + \epsilon_t \tag{3.1}$$

for

$$t \in [1, \ldots, T] \text{ where } T \text{ represents the number of observations}$$

and where

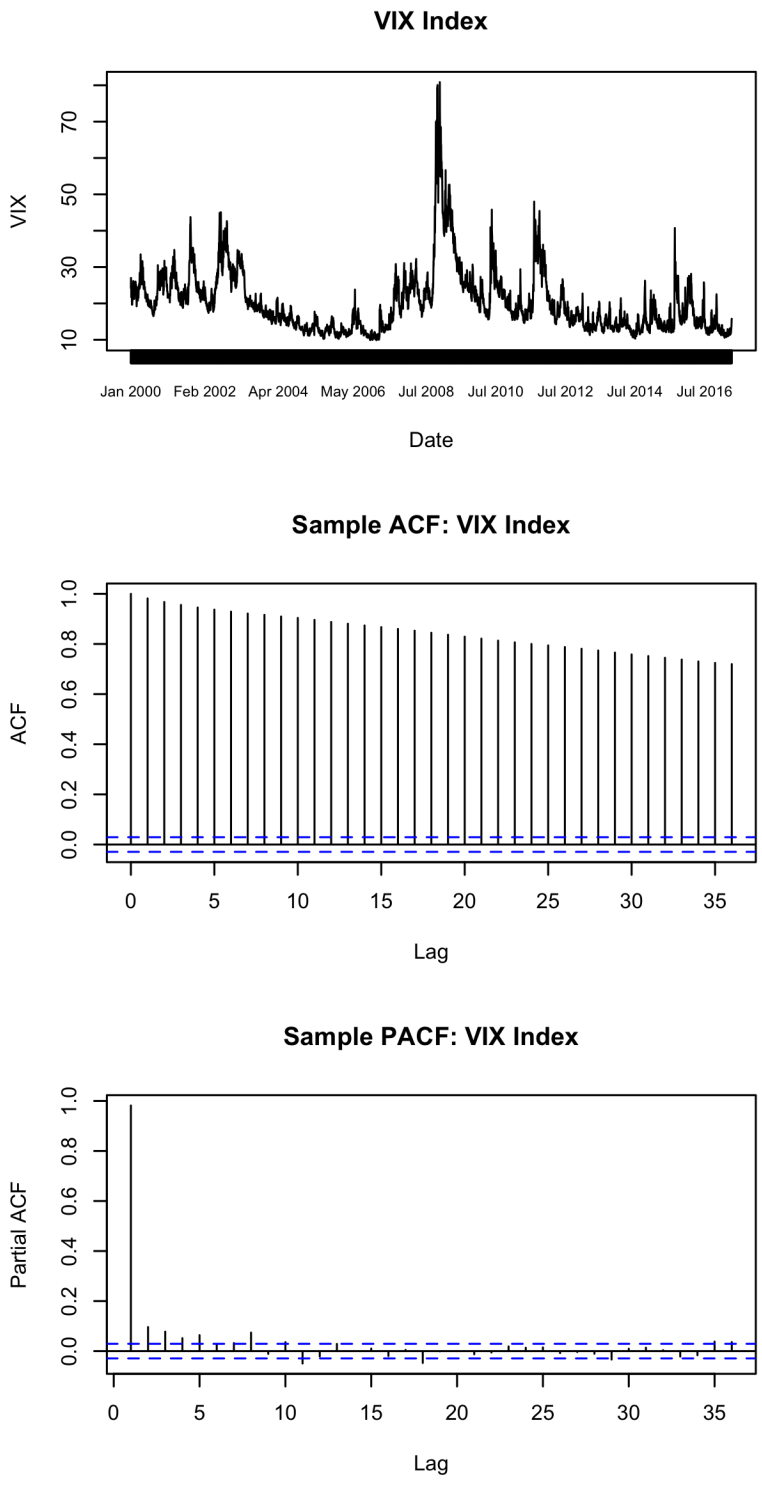

FIGURE 3.1: Time series plot (row 1), ACF (row 2) and PACF (row 3) of the VIX Index

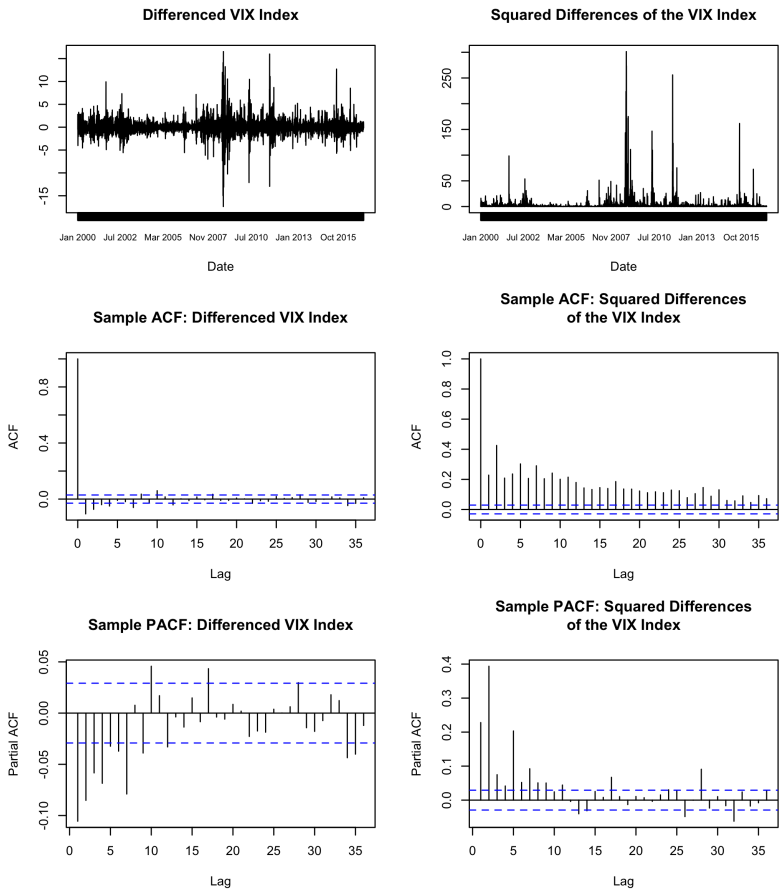

FIGURE 3.2: Panel 1 (left) and Panel 2 (right) contain the plot of the time series (first row), the sample ACF (second row) and the sample PACF (third row) for the differenced VIX index and the square of the differenced VIX index

$$\epsilon_t \text{ has some distribution with possibly switching parameters, } E[\epsilon_t] = 0 \text{ and } Cov(\epsilon_i, \epsilon_j) = 0 \text{ for all } i \neq j,$$

which is subject to stationarity and invertibility constraints, namely, the roots of the lag polynomial function $L(x)$ for both β and θ lie outside the unit circle. This chapter considers the case where ϵ_t is either Normally distributed with standard deviation σ_{S_t}, or t-distributed with scale parameter σ_{S_t} and degrees of freedom parameter δ_{S_t}. Furthermore, S is a process governed by an ergodic Markov chain with transition probability matrix A, with S_t taking on the value i for i in $\{1, 2, \ldots, N\}$. The parameters are indexed by the state of the Markov chain at time t, with $\alpha, \ldots, \sigma$ representing the vectors of the parameters across states for the Normally distributed error model, and $\alpha, \ldots, \delta$ for the t-distributed error model.

Let Y_t represent the information set up until time t and Ψ represent the parameters of the model, i.e. the vector $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{\sigma}, A)'$ (inclusive of $\boldsymbol{\delta}$ for the t-distributed error model). The prediction probability is defined as:

$$P(S_t = i | Y^{t-1}, \Psi),$$

the filtering probability as:

$$P(S_t = i | Y^{t}, \Psi),$$

and the smoothed probability as:

$$P(S_t = i | Y^{T}, \Psi).$$

Two models are nested within the Markov-switching ARMA (MS-ARMA) model described by equation 3.1: the "auto-regressive" (MS-AR) model where $\theta^j_{S_t} = 0$ for all j, and a "mean-only" model where $\beta^i_{S_t} = 0$ for all i and $\theta^j_{S_t} = 0$ for all j.

A battery of different variants of the three classes of models (mean-only, MS-AR and MS-ARMA), with both Normally distributed and t-distributed errors, were estimated. For each model class and error distribution, all parameters and a subset of different parameters were allowed to switch. Not all combinations of parameter switching were estimated, as this would be computationally unfeasible. For all of these model types, two, three and four-state models were estimated.

Model selection proceeds by selecting, within each class, the model with the lowest AIC (Akaike, 1998). The best models across the classes are then compared by analysing the pseudo-residuals of the model.

An assessment of the maximum likelihood estimator for the MS-ARMA model was also conducted, with a more detailed methodology, as well as the results, presented in Appendix C. A set of three MS-ARMA processes were simulated 100 times each, with samples of size 100, 250, 500 and 1000. The results suggest that the algorithm performs reasonably well in estimating the parameters of the model.

Model selection continues by parallel sampling (as described in Congdon (2006:349)) the model selected above, to select the correct number of states. The calculation is described below:

1. Sample $\theta^{(i)}_k$ B times from the posterior distribution of the parameters, $p(\theta|Y, M = k)$, where Y represents the data, $i \in [1, 2, ..., B]$, and M represents the number of states in the model, for $k = 1, 2, 3$.

2. For each of the M models, calculate the model probabilities $G_k^{(i)} = p(Y|\theta_k^{(i)}, M = k)p(\theta_k^{(i)}|M = k)p(M = k)$ for all $i = 1, 2, 3, ..., B$.
3. To avoid numerical underflow on the exponentiation of log-probabilities in the step above, instead find the maximum log model probability at each i, and subtract it from each of the k log model probabilities, at each i. One can now take the exponential of these scaled log model probabilities.
4. At each i, compute $S_k^{(i)} = \frac{G_k^{(i)}}{\sum_{j=1}^{M} G_j^{(i)}}$ for $k = 1, 2, 3$.
5. Calculate the posterior distribution of the states $p(M = k|Y) = \frac{1}{B}\sum_{i=1}^{B} S_k^{(i)}$.

The posterior distribution of the model parameters, required in the initial step mentioned above, was sampled initially using the Robust Adaptive Metropolis algorithm of Vihola (2012) for the burn-in period, implemented in the *adaptMCMC* (Scheidegger, 2018) package in R, followed by the Metropolis algorithm for the sampling. The Metropolis algorithm for sampling the posterior of a single variable, with a Normally distributed proposal distribution, is detailed below:

Algorithm 1 Metropolis Algorithm

Let B equal the desired number of samples from the posterior
Let X equal an empty vector of length B
Let x_0 equal a valid starting parameter value for the the likelihood $L(x)$ and prior $\pi(x)$
for i in the range $1 : B$ **do**
  Generate a proposal value $x^* \sim N(x_{i-1}, \sigma^2)$
  Generate $u \sim U(0, 1)$
  Evaluate the ratio $p = \frac{L(x^*)\pi(x^*)}{L(x_{i-1})\pi(x_{i-1})}$
  if $u < min(1, p)$ **then**
    Set $X[i] = x^*$
  else
    Set $X[i] = x_{i-1}$
  end if
end for

Convergence was assessed by trace plots (of the unthinned sample), and sample ACF plots (of the burned-in, thinned sample) of the individul parameters (see figures D.1 through D.6 in Appendix D) and the results checked by running another MCMC chain and verifying convergence.

Let T_{ij} be a random variable representing the time taken for first passage to state j from state i for the Markov chain. The expected first passage time to state j from state i for a Markov chain is then $E[T_{ij}]$. Expected first passage times of the estimated Markov chain of the selected model were calculated using the R package *markovchain* (Spedicato, 2017), along with the smoothed probabilities (calculated as in Kuan (2002)) of the states of the selected model, for each observation in the sample. In addition to this, the most likely *sequence* of states was calculated, using the Viterbi (Viterbi, 1967) algorithm.

3.5 Estimation and Posterior Sampling

3.5.1 Model Estimation

Estimation of the linear ARMA model was conducted using the method of maximum likelihood. For the AR and mean-only model, the formulae outlined below are simplified as necessary. Estimation of the Markov-switching ARMA models was non-trivial. The methodology described in Chen and Tsay (2007) was used to replace the path-dependent lagged error with the conditional expectation of the error. This is described below.

The model to be estimated is described in equation 3.1 above. The estimation described below describes the case for the Normally distributed error variant, however is trivially different for the t-distributed error variant.

One seeks to maximise the log likelihood function:

$$l(\Psi, Y) = \sum_{i=1}^{T} ln[f(y_t|Y^{t-1}, \Psi)], \tag{3.2}$$

where:

$$f(y_t|Y^{t-1}, \Psi) = \sum_{i}^{N} P(S_t = i, Y^{t-1}, \Psi) f(y_t|S_t = i, Y^{t-1}, \Psi) \tag{3.3}$$

and the second term in the summation is distributed normally:

$$f(y_t|S_t = i, Y^{t-1}, \Psi) \sim N(\boldsymbol{\mu}, \sigma_{S_t=i}),$$

where:

$$\boldsymbol{\mu} = y_t - \alpha_{S_t} - \sum_{i=1}^{p} \beta^i_{S_t} y_{t-i} - \sum_{j=1}^{q} \theta^j_{S_t} \epsilon_{t-j}. \tag{3.4}$$

However, the previous error ϵ_{t-j}, is not observed, so it is replaced with the conditional expectation, $\hat{\epsilon}_{t-j}$, of this quantity at time $(t-j)-1$. That is, as in Chen and Tsay (2007), the conditional expectation of the error at time $t-j$ is given by:

$$\hat{\boldsymbol{\epsilon}}_{t-j|Y^{(t-j)-1}} = \sum_{i=1}^{N} P(S_{t-j} = i|Y^{t-j-1})(y_{t-j} - \boldsymbol{E}[y_{t-j}|S_{t-j} = i, Y^{t-j-1}]), \tag{3.5}$$

where

$$P(S_{t-j} = i|Y^{(t-j)-1}) = \sum_{m=1}^{N} P(S_{t-j} = i|S_{(t-j)-1} = m) P(S_{(t-j)-1} = m|Y^{(t-j)-1}). \tag{3.6}$$

The filtering probabilities are:

$$P(S_t = i|Y^t, \Psi) = \frac{P(S_t = i|Y^{t-1}, \Psi) f(y_t|S_t = i, Y^{t-1}, \Psi)}{f(y_t|Y^{t-1}, \Psi)}, \tag{3.7}$$

and the prediction probabilities are:

$$P(S_{t+1} = i|Y^t, \Psi) = \sum_{j=1}^{N} P(S_t = i|S_{t-1} = j) P(S_t = j|Y^t, \Psi), \tag{3.8}$$

where the first term in the summation is the j, i^{th} entry in the transition probability matrix A.

With the above equations, one can obtain a value for each y_t by calculating the prediction probabilities (eq. 3.8) for time t, calculating y_t, calculating the filtering probabilities (eq. 3.7) and then calculating the prediction probabilities for time $t+1$. In this way one can iterate through all observations. This chapter used initial conditions for the prediction probabilities for time $t = 1$ as the prediction that the process would be in state 1 with probability 1, i.e. $(1, \ldots, 0)$. The initial values of the errors were set to 0.

The function above was numerically estimated in R using the BFGS algorithm of the *optim* function for the Mean-only model, while the L-BFGS-B algorithm was used for the AR and ARMA models to enable the constraining of the AR and MA parameters to the range $(-1, 1)$. The working parameters γ used for the optimisation were unconstrained (with the exception of the constraints placed on the AR and MA terms, as noted for those estimations), and were transformed appropriately to calculate the likelihood. For each of the $M^2 - M$ off-diagonal transition probability matrix elements p_{ij} (where M equals the number of states for $M \geq 2$), the transformation was as follows (with the diagonal working parameters γ_{ii} for $i = 1, 2, 3, ...M$ being equal to 0):

$$p_{ij} = \frac{e^{\gamma_{ij}}}{\sum_k^M e^{\gamma_{ik}}}. \tag{3.9}$$

For both the degrees-of-freedom parameter δ (where necessary) and the standard deviation/scale parameter σ (depending on the type of error), the working parameter was exponentiated, i.e. $\delta = e^{\gamma}$.

Parameters were initialised in the following manner:

1. For the mean parameter α, $\alpha_i = \frac{i}{2}\bar{Y}$, i represents the i^{th} state in the n state model, where $\bar{Y}$ is the sample mean of the data.
2. For the AR and MA parameters, each parameter was initialised uniform randomly between -1 and 1.
3. For the sigma/scale parameters, $\sigma_i = log(\frac{i}{2} sd(Y))$ with i again representing the i^{th} state in the n state model, and where $sd(Y)$ is the standard deviation of the data. The log transform ensures the parameter is in working parameter space.
4. All off-diagonal transition probability matrix elements were initialised as -3.

5. All degrees-of-freedom parameters δ were initialised uniform randomly between -1 and 1.

For each of the models considered, two estimations were made and the highest likelihood estimation kept.

Sampling from the posterior distribution of the parameters, $p(\theta|Y, M)$ where Y represents the data and M the number of states of the model, is required for the calculation of the posterior probability of the number of states. Initially the Robust Adaptive Metropolis algorithm (Vihola, 2012) was used during the burn-in period, with the standard Metropolis algorithm used for the sample.

In an attempt to avoid enforcing a strong prior on the parameters, the prior distribution for the off-diagonal transition probability parameters (constrained to be positive for this sampler, details are provided in the Results section) was chosen to be exponential with rate parameter 0.001. For all other parameters, a Normal distribution with mean 0 and a standard deviation of 1000 was selected.

The sampler was run for 350 000 iterations, with a burn-in period, using the Robust Adaptive Metropolis algorithm, of 250 000 iterations. Parameters were uniform randomly initialised between 0 and 0.1. The sampled parameters were thinned down to every 350^{th} sample.

3.6 Results

3.6.1 Model Selection

Model Estimation Results

The results from the battery of model estimations is provided in Table 3.1 & 3.2, and visualised in Figure 3.3 & 3.4. For the mean-only class, the three-state and four-state models with the lowest AIC are the t-distributed error models with all parameters switching. For this model class, the equivalent two-state model has the second lowest AIC among two-state models, narrowly beaten by the t-distributed error model without mean switching. For the autoregressive class, the lowest AIC model for two, three and four-state models was the t-distributed error model without mean or AR parameter switching. Finally, for the autoregressive moving-average model class, the two and three-state models with the lowest AIC were the t-distributed error models with all parameters switching. For four states, the model with the lowest AIC was the Normally distributed error model with all parameters switching (with a higher AIC than the best three-state model in this class). The same trend is observed with the table of BIC values, with the exception of the four-state ARMA model, where the Normally distributed error with sigma switching only had the lowest BIC.

A general trend across the results in Table 3.1 and Table 3.2 is a lower AIC and BIC as one increases the number of model states, with the AIC & BIC appearing to plateau as the total number of parameters increases. In addition, the results tend to indicate that models with t-distributed errors improve (lower) AIC & BIC, suggesting improvement in fit. This could be owing to the ability of a t-distributed error to capture "fat-tail" behaviour, a phenomenon common to financial data (Cont, 2002).

TABLE 3.1: AICs of varying Mean-only, AR and ARMA models. Results in bold face represent the lowest in model class for the number of states. Superscripts denote overall rank (ordered by lowest first) of the models according to AIC

	Two States	**Three States**	**Four States**
Mean-only model			
Normal error – all switching	14 656.81	14 126.61	14 023.76
Normal error – sigma switching only	14 659.54	14 126.06	14 021.91
Normal error – mean switching only	17 331.30	15 924.38	17 355.30
t error – all switching	14 266.26	**14 041.33**	**13 962.18**
t error – sigma & d.o.f switching	**14 264.27**	14 066.43	13 984.74
t error – mean switching only	15 064.89	15 074.89	15 088.89
AR model			
Normal error – all switching	14 634.98	14 113.06	14 010.44
Normal error – sigma switching only	14 641.25	14 112.52	14 006.59
Normal error – mean switching only	17 282.77	17 292.77	17 306.77
Normal error – AR switching only	16 903.55	16 616.72	16 571.44
t error – all switching	14 249.11	14 030.04	13 977.05
t error – sigma & d.o.f switching	**14 245.95**	**14 023.91**	**13 970.67**
t error – mean switching only	15 045.64	15 055.64	15 069.63
t error – AR switching	14 877.92	14 747.35	14 734.11
ARMA Model			
Normal error – all switching	14 515.81	14 008.65	**13 919.89**[2]
Normal error – sigma switching only	14 576.14	14 055.66	13 949.80
Normal error – mean switching only	16 220.25	15 710.85	17 219.66
Normal error – AR switching only	16 844.04	16 601.58	16 498.63
Normal error – MA switching only	16 662.94	16 268.57	16 211.12
t error – all switching	**14 111.92**	**13 910.27**[1]	13 920.76[3]
t error – sigma & d.o.f switching	14 174.77	13 963.48	13 930.75
t error – mean switching only	14 925.76	14 931.20	14 944.01
t error – AR switching only	14 816.07	14 747.04	14 706.46
t error – MA switching only	14 823.82	14 733.93	14 717.06

TABLE 3.2: BICs of varying Mean-only, AR and ARMA models. Results in bold face represent the lowest in model class for the number of states. Superscripts denote overall rank (ordered by lowest first) of the models according to BIC

	Two States	**Three States**	**Four States**
Mean-only model			
Normal error – all switching	14 695.29	14 203.58	14 152.04
Normal error – sigma switching only	14 691.61	14 190.20	14 130.95
Normal error – mean switching only	17 363.37	15 988.52	17 464.34
t error – all switching	14 317.57	**14 137.54**	**14 116.11**
t error – sigma & d.o.f switching	**14 309.17**	14 149.80	14 119.43
t error – mean switching only	15 103.37	15 145.44	15 204.34
AR model			
Normal error – all switching	14 686.29	14 209.27	14 164.37
Normal error – sigma switching only	14 679.73	14 183.07	14 122.04
Normal error – mean switching only	17 321.25	17 363.32	17 422.22
Normal error – AR switching only	16 942.03	16 687.27	16 686.89
t error – all switching	14 313.25	14 145.49	14 156.64
t error – sigma & d.o.f switching	**14 297.26**	**14 113.70**	**14 111.77**
t error – mean switching only	15 090.54	15 132.61	15 191.49
t error – AR switching	14 922.82	14 824.32	14 855.97
ARMA Model			
Normal error – all switching	14 579.95	14 124.10	14 099.48
Normal error – sigma switching only	14 621.04	14 132.63	**14 071.66**[3]
Normal error – mean switching only	16 265.15	15 787.82	17 341.52
Normal error – AR switching only	16 888.94	16 678.55	16 620.49
Normal error – MA switching only	16 707.84	16 345.54	16 332.98
t error – all switching	**14 188.89**	**14 044.96**[1]	14 126.00
t error – sigma & d.o.f switching	14 232.49	14 059.69[2]	14 078.27
t error – mean switching only	14 977.07	15 014.58	15 072.29
t error – AR switching only	14 867.38	14 830.42	14 834.74
t error – MA switching only	14 875.13	14 817.31	14 845.34

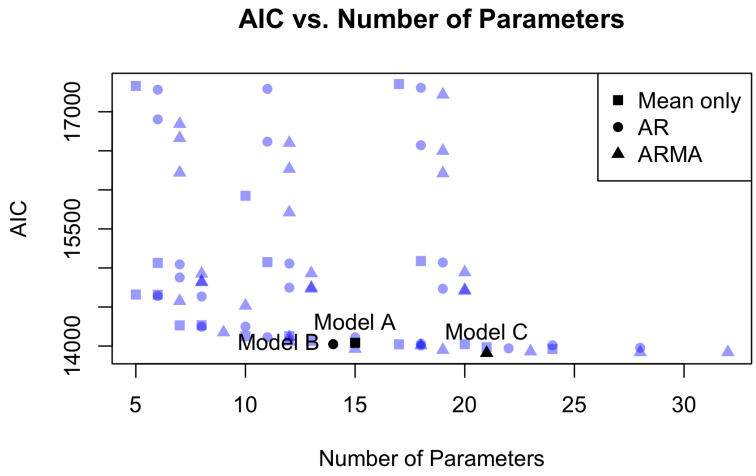

FIGURE 3.3: Scatter plot of AIC vs. parameter count for the results presented in Tables 3.1 & 3.2

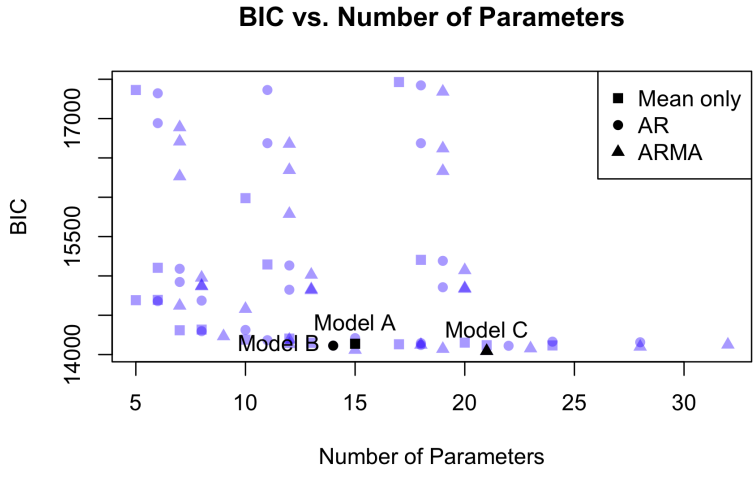

FIGURE 3.4: Scatter plot of BIC vs. parameter count for the results presented in Tables 3.1 & 3.2

Smith et al. (2006:567) showed that AIC tends to overestimate the number of within-state parameters, but is worse at over-selecting for the number of *states* in the model. This dynamic may explain, to some extent, the tendency for the improvement in t-distributed error fit, however the finding of over-selecting for the number of states, together with marginal gain in AIC and the increased complexity (from an interpretation standpoint) of four-state and higher models, leads this chapter to consider models of up to three-states only. Another finding from Table 3.1 & 3.2 is the seeming importance of allowing the error variance/scale parameter, and degrees of freedom parameter, to switch between regimes. This is evidenced by the large rise in AIC & BIC when constraining the switching of these parameters, in all classes and error distribution types, across both two and three-state models. Considering the results in tables 3.1 and Table 3.2, the best model for the mean-only class is the t-distributed error model with all parameters switching (Model A), for the AR model class it is the t-distributed error model with scale parameter and degrees of freedom switching (Model B) and for the ARMA model class it is the t-distributed error model with all parameters switching (Model C).

Zucchini, MacDonald and Langrock (2016:105) motivate the use of "pseudo-residuals" in assessing the fit of Markov-switching models. The uniform forecast pseudo-residual, using the notation of Zucchini, MacDonald and Langrock (2016:105), is

$$u_t = Pr(Y_t \leq y_t | Y^{(t-1)} = y^{(t-1)}) \tag{3.10}$$

with Y^t and $Y^{(t-1)}$ representing the likelihood of the first t and $t-1$ observations in the data resepctively. The normal forecast pseudo-residual, with Φ^{-1} representing the inverse CDF of the standard Normal distribution, is then

$$z_t = \Phi^{-1}(u_t) \tag{3.11}$$

with the calculation thereof provided in code in Appendix E. Figure 3.5 shows histograms of the uniform and normal forecast pseudo-residuals, and the Normal QQ-Plots of the normal forecast pseudo-residuals, for models A (mean-only, t-distributed error, all parameter switching), B (AR, t-distributed error, scale parameter and degrees of freedom switching only) & C (ARMA, t-distributed error, all parameters switching) selected above. The figures suggest that models A & B fit better than C, on the basis of the rather large deviation from Normality in the lower tails of the pseudo-residuals for Model C. Upon inspection of the upper tails of the Normal QQ-Plots for Models A & B, Model A appears to have a better fit. This, coupled with the simpler state-dependent structure of Model A, leads to it being selected for further analysis.

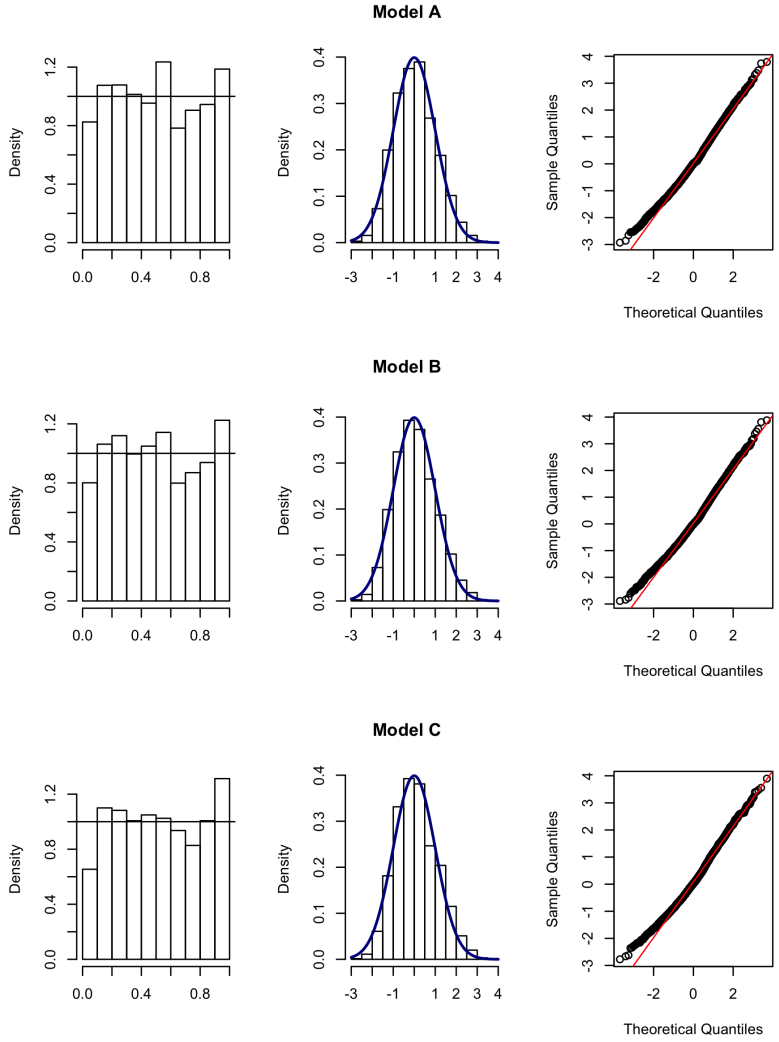

FIGURE 3.5: Column 1 shows the histogram of the uniform forecast pseudo-residuals for each of the best-in-class models. Columns 2 and 3 display the Normal forecast pseudo-residuals and Normal QQ-Plots for the same

3.6.2 Selecting Between Linear, Two and Three-state Models

The results of the MCMC sampling, using the Robust Adapted Metropolis algorithm for the burn-in period and the Metropolis algorithm thereafter, are presented in Table 3.3 below for the linear, two and three-state models (δ represents the degrees of freedom parameter). The working parameters making up the transition probability matrix were constrained to be positive only, with the transformation to the true parameters therefore being subtly different to the one used in the maximum likelihood estimation (each row of the matrix being the working parameters divided by their row sums). The estimate for each parameter is calculated by taking the mean of the burned-in, thinned, transformed sample from the posterior for that parameter, and the 95% credibility interval calculated from the same is provided in parentheses. Two chains were run for each of the linear, two-state and three-state models, with the results of the other runs, including the estimated parameters and credibility intervals, trace plots and sample autocorrelation plots, presented in Appendix D (Table D.1 and figures D.7 through D.12). Both chains, for each of the three models sampled, appear to have converged, owing to the lack of significant sample autocorrelation and stability of the trace plots for both runs. In addition, the similarity of the estimated parameters between chains, suggests the chains converged upon the same point. Finally, the multivariate potential scale reduction factors (Brooks & Gelman, 1998), derived from analysing the within-chain and between-chain variance for both of the runs for each of the three models, are provided in Table 3.4; all of the values are less than 1.1, suggesting convergence for all three models that were sampled.

The results of the parallel sampling calculation to estimate the posterior distribution of the number of states for the model, is shown in Table 3.5. The results suggest a strong preference for the three-state model, which matches the findings of Baba and Sakurai (2011) of three inherent states in the risk appetite process, as proxied for by the VIX index.

3.6.3 Discussion

The estimated parameters for state 2 of the three-state model show it is characterised by a low scale parameter of the error distribution (with high confidence), low degrees of freedom (heavy tails), negative mean (with high confidence) and the highest point estimate of the probability to remain in the state (however this is uncertain at the 95% credibility level). This intuitively suggests a lengthy "good" state, where the process is on average declining, with a low volatility (however with heavy tails). State 1, tentatively labelled as the “higher risk” state, has a lower point estimate of the probability to remain in the state (as compared to state 2), a negative mean (but the credibility interval is wide and includes positive values), higher error scale parameter (with high confidence) and the largest degrees of freedom parameter (with a wide credibility interval). State 3, labelled a "risky" state, has a positive mean (but with a credibility interval that includes negative values), the highest scale parameter value (with high confidence), the lowest degrees of freedom parameter (heaviest tails) and the lowest point estimate of the probability to remain in the state.

Table 3.6 displays the expected first passage times in days for the estimated Markov chain, with the 95% credibility interval in parentheses. Assuming one is currently in the "good" state, it would take an average of 31 days to move into the "higher risk" state, from which it takes an estimated average of 164 days to move into the "risky"

TABLE 3.3: Results of MCMC sampling from the posterior distribution (95% credibility interval in parentheses)

	State 1	State 2	State 3
Linear			
α	-0.0689 (-0.0976, -0.0408)		
σ	0.7637 (0.7284, 0.7980)		
δ	2.0846 (1.9228, 2.2403)		
Two-State			
α	-0.0523 (-0.1350, 0.0282)	-0.0564 (-0.0849, -0.0294)	
σ	1.5570 (1.4391, 1.6688)	0.5678 (0.5329, 0.6071)	
δ	3.4503 (2.9435, 4.1285)	5.4755 (4.2261, 7.3144)	
P_{1i}	0.9636 (0.9510, 0.9746)	0.0364 (0.0254, 0.0490)	
P_{2i}	0.0264 (0.0174, 0.0364)	0.9736 (0.9636, 0.9826)	
Three-State			
α	-0.0372 (-0.1010, 0.0266)	-0.0527 (-0.0808, -0.0257)	0.1016 (-0.2842, 0.4823)
σ	1.3439 (1.2138, 1.4805)	0.5116 (0.4762, 0.5493)	3.2501 (2.7176, 3.9126)
δ	13.2632 (6.9571, 29.5193)	5.8928 (4.3096, 8.0820)	4.8251 (2.9718, 8.0302)
P_{1i}	0.9521 (0.9353, 0.9664)	0.0369 (0.0248, 0.0524)	0.0110 (0.0048, 0.0185)
P_{2i}	0.0324 (0.0193, 0.0480)	0.9647 (0.9509, 0.9783)	0.0029 (0.0001, 0.0079)
P_{3i}	0.0681 (0.0338, 0.1167)	0.0084 (0.0004, 0.0251)	0.9236 (0.8780, 0.9578)

state, with a very wide interval ranging from 87 to 314 days. Once in this "risky" state, it takes a relatively short 18 days (up to 29 days at the 95% credibility level) to move down into the "higher risk" state, or 42 days to move down to the "good" state. The expected first passage from the "good" state to the "risky" state is 182 days (at worst 100 and at best 336 days at the 95% credibility level), combined with the expected 164 days for first passage assuming the market is in the "higher risk" state, means this state is quite rare.

Figure 3.6 displays the smoothed probabilities for the states for each day in the period 1 January 2006 until 1 January 2011. A subset was chosen to better visualise the switching between states, while this period was chosen as it contains the 2008/2009

TABLE 3.4: Multivariate potential scale reduction factors for the linear, two-state and three-state models sampled using MCMC. Values closer to 1 suggest convergence

Linear Model	1.00
Two-State Model	1.01
Three-State Model	1.07

TABLE 3.5: Estimated model probabilities

$P(M = 1\|Y)$	0.0000
$P(M = 2\|Y)$	0.0000
$P(M = 3\|Y)$	0.9999

financial crisis. Figure 3.7 shows the most likely path of states for the same period, as calculated using the Viterbi (Viterbi, 1967) algorithm (calculation provided in code in Appendix E). The two approaches identify roughly the same states in the data, neatly identifying State 3 ("risky") as the prevailing state during the period of volatility around the end of 2008 and beginning of 2009.

The three-state model estimation results, as well as the expected first passage results, suggests an almost hierarchical set of states; to move up the hierarchy the process needs to pass from "good" to "higher risk" before entering the "risky" state. The estimated probabilities and associated credibility intervals for the Markov chain also lend support to this: there is a very low probability (with high confidence) that the process moves from the "good" state to the "risky" state. It is instead more likely for it to shift into the "higher risk" state. This hierarchy appears to be slightly less rigid when moving out of the "risky" state, where the probability of shifting directly into the "good" state is roughly similar as in the converse, but with a wider credibility interval (0.0251 on the upper range). The parameters associated with these states were also consistent with the hierarchy described in Baba and Sakurai (2011), with “tranquil”, “turmoil” and “crisis” regimes. State 2 is a candidate for the tranquil regime, with a negative mean, low scale parameter value and high probability of remaining in the state. State 1 is a candidate for the turmoil regime; while the estimated mean is negative, this is with low confidence and it has a higher error scale parameter. However, it has the largest estimated degrees of freedom parameter, which is

TABLE 3.6: Expected First Passage Times (Days) for the Three-State Model (95% credibility (days) in parentheses)

	Arriving in		
	1	**2**	**3**
Starting in			
1		31 (22, 44)	164 (87, 314)
2	31 (21, 47)		182 (100, 336)
3	18 (10, 29)	42 (29, 56)	

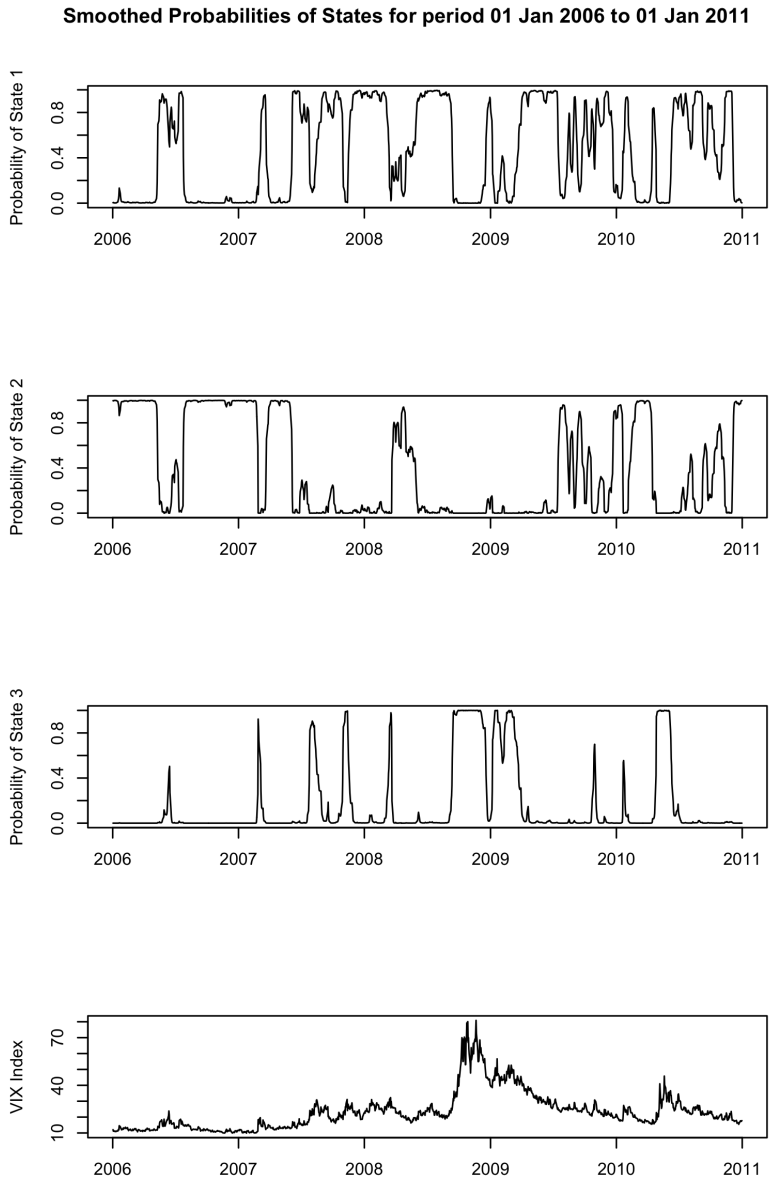

FIGURE 3.6: Smoothed probabilities of each state (rows 1, 2 & 3), as well as the VIX Index level (row 4), for each observation in the period 01 January 2006 to 01 January 2011

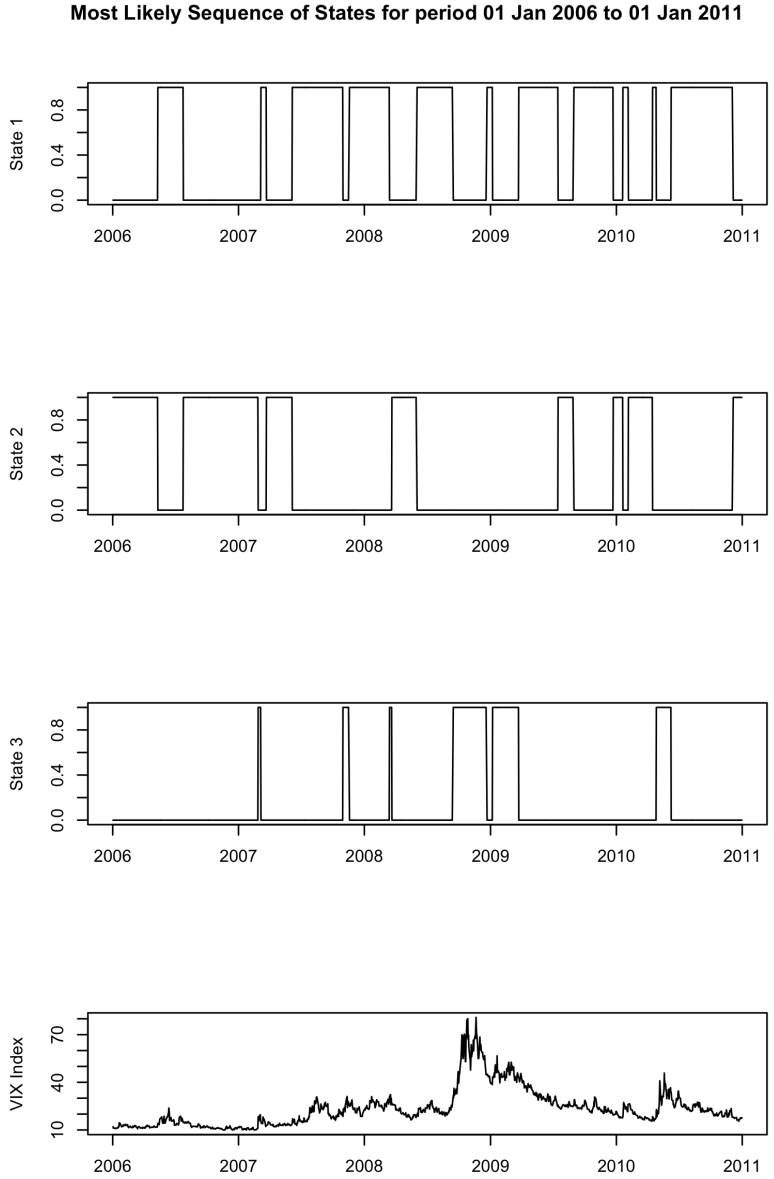

FIGURE 3.7: Most likely sequence of states (rows 1, 2 & 3), as well as the VIX Index level (row 4), for the observation period 01 January 2006 to 01 January 2011

a counter-intuitive result as one would expect the turmoil regime to have heavier tails than the tranquil regime. State 3 is a good candidate for a crisis regime – high error scale parameter, positive mean (with low confidence) and the lowest degrees of freedom parameter, meaning the heaviest tails.

The results suggest that global risk appetite, as proxied for by the VIX Index, can and does shift between different states of behaviour. This also suggests that linear models of the risk appetite are not adequately capturing the full dynamics of the process. Particularly noteworthy is the marked difference in scale parameters and degrees of freedom of the errors across the states, as well as the negative mean in tranquil regime (State 2). Considering the evidence provided for three states inherent in the risk appetite process, Chapter 4 repeats the analysis of Chapter 2, using the model-derived state classifications. It is left up to future research to elucidate the gain, if any, of considering four-state and higher models.

Chapter 4

Regime-Switching Model Validation

4.1 Introduction

This chapter provides a brief validation of the quality of states identified by the Markov-switching model of Chapter 3, by performing the analysis of Chapter 2, substituting the heuristic-based state identification methodology with the model-based approach of Chapter 3.

Similar results to those of Chapter 2 were found, namely, the broadly statistically significant interaction between the principal components and the risk states across currencies. Furthermore, the finding of increasing average R^2 for the Emerging/-Commodity currencies from the "low" to the "high" risk state, and the declining average R^2 from the "low" to the "high" risk state for Developed/European currencies, remained.

The chapter proceeds by briefly detailing the data and methodology adopted, noting any significant changes to the approach of Chapter 2, presenting the results and discussing them.

4.2 Data & Methodology

Daily returns for the 29 currencies of the analysis, as well as daily changes in the VIX Index, were used for the period 3 January 2000 until 13 April 2017 ($n = 4509$). This time period was selected to match the period used in the analysis of Chapter 3. As in chapters 2 and 3, all data were collected from Bloomberg (2017).

A PCA was run on the full panel of currencies and the first three components selected as in Chapter 2. This selection is motivated by the scree plot (Figure A.3 in Appendix A) which suggests that three principal components are appropriate. The proportion of total variance explained by these first five components is shown in Table 4.1. The loadings of the currencies on the first three principal components is shown in Table A.5 in Appendix A.

Currencies were grouped according to the clusters identified in Chapter 2, shown in Table 4.2.

To obtain the risk states, the most likely sequence of states was calculated using the Viterbi (Viterbi, 1967) algorithm. This approach to obtaining the model risk states

TABLE 4.1: Proportion of total variance explained by the first five principal components

Principal Component	Total Variance Explained
1	40.05%
2	8.53%
3	5.35%
4	3.58%
5	3.50%

was selected as it represents the most likely *sequence* of states given the data and the model chosen, rather than simply using, per observation in the sample, the state with the highest smoothed probability. The states identified in Chapter 3 were in no particular order, but for ease of comparison in this chapter, they have been mapped to those of Chapter 2 as follows: the "tranquil" regime (identified as State 2 in Chapter 3) is State 1 or the "low" risk state, the "turmoil" regime (identified as State 1 in Chapter 3) is State 2, or the "normal" risk state and the "crisis" regime (identified as State 3 in Chapter 3) is State 3, or the "high" risk state. 2029, 2145 and 335 observations were classified as State 1, 2 and 3 respectively. Figure 4.1 shows a time series plot of the VIX Index, shaded to indicate the state identified by the model. As noted in Chapter 3, the model identifies the financial crisis period around the end of 2008 and early 2009 as the "high risk" or "crisis" regime, which intuitively fits.

To test the hypothesis of no interaction between principal components and risk state for all currencies, the partial F-test was conducted, with the full model specified by Equation 4.1, and the restricted model specified by Equation 4.2 below.

$$\begin{aligned} Y_i &= \alpha + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_4 S_2 + \beta_5 S_3 \\ &+ \beta_6 PC_1 * S_2 + \beta_7 PC_2 * S_2 + \beta_8 PC_3 * S_2 + \beta_9 PC_1 * S_3 + \beta_{10} PC_2 * S_3 + \beta_{11} PC_3 * S_3 \end{aligned} \tag{4.1}$$

where:

Y_i = the returns of the ith currency in the sample
PC_1 = the first principal component
PC_2 = the second principal component
PC_3 = the third principal component
S_2 = the indicator/dummy variable denoting risk state 2 ("Normal")
S_3 = the indicator/dummy variable denoting risk state 3 ("High")

and

$$Y_i = \alpha + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_4 S_2 + \beta_5 S_3. \tag{4.2}$$

The null hypothesis for this test is as follows:

$$\beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = \beta_{11} = 0.$$

TABLE 4.2: Clusters and their constituent currencies

Cluster	**Currency**
Developed/European Currencies	ARS, CNY, EUR, JPY, SAR, GBP, SEK, HRK, CZK, BGN, DKK, RON, NOK, CHF, BAM
Emerging/Commodity Currencies	AUD, BRL, CAD, MXN, RUB, ZAR, TRY, HUF, PLN, NZD
Asian Currencies	INR, IDR, KRW, SGD

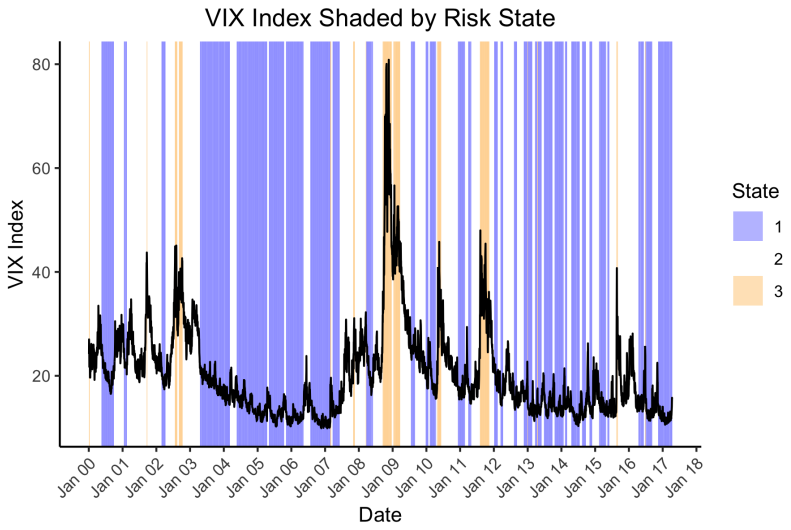

FIGURE 4.1: Time series plot of the VIX Index, shaded according to the state identified by the three-state, mean-only, t-distributed error and all parameter switching Markov-switching model of Chapter 3

The model specified by Equation 4.3 below was estimated to assess the impact of the carry trade factor on currency returns. The average R^2 of the models estimated for the currencies in the three clusters noted above, was calculated.

$$Y_i = \alpha + \beta_1 PC_2. \tag{4.3}$$

4.3 Results

4.3.1 Hypothesis Test of the Interaction Between Risk State and Principal Component

The p-values of the partial F-test of the null hypothesis are presented in Table 4.3. The findings are broadly statistically significant as in Chapter 2, with the exception of the Argentine Peso (ARS), the Croatian Kuna (HRK), the Norwegian Krone (NOK) and the Canadian Dollar (CAD), where in this analysis, the p-values for these currencies were considerably higher than in the analysis of Chapter 2. Furthermore, the results are different to the findings of Chapter 2, in that the null hypothesis was rejected for the Russian Rouble (RUB). These findings suggest that subtly different dynamics are being captured by this model (as compared to the heuristic of Chapter 2), or simply that natural statistical variability has resulted in such a change. An assessment of the uncertainty around the sequence of risk states could provide a better insight into the robustness of this finding.

4.3.2 Carry Trade Factor-Only Regressions

The results of the regressions of the currencies against the carry trade factor (second principal component) are displayed in Table A.6 in Appendix A. The average model R^2 across currencies within each cluster and risk state are visualised in Figure 4.2. One can clearly see a similar result to that of Chapter 2, namely the decline in average R^2 for the Developed/European currencies cluster as one moves from the "low" to "high" risk state, and the sharp contrast for the Emerging/Commodity currencies cluster. The table of average R^2 values is also provided in Table 4.4.

4.3.3 Discussion

The results of the partial F-test provide evidence that the finding of differing regression coefficients for the principal components across risk states, noted in Chapter 2, broadly persists. This is with the exception of the Argentine Peso, the Croatian Kuna, the Norwegian Krone and the Canadian Dollar, where the null hypothesis of no interaction between principal component regressors and risk state fails to be rejected, and the same null hypothesis for the Russian Rouble is rejected. This suggests that the Markov-switching model is capturing similar states and dynamics to those identified heuristically in Chapter 2.

The results of the carry trade factor-only regressions also remain similar to those of Chapter 2, with the finding that the carry trade factor (the second principal component) captures less of the variation in currencies in the Developed/European cluster

TABLE 4.3: p-Values for the hypothesis test of principal component-state interactions

Cluster	Currency	p-value
Developed/European Currencies	ARS	**0.6399**
	CNY	0.0000
	EUR	0.0000
	JPY	0.0000
	SAR	0.0000
	GBP	0.0003
	SEK	0.0003
	HRK	**0.1080**
	CZK	0.0000
	BGN	0.0000
	DKK	0.0000
	RON	0.0000
	NOK	**0.3616**
	CHF	0.0000
	BAM	0.0000
Emerging/Commodity Currencies	AUD	0.0000
	BRL	0.0005
	CAD	**0.1940**
	MXN	0.0000
	RUB	0.0000
	ZAR	0.0000
	TRY	0.0000
	HUF	0.0000
	PLN	0.0000
	NZD	0.0051
Asian Currencies	INR	0.0000
	IDR	0.0000
	KRW	0.0000
	SGD	0.0000

TABLE 4.4: Average R^2 per cluster by risk state

	Model Fit (R^2) (Chapter 2 equivalent in brackets)		
Cluster	**State 1**	**State 2**	**State 3**
Developed/European Currencies	0.136 (0.149)	0.077 (0.085)	0.055 (0.040)
Emerging/Commodity Currencies	0.055 (0.055)	0.119 (0.077)	0.306 (0.249)
Asian Currencies	0.040 (0.013)	0.058 (0.030)	0.133 (0.078)

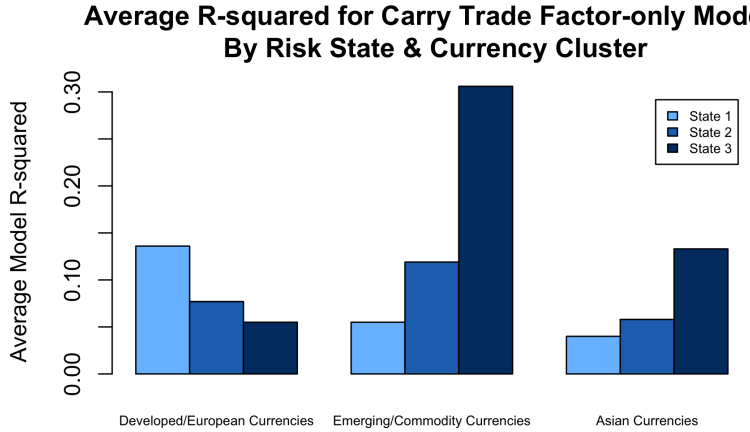

FIGURE 4.2: Average R^2 for the carry trade factor-only model, split by risk state and currency cluster

as one shifts from the "low" state to the "high" state. The opposite is observed for the currencies in the Asian cluster and quite sharply so for the currencies in the Emerging/Commodity cluster. Indeed, the states identified by the model of Chapter 3 tend to produce model fits that are superior to that of the heuristic from Chapter 2, as evidenced by Table 4.4. It is worth noting, however, that the sample underlying the regressions from Chapter 2 span a longer time period, suggesting the possibility of different findings when running this analysis on the full sample.

From a Bayesian perspective, the question of how robust this finding is in light of both model configuration & parameter uncertainty, has not been answered, and indeed is not within the scope of this project. A future area of research lies in quantifying and analysing the sensitivity of this result to different switching configurations and uncertainty in the most likely sequence of states inferred from the model. This could provide further evidence for, or against, the robustness of this finding and better aid practitioners in risk management decisions.

Chapter 5

Conclusion

The findings of Chapter 2 make a unique contribution to the literature by noting the significant joint interaction between the currency market factors (principal component regressors) and the risk state, suggesting the factors exert a different impact on currencies in different risk states of the world.

The findings also hint at the complex relationship, as noted by Polakow and Flint (2015), between the global carry trade, currencies and the "risk-on risk-off" signal. That is, when global markets are "risk-off" ("high" risk state), as proxied for by the VIX Index level, the carry trade drives variation in Emerging/Commodity currencies, while for Developed/European currencies it drives a smaller proportion. When markets are "risk-on" ("low" risk state) the relationship switches: the carry trade drives a smaller proportion of variation in Emerging/Commodity currencies (both compared to the Developed/European currencies cluster and relative to the "risk-off" state for the Emerging/Commodity currencies cluster) and a higher proportion of variation in Developed/European currencies (relative to the "risk-off" state).

A similar picture emerges for the unconditional means and standard deviations shown in Table 2.4 of Chapter 2: during the "risk-off" state, Emerging/Commodity cluster currencies display the highest standard deviation of returns and the lowest average return, while the opposite is seen for Developed/European currencies which have the highest average return and lowest standard deviation of returns.

One may ask whether these results are due to changes in interest rates stimulating carry trades. Lustig, Roussanov and Verdelhan (2011:3749) compared the covariation in returns between a carry portfolio (including the return due to the interest rate differential) and risk factors, and the covariation between the equivalent portfolio (comprising the spot currencies only) and risk factors. The authors concluded that it is the covariation between currencies and risk which drives the results of the regression. And so, the documented relationship between currencies and the carry trade factor is likely not due to variation in the interest rates of the currencies under consideration, but rather a direct relationship between these two variables.

Furthermore, in a scatterplot of observed returns against the carry trade factor, a majority of Emerging/Commodity cluster currencies span the origin. This finding suggests that the relationship between the carry trade and these currencies exists in both appreciating and depreciating directions. Furthermore, this finding holds even when considering the "High" risk state only (see Figure A.2 in Appendix A). That is, despite being in the "High" risk state, carry trade factor increases (decreases) are still on average associated with appreciations (depreciations) for emerging markets. This builds upon the literature documenting the effect of the global risk backdrop on

carry trade currencies with evidence that the high risk signal strengthens the impact of the carry trade on Emerging/Commodity currencies, but it doesn't drive or imply the direction of the change.

Chapter 3 contributes to the literature by modelling day-to-day changes in the VIX Index as a Markov-switching ARMA model (estimated in a frequentist framework using the approach described in Chen and Tsay (2007)) using t-distributed errors. However, when estimating a battery of models of differing class (mean-only, AR, ARMA), number of states, error distribution and switching schemes, and comparing on the basis of AIC & BIC and the pseudo-residuals, the mean-only class was selected. Further, the number of states in the model was selected as three, using the parallel sampling methodology described in Congdon (2006). This matches the number of states chosen in Baba and Sakurai (2011), where the model selection methodology was instead based on the Markov-switching criterion of Smith et al. (2006), and the states identified had similar characteristics and hierarchies, in the absence of an imposed constraint on the transition probabilities as in that study.

Chapter 4 presented a brief validation of the state identification model of Chapter 3, by reconsidering the analysis of Chapter 2, using the model-inferred states of Chapter 3 instead of heuristically identified states. The findings were encouraging, displaying similar results to those of Chapter 2, namely the statistically significant interaction between principal component regressors and risk state, as well as the tendency for the carry trade factor to explain more (less) variation in currencies as risk state increases for Emerging/Commodity currencies (Developed/European currencies). This suggests that the initial state identification method had some merit as a quick and easy approach. However, it is suggested that the approach of Chapter 3 is preferable, as it allows for the generation of forecasts for out-of-sample observations and provides a simple probabilistic model of the risk-state.

The results of this dissertation suggest that there is value in modelling the risk-appetite process (proxied for by the VIX Index) as a three-state, mean-only, Markov-switching model with t-distributed errors and switching in all parameters. The regimes identified provided a useful lens through which to analyse and begin to tease out the complex relationship between the carry trade factor (and other factors) and global currency market returns. Furthermore, the model used to describe the risk-appetite process could be used to generate forecasts of regimes, potentially aiding in risk management.

Future research could build on the findings and results of this dissertation by exploring and examining further the finding across most currencies of a statistically significant joint interaction between the currency market factors (principal components). Furthermore, quantifying the robustness of the general finding of carry trade sensitivity given risk state, across different clusters of currencies, could be achieved in a Bayesian framework by accounting for differing switching configurations, as well as uncertainty in the sequence of most likely states inferred from the model of risk appetite.

Appendix A

Tables & Figures

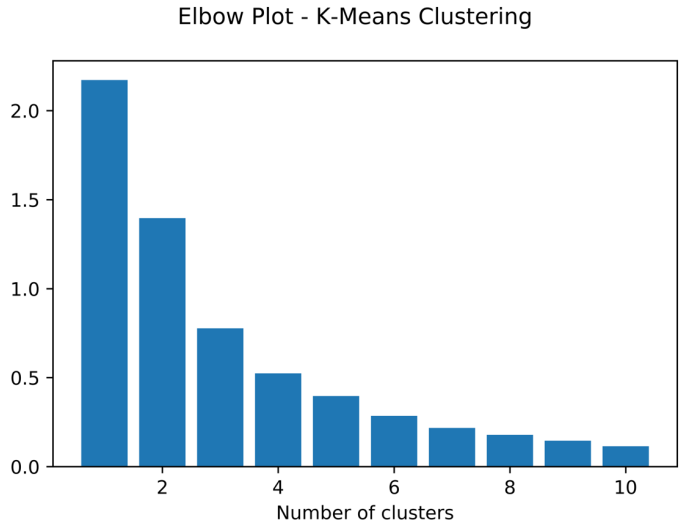

FIGURE A.1: Elbow plot for the k-means clustering analysis

TABLE A.1: Currency clusters found when $k = 4$

Cluster	Currency
1	AUD, BRL, CAD, MXN, ZAR, TRY, NZD
2	EUR, JPY, GBP, SEK, DKK, NOK, CHF
3	INR, IDR, KRW, SGD
4	ARS, CNY, RUB, SAR, HRK, CZK, HUF, PLN, BGN, RON, BAM

TABLE A.2: p-Values for the partial F-test – hypothesis of no interaction between risk state and principal component regressors

Cluster	**Currency**	**p-value**
Developed/European Currencies	ARS	0.0985
	CNY	0.0000
	EUR	0.0000
	JPY	0.0000
	SAR	0.0005
	GBP	0.0000
	SEK	0.0000
	HRK	0.0458
	CZK	0.0000
	BGN	0.0000
	DKK	0.0000
	RON	0.0001
	NOK	0.0000
	CHF	0.0000
	BAM	0.0000
Emerging/Commodity Currencies	AUD	0.0000
	BRL	0.0012
	CAD	0.0084
	MXN	0.0000
	RUB	**0.2004**
	ZAR	0.0000
	TRY	0.0706
	HUF	0.0000
	PLN	0.0605
	NZD	0.0000
Asian Currencies	INR	0.0000
	IDR	0.0000
	KRW	0.0000
	SGD	0.0000

TABLE A.3: Results from stratified regressions of Chapter 2 of the full model for each currency

Cluster	Currency	State	Model R^2	$n =$	Intercept	PC1	PC2	PC3
Developed / European Currencies	ARS	State 1	0.00	1070	-0.0001 (p=0.0801)	0.0001 (p=0.1409)	-0.0001 (p=0.0780)	0.0000 (p=0.6555)
		State 2	0.01	4541	-0.0005 (p=0.0004)	0.0001 (p=0.0055)	-0.0004 (p=0.0001)	0.0002 (p=0.1593)
		State 3	0.00	983	-0.0002 (p=0.2223)	0.0000 (p=0.6439)	-0.0001 (p=0.2253)	0.0001 (p=0.4974)
	CNY	State 1	0.06	1070	-0.0004 (p=0.2237)	0.0002 (p=0.1666)	-0.0009 (p=0.0011)	0.0035 (p=0.0000)
		State 2	0.01	4980	-0.0000 (p=0.3418)	0.0000 (p=0.0000)	-0.0001 (p=0.0112)	0.0001 (p=0.0000)
		State 3	0.01	1068	0.0000 (p=0.0294)	0.0000 (p=0.0119)	-0.0000 (p=0.3663)	0.0000 (p=0.0747)
	EUR	State 1	0.74	1070	-0.0000 (p=0.7505)	0.0018 (p=0.0000)	0.0013 (p=0.0000)	-0.0009 (p=0.0000)
		State 2	0.84	4980	0.0000 (p=0.6944)	0.0019 (p=0.0000)	0.0015 (p=0.0000)	-0.0003 (p=0.0000)
		State 3	0.89	1068	-0.0000 (p=0.8990)	0.0018 (p=0.0000)	0.0015 (p=0.0000)	-0.0000 (p=0.6658)
	JPY	State 1	0.57	1070	-0.0003 (p=0.0346)	0.0012 (p=0.0000)	0.0011 (p=0.0000)	0.0038 (p=0.0000)
		State 2	0.40	4980	0.0000 (p=0.6010)	0.0007 (p=0.0000)	0.0021 (p=0.0000)	0.0021 (p=0.0000)
		State 3	0.43	1068	0.0003 (p=0.2241)	0.0006 (p=0.0000)	0.0026 (p=0.0000)	0.0016 (p=0.0000)
	SAR	State 1	0.03	1070	-0.0000 (p=0.6283)	-0.0000 (p=0.5965)	-0.0000 (p=0.0813)	0.0001 (p=0.0000)

	State 2	0.02	4980	-0.0000 (p=0.7512)	0.0000 (p=0.8422)	0.0000 (p=0.5275)	0.0000 (p=0.0000)
	State 3	0.05	1068	0.0000 (p=0.6245)	-0.0000 (p=0.1118)	0.0000 (p=0.0016)	0.0000 (p=0.0000)
	State 1	0.58	1070	-0.0001 (p=0.6045)	0.0016 (p=0.0000)	0.0013 (p=0.0000)	0.0006 (p=0.0006)
GBP	State 2	0.47	4980	-0.0000 (p=0.9123)	0.0014 (p=0.0000)	0.0008 (p=0.0000)	0.0002 (p=0.0000)
	State 3	0.49	1068	0.0000 (p=0.9488)	0.0013 (p=0.0000)	0.0004 (p=0.0000)	0.0003 (p=0.0034)
	State 1	0.61	1070	0.0001 (p=0.3675)	0.0020 (p=0.0000)	0.0015 (p=0.0000)	-0.0004 (p=0.0244)
SEK	State 2	0.71	4980	0.0000 (p=0.7986)	0.0020 (p=0.0000)	0.0012 (p=0.0000)	-0.0001 (p=0.0159)
	State 3	0.75	1068	-0.0002 (p=0.1923)	0.0022 (p=0.0000)	0.0007 (p=0.0000)	-0.0002 (p=0.0214)
	State 1	0.79	632	0.0001 (p=0.6138)	0.0016 (p=0.0000)	0.0013 (p=0.0000)	-0.0005 (p=0.0003)
HRK	State 2	0.58	3763	-0.0001 (p=0.2334)	0.0016 (p=0.0000)	0.0008 (p=0.0000)	-0.0005 (p=0.0000)
	State 3	0.48	983	0.0001 (p=0.7195)	0.0015 (p=0.0000)	0.0008 (p=0.0000)	-0.0004 (p=0.0008)
	State 1	0.52	1000	0.0000 (p=0.7756)	0.0017 (p=0.0000)	0.0001 (p=0.5610)	-0.0013 (p=0.0000)
CZK	State 2	0.63	4238	0.0001 (p=0.2731)	0.0020 (p=0.0000)	0.0007 (p=0.0000)	-0.0005 (p=0.0000)
	State 3	0.72	983	0.0000 (p=0.9069)	0.0021 (p=0.0000)	0.0007 (p=0.0000)	-0.0006 (p=0.0000)
	State 1	0.36	1000	-0.0005 (p=0.0005)	0.0014 (p=0.0000)	-0.0000 (p=0.7897)	-0.0009 (p=0.0000)
BGN							

	State 2	0.15	4237	-0.0008 (p=0.0000)	0.0018 (p=0.0000)	0.0009 (p=0.0000)	-0.0010 (p=0.0000)
	State 3	0.49	983	0.0001 (p=0.7076)	0.0014 (p=0.0000)	0.0010 (p=0.0000)	0.0001 (p=0.5825)
	State 1	0.84	1070	0.0001 (p=0.4888)	0.0019 (p=0.0000)	0.0019 (p=0.0000)	-0.0005 (p=0.0000)
DKK	State 2	0.89	4980	0.0000 (p=0.4717)	0.0019 (p=0.0000)	0.0018 (p=0.0000)	-0.0001 (p=0.0000)
	State 3	0.90	1068	-0.0000 (p=0.9879)	0.0018 (p=0.0000)	0.0016 (p=0.0000)	-0.0001 (p=0.2373)
	State 1	0.13	1070	-0.0010 (p=0.0025)	0.0018 (p=0.0000)	-0.0003 (p=0.2346)	-0.0016 (p=0.0007)
RON	State 2	0.07	4980	-0.0011 (p=0.0000)	0.0017 (p=0.0000)	-0.0004 (p=0.0657)	-0.0019 (p=0.0000)
	State 3	0.31	1068	-0.0006 (p=0.0236)	0.0014 (p=0.0000)	0.0002 (p=0.2441)	-0.0004 (p=0.0085)
	State 1	0.71	1070	0.0001 (p=0.2234)	0.0020 (p=0.0000)	0.0015 (p=0.0000)	-0.0001 (p=0.5680)
NOK	State 2	0.70	4980	0.0000 (p=0.4200)	0.0021 (p=0.0000)	0.0011 (p=0.0000)	-0.0000 (p=0.4069)
	State 3	0.71	1068	-0.0003 (p=0.0877)	0.0020 (p=0.0000)	0.0006 (p=0.0000)	-0.0001 (p=0.5771)
	State 1	0.82	1070	0.0001 (p=0.4493)	0.0019 (p=0.0000)	0.0022 (p=0.0000)	0.0001 (p=0.4276)
CHF	State 2	0.70	4980	0.0001 (p=0.0729)	0.0017 (p=0.0000)	0.0027 (p=0.0000)	0.0005 (p=0.0000)
	State 3	0.69	1068	0.0001 (p=0.6211)	0.0017 (p=0.0000)	0.0024 (p=0.0000)	0.0004 (p=0.0000)
	State 1	0.79	632	0.0000 (p=0.6834)	0.0015 (p=0.0000)	0.0012 (p=0.0000)	-0.0003 (p=0.0150)
BAM							

		State 2	0.61	2986	-0.0001 (p=0.5067)	0.0015 (p=0.0000)	0.0010 (p=0.0000)	-0.0006 (p=0.0000)
		State 3	0.38	768	0.0002 (p=0.4729)	0.0013 (p=0.0000)	0.0009 (p=0.0000)	0.0004 (p=0.0257)
Emerging / Commodity Currencies	AUD	State 1	0.45	1070	0.0001 (p=0.3185)	0.0015 (p=0.0000)	-0.0019 (p=0.0000)	0.0010 (p=0.0000)
		State 2	0.55	4980	-0.0000 (p=0.5880)	0.0016 (p=0.0000)	-0.0020 (p=0.0000)	0.0005 (p=0.0000)
		State 3	0.71	1068	0.0003 (p=0.1910)	0.0018 (p=0.0000)	-0.0023 (p=0.0000)	0.0001 (p=0.3214)
	BRL	State 1	0.18	1070	-0.0025 (p=0.0000)	0.0014 (p=0.0000)	-0.0025 (p=0.0000)	-0.0003 (p=0.4486)
		State 2	0.19	4533	-0.0012 (p=0.0000)	0.0013 (p=0.0000)	-0.0032 (p=0.0000)	-0.0009 (p=0.0000)
		State 3	0.34	983	-0.0007 (p=0.0978)	0.0010 (p=0.0000)	-0.0033 (p=0.0000)	-0.0007 (p=0.0022)
	CAD	State 1	0.35	1070	0.0001 (p=0.5271)	0.0010 (p=0.0000)	-0.0013 (p=0.0000)	-0.0001 (p=0.4369)
		State 2	0.46	4980	0.0000 (p=0.9894)	0.0010 (p=0.0000)	-0.0015 (p=0.0000)	-0.0004 (p=0.0000)
		State 3	0.67	1068	-0.0001 (p=0.3923)	0.0011 (p=0.0000)	-0.0014 (p=0.0000)	-0.0003 (p=0.0002)
	MXN	State 1	0.36	1070	-0.0003 (p=0.3471)	0.0016 (p=0.0000)	-0.0057 (p=0.0000)	-0.0021 (p=0.0000)
		State 2	0.32	4980	-0.0002 (p=0.0039)	0.0009 (p=0.0000)	-0.0024 (p=0.0000)	-0.0007 (p=0.0000)
		State 3	0.54	1068	-0.0003 (p=0.1789)	0.0006 (p=0.0000)	-0.0028 (p=0.0000)	-0.0006 (p=0.0000)
	RUB	State 1	0.07	982	-0.0004 (p=0.0761)	0.0008 (p=0.0000)	-0.0010 (p=0.0000)	0.0001 (p=0.8089)

	State 2	0.14	4234	-0.0005 (p=0.0000)	0.0008 (p=0.0000)	-0.0016 (p=0.0000)	-0.0002 (p=0.0316)
	State 3	0.05	983	-0.0014 (p=0.0697)	0.0007 (p=0.0001)	-0.0013 (p=0.0003)	-0.0005 (p=0.2388)
	State 1	0.45	1070	-0.0002 (p=0.3706)	0.0022 (p=0.0000)	-0.0017 (p=0.0000)	-0.0004 (p=0.1113)
ZAR	State 2	0.42	4980	-0.0003 (p=0.0051)	0.0019 (p=0.0000)	-0.0025 (p=0.0000)	-0.0008 (p=0.0000)
	State 3	0.59	1068	0.0001 (p=0.8323)	0.0016 (p=0.0000)	-0.0026 (p=0.0000)	-0.0009 (p=0.0000)
	State 1	0.19	1070	-0.0009 (p=0.0001)	0.0015 (p=0.0000)	-0.0020 (p=0.0000)	-0.0013 (p=0.0003)
TRY	State 2	0.19	4980	-0.0010 (p=0.0000)	0.0014 (p=0.0000)	-0.0023 (p=0.0000)	-0.0015 (p=0.0000)
	State 3	0.26	1068	-0.0009 (p=0.0406)	0.0011 (p=0.0000)	-0.0023 (p=0.0000)	-0.0015 (p=0.0000)
	State 1	0.41	996	-0.0003 (p=0.0451)	0.0020 (p=0.0000)	-0.0004 (p=0.0142)	-0.0018 (p=0.0000)
HUF	State 2	0.66	4236	-0.0002 (p=0.0139)	0.0022 (p=0.0000)	-0.0004 (p=0.0000)	-0.0011 (p=0.0000)
	State 3	0.75	983	-0.0001 (p=0.4971)	0.0023 (p=0.0000)	-0.0000 (p=0.9268)	-0.0008 (p=0.0000)
	State 1	0.40	995	-0.0002 (p=0.2230)	0.0020 (p=0.0000)	-0.0007 (p=0.0000)	-0.0015 (p=0.0000)
PLN	State 2	0.57	4236	-0.0001 (p=0.2205)	0.0019 (p=0.0000)	-0.0008 (p=0.0000)	-0.0009 (p=0.0000)
	State 3	0.66	983	-0.0003 (p=0.2737)	0.0021 (p=0.0000)	-0.0008 (p=0.0000)	-0.0010 (p=0.0000)
	State 1	0.44	1070	0.0000 (p=0.9954)	0.0016 (p=0.0000)	-0.0014 (p=0.0000)	0.0012 (p=0.0000)
NZD							

		State 2	0.53	4980	0.0000 (p=0.6096)	0.0017 (p=0.0000)	-0.0019 (p=0.0000)	0.0006 (p=0.0000)
		State 3	0.65	1068	0.0001 (p=0.5004)	0.0017 (p=0.0000)	-0.0017 (p=0.0000)	0.0002 (p=0.1657)
Asian Currencies	INR	State 1	0.21	1070	-0.0001 (p=0.5021)	0.0001 (p=0.0005)	-0.0007 (p=0.0000)	0.0021 (p=0.0000)
		State 2	0.14	4980	-0.0002 (p=0.0031)	0.0003 (p=0.0000)	-0.0007 (p=0.0000)	0.0011 (p=0.0000)
		State 3	0.24	1068	-0.0004 (p=0.0009)	0.0003 (p=0.0000)	-0.0005 (p=0.0000)	0.0007 (p=0.0000)
	IDR	State 1	0.16	1070	-0.0001 (p=0.4075)	0.0002 (p=0.0000)	-0.0005 (p=0.0000)	0.0016 (p=0.0000)
		State 2	0.52	4584	-0.0003 (p=0.0150)	0.0004 (p=0.0000)	-0.0012 (p=0.0000)	0.0088 (p=0.0000)
		State 3	0.42	983	0.0004 (p=0.3298)	0.0005 (p=0.0000)	-0.0009 (p=0.0000)	0.0062 (p=0.0000)
	KRW	State 1	0.39	1070	0.0001 (p=0.4392)	0.0003 (p=0.0000)	-0.0010 (p=0.0000)	0.0034 (p=0.0000)
		State 2	0.38	4980	0.0000 (p=0.5379)	0.0004 (p=0.0000)	-0.0009 (p=0.0000)	0.0033 (p=0.0000)
		State 3	0.64	1068	-0.0003 (p=0.3606)	0.0005 (p=0.0000)	-0.0017 (p=0.0000)	0.0069 (p=0.0000)
	SGD	State 1	0.59	1070	0.0000 (p=0.6871)	0.0007 (p=0.0000)	-0.0004 (p=0.0000)	0.0019 (p=0.0000)
		State 2	0.56	4980	0.0001 (p=0.1572)	0.0007 (p=0.0000)	-0.0005 (p=0.0000)	0.0014 (p=0.0000)
		State 3	0.55	1068	0.0000 (p=0.9014)	0.0007 (p=0.0000)	-0.0003 (p=0.0000)	0.0008 (p=0.0000)

TABLE A.4: Results from stratified regressions of Chapter 2 of the carry trade factor-only model for each currency

Cluster	**Currency**	**State**	**Model** R^2	$n =$	**Intercept**	**PC2**
Developed / European Currencies	ARS	State 1	0.00	1070	-0.0001 (p=0.0990)	-0.0001 (p=0.1862)
		State 2	0.00	4541	-0.0005 (p=0.0004)	-0.0004 (p=0.0001)
		State 3	0.00	983	-0.0002 (p=0.2165)	-0.0001 (p=0.1459)
	CNY	State 1	0.00	1070	-0.0002 (p=0.4748)	-0.0001 (p=0.7164)
		State 2	0.00	4980	-0.0000 (p=0.3477)	-0.0000 (p=0.0268)
		State 3	0.00	1068	0.0000 (p=0.0321)	-0.0000 (p=0.0738)
	EUR	State 1	0.25	1070	0.0002 (p=0.2900)	0.0021 (p=0.0000)
		State 2	0.18	4980	0.0001 (p=0.5571)	0.0020 (p=0.0000)
		State 3	0.01	1068	-0.0000 (p=0.9094)	0.0004 (p=0.0007)
	JPY	State 1	0.24	1070	0.0000 (p=0.9768)	0.0024 (p=0.0000)
		State 2	0.19	4980	0.0000 (p=0.6337)	0.0022 (p=0.0000)
		State 3	0.29	1068	0.0002 (p=0.3605)	0.0022 (p=0.0000)
	SAR	State 1	0.00	1070	-0.0000 (p=0.8236)	-0.0000 (p=0.9138)
		State 2	0.00	4980	-0.0000 (p=0.7249)	0.0000 (p=0.8635)
		State 3	0.01	1068	0.0000 (p=0.6816)	0.0000 (p=0.0002)
	GBP	State 1	0.24	1070	0.0001 (p=0.3516)	0.0022 (p=0.0000)
		State 2	0.07	4980	0.0000 (p=0.8492)	0.0012 (p=0.0000)
		State 3	0.01	1068	-0.0000 (p=0.9728)	-0.0003 (p=0.0012)
	SEK	State 1	0.21	1070	0.0003 (p=0.0667)	0.0025 (p=0.0000)
		State 2	0.11	4980	0.0001 (p=0.6074)	0.0018 (p=0.0000)
		State 3	0.02	1068	-0.0002 (p=0.4784)	-0.0006 (p=0.0000)
	HRK	State 1	0.10	632	0.0005 (p=0.0108)	0.0017 (p=0.0000)
		State 2	0.03	3763	-0.0000 (p=0.9562)	0.0009 (p=0.0000)

	State 3	0.00	983	0.0001 (p=0.8616)	-0.0001 (p=0.2749)
	State 1	0.02	1000	0.0002 (p=0.1700)	0.0006 (p=0.0000)
CZK	State 2	0.03	4238	0.0001 (p=0.2197)	0.0010 (p=0.0000)
	State 3	0.02	983	-0.0000 (p=0.9717)	-0.0007 (p=0.0000)
	State 1	0.01	1000	-0.0003 (p=0.0668)	0.0004 (p=0.0009)
BGN	State 2	0.01	4237	-0.0008 (p=0.0002)	0.0012 (p=0.0000)
	State 3	0.00	983	0.0000 (p=0.8962)	0.0000 (p=0.7117)
	State 1	0.38	1070	0.0003 (p=0.0731)	0.0028 (p=0.0000)
DKK	State 2	0.23	4980	0.0001 (p=0.4904)	0.0023 (p=0.0000)
	State 3	0.01	1068	-0.0000 (p=0.9409)	0.0004 (p=0.0001)
	State 1	0.00	1070	-0.0008 (p=0.0178)	0.0003 (p=0.2135)
RON	State 2	0.00	4980	-0.0011 (p=0.0001)	0.0002 (p=0.3852)
	State 3	0.02	1068	-0.0006 (p=0.0558)	-0.0007 (p=0.0000)
	State 1	0.27	1070	0.0004 (p=0.0311)	0.0026 (p=0.0000)
NOK	State 2	0.09	4980	0.0001 (p=0.4091)	0.0016 (p=0.0000)
	State 3	0.02	1068	-0.0003 (p=0.3214)	-0.0007 (p=0.0000)
	State 1	0.42	1070	0.0003 (p=0.0518)	0.0033 (p=0.0000)
CHF	State 2	0.30	4980	0.0001 (p=0.1378)	0.0031 (p=0.0000)
	State 3	0.12	1068	0.0001 (p=0.8449)	0.0014 (p=0.0000)
	State 1	0.09	632	0.0004 (p=0.0123)	0.0016 (p=0.0000)
BAM	State 2	0.03	2986	0.0001 (p=0.5732)	0.0008 (p=0.0000)
	State 3	0.00	768	0.0002 (p=0.5496)	-0.0001 (p=0.4252)
	State 1	0.06	1070	0.0003 (p=0.0336)	-0.0010 (p=0.0000)
AUD	State 2	0.10	4980	-0.0000 (p=0.9266)	-0.0015 (p=0.0000)
	State 3	0.37	1068	0.0002 (p=0.4152)	-0.0034 (p=0.0000)
	State 1	0.07	1070	-0.0023 (p=0.0000)	-0.0018 (p=0.0000)
BRL					

	State 2	0.10	4533	-0.0011 (p=0.0000)	-0.0029 (p=0.0000)
	State 3	0.27	983	-0.0007 (p=0.1136)	-0.0039 (p=0.0000)
	State 1	0.07	1070	0.0002 (p=0.1361)	-0.0008 (p=0.0000)
CAD	State 2	0.12	4980	0.0000 (p=0.7546)	-0.0012 (p=0.0000)
	State 3	0.35	1068	-0.0001 (p=0.5288)	-0.0020 (p=0.0000)
	State 1	0.28	1070	-0.0002 (p=0.5958)	-0.0052 (p=0.0000)
MXN	State 2	0.17	4980	-0.0002 (p=0.0164)	-0.0021 (p=0.0000)
	State 3	0.46	1068	-0.0003 (p=0.2225)	-0.0032 (p=0.0000)
	State 1	0.01	982	-0.0003 (p=0.2490)	-0.0006 (p=0.0006)
RUB	State 2	0.06	4234	-0.0005 (p=0.0001)	-0.0015 (p=0.0000)
	State 3	0.03	983	-0.0014 (p=0.0714)	-0.0018 (p=0.0000)
	State 1	0.01	1070	0.0001 (p=0.6280)	-0.0006 (p=0.0011)
ZAR	State 2	0.08	4980	-0.0002 (p=0.0523)	-0.0020 (p=0.0000)
	State 3	0.34	1068	0.0001 (p=0.8587)	-0.0036 (p=0.0000)
	State 1	0.04	1070	-0.0008 (p=0.0017)	-0.0014 (p=0.0000)
TRY	State 2	0.05	4980	-0.0010 (p=0.0000)	-0.0019 (p=0.0000)
	State 3	0.15	1068	-0.0009 (p=0.0618)	-0.0030 (p=0.0000)
	State 1	0.00	996	-0.0001 (p=0.6336)	0.0002 (p=0.2517)
HUF	State 2	0.00	4236	-0.0001 (p=0.4080)	-0.0001 (p=0.3910)
	State 3	0.08	983	-0.0002 (p=0.6608)	-0.0015 (p=0.0000)
	State 1	0.00	995	0.0000 (p=0.8570)	-0.0001 (p=0.5464)
PLN	State 2	0.01	4236	-0.0000 (p=0.7974)	-0.0005 (p=0.0000)
	State 3	0.15	983	-0.0003 (p=0.4537)	-0.0021 (p=0.0000)
	State 1	0.01	1070	0.0002 (p=0.1563)	-0.0004 (p=0.0070)
NZD	State 2	0.08	4980	0.0001 (p=0.5209)	-0.0014 (p=0.0000)
	State 3	0.29	1068	0.0001 (p=0.6848)	-0.0028 (p=0.0000)

Asian Currencies	INR	State 1	0.01	1070	0.0000 (p=0.7482)	-0.0003 (p=0.0004)
		State 2	0.04	4980	-0.0002 (p=0.0054)	-0.0007 (p=0.0000)
		State 3	0.11	1068	-0.0004 (p=0.0014)	-0.0007 (p=0.0000)
	IDR	State 1	0.00	1070	0.0000 (p=0.8831)	-0.0001 (p=0.1219)
		State 2	0.02	4584	-0.0004 (p=0.0761)	-0.0016 (p=0.0000)
		State 3	0.02	983	0.0002 (p=0.6973)	-0.0012 (p=0.0000)
	KRW	State 1	0.01	1070	0.0002 (p=0.0507)	-0.0002 (p=0.0108)
		State 2	0.04	4980	0.0000 (p=0.6903)	-0.0009 (p=0.0000)
		State 3	0.07	1068	-0.0004 (p=0.3716)	-0.0019 (p=0.0000)
	SGD	State 1	0.03	1070	0.0002 (p=0.0333)	0.0004 (p=0.0000)
		State 2	0.02	4980	0.0001 (p=0.2820)	-0.0004 (p=0.0000)
		State 3	0.11	1068	-0.0000 (p=0.9295)	-0.0007 (p=0.0000)

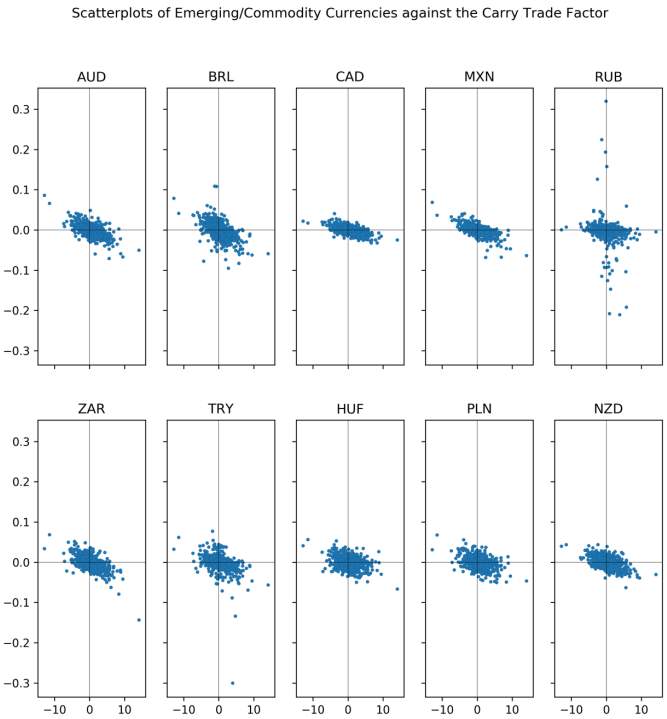

FIGURE A.2: Scatterplots of Emerging/Commodity currencies against the carry trade factor ("High" risk state of Chapter 2)

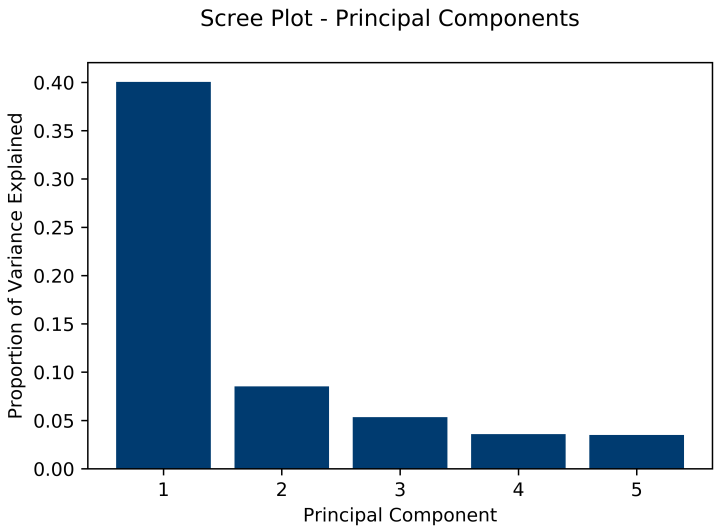

FIGURE A.3: Scree plot of the principal components (for the sub-sample considered in Chapter 4)

TABLE A.6: Results from stratified regressions of the carry trade factor-only model from Chapter 4, for each currency

Cluster	Currency	State	Model R^2	$n =$	Intercept	PC2
Developed / European Currencies	ARS	State 1	0.00	2029	-0.0004 (p=0.0090)	-0.0003 (p=0.0341)
		State 2	0.00	2145	-0.0007 (p=0.0152)	-0.0005 (p=0.0062)
		State 3	0.01	335	-0.0010 (p=0.0012)	-0.0001 (p=0.1243)
	CNY	State 1	0.01	2029	0.0000 (p=0.0351)	-0.0001 (p=0.0004)
		State 2	0.02	2145	0.0000 (p=0.2812)	-0.0001 (p=0.0000)
		State 3	0.01	335	0.0001 (p=0.0919)	-0.0000 (p=0.0273)
	EUR	State 1	0.25	2029	0.0004 (p=0.0007)	0.0024 (p=0.0000)
		State 2	0.16	2145	-0.0000 (p=0.7125)	0.0016 (p=0.0000)
		State 3	0.00	335	-0.0005 (p=0.3578)	-0.0000 (p=0.9011)
	JPY	State 1	0.10	2029	-0.0001 (p=0.6120)	0.0016 (p=0.0000)
		State 2	0.19	2145	0.0001 (p=0.6465)	0.0017 (p=0.0000)
		State 3	0.46	335	-0.0002 (p=0.6348)	0.0024 (p=0.0000)
	SAR	State 1	0.00	2029	0.0000 (p=0.8975)	-0.0000 (p=0.1584)
		State 2	0.00	2145	-0.0000 (p=0.8561)	0.0000 (p=0.2162)
		State 3	0.02	335	-0.0000 (p=0.9505)	0.0000 (p=0.0206)
	GBP	State 1	0.07	2029	0.0003 (p=0.0071)	0.0012 (p=0.0000)
		State 2	0.01	2145	-0.0001 (p=0.5153)	0.0003 (p=0.0000)
		State 3	0.06	335	-0.0010 (p=0.0676)	-0.0008 (p=0.0000)
	SEK	State 1	0.12	2029	0.0004 (p=0.0014)	0.0020 (p=0.0000)
		State 2	0.03	2145	-0.0001 (p=0.6711)	0.0009 (p=0.0000)
		State 3	0.06	335	-0.0007 (p=0.3249)	-0.0011 (p=0.0000)
	HRK	State 1	0.23	2029	0.0004 (p=0.0011)	0.0024 (p=0.0000)
		State 2	0.09	2145	-0.0000 (p=0.9681)	0.0013 (p=0.0000)

	State 3	0.00	335	-0.0007 (p=0.2971)	0.0001 (p=0.6908)
	State 1	0.17	2029	0.0005 (p=0.0001)	0.0023 (p=0.0000)
CZK	State 2	0.07	2145	0.0001 (p=0.6365)	0.0013 (p=0.0000)
	State 3	0.03	335	-0.0010 (p=0.1974)	-0.0008 (p=0.0011)
	State 1	0.23	2029	0.0003 (p=0.0018)	0.0023 (p=0.0000)
BGN	State 2	0.12	2145	-0.0000 (p=0.7217)	0.0014 (p=0.0000)
	State 3	0.00	335	-0.0005 (p=0.4025)	0.0002 (p=0.2528)
	State 1	0.25	2029	0.0004 (p=0.0008)	0.0024 (p=0.0000)
DKK	State 2	0.16	2145	-0.0000 (p=0.7247)	0.0016 (p=0.0000)
	State 3	0.00	335	-0.0005 (p=0.3681)	-0.0000 (p=0.9182)
	State 1	0.07	2029	0.0004 (p=0.0064)	0.0015 (p=0.0000)
RON	State 2	0.01	2145	-0.0004 (p=0.0024)	0.0004 (p=0.0000)
	State 3	0.02	335	-0.0010 (p=0.1502)	-0.0006 (p=0.0107)
	State 1	0.08	2029	0.0004 (p=0.0116)	0.0016 (p=0.0000)
NOK	State 2	0.01	2145	-0.0000 (p=0.8487)	0.0005 (p=0.0000)
	State 3	0.07	335	-0.0006 (p=0.3685)	-0.0012 (p=0.0000)
	State 1	0.27	2029	0.0003 (p=0.0028)	0.0027 (p=0.0000)
CHF	State 2	0.19	2145	0.0002 (p=0.2232)	0.0022 (p=0.0000)
	State 3	0.08	335	-0.0009 (p=0.1508)	0.0010 (p=0.0000)
	State 1	0.19	2002	0.0003 (p=0.0042)	0.0021 (p=0.0000)
BAM	State 2	0.09	2054	0.0000 (p=0.9104)	0.0013 (p=0.0000)
	State 3	0.00	330	-0.0006 (p=0.2942)	0.0002 (p=0.3556)
	State 1	0.00	2029	0.0002 (p=0.0868)	-0.0003 (p=0.0146)
AUD	State 2	0.08	2145	0.0002 (p=0.2596)	-0.0014 (p=0.0000)
	State 3	0.38	335	-0.0004 (p=0.5855)	-0.0035 (p=0.0000)
	State 1	0.11	2029	0.0002 (p=0.2153)	-0.0023 (p=0.0000)
BRL					

	State 2	0.21	2145	-0.0000 (p=0.8458)	-0.0030 (p=0.0000)
	State 3	0.42	335	-0.0010 (p=0.2698)	-0.0046 (p=0.0000)
	State 1	0.01	2029	0.0003 (p=0.0125)	-0.0004 (p=0.0000)
CAD	State 2	0.13	2145	-0.0000 (p=0.9854)	-0.0012 (p=0.0000)
	State 3	0.33	335	-0.0004 (p=0.4130)	-0.0020 (p=0.0000)
	State 1	0.19	2029	-0.0000 (p=0.8390)	-0.0021 (p=0.0000)
MXN	State 2	0.31	2145	-0.0001 (p=0.3276)	-0.0023 (p=0.0000)
	State 3	0.52	335	-0.0003 (p=0.5428)	-0.0034 (p=0.0000)
	State 1	0.05	2029	0.0002 (p=0.0778)	-0.0011 (p=0.0000)
RUB	State 2	0.13	2145	-0.0003 (p=0.0406)	-0.0020 (p=0.0000)
	State 3	0.13	335	-0.0008 (p=0.1303)	-0.0012 (p=0.0000)
	State 1	0.05	2029	0.0000 (p=0.8649)	-0.0019 (p=0.0000)
ZAR	State 2	0.16	2145	-0.0001 (p=0.6518)	-0.0026 (p=0.0000)
	State 3	0.41	335	-0.0003 (p=0.7129)	-0.0040 (p=0.0000)
	State 1	0.06	2029	-0.0000 (p=0.8450)	-0.0014 (p=0.0000)
TRY	State 2	0.11	2145	-0.0006 (p=0.0303)	-0.0028 (p=0.0000)
	State 3	0.34	335	-0.0004 (p=0.5222)	-0.0028 (p=0.0000)
	State 1	0.06	2029	0.0004 (p=0.0160)	0.0016 (p=0.0000)
HUF	State 2	0.00	2145	0.0001 (p=0.7360)	0.0004 (p=0.0037)
	State 3	0.09	335	-0.0014 (p=0.1174)	-0.0016 (p=0.0000)
	State 1	0.02	2029	0.0005 (p=0.0011)	0.0009 (p=0.0000)
PLN	State 2	0.00	2145	-0.0000 (p=0.9727)	-0.0002 (p=0.1233)
	State 3	0.12	335	-0.0013 (p=0.1515)	-0.0019 (p=0.0000)
	State 1	0.00	2029	0.0003 (p=0.0385)	-0.0002 (p=0.0808)
NZD	State 2	0.06	2145	0.0002 (p=0.3203)	-0.0013 (p=0.0000)
	State 3	0.32	335	-0.0006 (p=0.3459)	-0.0027 (p=0.0000)

Asian Currencies	INR	State 1	0.08	2029	0.0000 (p=0.7329)	-0.0009 (p=0.0000)
		State 2	0.09	2145	-0.0001 (p=0.0804)	-0.0007 (p=0.0000)
		State 3	0.19	335	-0.0004 (p=0.1833)	-0.0009 (p=0.0000)
	IDR	State 1	0.04	2029	-0.0001 (p=0.4933)	-0.0010 (p=0.0000)
		State 2	0.05	2145	-0.0001 (p=0.7376)	-0.0011 (p=0.0000)
		State 3	0.05	335	-0.0008 (p=0.1145)	-0.0007 (p=0.0000)
	KRW	State 1	0.04	2029	0.0002 (p=0.0207)	-0.0008 (p=0.0000)
		State 2	0.06	2145	0.0000 (p=0.9933)	-0.0009 (p=0.0000)
		State 3	0.14	335	-0.0005 (p=0.5325)	-0.0020 (p=0.0000)
	SGD	State 1	0.00	2029	0.0001 (p=0.0382)	-0.0001 (p=0.2525)
		State 2	0.03	2145	0.0001 (p=0.3505)	-0.0003 (p=0.0000)
		State 3	0.15	335	-0.0003 (p=0.2694)	-0.0007 (p=0.0000)

TABLE A.5: Loadings of each currency on the first three principal components found in Chapter 4

		Principal Component		
Cluster	**Currency**	**PC1**	**PC2**	**PC3**
Developed/European Currencies	ARS	0.010	-0.034	-0.032
	CNY	0.039	-0.073	-0.360
	EUR	0.265	0.208	0.018
	JPY	0.059	0.291	-0.224
	SAR	-0.014	0.030	-0.182
	GBP	0.204	0.030	-0.039
	SEK	0.251	0.074	0.039
	HRK	0.229	0.178	0.004
	CZK	0.253	0.116	0.078
	BGN	0.240	0.199	-0.039
	DKK	0.265	0.208	0.017
	RON	0.218	0.055	-0.011
	NOK	0.246	0.038	0.014
	CHF	0.193	0.263	-0.066
	BAM	0.205	0.170	-0.055
Emerging/Commodity Currencies	AUD	0.214	-0.213	0.031
	BRL	0.120	-0.311	0.153
	CAD	0.189	-0.223	0.095
	MXN	0.136	-0.366	0.193
	RUB	0.112	-0.208	-0.093
	ZAR	0.173	-0.257	0.140
	TRY	0.122	-0.228	0.116
	HUF	0.249	0.007	0.099
	PLN	0.238	-0.050	0.082
	NZD	0.205	-0.174	0.021
Asian Currencies	INR	0.080	-0.205	-0.394
	IDR	0.036	-0.141	-0.438
	KRW	0.061	-0.179	-0.501
	SGD	0.210	-0.116	-0.198

Appendix B

Joint Interaction Between Principal Components

In addition to the primary hypothesis of a joint interaction between the principal components and the risk state, the hypothesis of joint interaction between the principal components was tested, using the partial F-test to compare the full model of Equation B.1 below to the restricted model of Equation 2.1 (repeated here for convenience).

$$\begin{aligned} Y_i = \alpha + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_4 S_2 + \beta_5 S_3 \\ + \beta_6 PC_1 * S_2 + \beta_7 PC_2 * S_2 + \beta_8 PC_3 * S_2 + \beta_9 PC_1 * S_3 + \beta_{10} PC_2 * S_3 + \beta_{11} PC_3 * S_3 \\ + \beta_{12} PC_1 * PC_2 + \beta_{13} PC_1 * PC_3 + \beta_{14} PC_2 * PC_3. \end{aligned} \quad \text{(B.1)}$$

where:

Y_i = the returns of the i$^{\text{th}}$ currency in the sample
PC_1 = the first principal component
PC_2 = the second principal component
PC_3 = the third principal component
S_2 = the indicator/dummy variable denoting risk state 2 ("Normal")
S_3 = the indicator/dummy variable denoting risk state 3 ("High")

$$\begin{aligned} Y_i = \alpha + \beta_1 PC_1 + \beta_2 PC_2 + \beta_3 PC_3 + \beta_4 S_2 + \beta_5 S_3 \\ + \beta_6 PC_1 * S_2 + \beta_7 PC_2 * S_2 + \beta_8 PC_3 * S_2 + \beta_9 PC_1 * S_3 + \beta_{10} PC_2 * S_3 + \beta_{11} PC_3 * S_3. \end{aligned} \quad \text{(2.1)}$$

Formally, the null hypothesis is:

$$\beta_{12} = \beta_{13} = \beta_{14} = 0.$$

Table B.1 below displays the p-values of the hypothesis test detailed above, for each currency in the sample. The results indicate mostly significant results for currencies across all three clusters.

TABLE B.1: p-Values of partial F-test testing interactions between principal components of Chapter 2

Cluster	**Currency**	**p-value**
Developed/European Currencies	ARS	0.4845
	CNY	0.5620
	EUR	0.0000
	JPY	0.0000
	SAR	0.0000
	GBP	0.0061
	SEK	0.0010
	HRK	0.0010
	CZK	0.0618
	BGN	0.0000
	DKK	0.0000
	RON	0.0482
	NOK	0.0000
	CHF	0.0000
	BAM	0.0000
Emerging/Commodity Currencies	AUD	0.0000
	BRL	0.1385
	CAD	0.6004
	MXN	0.3166
	RUB	0.1883
	ZAR	0.0000
	TRY	0.0001
	HUF	0.1352
	PLN	0.0164
	NZD	0.0000
Asian Currencies	INR	0.0000
	IDR	0.7506
	KRW	0.0000
	SGD	0.0000

Appendix C

MS-ARMA Estimator Bias Calculation

To get an assessment of the bias of the full model presented in equation 3.1, simulations were run for three different models (models I, II & III), for sample sizes of $n =$ 100, 250, 500 and 1000. To obtain parameter estimates, three models were fit (with the likelihood function optimised using the *scipy* (Jones, Oliphant & Peterson, 2001) package's SLSQP algorithm) with random parameter initializations and the model (among those which converged) with the highest log likelihood was selected. An automated method needed to be employed to match the true state with the estimated state. A simple heuristic was used to match the states based on whichever state best matched the σ in the true parameter set as these were observed to be correctly disambiguated. In cases where the states could not be paired from the observed sigma, those simulations would be discarded, however, this did not occur in the experiment.

Each {model, sample size} pair was simulated 100 times, and the bias for a parameter calculated as the mean difference between the estimated and the true parameter. Hence, a negative bias indicates the estimator is underestimating the true value of the parameter, and vice versa. Standard deviations are reported, as well as m, the number of values which were used for the calculation. As noted above, almost all of the means used the full 100 simulations and it only occurred twice that a simulation needed to be discarded. The true parameters of the models simulated are shown in Table C.1. The results of the simulation are reported in tables C.2, C.3 & C.4.

The results of the simulations indicate a tendency for the absolute bias (and the standard deviation) of the transition probability estimates P to decline for each model as

TABLE C.1: True parameters for simulated models

Parameter	Model I		Model II		Model III	
	State 1	State 2	State 1	State 2	State 1	State 2
α_i	-1.00	0.50	0.20	-0.50	0.01	0.70
$\beta_{1,i}$	0.50	-0.50	0.10	-0.70	-0.10	-0.30
$\theta_{1,i}$	0.05	0.50	-0.15	0.05	0.60	0.15
σ_i	1.00	5.00	1.00	5.00	1.00	5.00
P_{1i}	0.90	0.10	0.70	0.3	0.6	0.40
P_{2i}	0.30	0.70	0.20	0.8	0.4	0.60

TABLE C.2: Bias for Model I simulated with various sample sizes. Standard deviation in parentheses

Parameter	n = 100 (m = 99)		n = 250 (m = 100)		n = 500 (m = 100)		n = 1000 (m = 100)	
	State 1	State 2	State 1	State 2	State 1	State 2	State 1	State 2
α_i	-0.15	0.05	-0.04	-0.26	-0.01	-0.05	-0.01	-0.08
	(0.43)	(1.73)	(0.19)	(0.88)	(0.12)	(0.62)	(0.08)	(0.48)
$\beta_{1,i}$	-0.09	0.19	-0.03	0.24	-0.01	0.35	-0.01	0.26
	(0.27)	(0.62)	(0.13)	(0.54)	(0.07)	(0.43)	(0.05)	(0.34)
$\theta_{1,i}$	0.08	-0.37	0.02	-0.33	-0.01	-0.41	-0.00	-0.31
	(0.32)	(0.67)	(0.15)	(0.55)	(0.09)	(0.44)	(0.06)	(0.34)
σ_i	-0.04	-0.58	-0.02	-0.23	-0.00	0.01	-0.00	0.03
	(0.11)	(0.99)	(0.07)	(0.50)	(0.05)	(0.35)	(0.03)	(0.27)
P_{1i}	-0.02	0.02	-0.01	0.01	-0.00	0.00	-0.00	0.00
	(0.06)	(0.06)	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
P_{2i}	0.09	-0.09	0.02	-0.02	0.02	-0.02	0.01	-0.01
	(0.18)	(0.18)	(0.09)	(0.09)	(0.06)	(0.06)	(0.04)	(0.04)

TABLE C.3: Bias for Model II simulated with various sample sizes. Standard deviation in parentheses

Parameter	n = 100 (m = 99)		n = 250 (m = 100)		n = 500 (m = 100)		n = 1000 (m = 100)	
	State 1	State 2	State 1	State 2	State 1	State 2	State 1	State 2
α_i	-0.09	-0.08	0.02	-0.01	0.02	-0.00	0.02	0.00
	(1.08)	(0.69)	(0.14)	(0.47)	(0.10)	(0.28)	(0.06)	(0.21)
$\beta_{1,i}$	-0.09	0.13	-0.05	-0.00	-0.04	-0.01	-0.05	0.00
	(0.33)	(0.33)	(0.09)	(0.17)	(0.05)	(0.12)	(0.03)	(0.08)
$\theta_{1,i}$	0.06	-0.19	0.08	-0.02	0.07	0.01	0.09	-0.01
	(0.34)	(0.39)	(0.09)	(0.19)	(0.06)	(0.16)	(0.04)	(0.10)
σ_i	-0.02	-0.19	-0.02	-0.10	0.00	-0.03	0.02	-0.03
	(0.43)	(0.54)	(0.14)	(0.32)	(0.09)	(0.24)	(0.08)	(0.15)
P_{1i}	-0.07	0.07	-0.01	0.01	-0.00	0.00	0.00	-0.00
	(0.20)	(0.20)	(0.07)	(0.07)	(0.05)	(0.05)	(0.03)	(0.03)
P_{2i}	0.05	-0.05	0.01	-0.01	0.01	-0.01	-0.00	0.00
	(0.15)	(0.15)	(0.06)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)

TABLE C.4: Bias for Model III simulated with various sample sizes. Standard deviation in parentheses

Parameter	n = 100 (m = 99)		n = 250 (m = 100)		n = 500 (m = 100)		n = 1000 (m = 100)	
	State 1	State 2	State 1	State 2	State 1	State 2	State 1	State 2
α_i	1.09	0.08	1.04	0.10	0.99	0.08	1.05	0.09
	(1.06)	(1.22)	(0.36)	(0.52)	(0.22)	(0.40)	(0.17)	(0.28)
$\beta_{1,i}$	-0.45	0.35	-0.45	0.40	-0.49	0.37	0.63	0.39
	(0.61)	(0.67)	(0.61)	(0.60)	(0.58)	(0.59)	(0.40)	(0.44)
$\theta_{1,i}$	0.32	-0.52	0.33	-0.55	0.38	-0.53	0.51	-0.53
	(0.63)	(0.67)	(0.59)	(0.61)	(0.57)	(0.58)	(0.40)	(0.43)
σ_i	0.07	-0.23	0.09	-0.22	0.16	-0.06	0.17	-0.08
	(0.34)	(0.66)	(0.20)	(0.38)	(0.12)	(0.30)	(0.10)	(0.19)
P_{1i}	-0.40	0.40	-0.39	0.39	-0.33	0.33	-0.33	0.33
	(0.24)	(0.24)	(0.16)	(0.16)	(0.10)	(0.10)	(0.07)	(0.07)
P_{2i}	0.13	-0.13	0.12	-0.12	0.10	-0.10	0.10	-0.10
	(0.15)	(0.15)	(0.12)	(0.12)	(0.08)	(0.08)	(0.05)	(0.05)

the sample size for the simulation is increased. This decline as simulation sample size increases is similarly observed in the standard deviations for all parameter estimates across models. For parameters other than the transition probabilities, there is also a tendency for the bias to decrease, however it was observed to increase slightly when moving from a sample size of 500 to 1000. The absolute bias for θ tends to be higher in State 2 compared to State 1, with the exception of the simulations with sample size 250, 500 and 1000 for Model II, and the simulation with sample size 1000 for Model III. For Model III, θ appears rather consistently large across the states and sample sizes.

The estimator typically gets the Markov chain transition probabilities correct in the higher sample sizes, except in the case of Model III for which the bias is consistently high. This could be due to the low persistence for each state in that particular configuration and could also be a reason for the poor bias on the other parameters. I suggest an intuition for this observation – the algorithm is attempting to disambiguate two states which have transition probabilities quite close to 0.5 and therefore it is difficult to obtain confidence that the process is in either of the regimes.

The results suggest that, for models with high persistence and a relatively large sample size ($n = 1000$), the bias is low enough that one could acceptably use these estimates.

Appendix D

Markov-Chain Monte Carlo Sampler Results

This Appendix presents the plots and results of the MCMC sampler run for the models in Chapter 3. Two chains were run, with the results of the first chain (first run) presented and used throughout Chapters 3 and 4, and the second chain (second run) is presented here. The Robust Adaptive Metropolis algorithm (Vihola, 2012) was used for the burn-in period, while the standard Metropolis algorithm was used after this. The point at which the burn-in ends (and the point at which the standard Metropolis algorithm is used) is indicated by a red line in the trace plots.

Figures D.1 through D.6 display the first run trace plots (of the unthinned sample) and sample ACFs (calculated from the thinned sample) for the linear, two-state and three-state models respectively.

Table D.1 displays the results of the second MCMC Samplers run for the linear, two-state and three-state models.

Figures D.7 through D.12 display the second run trace plots (of the full unthinned sample) and sample ACFs (calculated from the burned-in, thinned sample) for the linear, two-state and three-state models respectively.

TABLE D.1: Results of MCMC sampling from the posterior distribution (second run)

	State 1	**State 2**	**State 3**
Linear			
α	-0.0685 (-0.0959, -0.0436)		
σ	0.7630 (0.7341, 0.7981)		
δ	2.0761 (1.9221, 2.2565)		
Two-State			
α	-0.0533 (-0.1423, 0.0443)	-0.0571 (-0.0849, -0.0303)	
σ	1.5604 (1.4543, 1.6829)	0.5671 (0.5269, 0.6067)	
δ	3.4451 (2.8767, 4.0828)	5.4155 (4.1481, 7.1417)	
P_{1i}	0.9635 (0.9503, 0.9749)	0.0365 (0.0251, 0.0497)	
P_{2i}	0.0269 (0.0185, 0.0357)	0.9731 (0.9643, 0.9815)	
Three-State			
α	-0.0357 (-0.1042, 0.0390)	-0.0547 (-0.0772, -0.0286)	0.1150 (-0.3109, 0.5561)
σ	1.3556 (1.2196, 1.4825)	0.5131 (0.4747, 0.5493)	3.3327 (2.7472, 4.0908)
δ	12.2172 (6.8555, 24.0163)	5.9346 (4.3369, 8.5115)	5.0148 (2.9995, 9.8896)
P_{1i}	0.9525 (0.9370, 0.9669)	0.0369 (0.0229, 0.0503)	0.0105 (0.0045, 0.0182)
P_{2i}	0.0333 (0.0226, 0.0450)	0.9642 (0.9515, 0.9738)	0.0026 (0.0002, 0.0065)
P_{3i}	0.0626 (0.0311, 0.1054)	0.0093 (0.0004, 0.0285)	0.9281 (0.8825, 0.9600)

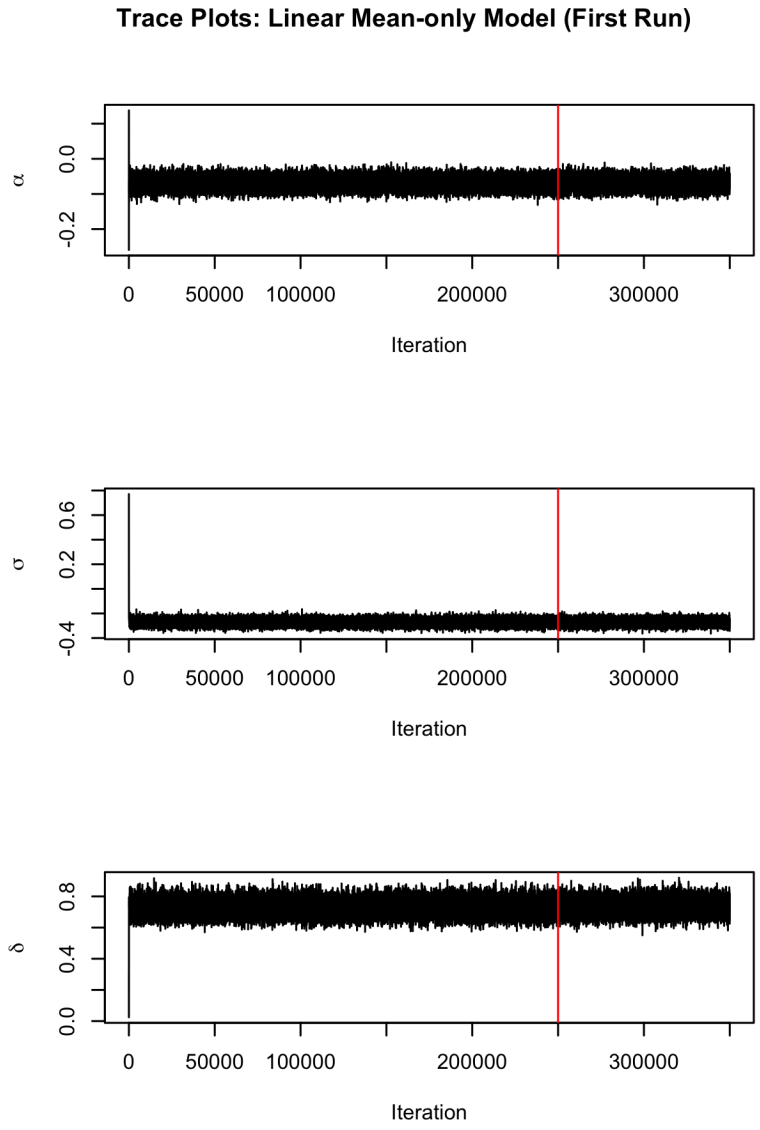

FIGURE D.1: Trace plots for the first run chain of the linear model. The red line indicates the end of the burn-in period

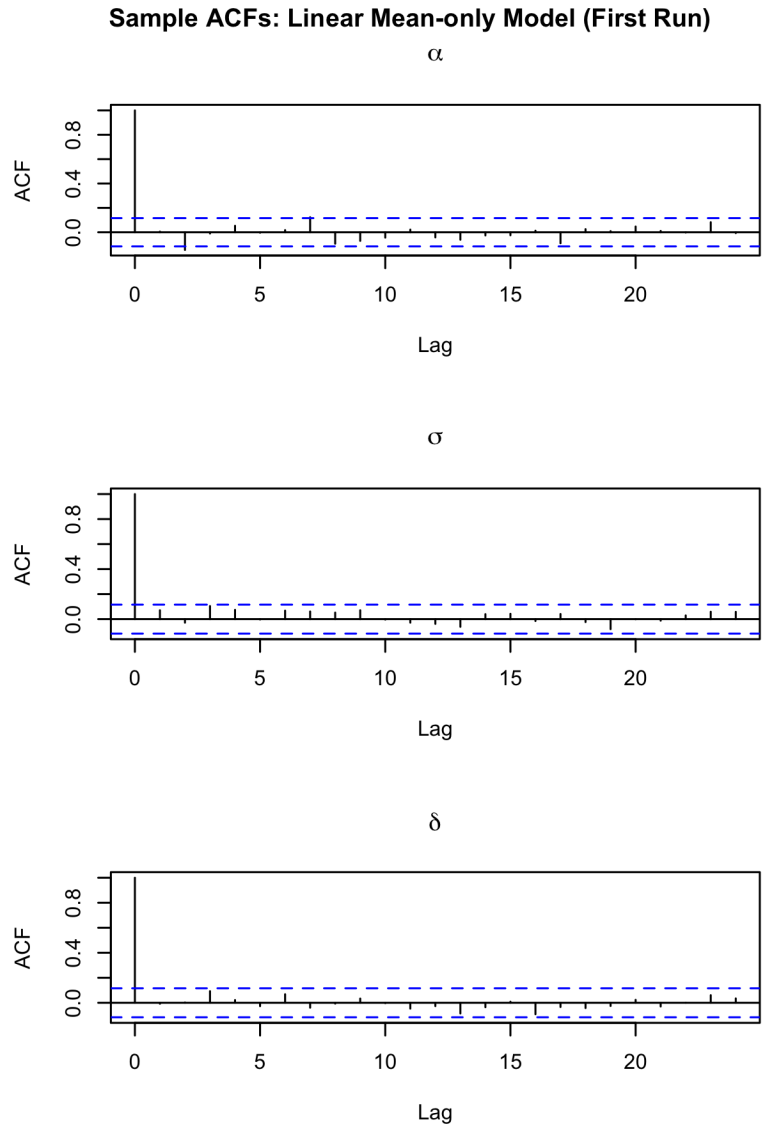

FIGURE D.2: Sample ACFs for the thinned posterior sample of parameters in the first run chain of the linear model

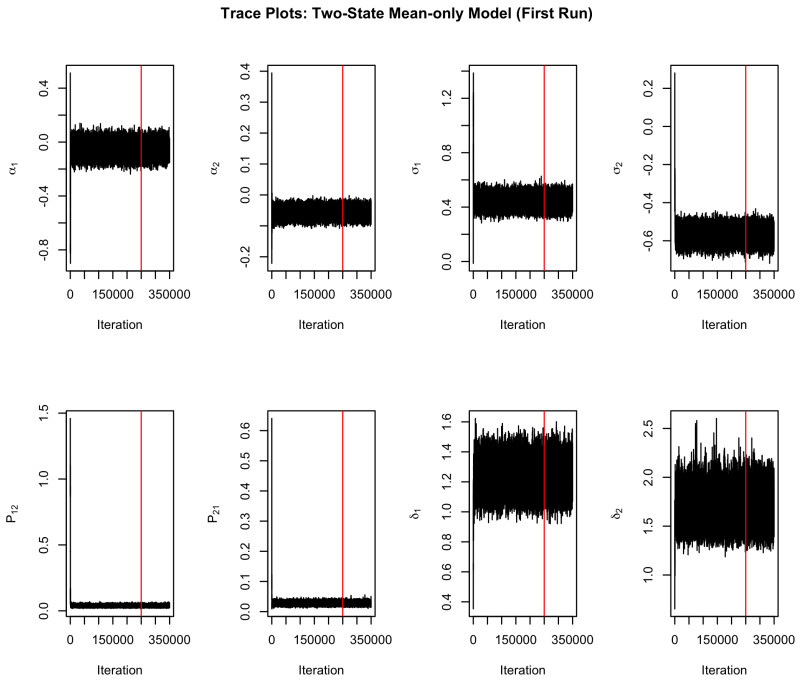

FIGURE D.3: Trace plots for the first run chain of the two-state model. The red line indicates the end of the burn-in period

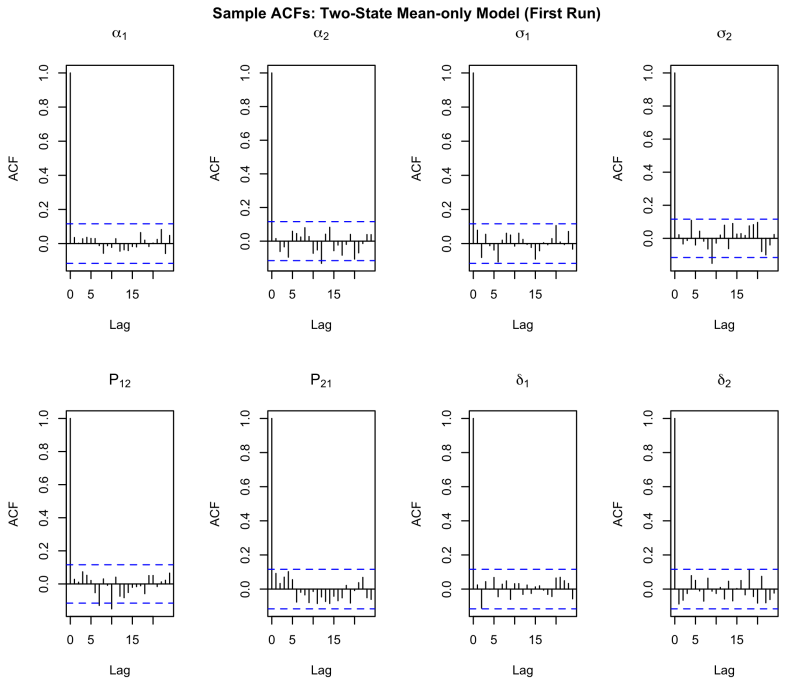

FIGURE D.4: Sample ACFs for the thinned posterior sample of parameters in the first run chain of the two-state model

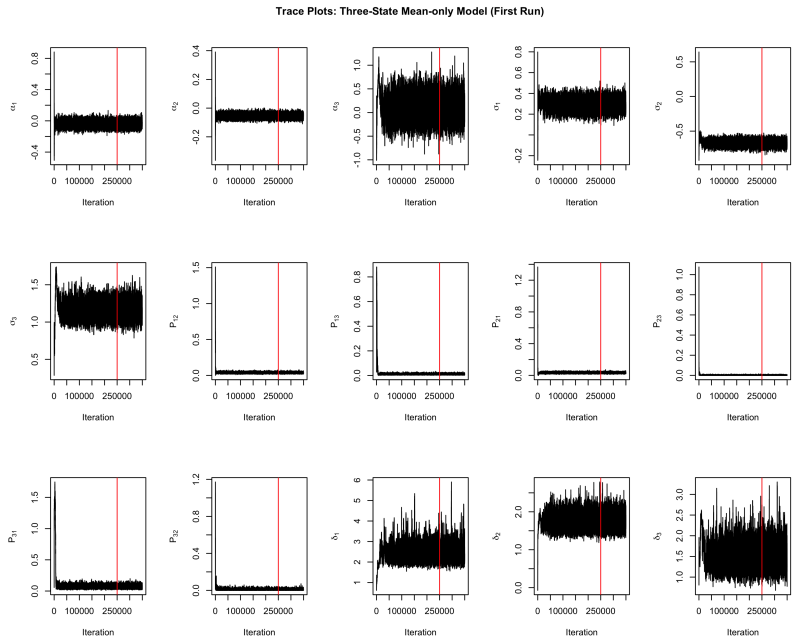

FIGURE D.5: Trace plots for the first run chain of the three-state model. The red line indicates the end of the burn-in period

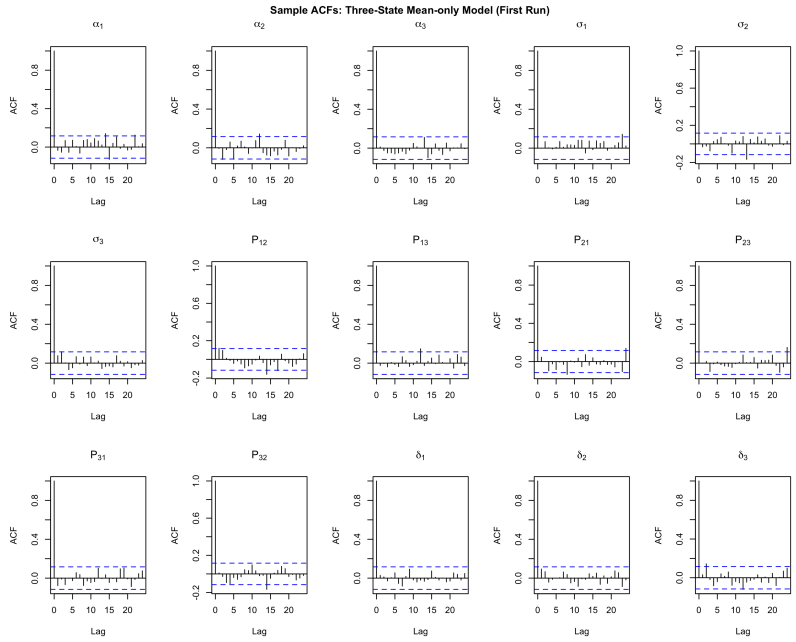

FIGURE D.6: Sample ACFs for the thinned posterior sample of parameters in the first run chain of the three-state model

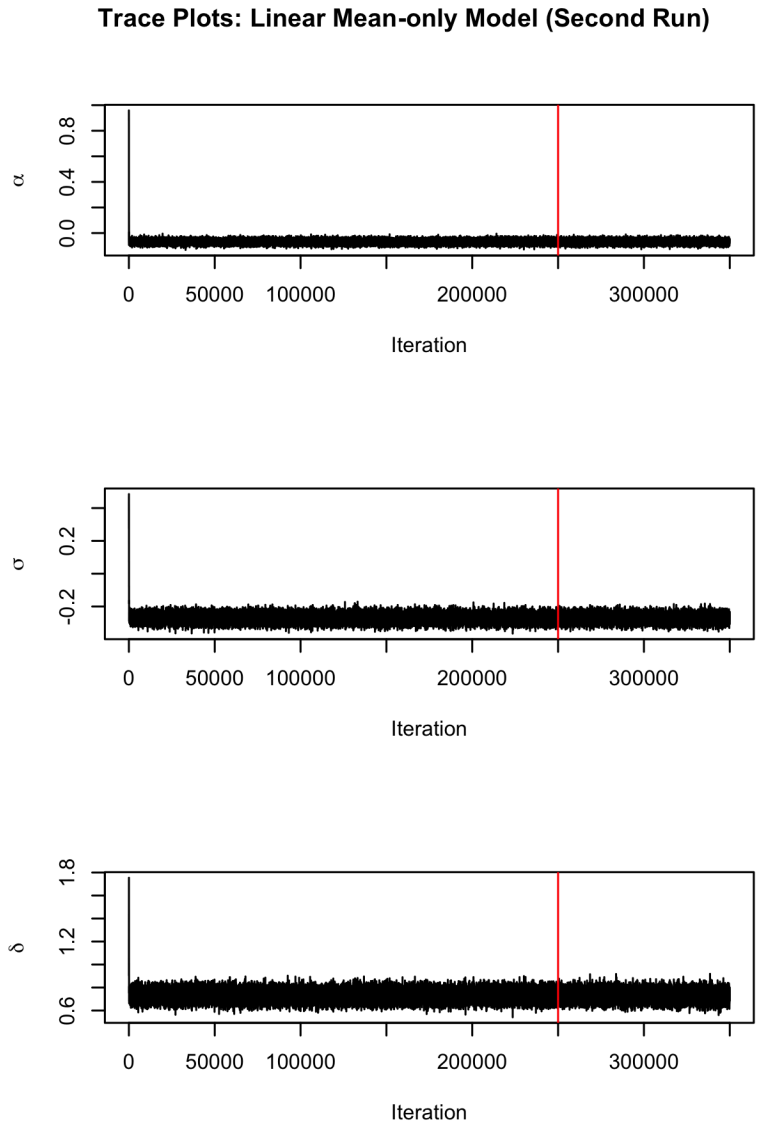

FIGURE D.7: Trace plots for the second run chain of the linear model. The red line indicates the end of the burn-in period

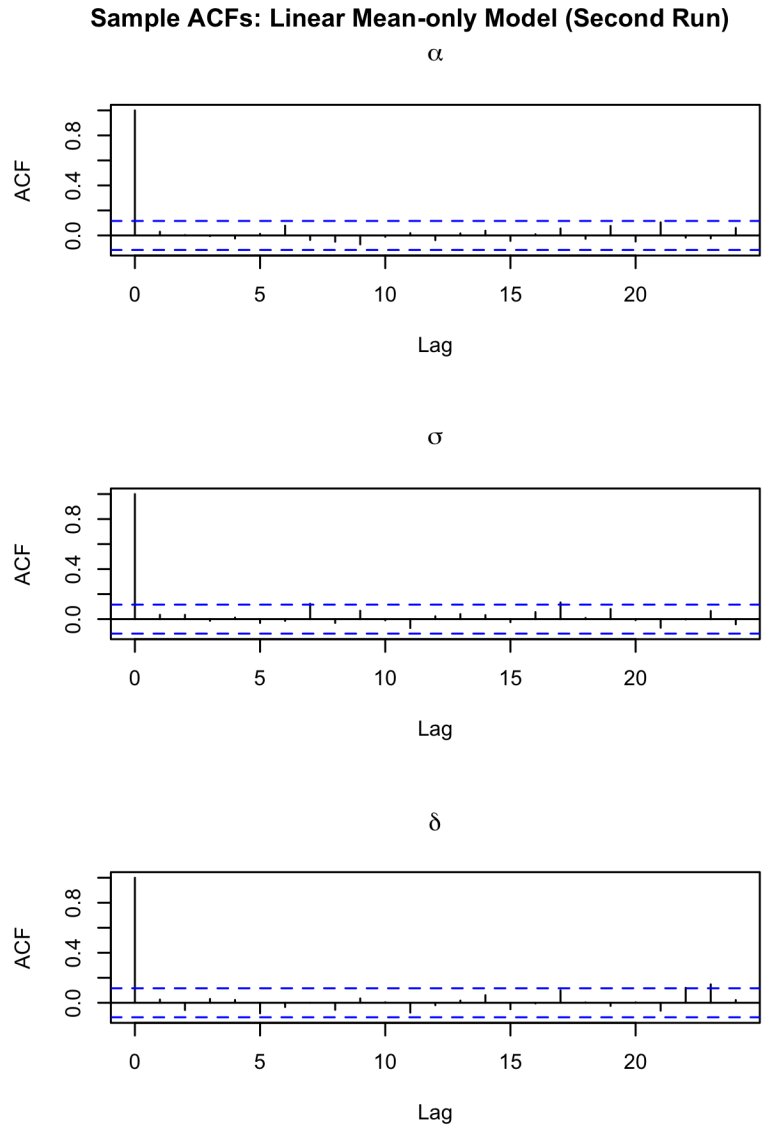

FIGURE D.8: Sample ACFs for the thinned posterior sample of parameters in the second run chain of the linear model

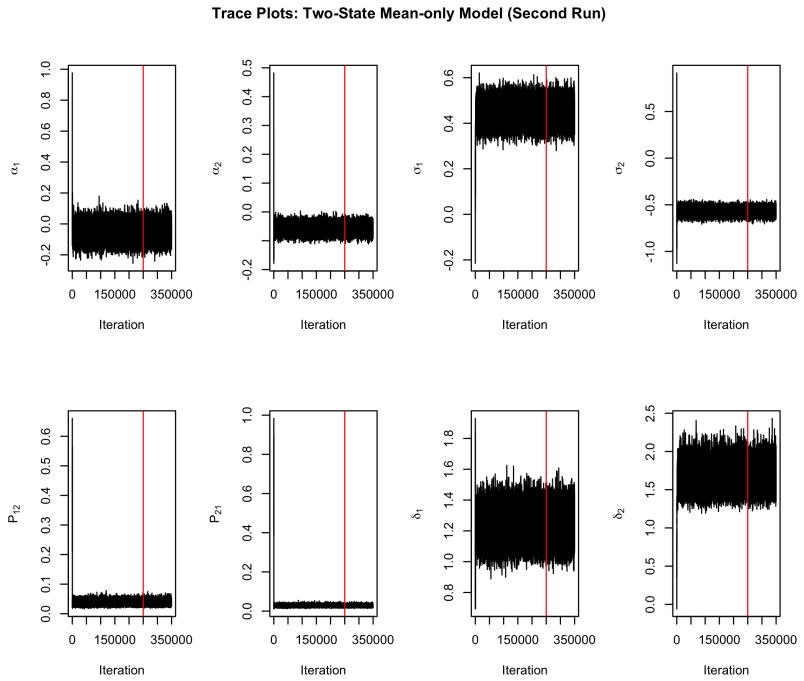

FIGURE D.9: Trace plots for the second run chain of the two-state model. The red line indicates the end of the burn-in period

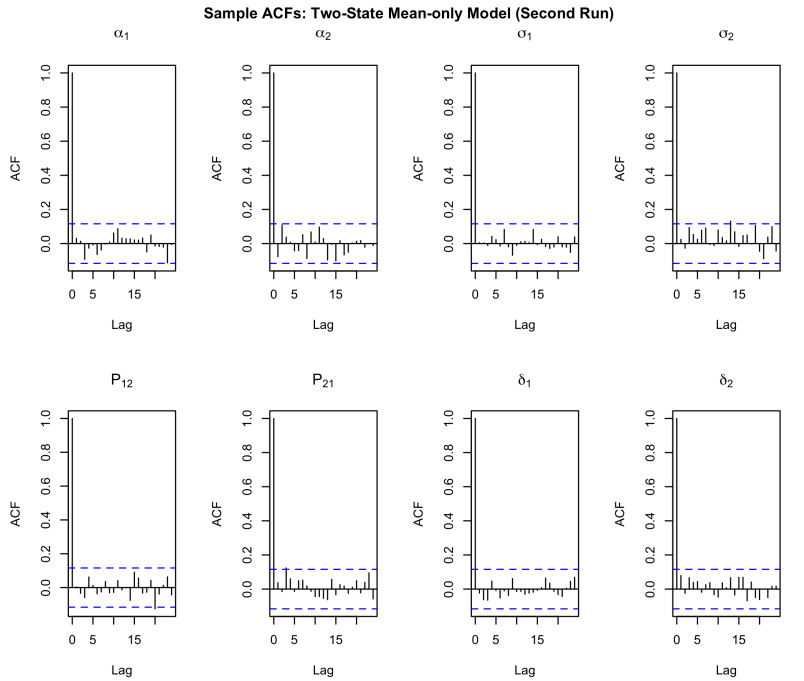

FIGURE D.10: Sample ACFs for the thinned posterior sample of parameters in the second run chain of the two-state model

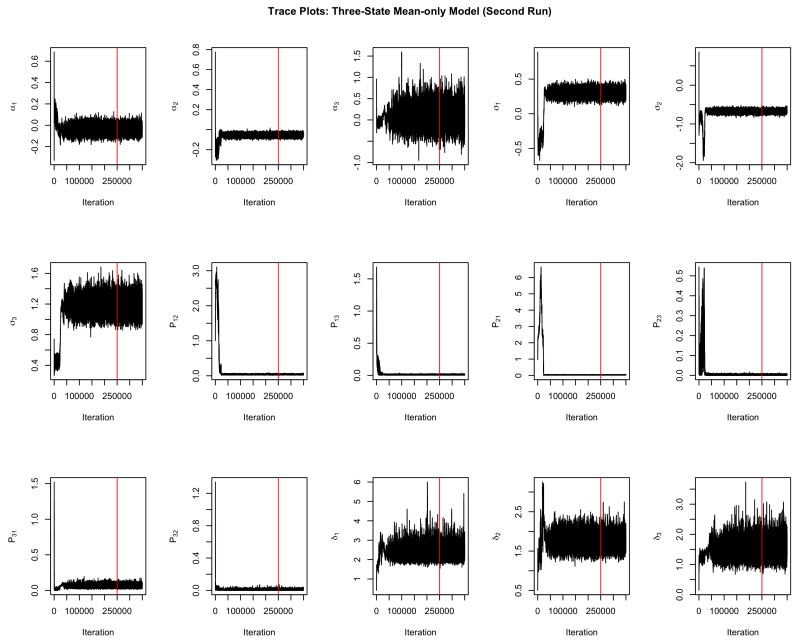

FIGURE D.11: Trace plots for the second run chain of the three-state model. The red line indicates the end of the burn-in period

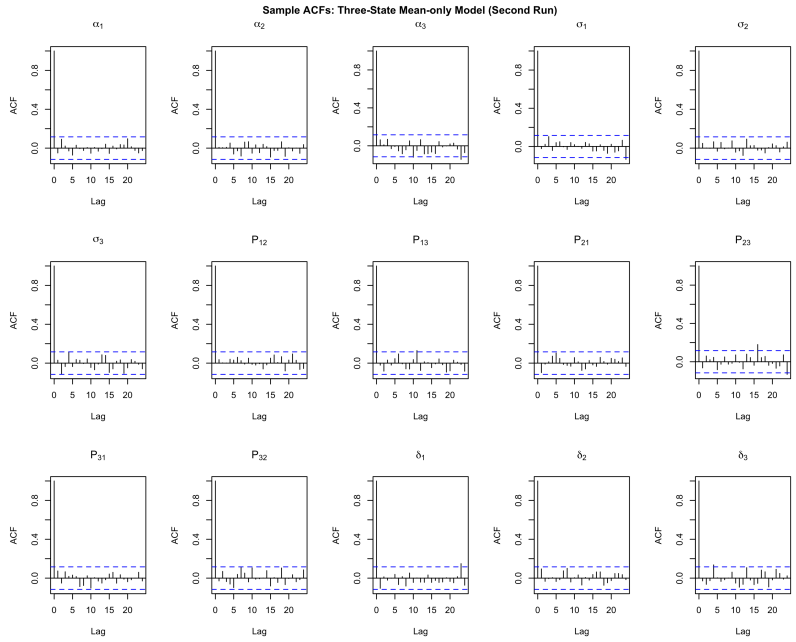

FIGURE D.12: Sample ACFs for the thinned posterior sample of parameters in the second run chain of the three-state model

Appendix E

Code

E.1 Chapter 3 Code

E.1.1 Markov-switching ARMA Model

MS-ARMA Model Likelihood

```
likelihood_func_arma <- function(params, Y, n_states, switching, use_t)
    {

  ###Set the parameters###
  N = length(Y)
  if(switching[1] == TRUE){
    alpha = params[1:n_states]
    param_index = n_states + 1
  }
  else{
    alpha = rep(params[1], n_states)
    param_index = 2
  }
  if(switching[2] == TRUE){
    beta = params[param_index:(param_index+n_states-1)]
    param_index = param_index + n_states
  }
  else{
    beta = rep(params[param_index], n_states)
    param_index = param_index + 1
  }
  if(switching[3] == TRUE){
    theta = params[param_index:(param_index+n_states-1)]
    param_index = param_index + n_states
  }
  else{
    theta = rep(params[param_index], n_states)
    param_index = param_index + 1
  }
  if(switching[4] == TRUE){
    sigma = exp(params[param_index:(param_index+n_states-1)])
    param_index = param_index + n_states
  }
  else{
    sigma = rep(exp(params[param_index]), n_states)
    param_index = param_index + 1
  }
  #Probability matrix
  initial_prob_param_index = param_index - 1
```

```
working_prob_matrix = diag(x=0, nrow=n_states)
for(i in 1:n_states){
  for(j in 1:n_states){
    if(i != j){
      if(i>j){
        offset = 0
      }
      else{
        offset = 1
      }
      index = initial_prob_param_index + (n_states-1)*(i-1)+j -
          offset
      working_prob_matrix[i,j] = params[index]
    }
  }
}
probs = exp(working_prob_matrix)/rowSums(exp(working_prob_matrix))
param_index = initial_prob_param_index + (n_states-1)*n_states + 1
if(use_t == TRUE){
  if(switching[5] == TRUE){
    df = exp(params[param_index: (param_index + n_states - 1)])
  }
  else{
    df = exp(rep(params[param_index], n_states))
  }
}

##############
#Start recursive likelihood calculation
prediction = c(1, rep(0, (n_states-1))) #initialise pred. probs
previous_y = 0
previous_e = 0
likelihood = 0
for(i in 1:N){
  diagonal = c()
  cond_exp = c()
  if(use_t == TRUE){
    for(k in 1:n_states){
      mean = alpha[k] + beta[k]*previous_y + theta[k]*previous_e
      diagonal = c(diagonal, dt((Y[i]-mean)/sigma[k], df = df[k])/
          sigma[k])
      cond_exp = c(cond_exp, mean)
    }
  }
  else{
    for(k in 1:n_states){
      mean = alpha[k] + beta[k]*previous_y + theta[k]*previous_e
      diagonal = c(diagonal, dnorm(Y[i], mean = mean, sd = sigma[k]))
      cond_exp = c(cond_exp, mean)
    }
  }
  P_x_i = diag(diagonal)
  numerator = prediction %*% P_x_i

  #Calculate conditional error (not done for mean-only and AR models)
  previous_e = sum(prediction*(Y[i] - cond_exp))

  filtering_probs = numerator/sum(numerator)
  prediction = filtering_probs %*% probs
  likelihood = likelihood + log(sum(numerator))

  previous_y = Y[i]
}
```

```
  return (-1*likelihood) #Return negative ll as optimiser minimizes
}
```

MS-ARMA Pseudo-Residual Calculation

```
#Pseudo residual calc for this model class
calculate_pr_arma <- function(params, Y, n_states, switching, use_t){

  ###Set the parameters###
  N = length(Y)
  if(switching[1] == TRUE){
    alpha = params[1:n_states]
    param_index = n_states + 1
  }
  else{
    alpha = rep(params[1], n_states)
    param_index = 2
  }
  if(switching[2] == TRUE){
    beta = params[param_index:(param_index+n_states-1)]
    param_index = param_index + n_states
  }
  else{
    beta = rep(params[param_index], n_states)
    param_index = param_index + 1
  }
  if(switching[3] == TRUE){
    theta = params[param_index:(param_index+n_states-1)]
    param_index = param_index + n_states
  }
  else{
    theta = rep(params[param_index], n_states)
    param_index = param_index + 1
  }
  if(switching[4] == TRUE){
    sigma = exp(params[param_index:(param_index+n_states-1)])
    param_index = param_index + n_states
  }
  else{
    sigma = rep(exp(params[param_index]), n_states)
    param_index = param_index + 1
  }
  #Probability matrix
  initial_prob_param_index = param_index - 1
  working_prob_matrix = diag(x=0, nrow=n_states)
  for(i in 1:n_states){
    for(j in 1:n_states){
      if(i != j){
        if(i>j){
          offset = 0
        }
        else{
          offset = 1
        }
        index = initial_prob_param_index + (n_states-1)*(i-1)+j -
            offset
        working_prob_matrix[i,j] = params[index]
      }
    }
  }
  probs = exp(working_prob_matrix)/rowSums(exp(working_prob_matrix))
  param_index = initial_prob_param_index + (n_states-1)*n_states + 1
```

```
  if(use_t == TRUE){
    if(switching[5] == TRUE){
      df = exp(params[param_index: (param_index + n_states - 1)])
    }
    else{
      df = exp(rep(params[param_index], n_states))
    }
  }

  #Start recursive likelihood calculation
  prediction = c(1, rep(0, (n_states-1)))
  previous_y = 0
  previous_e = 0
  pr = c()
  for(i in 1:N){
    diagonal = c()
    pseudo_residual_diagonal = c()
    cond_exp = c()
    if(use_t == TRUE){
      for(k in 1:n_states){
        mean = alpha[k] + beta[k]*previous_y + theta[k]*previous_e
        diagonal = c(diagonal, dt((Y[i]-mean)/sigma[k], df = df[k])/
            sigma[k])
        pseudo_residual_diagonal = c(pseudo_residual_diagonal,
                                     pt((Y[i]-mean)/sigma[k], df = df[k
                                         ]))
        cond_exp = c(cond_exp, mean)
      }
    }
    else{
      for(k in 1:n_states){
        mean = alpha[k] + beta[k]*previous_y + theta[k]*previous_e
        diagonal = c(diagonal, dnorm(Y[i], mean = mean, sd = sigma[k]))
        pseudo_residual_diagonal = c(pseudo_residual_diagonal,
                                     pnorm(Y[i], mean = mean, sd =
                                         sigma[k]))
        cond_exp = c(cond_exp, mean)
      }
    }
    P_x_i = diag(diagonal)
    pr = c(pr, sum(prediction %*% diag(pseudo_residual_diagonal)))
    numerator = prediction %*% P_x_i

    #Calculate conditional error
    previous_e = sum(prediction*(Y[i] - cond_exp))

    filtering_probs = numerator/sum(numerator)
    prediction = filtering_probs %*% probs
    previous_y = Y[i]
  }
  return (pr)
}
```

E.1.2 Smoothing Probabilites

```
#Smoothing probabilities calculation for 3-state, all switching, t-dbn
#As calculated in Kuan (2002)

smoothing_probabilities <- function(params, Y, n_states, use_t){

  ##############
  ###Set the parameters###
```

```
  N = length(Y)
  #Convert the parameters
  alpha = params[1:3]
  sigma = params[4:6]
  probs = matrix(c(1-sum(params[7:8]), params[7:8],
                   params[9], 1-sum(params[9:10]), params[10],
                   params[11:12], 1-sum(params[11:12])),
                 byrow=TRUE, ncol=3)
  df = params[13:15]

  #Start recursive likelihood calculation to get filtering probs
  prediction = c(1, rep(0, (n_states-1)))
  filtering_probs_matrix = matrix(data = NA, nrow = N, ncol = n_states)
  prediction_probs_matrix = matrix(data = NA, nrow = N, ncol = n_states
      )
  for(i in 1:N){

    if(use_t == TRUE){
      diagonal = c()
      for(k in 1:n_states){
        diagonal = c(diagonal,
                     dt((Y[i]-alpha[k])/sigma[k], df = df[k])/sigma[k])
      }
      P_x_i = diag(diagonal)
    }
    else{
      diagonal = c()
      for(k in 1:n_states){
        diagonal = c(diagonal, dnorm(Y[i], mean = alpha[k], sd = sigma[
            k]))
      }
      P_x_i = diag(diagonal)
    }
    numerator = prediction %*% P_x_i

    filtering_probs = numerator/sum(numerator)
    filtering_probs_matrix[i,] = filtering_probs

    prediction = filtering_probs %*% probs
    prediction_probs_matrix[i,] = prediction
  }

  #Calculate smoothing probabilities
  #Start at the end
  smoothed_probs_matrix = matrix(data = NA, nrow = N, ncol = n_states)
  smoothed_probs_matrix[N,] = filtering_probs_matrix[N,]
  for(i in rev(1:(N-1))){
    ith_smoothed = filtering_probs_matrix[i,] * ((smoothed_probs_matrix
        [(i+1),]/prediction_probs_matrix[i,]) %*% t(probs))
    smoothed_probs_matrix[i,] = ith_smoothed
  }
  return(smoothed_probs_matrix)
}
```

E.1.3 Parallel Sampling

```
#Parallel sampling calculation
#post_dbn_n_thin refers to the relevant
#thinned sample from the posterior, obtained from MCMC
#I multiply the log likelihoods by -1 because the
#functions return the negative loglikelihood
```

```
Y = read.csv("data/Differenced_VIX_2000.csv")
Y = Y[, "Delta"]
switching_vector = c(TRUE, TRUE, TRUE)
USE_T = TRUE

prior_1 = 1/3
prior_2 = 1/3
prior_3 = 1/3

G_1 = c()
for(l in 1:nrow(post_dbn_1_thin)){
  prior_log_prob = sum(dnorm(post_dbn_1_thin[l,], sd=1000, log=TRUE))
  ll = -1*likelihood_func_single_state_mean(params = post_dbn_1_thin[l
      ,],
                                            Y = Y,
                                            use_t = USE_T)
  ith = ll+log(prior_1)+prior_log_prob
  G_1 = c(G_1, ith)
}

G_2 = c()
for(l in 1:nrow(post_dbn_2_thin)){
  prior_log_prob = 0
  for(j in 1:length(post_dbn_2_thin[l,])){
    if(j == 5 || j == 6){
      prior_log_prob = prior_log_prob + log(dexp(post_dbn_2_thin[l,j],
                                                 rate = 0.001))
    }
    else{
      prior_log_prob = prior_log_prob + log(dnorm(post_dbn_2_thin[l,j],
                                                  sd=1000))
    }
  }
  ll = -1*likelihood_func_mean(params = post_dbn_2_thin[l,],
                               Y = Y,
                               n_states = 2,
                               switching = switching_vector,
                               use_t = USE_T)
  ith = ll+log(prior_2)+prior_log_prob
  G_2 = c(G_2, ith)
}

G_3 = c()
for(l in 1:nrow(post_dbn_3_thin)){
  prior_log_prob = 0
  for(j in 1:length(post_dbn_3_thin[l,])){
    if(j >= 7 && j <= 12){
      prior_log_prob = prior_log_prob + log(dexp(post_dbn_3_thin[l,j],
                                                 rate=0.001))
    }
    else{
      prior_log_prob = prior_log_prob + log(dnorm(post_dbn_3_thin[l,j],
                                                  sd=1000))
    }
  }
  ll = -1*likelihood_func_mean(params = post_dbn_3_thin[l,],
                               Y = Y,
                               n_states = 3,
                               switching = switching_vector,
                               use_t = USE_T)
  ith = ll+log(prior_3)+prior_log_prob
  G_3 = c(G_3, ith)
}
```

```
max_ll = apply(matrix(c(G_1, G_2, G_3), byrow = FALSE, ncol=3), 1, max)
G_1_scaled = exp(G_1 - max_ll)
G_2_scaled = exp(G_2 - max_ll)
G_3_scaled = exp(G_3 - max_ll)

denominator = G_1_scaled + G_2_scaled + G_3_scaled
num = length(G_1_scaled)
#Calculate
final_prob_1 = sum(G_1_scaled/denominator, na.rm = TRUE)/num
final_prob_2 = sum(G_2_scaled/denominator, na.rm = TRUE)/num
final_prob_3 = sum(G_3_scaled/denominator, na.rm = TRUE)/num
```

E.1.4 Viterbi Algorithm to Calculate Most Likely States

```
#Viterbi Algorithm for 3 state chain, all parameter switching and t
    errors

emission_log_probability <- function(Y, mean, sigma,
                                     use_t = FALSE, dof = NULL){
  if(use_t == TRUE){
    prob = dt((Y-mean)/sigma, df = dof)/sigma
  }
  else{
    prob = dnorm(Y, mean = mean, sd = sigma)
  }
  return(log(prob))
}

calculate_viterbi <- function(data, params, n_states, use_t){

  #Convert the parameters
  alpha = params[1:3]
  sigma = params[4:6]
  probs = matrix(c(1-sum(params[7:8]), params[7:8],
            params[9], 1-sum(params[9:10]), params[10],
            params[11:12], 1-sum(params[11:12])),
            byrow=TRUE, ncol=3)
  df = params[13:15]
  #Initialise the tables
  T_1 = matrix(data = 0, nrow=n_states, ncol = length(data))
  T_2 = matrix(data = 0, nrow=n_states, ncol = length(data))
  initial_prediction_probs = c(1, rep(0, (n_states-1)))
  T_1[1,1] = emission_log_probability(data[1], alpha[1],
                                      sigma[1], use_t = use_t, df[1])
  T_1[2,1] = log(0)
  T_1[3,1] = log(0)
  A = log(probs)
  for (j in 2:length(data)){
    for(i in 1:n_states){
      B_iyj = emission_log_probability(data[j],
                                       alpha[i],
                                       sigma[i], use_t = use_t, df[i])
      f = T_1[, j-1] + A[, i]
      T_1[i,j] = max(f + B_iyj)
      T_2[i,j] = which.max(f)
    }
  }
  X = c(rep(0, length(data)))
  x_T = which.max(T_1[,length(data)])
  X[length(data)] = x_T
  for(j in rev(2:length(data))){
```

```
    X[j-1] = T_2[X[j],j]
  }
  return(X)
}
```

References

Akaike, H. (1998). "Information Theory and an Extension of the Maximum Likelihood Principle". *Selected Papers of Hirotugu Akaike*. Ed. by E. Parzen, K. Tanabe & G. Kitagawa. New York, NY: Springer New York, pp. 199–213. ISBN: 978-1-4612-1694-0. DOI: 10.1007/978-1-4612-1694-0_15.

Alvarez-Plata, P. & Schrooten, M. (2006). "The Argentinean Currency Crisis: A Markov-Switching Model Estimation". *The Developing Economies* 44.1, pp. 79–91. DOI: 10.1111/j.1746-1049.2006.00004.x.

Ang, A. & Timmermann, A. (2012). "Regime Changes and Financial Markets". *Annual Review of Financial Economics* 4.1, pp. 313–337. DOI: 10.1146/annurev-financial-110311-101808. URL: https://doi.org/10.1146/annurev-financial-110311-101808.

Baba, N. & Sakurai, Y. (2011). "Predicting Regime Switches in the VIX Index with Macroeconomic Variables". *Applied Economics Letters* 18.15, pp. 1415–1419. DOI: 10.1080/13504851.2010.539532.

Bloomberg L.P. (2017). *Exchange Rates and various indices*. Available: Bloomberg database [2017, April 13].

Brooks, S.P. & Gelman, A. (1998). "General Methods for Monitoring Convergence of Iterative Simulations". *Journal of Computational and Graphical Statistics* 7.4, pp. 434–455. DOI: 10.1080/10618600.1998.10474787.

Brunnermeier, M.L., Nagel, S. & Pedersen, L.H. (2008). "Carry Trades and Currency Crashes". *NBER Macroeconomics Annual* 23.1, pp. 313–348. ISSN: 0889-3365. DOI: 10.1086/593088.

Burnside, C. (2012). "Carry Trades and Risk". *Handbook of Exchange Rates*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Chap. 10, pp. 283–312. ISBN: 9780470768839. DOI: 10.1002/9781118445785.ch10. URL: http://doi.wiley.com/10.1002/9781118445785.ch10.

Burnside, C. et al. (2011). "Do Peso Problems Explain the Returns to the Carry Trade?" *Review of Financial Studies* 24.3, pp. 853–891. ISSN: 08939454. DOI: 10.1093/rfs/hhq138.

Carrasco, M., Hu, L. & Ploberger, W. (2014). "Optimal Test for Markov Switching Parameters". *Econometrica* 82.2, pp. 765–784. DOI: 10.3982/ECTA8609.

Chen, C.-C. & Tsay, W. (2007). "Estimating Markov-Switching ARMA Models with Extended Algorithms of Hamilton". Available: https://www.researchgate.net/publication/228432191_Estimating_Markov-Switching_ARMA_Models_with_Extended_Algorithms_of_Hamilton [2018, June 2].

Christiansen, C., Ranaldo, A. & Söderlind, P. (2011). "The Time-Varying Systematic Risk of Carry Trade Strategies". *Journal of Financial and Quantitative Analysis* 46.04, pp. 1107–1125. ISSN: 0022-1090. DOI: 10.1017/S0022109011000263.

Congdon, P. (2006). "Bayesian Model Choice Based on Monte Carlo Estimates of Posterior Model Probabilities". *Computational Statistics & Data Analysis* 50, pp. 346–357. DOI: 10.1016/j.csda.2004.08.001.

Cont, R. (Mar. 2002). "Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues". *Quantitative Finance* 1, pp. 223–236. DOI: 10.1088/1469-7688/1/2/304.

Curcuru, S., Vega, C. & Hoek, J. (2010). "Measuring carry trade activity". *Proceedings of the IFC Conference on "Initiatives to Address Data Gaps Revealed by the Financial Crisis"* 34. Ed. by Bank for International Settlements, pp. 436–453. URL: http://www.bis.org/ifc/events/5ifcconf/curcuru.pdf.

Davies, R.B. (1977). "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative". *Biometrika* 64.2, pp. 247–254. ISSN: 00063444. DOI: 10.2307/2335690.

Di Sanzo, S. (2009). "Testing for Linearity in Markov Switching Models: a Bootstrap Approach". *Statistical Methods and Applications* 18.2, pp. 153–168. ISSN: 1613-981X. DOI: 10.1007/s10260-007-0080-6.

Francq, C. & Zakoian, J.-M. (2001). "Stationarity of Multivariate Markov–switching ARMA models". *Journal of Econometrics* 102.2, pp. 339–364. ISSN: 0304-4076. DOI: 10.1016/S0304-4076(01)00057-4.

Froot, K.A. & Thaler, R.H. (1990). "Anomalies: Foreign Exchange". *Journal of Economic Perspectives* 4.3, pp. 179–192. ISSN: 0895-3309. DOI: 10.1257/jep.4.3.179.

Goldfeld, S.M. & Quandt, R.E. (1973). "A Markov model for Switching Regressions". *Journal of Econometrics* 1.1, pp. 3–15. ISSN: 03044076. DOI: 10.1016/0304-4076(73)90002-X.

Gray, S.F. (1996). "Modeling the Conditional Distribution of Interest Rates as a Regime-switching Process". *Journal of Financial Economics* 42.1, pp. 27–62. ISSN: 0304-405X. DOI: 10.1016/0304-405X(96)00875-6.

Guttman, L. (1954). "Some Necessary Conditions for Common Factor Analysis". *Psychometrica* 19, pp. 149–161.

Hamilton, J.D. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle". *Econometrica* 57.2, pp. 357–384. DOI: 10.2307/1912559.

Hansen, B.E. (1992). "The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP". *Journal of Applied Econometrics* 7, S61–S82. ISSN: 08837252, 10991255. URL: http://www.jstor.org/stable/2284984.

Jeanne, O. & Masson, P. (2000). "Currency Crises, Sunspots and Markov-switching Regimes". *Journal of International Economics* 50.2, pp. 327–350. ISSN: 0022-1996. DOI: 10.1016/S0022-1996(99)00007-0.

Jones, E., Oliphant, T., Peterson, P., et al. (2001). *SciPy: Open source scientific tools for Python*. Available: http://www.scipy.org/.

Katzke, N.F. & Polakow, D.A. (2017). Carry and Consequence: Understanding the Recent Resilience of Emerging Market Currencies. (Unpublished).

Kim, C.-J. & Kim, Jaeho (2015). "Bayesian Inference in Regime-Switching ARMA Models With Absorbing States: The Dynamics of the Ex-Ante Real Interest Rate Under Regime Shifts". *Journal of Business & Economic Statistics* 33.4, pp. 566–578. DOI: 10.1080/07350015.2014.979995.

Kuan, C.-M. (2002). *Lecture on the Markov Switching Model*. Available: http://homepage.ntu.edu.tw/~ckuan/pdf/Lec-Markov_slide_Spring_2011.pdf [2018, September 17]. Institute of Economics, Academia Sinica.

Lustig, H., Roussanov, N. & Verdelhan, A. (2011). "Common Risk Factors in Currency Markets". *Review of Financial Studies* 24.11, pp. 3731–3777. ISSN: 08939454. DOI: 10.1093/rfs/hhr068.

MacQueen, J. (1967). "Some Methods for classification and Analysis of Multivariate Observations". *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press* 1.14, pp. 281–297. ISSN: 00970433.

Menkhoff, L. et al. (2012). "Carry Trades and Global Foreign Exchange Volatility". *Journal of Finance* 67.2, pp. 681–718. ISSN: 00221082. DOI: 10.1111/j.1540-6261.2012.01728.x.

Organisation for Economic Co-operation and Development [OECD] (2017). *Short-term Interest Rates (Indicator)*. Available: https://data.oecd.org/interest/short-term-interest-rates.htm [2017, September 5].

Polakow, D.A. & Flint, E.J. (2015). "Global Risk Factors and South African Equity Indices". *South African Journal of Economics* 83.4, pp. 598–616. ISSN: 18136982. DOI: 10.1111/saje.12065.

Romo, J.M. (2011). "Volatility Regimes for the VIX index". Available: https://www.researchgate.net/publication/215689540_Volatility_Regimes_for_the_VIX_index [2018, August 18].

Scheidegger, A. (2018). *adaptMCMC: Implementation of a Generic Adaptive Monte Carlo Markov Chain Sampler*. R package version 1.3. URL: https://CRAN.R-project.org/package=adaptMCMC.

Schwarz, G. (1978). "Estimating the Dimension of a Model". *The Annals of Statistics* 6.2, pp. 461–464. DOI: 10.1214/aos/1176344136.

Sclove, S.L. (1983). "Time-series Segmentation: A Model and a Method". *Information Sciences* 29.1, pp. 7–25.

Smith, A., Naik, P.A. & Tsai, C.-L. (2006). "Markov-switching Model Selection using Kullback–Leibler Divergence". *Journal of Econometrics* 134.2, pp. 553–577. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2005.07.005.

Spedicato, G.A. (July 2017). *Discrete Time Markov Chains with R*. R package version 0.6.9.7. URL: https://cran.r-project.org/web/packages/markovchain/index.html.

Vihola, M. (2012). "Robust Adaptive Metropolis Algorithm with Coerced Acceptance Rate". *Statistics and Computing* 22.5, pp. 997–1008. ISSN: 1573-1375. DOI: 10.1007/s11222-011-9269-5.

Viterbi, A. (1967). "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm". *IEEE Transactions on Information Theory* 13.2, pp. 260–269. ISSN: 0018-9448. DOI: 10.1109/TIT.1967.1054010.

Zucchini, W., MacDonald, I. & Langrock, R. (2016). *Hidden Markov Models for Time Series : an Introduction using R*. eng. Second edition. Monographs on Statistics and Applied Probability ; 150. Boca Raton, Florida: CRC Press. ISBN: 1-315-37248-7.

Template

This dissertation was typeset using LaTeX and made use of version 2.5 (27/8/17) of the template created by Sunil Patel (available here), which is itself a modified version of the original template by Steve Gunn (available here).