

Synchronization of Markov Chains in Multivariate Regime-Switching Models

DISSERTATION
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor of Philosophy in Management

submitted by

Raphael Vial

from

Lucerne and Le Crêt (Fribourg)

Approved on the application of

Prof. Dr. Karl Frauendorfer

and

Prof. Dr. Pascal Gantenbein

Dissertation no. 4380

Difo-Druck GmbH, Bamberg 2015

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, October 22, 2014

The President:

Prof. Dr. Thomas Bieger

Acknowledgement

This dissertation was carried out as part of my doctoral studies at the University of St. Gallen. I gratefully acknowledge the assistance of all those people who have supported me throughout.

I would like to thank Prof. Dr. Karl Frauendorfer for the great opportunity of writing my thesis under his supervision. I am very grateful for his trust, continuous support, and valuable advice during my entire doctoral studies. His confidence in me and my project allowed me to freely conduct my research while still benefitting from his invaluable experience when needed.

My gratitude also goes to Prof. Dr. Pascal Gantenbein, who kindly agreed to serve as co-supervisor. I thank him for his trust and valuable input. Our challenging discussions influenced my thesis in many ways.

I am also grateful to all the lecturers and colleagues whose advice markedly contributed to the quality of this thesis. Special thanks go to Dr. Mark Kyburz for teaching me the essentials of academic writing and for editing this thesis.

Finally, I would like to thank my family for their confidence, unwavering support, and unconditional love. My brother Christian, who has yet to experience this endeavor and who has always encouraged me. And my parents, who have enabled me to write this thesis and never uttered the slightest doubt about my decision to pursue this journey from start to finish. They all deserve my highest gratitude. I dedicate this thesis to them.

Küssnacht, August 2014

Raphael Vial

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Research Focus	2
1.3	Structure of the Thesis	3
2	Theory of Markov Regime-Switching	5
2.1	Literature Review	6
2.2	Regime-Switching Framework	10
2.2.1	Data-Generating Process	10
2.2.2	Markov Chains	12
2.2.3	Likelihood Function and Parameter Estimation	14
2.2.4	Regime Inference	17
2.3	Extension to Multivariate Regime-Switching	20
3	Synchronization of Markov Chains	27
3.1	Definition of Synchronization	29
3.2	Measuring In-Phase Synchronization	30
3.2.1	Pairwise Correlation	31
3.2.2	Concordance Index	31
3.2.3	Strong Perfect Non-Synchronization	32
3.2.4	Strong Perfect Positive Synchronization	34
3.3	Measuring Phase-Shifted Synchronization	35
3.3.1	Cross-Correlation	36
3.3.2	Dynamic Time Warping	37
3.3.3	Longest Common Subsequence	39
3.4	Measuring Time-Varying Synchronization	40

4	Data Analysis	43
4.1	Dataset	43
4.2	Descriptive Statistics	44
4.3	Regime-Switching Analysis	46
4.3.1	Univariate Regime-Switching Results	46
4.3.2	Multivariate Regime-Switching Results	50
4.3.3	Likelihood Ratio Tests	53
4.4	Regime Synchronization Analysis	56
4.4.1	Analysis of In-Phase Synchronization	58
4.4.2	Analysis of Phase-Shifted Synchronization	60
4.4.3	Analysis of Time-Varying Synchronization	66
4.5	Summary	68
5	Flexible Regime-Switching Models	71
5.1	General Model Specification	75
5.1.1	Data-Generating Process	75
5.1.2	Bivariate Model with Multiple Markov Chains	76
5.1.3	Multivariate Model with Multiple Markov Chains	81
5.2	State Inference	86
5.3	Parameter Estimation	88
5.3.1	Factorization of Correlation Matrices	89
5.3.2	Clustering of Assets	95
5.3.3	Use of Prior Information	97
5.4	Summary	101
6	Empirical Analysis	103
6.1	Model Comparison	103
6.2	Empirical Results of the Flexible Model	110
6.2.1	Clustering of Data	110
6.2.2	Analysis of the Core Sample	112
6.2.3	Analysis of the Extended Sample	120
6.2.4	Model Robustness	127

6.3	Accuracy of the Synchronization Factor	129
6.4	Model Forecasting Power	131
6.5	Summary	134
7	Asset Allocation under Regime-Switching	137
7.1	Asset Allocation Problem	138
7.1.1	Buy-and-Hold Allocation Strategy	138
7.1.2	Asset Allocation with Rebalancing	140
7.2	Asset Allocation Results	144
7.2.1	Portfolio Weights	144
7.2.2	Performance of the Strategies	149
8	Conclusion	163
	References	165

List of Tables

4.1	Descriptive statistics of the sample	45
4.2	Unconditional correlation parameters of the sample	45
4.3	Univariate regime-switching model with regime-dependent means .	47
4.4	Univariate regime-switching model with regime-independent means	48
4.5	Multivariate regime-switching model with regime-independent means	51
4.6	Regime-dependent correlations of the multivariate model	52
4.7	Likelihood ratio statistics for the core and the extended sample . . .	55
4.8	Correlation and concordance between univariate regime processes .	58
4.9	Tests for SPPS and SNS between univariate regime processes	59
4.10	Pairwise cross-correlation between two regime processes	61
4.11	Dynamic time warping between two regime processes	62
4.12	Longest common subsequence between two regime processes	64
5.1	Literature overview of dynamic regime-switching approaches	73
5.2	State space of the bivariate switching model	76
5.3	Extended state space with an additional synchronization factor . . .	81
6.1	Comparison of different model specifications	106
6.2	Comparison of different linear and non-linear model setups	109
6.3	k -means clusters of the regime processes	111
6.4	Moment parameters FLEX_2C (core sample)	114
6.5	Correlation matrices FLEX_2C (core sample)	116
6.6	Moment parameters FLEX_2C (extended sample)	121
6.7	Correlation matrices FLEX_3C (extended sample)	125
6.8	Turning-point prediction of the models	134
7.1	Model performance figures (core sample, buy-and-hold)	153

7.2	Model performance figures (core sample, rebalancing)	154
7.3	Model performance figures (extended sample, buy-and-hold)	157
7.4	Model performance figures (extended sample, rebalancing)	158
7.5	Performance figures of the flexible model with restricted correlation matrices (core sample)	161

List of Figures

1.1	Structure of the thesis	4
2.1	Comparison of univariate and multivariate regime processes	25
3.1	Possible forms of synchronization	36
3.2	Dynamic time warping synchronization path	38
3.3	Longest common subsequence synchronization path	40
4.1	Univariate regime probability processes	49
4.2	Multivariate regime probability processes	50
4.3	Comparison of univariate and multivariate regime-switching results	53
4.4	Scatterplots of pairwise univariate regime probabilities	57
4.5	Detection of phase-shifts using dynamic time warping	62
4.6	Matching of two regime processes using dynamic time warping	63
4.7	Longest common subsequence of two regime processes	65
4.8	Matching of two regime processes using longest common subsequence	66
4.9	Time-varying correlation between two regime processes	67
4.10	Squared differences between univariate regime processes	68
5.1	Factorization of the composite correlation matrices	92
5.2	Factorization algorithm for the correlation matrices	93
6.1	Silhouette values of k -means clustering	112
6.2	Smoothed regime probabilities of state 3 in FLEX_2C (core sample)	117
6.3	Smoothed state probabilities of FLEX_2C (core sample)	118
6.4	Synchronization processes of FLEX_2C and FLEX_3C (core sample)	120
6.5	Smoothed state probabilities of FLEX_2C (extended sample)	123

6.6	Smoothed synchronization processes of FLEX_3C (extended sample)	126
6.7	Robustness of the synchronization factor in subperiods	127
6.8	Robustness of the synchronization factor for alternative clusters . .	129
6.9	Prediction accuracy of the synchronization factor	131
6.10	Rolling-window forecasts of standard deviations	133
7.1	Portfolio weights for different holding periods	146
7.2	Portfolio weights for different holding periods (equity-only)	147
7.3	Portfolio weights for different risk aversions	148
7.4	Performance comparison of the models (core sample)	152
7.5	Performance comparison of the models (extended sample)	156
7.6	Performance comparison of the flexible model with restricted and unrestricted correlation matrices (core sample)	160

Abstract

Multivariate regime-switching presents an efficient way of jointly modeling the cyclical behavior of financial time series. Standard regime-switching models thereby *a priori* determine the relationship between the regime-switches of individual assets. These switches are usually assumed to be either perfectly synchronized or fully independent. However, neither assumption seems realistic in practice.

This thesis develops a multivariate Markov regime-switching model to infer the actual degree of synchronization from the underlying data. This flexible model allows subgroups of assets to be driven by individual Markov chains. At the same time, these Markov chains underlie a dynamically changing degree of synchronization. In comparison to most existing solutions, this model is not restricted to bivariate analysis. To keep the model traceable, a novel factorization algorithm for the regime-dependent correlation matrix is formulated. This algorithm scales down the increase in parameters and presents an efficient way of ensuring positive semi-definite correlation matrices.

The structure of the flexible regime-switching model is motivated by the initial synchronization analysis conducted in this thesis. The analysis of univariate regime-switching results shows that neither perfectly synchronized nor fully independent regime cycles are empirically observable. The synchronization of regime cycles tends to dynamically change over time. Some assets, however, might show more contemporaneous switching dynamics and can therefore be governed by a joint regime process.

The empirical results for a sample of six international equity markets confirm the assumptions underlying this thesis. The flexible model reveals a stable synchronization factor, marked by one particular change in synchronization. The estimated parameters of this model closely cover the individual dynamics of their underlying assets and confirm the model's validity. Moreover, in some states, correlation matrices

show very low or even negative parameters. This observation can be of particular value for portfolio allocation analysis.

Finally, to evaluate the performance of flexible regime-switching, an asset allocation problem is introduced. The comparison of the different models reveals the favorable dynamics of flexible regime-switching for medium-term investment horizons.

Zusammenfassung

Multivariate Regime-Switching Ansätze stellen eine effiziente Möglichkeit dar, um das zyklische Verhalten von mehreren Finanzmarktzeitreihen gemeinsam zu modellieren. Gängige Regime-Switching Modelle unterstellen dabei *a priori* in welcher Abhängigkeit einzelne Zeitreihen ihre Regime wechseln. Diese Regimewechsel werden üblicherweise als perfekt synchronisiert oder vollständig unabhängig vermutet. Keine dieser Annahmen scheint jedoch realistisch.

Die vorliegende Arbeit entwickelt ein multivariates Markov Regime-Switching Modell, welches die Abhängigkeit dieser Regimewechsel von den zugrundeliegenden Daten herleiten kann. Dieses flexible Modell erlaubt es, dass einzelne Subgruppen von Anlageinstrumenten durch individuelle Markov-Ketten gesteuert werden. Gleichzeitig weisen diese individuellen Regime-Prozesse eine sich dynamisch verändernde Synchronisation auf. Im Vergleich zu bestehenden Lösungsansätzen ist das Modell jedoch nicht auf bivariate Zeitreihen beschränkt. Um das Modell lösbar auszugestalten, wurde zudem ein neuer Faktorisierungsalgorithmus entwickelt, welcher es ermöglicht die regimeabhängige Korrelationsmatrix zu berechnen. Dieser Algorithmus reduziert die Zunahme an Modellparametern und stellt eine effiziente Lösung zur Sicherstellung von positiv semi-definiten Korrelationsmatrizen dar.

Der Aufbau des flexiblen Regime-Switching Modells ist dabei durch die im Rahmen dieser Arbeit durchgeführte Synchronisationsanalyse bestimmt. Die Analyse von univariaten Regime-Switching Modellen hat gezeigt, dass weder perfekt synchronisierte, noch vollständig unabhängige Regime-Zyklen empirisch beobachtbar sind. Vielmehr verändert sich die Synchronisation dieser Zyklen dynamisch über die Zeit hinweg. Einzelne Zeitreihen können dabei jedoch sehr ähnliche Eigenschaften aufweisen, was deren Modellierung über einen gemeinsamen Regime-Prozess rechtfertigt.

Die empirische Untersuchung des entwickelten Modells stützte sich dabei auf eine

Stichprobe von sechs internationalen Aktienmärkten. Die Analyse dieser Märkte bestätigte die getroffenen Annahmen. Das flexible Regime-Switching Modell weist einen stabilen Synchronisationsfaktor auf, welcher sich über die Untersuchungsperiode hinweg einmal wesentlich veränderte. Die berechneten Parameter für das Modell widerspiegeln dabei die individuellen Eigenschaften der zugrundeliegenden Zeitreihen und bestätigen damit dessen Validität. Des Weiteren weisen die Korrelationsmatrizen in gewissen "States" sehr tiefe oder sogar negative Werte auf. Diese Eigenschaften können von besonderem Wert für die Portfolio-Analyse sein.

Um schlussendlich die Performance des flexiblen Regime-Switching Modells zu bestimmen, wurde dessen Asset Allocation untersucht. Der Vergleich von verschiedenen linearen und Regime-Switching Modellen hat dabei deutlich die positiven Eigenschaften des flexiblen Modells für mittelfristige Investitionszeiträume hervorgehoben.

Chapter 1

Introduction

1.1 Motivation

Most financial markets reveal a cyclical behavior in their dynamics. This cyclical behavior is marked by occasional shifts, which are often shared by different individual markets. To analyze the joint dynamics of these market cycles, multivariate regime-switching models have become a common econometric technique. However, multivariate regime-switching models generally define the dependence structure among individual market cycles *a priori*.

Two polar assumptions emerge from this specification. The cycles in a multivariate model are either assumed to be fully independent or to be perfectly synchronized. In the independent case, each time series follows its individual regime dynamics. Consequently, Markov processes of individual markets are presumed independent, and regime-switches across time series are therefore purely idiosyncratic.

In the fully dependent case, on the other hand, all time series are governed by a common Markov chain and switch regimes contemporaneously. Shifts in the dynamics of individual time series are presumed to be perfectly synchronized across a sample. Phillips (1991) was among the first to analyze these extreme cases of synchronization. He remarked that the true model solution might lie somewhere in between full synchronization and perfect independence. Based on Phillips's remarks, two different lines of research have emerged. One line of research assumes a single Markov chain process, which governs all underlying assets. The regime dynamics of individ-

ual assets, however, might lead or lag this common Markov process (see Cakmakli, Paap & van Dijk, 2011).

The other line of research assumes that each asset follows a distinct regime cycle. And yet these individual cycles can show some intermediate degree of synchronization. This specification rejects the *a priori* assumption of full independence or perfect synchronization and infers the actual degree of dependence from the data (Camacho & Perez-Quiros, 2006). Leiva-Leon (2012a, 2012b) extended this idea of coherent cycles by allowing the degree of synchronization to vary over time.

The structures presented in the previous paragraphs come at the price of an extended parameter set and an increased model complexity. These shortcomings motivate the bivariate focus encountered in most research. By contrast, this thesis aims to extend the focus to multivariate models. Model complexity is thereby kept as low as possible by clustering time series with similar characteristics. Each cluster is driven by an individual Markov chain process. Consequently, synchronization is no longer measured among the Markov chains of two individual time series, but instead among the Markov chain dynamics of multivariate clusters. To keep the model traceable, a novel factorization algorithm for the regime-dependent correlation matrix is formulated. This algorithm scales down the increase in parameters and presents an efficient way of ensuring positive semi-definite correlation matrices.

1.2 Research Focus

This thesis aims to develop a multivariate regime-switching model governed by multiple Markov chains. Each Markov chain drives the regime-dynamics of a homogeneous subgroup of assets. The degree of synchronization across these Markov chains is time-varying and can lie anywhere between full independence and perfect synchronization. This structure allows the model to dynamically analyze higher-dimensional regime-switching problems and to determine their degree of synchronization. In addition, the model is able to reproduce most stylized facts of financial markets, which are mimicked by standard regime-switching models.

This thesis further aims to analyze the out-of-sample performance of different regime-

switching specifications. To quantify this performance, a dynamic asset allocation model is introduced. This asset allocation model accounts for regime-switches and enables the investigation of multistage portfolio optimization problems. The multi-period focus is thereby intended to capture the underlying regime dynamics.

This dissertation makes four main contributions to existing research. First, it presents different linear and nonlinear methods for measuring the synchronization of regime processes. Second, it introduces a flexible regime-switching model capable of capturing switching dynamics among individual groups of assets. Third, it develops a correlation factorization technique to keep the number of model parameters low and to guarantee positive semi-definiteness of the correlation matrices. Finally, it contributes to existing research on asset allocation under regime-switching.

1.3 Structure of the Thesis

Figure 1.1 surveys the three parts of this study. The *first part* consists of Chapters 2 to 4 and examines the standard regime-switching framework. Chapter 2 reviews the relevant literature and specifies the structure of univariate and multivariate regime-switching models. In Chapter 3, methods for measuring the degree of synchronization among regime processes are outlined. Chapter 4 then provides data analysis and presents initial regime-switching results. Further, the synchronization of univariate Markov regime processes is investigated.

The *second part* consists of Chapters 5 and 6. This part develops the flexible regime-switching model, which allows for asynchronous regime shifts. Chapter 5 starts by specifying the data- and the regime-generating process. It then presents inference techniques to reduce the number of model parameters. Finally, possible model extensions are analyzed. Chapter 6 illustrates empirical results for the previously specified models.

Chapter 7, the *third part* of this study, focuses on asset allocation under regime-switching. This chapter evaluates the performance and the value added of different regime-switching models presented throughout this work. The models are compared

by means of myopic and multistage dynamic asset allocation problems. Chapter 8 summarizes the main results and concludes this thesis.

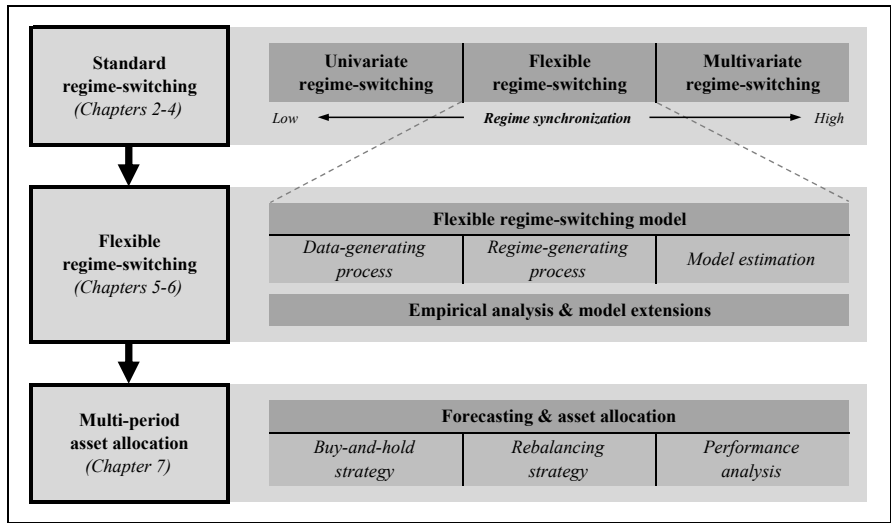


Figure 1.1: Structure of the thesis

Chapter 2

Theory of Markov Regime-Switching

The behavior of most economic and financial time series is marked by episodes of structural changes. These changes in behavior can be associated with events such as financial crises, significant changes in government policies, natural disasters, or wars. A simple way to model such structural changes could be to assume their occurrence as a foreseeable, deterministic event (Hamilton, 1994). Data prior to the event would be described by model 2-1 whereas data after the event would be described by model 2-2

$$y_t = \mu_1 + \varepsilon_t \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_1^2) \quad (2-1)$$

$$y_t = \mu_2 + \varepsilon_t \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_2^2), \quad (2-2)$$

where y_t is a return series, and μ_i and σ_i^2 define the mean and variance in period i (further $\mu_1 > \mu_2$ or $\sigma_1 < \sigma_2$ is assumed). However, if the time series dynamics have changed in the past, they might change again in the future. This aspect should be considered in the model specification. Moreover, the changes in dynamics should not be regarded as perfectly predictable, but rather as random events (Hamilton, 1994).

A suitable way of accounting for these features is to presume a stochastic process governing the transition from one market state to another. This stochastic process is described by the discrete-valued and unobservable random variable s_t , which indi-

cates the prevailing regime at time t

$$y_t = \mu_{s_t} + \varepsilon_t \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_{s_t}^2). \quad (2-3)$$

The distributional moments of this model no longer depend on deterministic events. Rather, they are conditioned on the underlying regime-variable. The variable itself follows a Markov chain process and needs to be inferred from the data.

The next sections will introduce this so-called regime-switching model in more detail. Section 2.1 surveys the literature on regime-switching models and their application to portfolio optimization. Section 2.2 then presents the regime-switching framework, which is specified by a data-generating process and a Markov chain. The last section extends the basic univariate framework to a multivariate regime-switching model and discusses the implications of this extension.

2.1 Literature Review

The origins of regime-switching models date back more than 50 years. Early research focused on a simple model that incorporates a single nonrecurring structural break. Quandt (1958) first attempted to estimate the parameters of this linear regression system and assumed that such a system obeys two separate regimes. Similar to the model in 2-1 and 2-2, he assumed a point in time at which the system switches the regime and data are described by a different regression equation. Using maximum likelihood estimation, Quandt was able to infer the corresponding turning point and to determine the regression parameters. However, his model considers only the possibility of a single switch and was tested in a hypothetical sampling experiment. In 1972, Quandt extended his earlier model to allow for multiple regime-switches. He introduced the λ -method, where λ and $1 - \lambda$ represent unknown probabilities for the observed data points being driven by regime 1 or 2. He applied this model to the US housing market between June 1959 and November 1969, and obtained significant results for the two regimes.

Goldfeld and Quandt (1973) relaxed the assumption of constant regime probabilities and introduced the first Markov-switching model. They assumed that the current regime depends on its preceding state through a Markov chain process. In order to capture state-dependency, they introduced a transition probability matrix, which governs the transitions across states. They applied this so-called τ -method to the same data sample as Quandt (1972) and obtained similar results as for the λ -method. Over time, models became more complex and sophisticated due to computational improvements and data availability. As a result, Hamilton (1989) presented a new approach to proceed with regime-switching analysis. He showed that the regimes – up until then often considered to be exogenous – cannot be observed directly. Instead, a probabilistic inference on their existence must be drawn. Hamilton (1989, 1994, 2005) proposed a nonlinear iterative filter, in which the likelihood function is calculated as a by-product of the algorithm. This algorithm is similar to the Kalman filter, but underlies a nonlinear process. Hamilton (1989) applied his model to US real GNP data between March 1951 and December 1984. He found two distinct regimes of positive and negative GNP growth.

Regime-switching models have received considerable interest in different areas of research. Some important contributors are Hamilton (1989) on macroeconomic variables; Ang and Bekaert (1999, 2002a, 2004) and Ang and Chen (2002) on equities; and Ang and Bekaert (2002b, 2002c), Garcia and Perron (1996), and Gray (1996) on interest rates. Moreover, Guidolin and Timmermann (2005a, 2005b, 2006, 2007) and Schwendener (2010) researched multiple asset classes.

Early research on Markov regime-switching mainly focused on univariate time series. Phillips (1991) extended this research by combining the business cycles of two countries in a joint setup. He aimed to evaluate the transmission of the two business cycles, where each cycle was represented through an individual regime process. Phillips was the first to analyze this transmission of cycles. He examined the two extreme cases of fully independent and perfectly correlated regime-switches. Moreover, he analyzed intermediate cases, where one regime process either leads or lags the other. Phillips showed that the consideration of the business cycles of two countries improves both the forecast and the fit of the model. However, his results for

the pairwise synchronization of the US, Canada, Germany, and the UK turned out to be insignificant. The results revealed that the international transmission of business cycles is mainly due to worldwide shocks and not to inter-country transmissions.

In 2006, Bengoechea et al. revisited Phillips's approach (1991). Similar to Phillips, they allowed for an intermediate solution between the two extreme cases of independent and perfectly synchronized regime cycles. Bengoechea et al. (2006) linearly combined these two cases in a bivariate framework to determine the degree of pairwise synchronization.

Camacho and Perez-Quiros (2006) readopted this approach. They analyzed the real GDP data for a sample of G7 countries between March 1980 and June 2004 to determine the level of regime synchronization. Camacho and Perez-Quiros found significant results for the business cycle synchronization inferred from the GDP data. The results indicate that European and Anglo-Saxon countries form two distinct groups with regard to their business cycles. Business cycles within a group are synchronized, whereas business cycles across groups are desynchronized.

Cakmakli et al. (2011) made another important contribution to the analysis of regime synchronization. They developed a Markov-switching vector autoregressive model (MS-VAR) that allows for the imperfect synchronization of regimes. Cakmakli et al. applied their model to a sample of size-based portfolios and showed that large- and small-cap equity portfolios switch contemporaneously into extreme regimes (bull and bear states). However, for switches to moderate regimes (transition states), the large-cap portfolio leads the small-cap portfolio by one month.

Further contributions in the area of regime synchronization were made by Krolzig (1997) on perfectly synchronized regime cycles, and by Ang and Bekaert (2002b), Krolzig (2001), Smith and Summers (2005), and Otranto (2010) on fully independent cycles. Hamilton and Perez-Quiros (1996), Paap, Segers and van Dijk (2009), and Cakmakli et al. (2011) considered the case of a common but phase-shifted regime cycle. Finally, Bengoechea et al. (2006), Camacho and Perez-Quiros (2006), and Leiva-Leon (2012a, 2012b) considered the case of limited synchronization among individual regime processes. The most important contributors and their models are also presented in Table 5.1 (see Chapter 5).

The emergence of bivariate and multivariate regime-switching models has also drawn attention to portfolio selection. Investors are faced with a time-varying portfolio choice problem, given the regime-dependent means and potentially changing comovement structures.

Ang and Bekaert (1999, 2002a) were the first to analyze these dynamics in a portfolio optimization context. They applied a multivariate regime-switching model to an international equity portfolio. The resulting regime-probabilities were used as input for a multi-period portfolio optimization problem of an investor with constant relative risk aversion (CRRA). Ang and Bekaert compared their results to a buy-and-hold strategy and found significant differences. A regime-switching allocation thereby proved to be of particular economic value for longer investment horizons. Moreover, the inclusion of a risk-free asset in the investment universe added further value to the regime-switching strategy.

Guidolin and Timmermann (2007) extended the research on asset allocation under regime-switching. In contrast to Ang and Bekaert (1999, 2002a), however, they applied the regime-switching model to a portfolio of stocks, bonds, and cash. Further, they assumed the regimes to be unobservable, whereas Ang and Bekaert (1999, 2002a) required the prevailing regime to be known with certainty. The results of Guidolin and Timmermann (2007) indicate that portfolio allocation varies considerably among regimes. Their results further confirm the importance of regime-switching information for asset allocation decisions.

Further research in this area has been conducted by Morger (2006), Frauendorfer, Jacoby and Schwendener (2007), Leippold and Morger (2007), Schwendener (2010), and Angeloni and Sverrisson (2012). A dynamic asset allocation model is also introduced below (Chapter 7). This model will evaluate different regime-switching specifications in terms of out-of-sample performance. The results of this evaluation will finally help to determine the power of the flexible regime-switching model developed in this thesis.

2.2 Regime-Switching Framework

The above literature review has revealed the versatility of regime-switching models. Despite this versatility, most of these models share a common underlying structure. The following sections present the elements of this structure in a general regime-switching framework.¹ This framework first describes the data-generating process (DGP) and the Markov chain. It then examines the likelihood function and the regime-inference procedure.

2.2.1 Data-Generating Process

The data-generating process of a regime-switching model underlies the general structure²

$$\begin{aligned}
 y_t &= f(\tilde{y}_{t-1}, \tilde{z}_{t-1}, \tilde{s}_t) + \varepsilon_{yt} \\
 z_t &= c(s_t) + A(s_t) \tilde{z}_{t-1} + \varepsilon_{zt} \\
 \varepsilon_t &= (\varepsilon'_{yt}, \varepsilon'_{zt})' \\
 \varepsilon_t \mid \tilde{s}_t &\sim \mathcal{N}(0, \Sigma(\tilde{y}_{t-1}, \tilde{z}_{t-1}, \tilde{s}_t)),
 \end{aligned} \tag{2-4}$$

where y_t is the $(N \times 1)$ vector of observable returns for the N assets at time t , z_t is a $(M \times 1)$ vector of predictive instruments, and s_t indicates the regime at time t . Further, $f(\cdot)$ and $\Sigma(\cdot)$ define the conditional mean vector and covariance matrix, and the tilde specifies a time series process, for instance, $\tilde{y}_{t-1} = \{y_i\}_{i=1}^{t-1}$.³ Finally, c and A describe the (regime-dependent) intercept and factor loadings of a VAR specification.

Based on this general structure, the DGP of a particular regime-switching model can be specified in different ways. For example, the conditional mean of y_t might simply

¹This framework follows the remarks in Hamilton (1994, 2005) and in Guidolin (2013).

²This structure follows the illustrations in Ang and Bekaert (1999) and might ignore special cases.

³In most specifications, the function will depend only on the p most recent observations of a time series $\tilde{y}_{t-p,t-1} = \{y_i\}_{i=t-p}^{t-1}$. However, for ease of exposition, the $t-p$ subscript is omitted.

depend on the current regime s_t . The mean could additionally depend on the most recent realization of the predictive instruments z_{t-1} (for instance, dividend yields, earnings yields, or short rates).

Likewise, a more complex model might follow a Markov switching vector autoregressive (MS-VAR) structure. In this case, $f(\cdot)$ would depend on the p most recent return observations of all N assets $\{y_i\}_{i=t-p}^{t-1}$. If the VAR specification further implies a switching in means, f would also depend on the p most recent regime observations $\{s_i\}_{i=t-p}^t$.⁴ Krolzig (1997) provides a good overview of the specification of such MS-VAR models and of their limitations.

Similar flexibility applies to the conditional covariance matrix Σ . The conditionality of the covariance matrix allows for regime-dependent ARCH and GARCH structures in the error terms ε_t . Examples of these structures are presented by Gray (1996), Ramchand and Susmel (1998b, 1998a), Hamilton and Susmel (1994), and Ang and Bekaert (1999, 2002a).⁵ To keep a model parsimonious, orthogonality between return errors ε_{yt} and instrument errors ε_{zt} is generally assumed. This results in a block-diagonal structure of the covariance matrix Σ .

Formula 2-4 reveals the diversity of regime-switching models. Each individual parameter can theoretically be defined as either regime-dependent or regime-independent. Despite this flexibility, DGPs of standard regime-switching models generally follow a simple structure. This also applies to the benchmark models specified in the following paragraphs.

In standard univariate regime-switching models, the conditional mean and the conditional variance are often specified as $f(\tilde{y}_{t-1}, \tilde{z}_{t-1}, \tilde{s}_t) = \mu_{s_t}$ and as $\Sigma(\tilde{y}_{t-1}, \tilde{z}_{t-1}, \tilde{s}_t) = \sigma_{s_t}^2$. These assumptions result in the following data-generating process

$$y_t = \mu_{s_t} + \sigma_{s_t} \varepsilon_t \quad \varepsilon_t \sim \mathcal{N}(0, 1), \quad (2-5)$$

where ε_t is a Gaussian white noise process.⁶

⁴Krolzig (1997) shows that this does not apply if the model underlies a switching in the intercept. In this case, f would only depend on the current regime s_t .

⁵Similarly, the correlation matrices might follow a (regime-dependent) DCC structure as presented by Billio, Caporin and Gobbo (2006) and Otranto (2010).

⁶This specification ignores predictive instruments and AR dynamics to keep the model complexity as

Equation 2-5 shows that both mean and variance terms are regime-dependent. Their specification is equal to that in the introductory example and is intentionally kept simple. In what follows, this setup is referred to as the *fully specified benchmark model* for univariate time series.

Standard multivariate models are structured in a similar way. The conditional mean vector and the conditional covariance matrix are defined as $f(\tilde{y}_{t-1}, \tilde{z}_{t-1}, \tilde{s}_t) = \boldsymbol{\mu}_{s_t}$ and as $\Sigma(\tilde{y}_{t-1}, \tilde{z}_{t-1}, \tilde{s}_t) = \Sigma_{s_t}$. The data-generating process is then specified by

$$\mathbf{y}_t = \boldsymbol{\mu}_{s_t} + \Sigma_{s_t}^{1/2} \varepsilon_t \quad \varepsilon_t \sim \mathcal{N}(0, I), \quad (2-6)$$

where \mathbf{y}_t indicates the $(N \times 1)$ return vector for the N assets at time t , $\boldsymbol{\mu}_{s_t}$ is the $(N \times 1)$ vector of mean returns conditional on the prevailing regime s_t , Σ_{s_t} is the $(N \times N)$ regime-dependent covariance matrix, and I indicates the identity matrix. Below, this DGP is referred to as the *fully specified benchmark model* for multivariate time series.

2.2.2 Markov Chains

The DGP specification revealed the conditionality of the moment and covariance parameters on the regime variable s_t . A Markov chain represents one of the simplest processes for describing the probability law of this random variable. This discrete-state stochastic process consists of a sequence of random state variables. The state variable s_t of a first-order Markov chain is thereby defined as being only serially dependent on its preceding state

$$\begin{aligned} \Pr(s_t = j \mid s_{t-1} = i, s_{t-2} = k, \dots, y_{t-1}, y_{t-2}, \dots, y_1) \\ = \Pr(s_t = j \mid s_{t-1} = i) = p_{ij}, \end{aligned} \quad (2-7)$$

low as possible. Hamilton and Susmel (1994) showed that the latter are not observed for monthly data but only for data with higher frequency.

where p_{ij} gives the probability that the regime will change from state i to state j . Equation 2-7 shows that the current regime depends on the past only through its most recent state. This process is defined as a K -state first-order Markov chain with transition probabilities $\{p_{ij}\}_{i,j=1,2,\dots,K}$. These probabilities are usually collected in a $(K \times K)$ transition probability matrix

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{K1} \\ p_{12} & p_{22} & \cdots & p_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1K} & p_{2K} & \cdots & p_{KK} \end{bmatrix},$$

where the element in row j and column i of matrix \mathbf{P} equals the transition probability p_{ij} . It holds that each column of \mathbf{P} sums up to unity $\sum_{j=1}^N p_{ij} = 1$ and that each element in \mathbf{P} is non-negative ($p_{ij} \geq 0, \forall i, j$).

In general, this thesis follows a two-state first-order Markov chain. It further presumes the Markov chain to be *irreducible*, *ergodic*, and *stationary*. These assumptions support the estimation of stable and efficient models by avoiding singularities and absorbing states.

A Markov chain is called *irreducible* if there is only a single communicating class in the state space. This communicating class exists if every state j is accessible from every state i within finite time ($\mathbf{P}_{ij}^n > 0$ for some exponent $n \geq 0; i \leftrightarrow j$). Stated differently, it is possible to proceed from every state to every other state, however, not necessarily in a single step (Hamilton, 1994; Grinstead & Snell, 2012).

Ergodicity extends the definition of irreducibility by imposing the existence of a stationary vector of probabilities such that

$$\pi = \mathbf{P}\pi,$$

where π is the vector of ergodic state probabilities. These probabilities are consistent with the unconditional or steady-state probabilities (Hamilton, 1994). Ergodic

probabilities can therefore be interpreted as long-run or unconditional state probabilities.⁷

Finally, a Markov chain is called *stationary* or *time-homogeneous* if the transition probabilities are time-invariant ($p_{t,ij} = p_{ij}$).⁸

2.2.3 Likelihood Function and Parameter Estimation

The regime-switching example at the beginning of this chapter summarized what the last two sections have reviewed in greater detail: At every point of time, the observed variable y_t is drawn from one of K different distributions (Hamilton, 1994). Each of these distributions is specified by a density function, which is conditional on the regime variable s_t . In case the DGP follows a normal distribution, the K conditional density functions are specified as

$$\eta_{jt} = f(y_t | s_t = j, \Omega_{t-1}; \theta) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp \left[-\frac{(y_t - \mu_j)^2}{2\sigma_j^2} \right], \quad (2-8)$$

for $j = 1, \dots, K$, where θ is the vector of population parameters, and where Ω_{t-1} represents information up to time $t - 1$.⁹ In this simple case, the parameter vector consists of individual mean μ_j and variance σ_j^2 parameters for each single regime. Equation 2-8 shows the distribution's conditionality on the unobservable state variable s_t . This variable indicates which of the K regimes the process is in at time t .

⁷Similar to ergodicity, irreducibility implies the existence of only one eigenvalue equal to unity (Hamilton, 1994). It does not, however, limit the remaining eigenvalues to lying inside the unit circle. As a result, more than one eigenvalue might be located on the unit circle. Therefore, the matrix \mathbf{P}^m does not necessarily converge with $m \rightarrow \infty$ and no stable unconditional probabilities can be derived. Hence, whereas ergodicity implies irreducibility, this relation cannot be reversed. Not all irreducible Markov chains are also ergodic.

⁸This thesis assumes time-invariant transition probabilities to keep model complexity low. For example, Diebold, Lee and Weibach (1994) and Kim and Nelson (1999) presented alternative model specifications with time-varying transition probabilities (TVTP).

⁹Note that Ω_{t-1} might have a broader focus than the time series process $\tilde{y}_{t-1} = \{y_i\}_{i=1}^{t-1}$. Ω_{t-1} summarizes all exogenous time series information that is available up until time $t - 1$.

However, the state variable is unobservable and therefore needs to be inferred from the data.

Consequently, the specification of a regime-switching model consists of two steps. In a first step, the model parameters are estimated. Given these estimates, the regime probabilities are then inferred in a second step.¹⁰

To investigate the two steps individually, the regime-generating process needs to be simplified. Hamilton (1994) proposes the unobserved regime variable s_t to be drawn from an unconditional probability distribution

$$\Pr(s_t = j; \theta) = \pi_j, \quad (2-9)$$

where π_j is the time-invariant probability of state j . The regime probabilities are independent of past regimes and of past information. This assumption simplifies the regime inference, as the state probabilities are constant and do not follow a Markov chain process. The resulting model represents an i.i.d. mixture model. The corresponding parameter set θ consists of regime-dependent means, regime-dependent variances, and unconditional regime probabilities: $\theta = \{\mu_1, \dots, \mu_K, \sigma_1^2, \dots, \sigma_K^2, \pi_1, \dots, \pi_{K-1}\}$. These parameters can be estimated by maximizing the likelihood function through numerical optimization (Hamilton, 2005).

The specification of this likelihood function relies on the previously defined unconditional state probabilities π_j and on the conditional density functions η_{jt} for $j = 1, \dots, K$.

Note the definition of conditional probability

$$\Pr(A|B) = \frac{\Pr(A \text{ and } B)}{\Pr(B)},$$

¹⁰This approach is valid for the maximum likelihood estimation (MLE) and for the expectation maximization (EM) algorithm (see Dempster, Laird & Rubin, 1977). For example, the Bayesian procedure uses a different approach. It estimates both the parameters and the regimes in the same step. However, in general, the present study relies on the two-step MLE approach.

which can be restated as

$$\Pr(A \text{ and } B) = \Pr(A|B) \cdot \Pr(B).$$

Consequently, the product of the regime-dependent probability function η_{jt} and of the marginal regime probability π_j results in the joint probability density function

$$p(y_t, s_t = j | \Omega_{t-1}; \theta) = f(y_t | s_t = j, \Omega_{t-1}; \theta) \cdot \Pr(s_t = j; \theta). \quad (2-10)$$

This function defines the density of the joint event of y_t being drawn from the regime-dependent density function η_{jt} and of s_t being in state j .

The unconditional density then results from summing the joint functions over all K possible states

$$f(y_t | \Omega_{t-1}; \theta) = \sum_{j=1}^K p(y_t, s_t = j | \Omega_{t-1}; \theta).$$

Finally, the log-likelihood $\mathcal{L}(\theta)$ of the time series can be found by summing the log-densities over time

$$\mathcal{L}(\Omega_T; \theta) = \log(f(y_1, y_2, \dots, y_T | \Omega_0; \theta)) = \sum_{t=1}^T \log f(y_t | \Omega_{t-1}; \theta),$$

given that the regime variable is i.i.d. across time. Fortunately, this i.i.d. assumption holds for the simple time-invariant probability distribution defined in 2-9 as well as for the Markov chain process. The parameter set θ is then estimated by maximizing the likelihood function, given the constraints $\sum_j \pi_j = 1$ and $\pi_j \geq 0, \forall j$.

2.2.4 Regime Inference

The previous section assumed time-invariant regime probabilities. This assumption simplified the description of the likelihood function. In a full-scale regime-switching model, however, the regimes follow a Markov chain process. This process defines the current state as depending on its history only through the most recent state. Because the regime process is unobservable, it needs to be inferred from the data. This inference requires the full parameter set, which consists of regime-dependent means μ_j , of regime-dependent variances σ_j^2 , and of the matrix of transition probabilities \mathbf{P} .¹¹

Given the set of population parameters $\theta = \{\mu_1, \dots, \mu_K, \sigma_1^2, \dots, \sigma_K^2, p_{11}, \dots, p_{K(K-1)}\}$, the probability of an individual regime being responsible for the observation of y_t can be inferred. This inference is now expressed in a conditional state probability $\Pr(s_t = j | \Omega_\tau; \theta)$, where the information set Ω_τ includes y_t up to time τ , $\{y_t\}_{t=1}^\tau$. Conditional on the position of τ relative to the current time t , the estimation of state probabilities is either a forecasting ($\tau < t$), filtering ($\tau = t$), or smoothing ($\tau > t$) procedure. These three procedures will be analyzed below.

Forecasting

Let $\Pr(s_t = j | \Omega_t; \theta)$ define the probability that the regime is in state j , given information up to time t . The probabilities for the K individual regimes are then collected in a $(K \times 1)$ vector $\xi_{t|t}$. One-period-ahead regime forecasts result when multiplying the probabilities at time t by the transition probability matrix \mathbf{P}

$$\xi_{t+1|t} = \mathbf{P}\xi_{t|t}.$$

Hence, the forecasted regime-probabilities $\xi_{t+1|t}$ are still conditioned on time t in-

¹¹The multivariate model follows the same approach. However, its parameter set additionally contains correlation terms.

formation. Similarly, an m -period-ahead forecast is specified by

$$\xi_{t+m|t} = \mathbf{P}^m \xi_{t|t},$$

where \mathbf{P}^m represents the transition probability matrix to the m th power. Element p_{ij} in this extended transition matrix defines the probability that regime i at time t will be followed by regime j , m periods in the future. However, the forecasted probabilities again depend on information at time t .

The forecasting or prediction procedure is commonly used to form expectations about future events. Examples include the application of regime-switching to asset allocation problems (Ang & Bekaert, 1999, 2002a; Morger, 2006; Ammann & Verhofen, 2006; Verhofen, 2006; Frauendorfer et al., 2007; Guidolin & Timmermann, 2007) or to out-of-sample regime-predictability tests (Hamilton & Perez-Quiros, 1996; Paap et al., 2009).

Chapter 6 will apply the forecasting procedure to analyze the models in this thesis for their out-of-sample predictability power. Further, in Chapter 7 the forecasting procedure is applied to an asset allocation problem. Moreover, prediction also constitutes the first phase of the filtering procedure.

Filtering

Hamilton (1989) was the first to propose a non-linear iterative filtering technique to infer state probabilities from the data. This filtering procedure – known as Hamilton filter – consists of two steps: a prediction step and an updating step. Following these two steps, the filtered state probabilities are inferred from the recursive algorithm

$$\xi_{t|t-1} = \mathbf{P} \xi_{t-1|t-1} \tag{2-11}$$

$$\xi_{t|t} = \frac{\xi_{t|t-1} \odot \eta_t}{\mathbf{1}' (\xi_{t|t-1} \odot \eta_t)} = \frac{\mathbf{P} \xi_{t-1|t-1} \odot \eta_t}{\mathbf{1}' (\mathbf{P} \xi_{t-1|t-1} \odot \eta_t)}, \tag{2-12}$$

where \odot defines an element-by-element multiplication, and where $\mathbf{1}$ is a $(K \times 1)$

vector of ones. Equation 2-11 depicts the known prediction or forecasting step. On the other hand, equation 2-12 defines the updating step and relies on Bayes' theorem.¹²

The numerator in 2-12 describes element-wise products of the K predicted state probabilities $\xi_{j,t|t-1} = \Pr(s_t = j | \Omega_{t-1}; \theta)$ and their corresponding conditional density functions η_{jt} . The resulting elements specify joint densities for the K states (see formula 2-10). The denominator of equation 2-12 depicts the sum of all joint densities, which results in the unconditional density of y_t

$$f(y_t | \Omega_{t-1}; \theta) = \mathbf{1}' (\mathbf{P} \xi_{t-1|t-1} \odot \eta_t). \quad (2-13)$$

The filtering algorithm is initialized with the state probabilities $\xi_{1|0}$. These starting values are either predetermined or treated as unknown random parameters. In the latter case, they represent part of the parameter set.

After the parameter set θ is estimated (see Section 2.2.3), the nonlinear filtering procedure in equations 2-11 and 2-12 is run iteratively. Equation 2-11 predicts the one-period-ahead regime probabilities. In a next step, equation 2-12 filters the probabilities as new information about y_t becomes available.

Formula 2-14 shows that the log-likelihood function \mathcal{L} – which is subject to maximization – is a by-product of the filtering process

$$\mathcal{L}(\Omega_T; \theta) = \sum_{t=1}^T \log f(y_t | \Omega_{t-1}; \theta). \quad (2-14)$$

Smoothing

The smoothed regime inference at time t considers information available until a later date T . Kim (1994) presents the following recursive algorithm to solve this problem

$$\xi_{t|T} = \xi_{t|t} \odot (\mathbf{P}' \cdot (\xi_{t+1|T} \oslash \xi_{t+1|t})), \quad (2-15)$$

¹²Note that Bayes' theorem is defined by $\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)}$.

where \oslash denotes an element-by-element division. The algorithm starts with the inferred regime-probabilities at time T and works backwards until time t . Formula 2-15 further shows that the smoothing procedure requires the filtered state probabilities as input variables.

The model estimation presented so far underlies a two-step approach. In a first step, population parameters are estimated. Maximum likelihood estimation (MLE) thereby presents a common technique for this task. Based on the estimated parameters, the regime probabilities are then inferred in a second step. This estimation technique belongs to the category of frequentist inference.

The Gibbs sampling procedure represents an alternative estimation technique (see, for example, Geman & Geman, 1984; Albert & Chib, 1993; McCulloch & Tsay, 1994; Chib, 1996). This technique is a Markov chain Monte Carlo (MCMC) algorithm, which belongs to the category of Bayesian inference. Gibbs sampling considers both the parameters θ and the regime process s_t as random variables.¹³ Consequently, the regime process itself becomes part of the parameter set and is no longer inferred from the estimated parameters. The Gibbs sampler is thereby able to estimate higher dimensional problems where frequentist approaches are less feasible. This thesis, however, relies on maximum likelihood estimation due to its favorable properties regarding the DGP specification (see also Kim & Nelson, 1999).

2.3 Extension to Multivariate Regime-Switching

The baseline framework is easily extended to multivariate time series. This extension affects only the data-generating process and the conditional density function. The random variable s_t , on the other hand, defines the prevailing regime irrespective of

¹³Below, the parameter space θ is specified according to the frequentist approach. The parameter set contains all population parameters, but excludes the T regime variables, where T refers to the length of the observation period.

the number of underlying assets. Consequently, the Markov chain structure remains unchanged.

Section 2.2.1 already extended the data-generating process to multivariate time series. Its covariance matrix Σ_{s_t} is subsequently decomposed into standard deviation and correlation terms

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu}_{s_t} + \Sigma_{s_t}^{1/2} \varepsilon_t & \varepsilon_t &\sim \mathcal{N}(0, I) \\ \Sigma_{s_t} &= D_{s_t} R_t D_{s_t} \\ R_t &= R_{s_t}, \end{aligned} \tag{2-16}$$

where D_{s_t} is a $(N \times N)$ diagonal matrix composed of regime-dependent standard deviations, and where R_{s_t} is a regime-dependent $(N \times N)$ correlation matrix.

The corresponding conditional probability density function (pdf) follows a standard normal distribution. Similar to the univariate conditional pdf in 2-8, the multivariate conditional pdf is specified as

$$\begin{aligned} \eta_{jt} &= f(\mathbf{y}_t | s_t = j, \Omega_{t-1}; \theta) \\ &= \frac{1}{\sqrt{(2\pi)^N |D_j R_j D_j|}} \exp \left(\left[-\frac{1}{2} (\mathbf{y}_t - \boldsymbol{\mu}_j) D_j^{-1} R_j^{-1} D_j^{-1} (\mathbf{y}_t - \boldsymbol{\mu}_j)' \right] \right) \\ &\quad \text{for } j = 1, \dots, K, \end{aligned} \tag{2-17}$$

where $|\cdot|$ describes the determinant. Formula 2-16 showed the additional covariance parameters that need to be estimated in the multivariate model. In 2-17, these are again decomposed into standard deviations and correlations.

The resulting parameter sets of the univariate θ_{uv} and the multivariate models θ_{mv} are then defined by

$$\theta_{uv} = \{\mu_1, \dots, \mu_K, \sigma_1^2, \dots, \sigma_K^2, p_{11}, \dots, p_{K(K-1)}\},$$

and by

$$\theta_{mv} = \{\mu_{1,1}, \dots, \mu_{1,K}, \dots, \mu_{N,1}, \dots, \mu_{N,K}, \sigma_{1,1}^2, \dots, \sigma_{N,K}^2, \\ \rho_{21,1}, \dots, \rho_{N(N-1),K}, p_{11}, \dots, p_{K(K-1)}\},$$

where $\mu_{i,k}$ and $\sigma_{i,k}^2$ are the mean and variance of asset i in regime k , and where $\rho_{i,j,k}$ refers to the correlation between asset i and j in regime k .

Estimation of the Covariance Matrix

The presented covariance decomposition enables an individual estimation of standard deviation and correlation parameters. Bollerslev (1990), Tse and Tsui (2002), Engle (2002), and Barnard, McCulloch and Meng (2000) applied this covariance decomposition to constant conditional correlation (CCC) and to GARCH structures. Pelletier (2006) further extended this technique to regime-switching models.

The covariance decomposition simplifies the parameter estimation in three ways: First, it allows one to separately specify the regime-dependence of standard deviations and correlations. For example, a restriction of the switching in correlations can significantly reduce the parameter set of high-dimensional time series.¹⁴

Second, the model estimation requires the specification of parameter boundaries. These boundaries are easier to define for standard deviations and correlations than for covariances. In case the correlations are additionally specified to be regime-dependent, the variance and covariance estimates are almost unbounded (Geweke & Amisano, 2003; Pelletier, 2006). The only limitation applies to variances, which are restricted to be non-negative. Any further boundary conditions are imposed on the correlation matrix.

Finally, the decomposition enables the estimation of positive semi-definite covariance matrices. A covariance matrix needs to be positive semi-definite (PSD) at every point in time (Pelletier, 2006). However, maximum likelihood estimation optimizes

¹⁴Pelletier (2006), however, clearly demonstrated the importance of both switching standard deviation and correlation parameters. Therefore, such a restriction is only useful for very large samples (see, for example, Hamilton & Owyang, 2012).

the parameter set element-by-element. Consequently, the resulting matrices are not necessarily positive semi-definite. A non-PSD covariance results in a negative determinant and is non-invertible. Accordingly, the conditional pdf in equation 2-17 would create infeasible results or values near infinity. This would cause the optimization algorithm to stop at a local optimum, away from the global maximum. Covariance decomposition simplifies this problem. Only the correlation matrix needs to be PSD. Standard deviations are simply restricted to be non-negative.

Again, different ways exist to guarantee the PSD structure of a correlation matrix. The most obvious way is to test the matrix for positive semi-definiteness after each iteration of the MLE optimization. An estimate is rejected if the correlation matrix has a non-positive determinant. Despite its simplicity, this approach has a very slow convergence rate.

Pelletier (2006) alternatively proposed a Cholesky decomposition to ensure PSD results. The Cholesky decomposition presents a more efficient way of imposing the properties of a correlation matrix on R_{s_t} . It is given by

$$R_{s_t} = L_{s_t} L_{s_t}',$$

where L_{s_t} is a lower triangular matrix, and where L_{s_t}' is its transpose. Pelletier (2006) showed that the constraints imposed on this matrix are very loose. The elements of L_{s_t} should lie between -1 and $+1$. Further, the diagonal elements of the resulting correlation matrix should equal unity. These simple constraints will automatically assure a correlation matrix with PSD properties. Moreover, the constraints will guarantee off-diagonal correlation values to lie between -1 and $+1$.

The off-diagonal elements of $L_{s_t,ij}$, where $i > j$, are the parameters to be estimated. Following this approach, the diagonal elements of L_{s_t} can be calculated as

$$L_{ii} = \sqrt{R_{ii} - \sum_{m=1}^{i-1} L_{im}^2} \quad \text{for } i = 1, \dots, K, \quad (2-18)$$

where R_{ii} refers to the i th diagonal element of the correlation matrix and hence

equals unity. For ease of exposition, the corresponding regime s_t for L and R is omitted. The structure of 2-18 clearly reveals the previously stated constraints on the matrix L_{s_t} . Further, the resulting diagonal elements L_{ii} are restricted to be real-valued, which guarantees that the solution is unique.

The parameter estimation itself is very intuitive. MLE optimizes for the parameters in L_{s_t} instead of the correlations in R_{s_t} . Thereby, the $K \cdot N(N - 1)/2$ correlation parameters in θ_{mv} are replaced by the $K \cdot N(N - 1)/2$ lower triangular parameters in L_{s_t} for $s_t = 1, \dots, K$, where N defines the number of assets. Once these parameters are estimated, the correlation matrix is reassembled within the likelihood function. For this purpose, the lower triangular matrix is multiplied by its transpose. Because MLE optimizes for the lower triangular values of L_{s_t} and because the solution is unique ($L_{s_t^*;ii} \notin \mathbb{R}^+, \forall i$), the resulting correlation matrix will always be PSD.

Markov Chains in Multivariate Models

The Markov chain of a univariate model governs a single underlying asset. Multiple assets result in multiple univariate Markov chain processes, and any joint dynamics among these processes are not considered.¹⁵ Nevertheless, there is still a chance that individual Markov chains will reside in the same regime. For example, the joint regime probability of two assets is defined by

$$\Pr(s_t^a = s_t^b) = \sum_{k=1}^K \Pr(s_t^a = k) \Pr(s_t^b = k) > 0.$$

In a standard multivariate model, by way of contrast, all assets are driven by a common Markov chain. At any point, the individual time series jointly reside in the same regime. Switches in regimes happen contemporaneously across time series and are therefore perfectly correlated. Such a common regime process increases model stability. At the same time, however, this process might be less representative

¹⁵By construction, the univariate specification also ignores any covariance in returns across individual assets. This implication is contradictory, as independent regime processes do not necessarily imply independent return processes.

for individual time series. Regime dynamics of individual assets are aggregated in a common cycle and might fade out (Dueker & Sola, 2008). A univariate process, on the other hand, infers only the regime dynamics of its underlying asset. The process therefore specifically resembles the switching-behavior of this asset.

Figure 2.1 depicts an example of regime-switching in equity data. This example comprises three individual country indices. The charts show smoothed regime probabilities for state 1, which resembles the properties of a bull market. The top graph presents univariate state probabilities for the US, the UK, and Germany. These countries show slightly different regime cycles compared to each other. Visual inspection reveals that the regimes are neither fully independent nor perfectly synchronized.

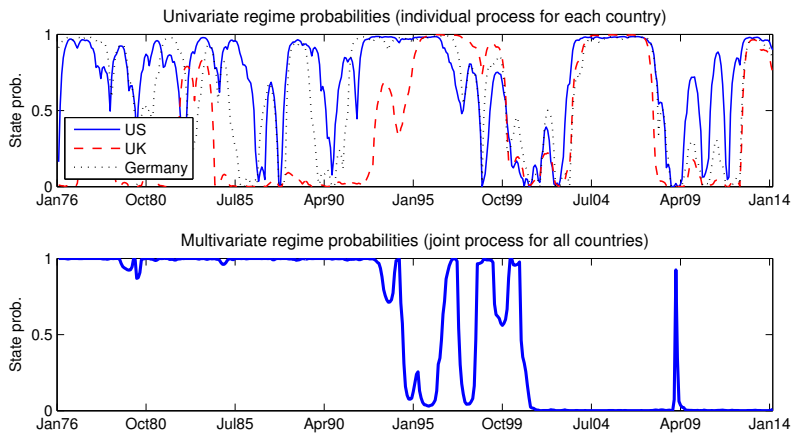


Figure 2.1: *Smoothed state probabilities of the univariate and multivariate benchmark regime-switching models for the period from January 1976 to February 2014. The regime processes depict the probabilities of residing in a bull market regime.*

The bottom graph depicts a multivariate regime process for the time series of the three countries. The figure shows that the multivariate process usually switches regimes when the univariate regime processes behave similarly (for example, when they jointly switch states). However, when the univariate regime processes differ

from each other, the multivariate process reveals a more stable behavior and does not necessarily cover the regime changes of individual assets. Hence, the multivariate regime process constitutes only an intermediate solution. It does not necessarily explain the switching dynamics of any underlying asset appropriately.

The previous example illustrates possible differences among the regime processes of individual assets and highlights the principal research questions of this thesis: How do the Markov chains of individual assets interact? Moreover, how can these dynamics be modeled in a multivariate setup without aggregating all information in a single regime process? To determine this interaction, the next chapters will analyze the degree of synchronization among individual regime processes. The results of this analysis will provide further insights about the best model structure. Chapters 5 and 6 will then present a hybrid model solution, which accounts for these features.

Chapter 3

Synchronization of Markov Chains

The previous analysis revealed that univariate and multivariate regime-switching models underlie the same Markov chain specification. However, analysis also showed that the Markov chain in a univariate model governs a single underlying asset, whereas it jointly governs a multitude of assets in a multivariate model.

In the univariate case, the analysis of multiple assets results in multiple regime processes. The univariate specification, however, reveals no further information about the joint dynamics of these individual processes.¹ Each asset can switch its regime irrespective of the other assets, and hence any joint dynamics are ignored. Common multivariate regime-switching models, in contrast, imply that underlying assets switch their regimes contemporaneously or fully independently.² However, these two polar cases seem unrealistic.

In practice, the degree of synchronization is expected to lie somewhere in between. Figure 2.1 exemplified this intermediate dependence among country-specific regime cycles. The analysis of univariate regime-switching models therefore allows for more flexibility as these models do not impose any *a priori* assumption about the

¹Note that an independent estimation of different Markov chains does not imply that the empirically observed processes are in fact independent.

²Note that the assumption of independently switching regimes results in a multivariate model with 2^N states (see Ang & Bekaert, 2002b). Each of these states represents a combination of the regimes of individual assets. However, this specification is infeasible to estimate for larger samples, given the number of states and the corresponding number of covariance matrices.

dependence of two cycles. Consequently, the true degree of synchronization can be derived empirically.

These observations raise the question of how the degree of synchronization among two regime probability processes, S^x and S^y , should be measured

$$\text{sync}(S^x, S^y),$$

where S^i indicates the probability process of asset i being in a bull market state, $\Pr(s_t^i = 1 | \Omega_T; \theta)$ for $t = 1, \dots, T$.

The following sections present different methods for measuring this degree of synchronization. The individual regime processes S^i are thereby inferred from univariate switching models. The empirical analysis of the synchronization of these cycles will then reveal the actual dependence structure and will provide further insights for the subsequent specification of an alternative model setup.³

Most research on the synchronization of cycles has focused on economic business cycles (see Bordo & Helbling, 2003; Artis, Marcellino & Proietti, 2004; Owyang, Piger & Wall, 2005; Smith & Summers, 2005; Canova, Ciccarelli & Ortega, 2007; Kose, Otrok & Whiteman, 2003, 2008, among others). In contrast, this thesis focuses on the synchronization of regime cycles (see, for example, Edwards, Biscarri & Perez de Gracia, 2003; Bekaert, Harvey & Ng, 2005; Camacho & Perez-Quiros, 2006; Hamilton & Owyang, 2012). Furthermore, the present study contributes to the existing literature by introducing additional measures from related disciplines.

This chapter begins with a short definition of synchronization. Section 3.2 then introduces different measures to test for the extreme cases of full independence and perfect synchronization. Finally, Sections 3.3 and 3.4 present methods for measuring intermediate degrees of synchronization. These measures will support the

³Camacho and Perez-Quiros (2006) remarked that univariate regime-switching models accurately define independent cycles. However, in case the two univariate cycles are dependent, typical synchronization measures underestimate this degree. For this reason, the current chapter introduces more advanced techniques to correct for serial correlation inherent in Markov chains. These techniques promise to measure the degree of synchronization between two dependent cycles more exactly.

development of a regime-switching model that appropriately accounts for the synchronization dynamics between underlying regime cycles (see Chapter 5).

3.1 Definition of Synchronization

The term synchronization derives from Greek and means *the same time* or *common time*. Synchronization implies that things happen at the same time and speed. The degree of synchronization is thereby a measure of similarity across two or more cycles. However, there exists no unique approach to measuring this degree.

Harding and Pagan (2006), for example, assumed three basic degrees of synchronization. The first case describes identical cycles and is called strong perfect positive synchronization (SPPS). The second case describes cycles that behave exactly opposite to each other and is referred to as strong perfect negative synchronization (SPNS). Finally, the third case describes two independent cycles that are strongly non-synchronized (SNS).

Candelon, Piplack and Straetmans (2008, 2009) remarked that these cases might be too restrictive. They proposed an intermediate degree of synchronization between SNS and SPPS. Their test of imperfect multivariate synchronization, however, provides only a joint synchronization value (Candelon et al., 2009). This value defines the average synchronization among all cycle pairs. The current chapter, by way of contrast, focuses on pairwise synchronization. The emphasis on pairwise structures prevents a possible dilution of individual synchronization effects. Accordingly, alternative measures, which test for intermediate synchronization across two cycles, need to be specified.

Further analysis thereby distinguishes three forms of synchronization. First, in-phase synchronization focuses on the contemporaneous behavior of two cycles (Section 3.2). The measures presented by Harding and Pagan (2006) belong to this category. Second, phase-shifted synchronization assumes common but possibly time-shifted cycle dynamics (Section 3.3). Finally, time-varying synchronization focuses on dependence dynamics that change over time (Section 3.4). The next sections introduce different measures to test for these forms of synchronization. In Chapter 4, these

measures will then help to determine the dynamics among assets and support the subsequent model development and calibration.

Most research on regime-cycle synchronization has emerged from business-cycle analysis. However, business cycles are characterized by their binary specification. A cycle resides either in an expansion or in a contraction phase.⁴ In contrast, the state probabilities in a regime-switching model show a continuous structure.

Despite the favorable structure of continuous state probabilities, some synchronization measures in this chapter rely on binary probabilities. For this reason, a marginal transform function $F(\cdot)$ transforms continuous state probabilities into binary probabilities $F(S_t^i) = \check{S}_t^i, \forall i$ where necessary. Here, S_t^i is a process of filtered or smoothed state probabilities (the original regime process) and \check{S}_t^i is a regime process with binary probabilities. This thesis uses the indicator function as a marginal transform function. Regime probabilities equal or greater than 0.5 transform to unity and regime probabilities smaller than 0.5 transform to zero $F(\cdot) = \mathbb{I}_{\xi_t | t \geq 0.5}$. As this transformation causes a loss of information, binary state probabilities are used only where appropriate.

3.2 Measuring In-Phase Synchronization

This section focuses on synchronization measures that analyze the contemporaneous behavior of two cycles. Hereby, measures that test for full independence or perfect synchronization are of particular interest (SNS and SPPS). A rejection of these polar cases justifies the search for alternative specifications of multivariate regime-switching models, which do not follow one of the polar dependence assumptions.

⁴This binary classification is adopted by research institutes such as the National Bureau of Economic Research (NBER), and is inherent in certain tests of synchronization.

3.2.1 Pairwise Correlation

Pairwise correlation is one of the simplest and most intuitive approaches to measuring the relationship between two regime cycles

$$\rho_{S^{xy}} = \frac{E\left(\check{S}_t^x \check{S}_t^y\right) - \mu_x \mu_y}{\sqrt{\mu_x (1 - \mu_x) \mu_y (1 - \mu_y)}} \quad \text{for } t = 1, \dots, T, \quad (3-1)$$

where $\mu_i = E\left(\check{S}_t^i\right)$ for $i = x, y$. This measure relies on the density of its underlying regime processes. It is therefore applicable to either binary $\left(\check{S}_t^i\right)$ or smoothed state probabilities $\left(S_t^i\right)$.⁵

3.2.2 Concordance Index

The concordance index measures the fraction of time of two cycles residing in the same phase. This measure was advocated by Harding and Pagan (2002, 2006) and can be specified by

$$\hat{I} = \frac{1}{T} \left\{ \sum_{t=1}^T \check{S}_t^x \check{S}_t^y + \sum_{t=1}^T \left(1 - \check{S}_t^x\right) \left(1 - \check{S}_t^y\right) \right\}. \quad (3-2)$$

Equation 3-2 shows that the concordance index relies on regime phases rather than on regime cycle densities. Consequently, the index requires binary state probabilities as input variables. It returns a maximum value of unity when two cycles are identical. The index equals zero in case two cycles are strongly negatively synchronized.

One of the problems that Harding and Pagan (2006) raised is the event of independent cycles. In this case, \hat{I} is expected to be close to 0.5. In reality, however, \hat{I} turns out to be much higher. This misspecification is due to the structure of the regime cycles, which are relatively time-consistent (serially correlated). Consequently, their

⁵Note that the specification in 3-1 focuses solely on binary probabilities. For smoothed state probabilities, the corresponding function resembles the definition of Pearson's correlation coefficient.

expected probabilities for one of the two regimes are more likely to be close to unity than to 0.5. Harding and Pagan (2006) demonstrated that this property results in an overestimation of the concordance index value.⁶

3.2.3 Strong Perfect Non-Synchronization

To test for strong perfect non-synchronization, Harding and Pagan (2006) analyzed the correlation between two business cycles. This pairwise analysis can be conducted for each cycle pair in the sample. However, whereas Harding and Pagan relied on binary state probabilities, this thesis uses smoothed state probabilities. The latter are expected to provide more accurate results for regime cycles.

Harding and Pagan (2006) proposed a generalized method of moments (GMM) estimation to calculate the pairwise correlation between two cycles. The advantage of GMM estimation is its robust standard errors. By definition, Markov chains show a high degree of serial correlation. To account for this serial correlation, GMM provides heteroskedastic and autocorrelation consistent (HAC) covariance matrix estimates. Therefore, GMM helps to estimate unbiased moments and standard errors. The HAC standard errors in this thesis are thereby specified with Bartlett weights and with five Newey and West (1994) lags. This correction is especially useful when testing the null hypothesis of $\rho_S = 0$. Harding and Pagan (2006) showed that positive serial correlation in cycles increases chances that this null hypothesis is incorrectly rejected when HAC is not considered.

To estimate the correlation between two regime processes, the moment condition is specified as

$$E \left[\frac{(S_t^x - \mu_x)(S_t^y - \mu_y)}{\sigma_x \sigma_y} - \rho_S \right] = 0, \quad (3-3)$$

⁶ Artis et al. (2004) proposed a standardization of this index to test the null hypothesis of independent regime cycles. The standardized index \hat{I}^* is thereby defined as $\hat{I}^* = \hat{I} - \bar{I}$, where \bar{I} is an estimate of the expected value of \hat{I} . This estimate relies on the assumption that the two cycles are independent. To test for the null of independent cycles, \hat{I}^* is divided by its asymptotic standard error.

where $\mu_i = E(S_t^i)$ is the expectation for being in regime 1 and where $\sigma_i^2 = E[(S_t^i - \mu_i)^2]$ denotes the variance of this estimate for $i = x, y$. The corresponding sample moments are then defined by

$$\frac{1}{T} \sum_{t=1}^T S_t^i - \hat{\mu}_i = 0, \quad (3-4)$$

$$\frac{1}{T} \sum_{t=1}^T (S_t^i - \hat{\mu}_i)^2 - \hat{\sigma}_i^2 = 0, \quad (3-5)$$

and

$$\frac{1}{T} \sum_{t=1}^T \frac{(S_t^x - \hat{\mu}_x)(S_t^y - \hat{\mu}_y)}{\hat{\sigma}_x \hat{\sigma}_y} - \hat{\rho}_S = 0, \quad (3-6)$$

where the hat operator indicates a parameter estimate. Harding and Pagan (2006) remarked that equation 3-6 requires each cycle's mean and variance estimates (3-4 and 3-5) to calculate the correlation $\hat{\rho}_S$. To test the null of $\rho_S = 0$, however, mean and variance terms can be estimated in a preceding step. The correlation parameter is then separately estimated in a subsequent step. This correlation presents a sequential method of moments estimator (Newey, 1984; Harding & Pagan, 2006). Accordingly, the moment condition in 3-3 can be restated as

$$E[m_t(\theta, S_t^x, S_t^y) - \rho_S] = 0, \quad (3-7)$$

where m_t is a separate method of moments function. The mean and variance parameters are estimated in this separate function and have the following moment conditions

$$E \left[\begin{array}{c} S_t^j - \mu_j \\ (S_t^j - \mu_j)^2 - \sigma_j^2 \end{array} \right] = 0. \quad (3-8)$$

The estimation of the correlation parameter in 3-7 is then conditioned on these input parameters $\theta = \{\mu_x, \mu_y, \sigma_x, \sigma_y\}$.⁷

⁷The separate estimation of the mean and variance parameters does not influence the asymptotic $\mathcal{N}(0, 1)$ distribution of $T^{1/2} \hat{\nu}^{-1/2} (\hat{\rho}_S - \rho_S)$, where T is the sample length, and where $\hat{\nu}$ is the asymp-

3.2.4 Strong Perfect Positive Synchronization

Strong perfect positive synchronization (SPPS) can be tested in a similar vein to SNS. However, the null hypothesis of $\rho_S = 1$ implies matching cycle means. Therefore, a test for equal means is initially conducted

$$E(S_t^y - S_t^x) = 0.$$

Again, the method of moments procedure delivers HAC standard errors to determine the statistical significance of results. SPPS is already rejected if the means significantly differ from each other.

If the null of equal means is not rejected, the regime correlation is analyzed in a second step. Thereby, the test for perfect correlation is structured similarly as in the SNS case. However, the null hypothesis is now $\rho_S = 1$. The sequential method of moments estimator no longer holds under this assumption because $E[\delta m_t / \delta \theta] \neq 0$. Accordingly, the moment conditions need to be jointly estimated. They are specified as

$$\begin{aligned} E[S_t^j - \mu_j] &= 0 \\ E\left[\left(S_t^j - \mu_j\right)^2 - \sigma_j^2\right] &= 0 \quad \text{for } j = x, y \end{aligned}$$

and as

$$E\left[\frac{(S_t^x - \mu_x)(S_t^y - \mu_y)}{\sigma_x \sigma_y} - \rho_S\right] = 0.$$

otic variance of $T^{1/2}(\hat{\rho}_S - \rho_S)$. The independence of parameters and asymptotic distribution is due to $E[\delta m_t / \delta \theta] = 0$, given the null hypothesis of $\rho_S = 0$.

The corresponding sample moments are defined by

$$h_t(\theta, S_t^x, S_t^y) = \begin{bmatrix} S_t^i - \hat{\mu}_i \\ (S_t^i - \hat{\mu}_i)^2 - \hat{\sigma}_i^2 \\ \frac{(S_t^x - \hat{\mu}_x)(S_t^y - \hat{\mu}_y)}{\hat{\sigma}_x \hat{\sigma}_y} - \hat{\rho}_S \end{bmatrix}$$

and by

$$f(\theta, S^x, S^y) = \frac{1}{T} \sum_{t=1}^T h_t(\theta, S_t^x, S_t^y).$$

The test for $\rho_S = 1$ reveals another problem because the correlation bounds are $|\rho_S| \leq 1$. Consequently, the test of the null hypothesis is conducted on the “boundary of the parameter space” (Harding & Pagan, 2006, p. 68). As a result, $T^{1/2}\hat{\nu}^{-1/2}\hat{\rho}_S$ is no longer asymptotically $\mathcal{N}(0, 1)$ distributed. Again, T defines the sample length and $\hat{\nu}$ is the variance of $T^{1/2}\hat{\rho}_S$.

A solution to this problem is to use $T^{1/2}\hat{\nu}^{-1/2}(\hat{\rho}_S - 1)$. Chant (1974), Andrews (2001), and Harding and Pagan (2006) pointed out that this solution asymptotically follows a half-normal distribution.

3.3 Measuring Phase-Shifted Synchronization

Measures of in-phase synchronization determine the contemporaneous resemblance of two cycles. However, a particular cycle might lead or lag another by a certain period. In this case, in-phase analysis would reveal biased synchronization values. The approaches in this section take into account different forms of such phase-shifts.

Figure 3.1 illustrates three possible forms of synchronization. The top-left graph shows the basic in-phase analysis. Corresponding measures conduct a one-to-one point comparison of contemporaneous events. Further, the top-right graph shows a form of phase-shifted synchronization analysis, which also applies a one-to-one point comparison. However, the phase-shifted comparison might allow for a fixed

lag between two underlying cycles. The cross-correlation represents a particular technique that belongs to the class of phase-shifted synchronization measures. Finally, the bottom graphs represent non-metric functions of phase-shifted synchronization. These functions allow for one-to-many (and many-to-one) point comparisons. They imply that a particular observation in the first cycle might match multiple observations in the second cycle and vice versa.⁸ Likewise, some observations might reside fully unmatched. Dynamic time warping and longest common subsequence represent corresponding measures, which are presented below. Given that the polar cases of in-phase synchronization can be rejected, these techniques will help to detect alternative synchronization dynamics between regime cycles and will therefore support the definition of a more appropriate regime-switching model.

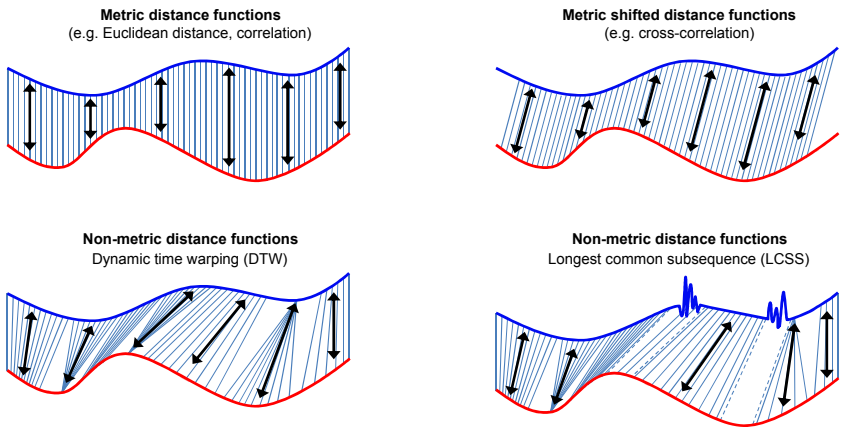


Figure 3.1: *Possible forms of synchronization*

3.3.1 Cross-Correlation

In a simple form, pairwise cross-correlation relies on the correlation function in 3-1, but applies a time-lag to one of the input series S^x or S^y . This lag-structure allows one to account for phase-shifts when measuring the pairwise correlation of two

⁸This structure accounts for symmetric or asymmetric shifts in regime cycles.

cycles. For example, the analysis of monthly lags between zero and six months provides thirteen different correlations for these cycles. If a particular correlation turns out significantly higher, its corresponding lag might indicate the length of the shift across cycles.

3.3.2 Dynamic Time Warping

Dynamic time warping (DTW) is another way of measuring synchronization among phase-shifted regimes. The DTW algorithm is used in time series analysis to find the optimal alignment of two time-dependent sequences. These sequences are warped in a nonlinear way to optimally match each other.⁹ The DTW algorithm is thereby specified as

$$DTW(i, j) = \begin{cases} d(S_i^x, S_j^y) + \min \{DTW(i-1, j-1), \dots \\ DTW(i-1, j), DTW(i, j-1)\} & \text{if } |i-j| < \delta \\ \infty & \text{otherwise,} \end{cases} \quad (3-9)$$

where $d(S_i^x, S_j^y)$ measures the metric difference between S_i^x and S_j^y , and where δ is a global constraint, which limits the period of phase-shift (Oates et al., 1999; Morse & Patel, 2007; Cassisi et al., 2012). Further, $DTW(\cdot, 0) = DTW(0, \cdot) = 0$ holds, and $1 \leq i, j \leq T$. The δ factor thereby constrains the lead and lag length among cycles (known as Sakoe-Chiba band) to avoid warping paths with large phase-shifts.

The DTW function aims to find the optimal warping path. This path indicates the smallest distance between S^x and S^y along time. As a result, the final observation $DTW(T, T)$ indicates the minimum distance among the – possibly phase-shifted – cycles.¹⁰ This value is important for assessing the similarity of two processes. However, the optimal warping path itself is of comparable interest. The optimal path

⁹For a detailed analysis of DTW and its estimation algorithm, see Oates, Firoiu and Cohen (1999) and Cassisi, Montalto, Aliotta, Cannata and Pulvirenti (2012).

¹⁰The structure in 3-9 reveals that DTW follows a dynamic programming approach.

indicates the time-varying phase-shifts along time, which do not become apparent from the $DTW(T, T)$ value.

To determine the distance between S^x and S^y , metric distance measures can be applied (Oates et al., 1999; Morse & Patel, 2007; Cassisi et al., 2012). The Euclidean distance presents a common measure for this task¹¹

$$d(S_i^x, S_j^y) = |S_i^x - S_j^y|.$$

Figure 3.2 illustrates a sequence of an optimal warping path between the US and the UK regime process. The solid blue line and the dotted red line represent paths with different lead-lag constraints (three and six months). The warping paths reveal the time-varying lead and lag relationship between the cycles of the US and the UK.

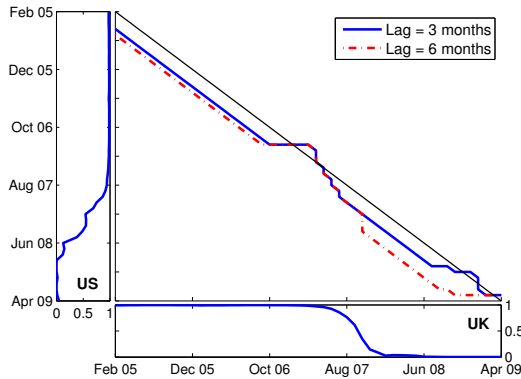


Figure 3.2: *DTW synchronization path for the regime processes of the US and the UK between February 2005 and April 2009. The two subplots depict the univariate regime processes of the respective country. The main plot illustrates the synchronization of the two countries. The continuously decreasing line indicates perfect synchronization without any phase shifts. Values below this line describe a leading US cycle, values above describe a leading UK cycle.*

The advantage of DTW is its one-to-many point comparison. Traditional measures, as for example the correlation, conduct only one-to-one point comparisons. In con-

¹¹ Alternative distance measures, such as the Manhattan distance or the Minkowski distance, can also be applied.

trast, DTW recognizes similar cycle shapes even if their signal is transformed, that is, if it is shifted or scaled (Cassisi et al., 2012). This feature allows DTW to detect possible asymmetries in regime cycles.¹²

3.3.3 Longest Common Subsequence

Longest common subsequence (LCSS) is another similarity measure that allows for one-to-many point comparisons. However, whereas DTW matches all sample points, LCSS ignores certain outliers (see also Figure 3.1). This feature makes LCSS less resilient to noise. As a consequence, its similarity measure is less distorted.

Figure 3.1 shows that the LCSS specification resembles that of DTW. However, LCSS measures the similarity of two cycles in terms of relative and not in terms of absolute distance. Two cycles S^x and S^y are considered to be similar as long as their distance is below a certain threshold value ϵ

$$LCSS(i, j) = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0 \\ LCSS(i-1, j-1) + 1 & \text{if } |S_i^x - S_j^y| < \epsilon \\ \max \{LCSS(i, j-1), LCSS(i-1, j)\} & \text{otherwise,} \end{cases}$$

where $0 \leq i, j \leq T$ (Morse & Patel, 2007; Cassisi et al., 2012). The recurrence function $LCSS(\cdot, \cdot)$ measures the period length during which the distance of individual cycle observations is below the value ϵ .¹³ Adding the side condition $|i - j| < \delta_{LCSS}$ further constrains the shift size. Otherwise, large cycle shifts could result in arbitrary matches due to similar but unrelated events of two cycles.

Figure 3.3 exemplifies the LCSS path between the US and the UK process. Similar to DTW, the LCSS path shows phases where the US regime cycle leads the UK regime cycle and vice versa. As expected, DTW and LCSS reveal similar patterns.

¹²Cakmakli et al. (2011) presented a regime-switching model that relies on a similar synchronization assumption as the DTW algorithm.

¹³Similar to DTW, the longest common subsequence algorithm is based on dynamic programming.

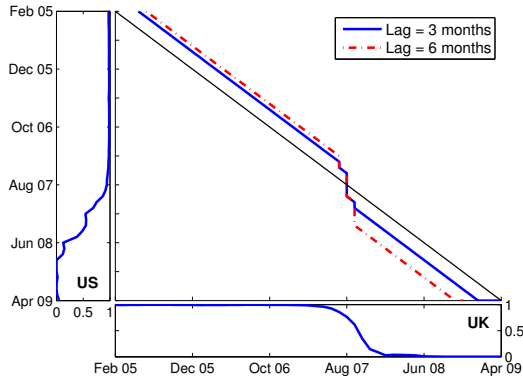


Figure 3.3: *LCSS synchronization path for the regime processes of the US and the UK between February 2005 and April 2009. The two subplots depict the univariate regime processes of the respective country. The main plot illustrates the synchronization of the two countries. The continuously decreasing line indicates perfect synchronization without any phase shifts. Values below this line describe a leading US cycle, values above describe a leading UK cycle.*

In the DTW algorithm, the final value $DTW(T, T)$ returns the minimum cycle distance. In contrast, the value $LCSS(T, T)$ only indicates the length of the common cycle. However, the transformation $LCSS(T, T)/T$ returns a relative measure of similarity. A value closer to unity reveals a better match.

3.4 Measuring Time-Varying Synchronization

The measures specified so far determine synchronization across the full sample history. The focus on the full history follows the specification of standard univariate and multivariate regime-switching models. These models infer the regime process for the full length of the time series. Accordingly, corresponding tests for fully independent or perfectly synchronized cycles should also consider the full length of the time series.

However, alternative synchronization assumptions need to be tested for if these two

polar cases are rejected. The previous section presented one particular case that assumed that the regime cycles exhibit a phase-shifted dependence structure.

The current section proposes an alternative possibility. Individual cycles might show time-varying degrees of synchronization. To test for this hypothesis, no additional synchronization measures are required. Existing techniques can be applied on a rolling-window basis.¹⁴ This approach will nest all previous model specifications.

Hamilton and Owyang (2012) and Leiva-Leon (2012a, 2012b) proposed similar assumptions with regard to the dynamic structure of regime cycles. They showed that the synchronization dynamics change over time and that a time-varying synchronization factor might be considered.

¹⁴This specification raises the question if filtered or smoothed regime probabilities should be used. The analysis can in fact rely on smoothed regime probabilities. However, the probabilities may only be smoothed over the rolling-window period $t - p$ to t ($\xi_{t-p|t}$), but not over the entire sample period $t - p$ to T ($\xi_{t-p|T}$), where p indicates the length of the rolling-window.

Chapter 4

Data Analysis

The aim of this thesis is to develop a flexible multivariate regime-switching model that allows individual assets to follow their own switching dynamics. This specification requires multiple regime processes, which follow individual – possibly synchronized – Markov chain dynamics. To specify such a model, it is necessary to understand the synchronization dynamics in detail. The previous chapter presented different techniques to measure these dynamics. Below, these techniques will be applied to empirical data.

The results of this analysis shed light on three different areas. First, they determine the empirically observable regime synchronization. Second, the results point out potential shortcomings of the univariate and multivariate benchmark models. Finally, they define the structural guidelines for a more flexible regime-switching model.

This chapter first describes the dataset and conducts a descriptive analysis. Subsequently, the dataset is used to estimate the benchmark regime-switching models. This step includes the parameter estimation and the inference of regime processes. Finally, the regime processes of the univariate models are analyzed with respect to their pairwise synchronization.

4.1 Dataset

This thesis uses gross total return indices from Datastream for the following six markets: United States, United Kingdom, Germany, Japan, Pacific ex Japan, and

Switzerland. The US, Germany, and the UK build the core sample and match that of Ang and Bekaert (1999, 2002a, 2002b). Japan is excluded from the core sample despite its higher market capitalization compared to the UK and Germany. This exclusion is due to the liberalization process Japan underwent in the 1980s.

Japan, Pacific ex Japan¹, and Switzerland extend the core sample. They lead to a balanced representation of the three world regions North America, Europe, and Asia-Pacific (see Schwendener, 2010). The core and the extended sample represent approximately 60.0% and 81.6% of the developed-country market capitalization. The separation into core and extended sample will better illustrate the drawbacks of standard multivariate regime-switching models.

The sample period ranges from December 1975 to February 2014, for a total of 458 monthly return observations. All return series are based on USD denominated indices and are continuously compounded. The starting date was selected due to the 1973/1974 stock market crash, which ended in December 1974. The exclusion of this crash results in a balanced occurrence of bull and bear market cycles.

4.2 Descriptive Statistics

Table 4.1 presents the descriptive statistics for the six market indices. The annualized mean of the monthly returns lies between 8.2% for Japan and 12.4% for Switzerland. Mean returns turned out relatively similar across the sample. Japan represents the only exception. Its low mean is due to the Japanese asset price bubble in the 1990s. This price bubble is also one of the reasons for Japan's exclusion from the core sample. The annualized standard deviations range from 15.2% for the US to 23.1% for Pacific ex Japan.

Japan and Pacific ex Japan further show relatively low values in terms of return per unit of risk. For Japan, this is in line with the previous argumentation. For the Pacific area, the low return per unit of risk is mainly due to Black Monday. In October 1987, the Pacific area showed an extreme monthly return minimum of -48.1%. In fact, the

¹Note that the Pacific index does not represent a country, but a regional index.

return minima of most analyzed markets date back to this event. For the return maxima the dates are less conclusive.

Table 4.1: *Descriptive statistics of the sample, covering monthly returns from December 1975 to February 2014.*

	Mean	Med.	Min.	Max.	Std.	Skew.	Kurt.	JB-stat.
US	11.3%	15.6%	-23.3%	12.6%	15.2%	-0.85	5.88	213.0***
UK	12.2%	15.3%	-23.9%	18.8%	19.2%	-0.51	4.87	86.7***
Germany	10.1%	17.3%	-23.1%	17.7%	20.7%	-0.70	4.59	86.1***
Japan	8.2%	8.5%	-19.3%	24.0%	20.8%	0.06	3.75	11.0***
Pacific ex Japan	11.3%	14.7%	-48.1%	21.9%	23.1%	-1.31	10.18	1113.7***
Switzerland	12.4%	16.5%	-20.1%	14.5%	17.2%	-0.61	4.39	65.5***

Values are shown on an annualized basis; JB-stat. shows the Jarque-Bera test statistics.

*** indicates the significance of the Jarque-Bera test statistics at the 1% level.

Most markets show returns that are skewed to the left. Japan represents an exception with a positive skewness of 0.06. Moreover, all the markets are leptokurtic, with Japan showing the lowest kurtosis of 3.75 and Pacific ex Japan showing the highest kurtosis of 10.18. The Jarque-Bera test for normally distributed returns is rejected for all markets at the 1% level. Especially the statistics for the US and the Pacific area are clearly at odds with the normal distribution. These initial results indicate that a linear return model would describe the behavior of the underlying time series only insufficiently.

Table 4.2: *Unconditional correlation parameters of the sample, covering monthly returns from December 1975 to February 2014.*

	US	UK	GE	SW	JP	PA
US	1					
UK	0.65	1				
Germany	0.60	0.63	1			
Switzerland	0.57	0.66	0.77	1		
Japan	0.36	0.46	0.40	0.48	1	
Pacific ex Japan	0.60	0.63	0.55	0.43	0.53	1

Legend: GE = Germany; SW = Switzerland; JP = Japan; PA = Pacific ex Japan.

Table 4.2 presents the unconditional correlation matrix for the six markets. For better interpretation, the data are sorted by their geographical region. The correlations show a connection between European markets and the US, as well as between the US and the other markets. In contrast, Japan displays a more idiosyncratic behavior. To a lesser extent, this behavior is also observed for the Pacific area.

4.3 Regime-Switching Analysis

The subsequent regime-switching analysis relies on the benchmark models of Chapter 2. Analysis will reveal additional details about the switching behavior and the structure of the sample. These insights will support the model specification in Chapter 5.

4.3.1 Univariate Regime-Switching Results

Table 4.3 presents the parameter estimates for the six markets. Each market is specified by a univariate model and follows two distinct regimes. In regime 1, the markets show higher mean returns and lower volatilities than in regime 2.² Moreover, regime 1 depicts higher state transition probabilities than regime 2. The corresponding duration values indicate that the country and region indices reside between 8 and 50 months in regime 1. In contrast, they only last between 3 and 25 months in regime 2.

The two regimes resemble the typical behavior of equity data. Regime 1 can be interpreted as a bull market state (lower volatility and correlations, higher returns) whereas regime 2 resembles the behavior of a bear market state (higher volatilities and correlations, lower returns; see, for instance, Ang & Bekaert, 2002a). Despite

²The regime-switching models in this thesis were generally specified as having lower volatilities in regime 1 than in regime 2. This specification was necessary to identify the two regimes and to avoid label switching problems (see Hamilton & Owyang, 2012).

these observable stylized features, there is no consensus in academic literature about the exact definition of bull and bear market states (Candelon et al., 2008).

Table 4.3: *Parameter estimates for the six sample markets based on a univariate regime-switching model, covering monthly returns from December 1975 to February 2014. Means and volatilities are treated as regime-dependent.*

	μ_1	μ_2	σ_1	σ_2	p_{11}	p_{22}	Dura ₁	Dura ₂	Lik
US	0.17*** (0.035)	-0.15 (0.142)	0.12*** (0.010)	0.24*** (0.002)	0.96*** (0.026)	0.84*** (0.083)	22.3	6.1	859.0
UK	0.15*** (0.030)	-0.14 (0.207)	0.16*** (0.007)	0.42*** (0.002)	0.98*** (0.010)	0.86*** (0.072)	49.5	7.1	718.7
Germany	0.17*** (0.034)	0.02 (0.070)	0.14*** (0.009)	0.27*** (0.002)	0.95*** (0.024)	0.94*** (0.038)	21.4	16.1	714.9
Japan	0.12*** (0.041)	0.02 (0.060)	0.14*** (0.008)	0.25*** (0.001)	0.96*** (0.021)	0.96*** (0.020)	25.7	25.4	702.5
Pacific ex Japan	0.23*** (0.032)	-0.18*** (0.007)	0.14*** (0.008)	0.37*** (0.002)	0.95*** (0.018)	0.89*** (0.036)	19.1	8.9	666.1
Switzerland	0.21*** (0.033)	-0.10 (0.107)	0.12*** (0.009)	0.25*** (0.002)	0.87*** (0.053)	0.70*** (0.045)	7.7	3.3	790.4

Values are shown on an annualized basis; standard errors are depicted in parentheses; Dura_{*i*} defines the duration of regime *i* (in months), where $\text{Dura}_i = \frac{1}{1-p_{ii}}$; p_{11} and p_{22} indicate the probability of remaining in regime 1 or 2, respectively; and Lik is the log-likelihood.

*** indicates the significance of the parameter estimates (different from zero) at the 99% level.

Most of the estimated parameters are significantly different from zero at the 99% level (indicated by ***). In regime 2, however, all means except for Pacific ex Japan show insignificant deviations from zero. There are several reasons for this behavior. On the one hand, regime 2 might describe a state in which returns do not significantly differ from zero. However, the relatively high magnitude of the means of regime 2 indicate that this assumption can be rejected.

On the other hand, the data might not underlie any switching in means. In fact, the standard errors for the means in regime 2 are very high. Such high standard errors might be due to few observations in this regime. Under these circumstances, the same problem would be observed for volatilities in regime 2. This is, however, not the case and indicates the irrelevance of switching means for most markets. Therefore, the benchmark models subsequently specify their means as non-switching (*ceteris paribus*). Undisclosed likelihood ratio tests further supported this decision (see also Section 4.3.3 for the multivariate likelihood ratio test).

Table 4.4 presents the univariate specification without switching means. The standard deviations and the regime durations remain almost unchanged. Even the decrease in marginal likelihood is very low. This outcome clearly supports the decision of restricting the means, as they do not significantly influence the parameter set.

The non-switching mean values are a weighted average of the previously observed regime-dependent means. In Table 4.4, they range from 9% for Japan to 18% for Pacific ex Japan. All of these means are highly significant and even slightly higher than in the descriptive statistics (Table 4.1).³

Table 4.4: *Parameter estimates for the six sample markets based on a univariate regime-switching model, covering monthly returns from December 1975 to February 2014. Means are treated as regime-independent.*

	μ	σ_1	σ_2	p_{11}	p_{22}	Dura ₁	Dura ₂	Lik
US	0.14*** (0.023)	0.11*** (0.009)	0.22*** (0.002)	0.95*** (0.025)	0.87*** (0.087)	19.3	7.6	808.0
UK	0.16*** (0.028)	0.11*** (0.014)	0.23*** (0.004)	0.96*** (0.032)	0.98*** (0.017)	26.6	41.5	700.5
Germany	0.13*** (0.029)	0.14*** (0.010)	0.29*** (0.002)	0.95*** (0.027)	0.92*** (0.056)	19.7	12.0	667.7
Japan	0.09*** (0.031)	0.14*** (0.013)	0.26*** (0.002)	0.96*** (0.023)	0.96*** (0.024)	24.5	23.5	653.7
Pacific ex Japan	0.18*** (0.031)	0.14*** (0.010)	0.34*** (0.002)	0.94*** (0.022)	0.88*** (0.045)	16.5	8.5	634.0
Switzerland	0.15*** (0.025)	0.12*** (0.010)	0.27*** (0.002)	0.89*** (0.049)	0.69*** (0.104)	9.0	3.2	744.9

Values are shown on an annualized basis; standard errors are depicted in parentheses; Dura_{*i*} defines the duration of regime *i* (in months), where $\text{Dura}_i = \frac{1}{1-p_{ii}}$; p_{11} and p_{22} indicate the probability of remaining in regime 1 or 2, respectively; and Lik is the log-likelihood.

*** indicates the significance of the parameter estimates (different from zero) at the 99% level.

The standard deviations for the two regimes are very distinct. Regime 2 depicts values twice as high as those in regime 1. In the first regime, annualized standard deviations range from 11% for the US and the UK to 14% for Germany, Japan, and the Pacific area. In the second regime, they range from 22% for the US to 34% for Pacific ex Japan. This relation between the two regimes persists across all markets.

³This might be attributed to the dynamic structure of the regime-switching model.

Table 4.4 further depicts the transition probabilities p_{11} and p_{22} . Most markets show probabilities above 0.90, which indicates long cycle durations. The persistence of regime 1 is generally higher than that of regime 2. This persistence confirms previous observations that assets spend relatively more time in bull market regimes than in bear market regimes (see Ang & Bekaert, 2002a; Schwendener, 2010). However, a detailed analysis of individual transition probabilities reveals some idiosyncrasies among markets. The US, Pacific ex Japan, and Switzerland, for example, show relatively low persistent transition probabilities in bear markets. Consequently, these countries experience shorter bear phases than other countries. Overall, Switzerland shows relatively low transition values in both states.

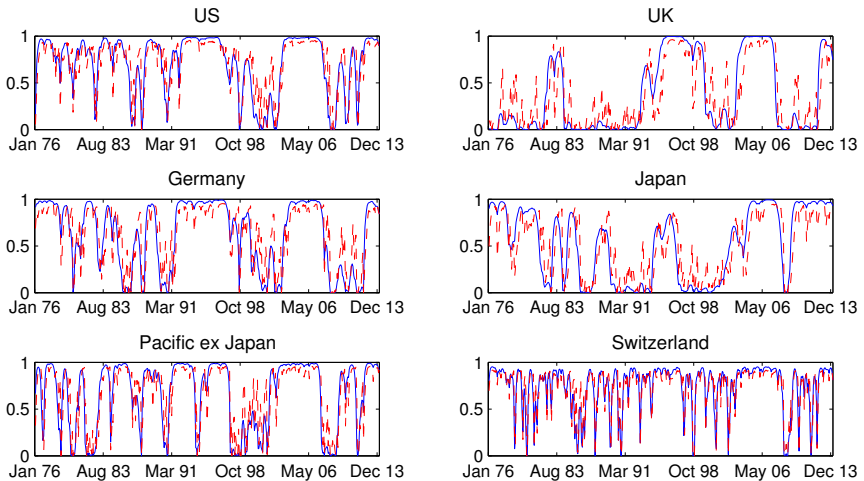


Figure 4.1: *Filtered and smoothed regime probability processes for the six sample markets based on the univariate regime-switching model (dashed red line: filtered probabilities; solid blue line: smoothed probabilities). Probabilities are depicted for regime 1. The period under study is from December 1975 to February 2014.*

The six graphs in Figure 4.1 visualize the previous observations. The solid blue lines illustrate smoothed probabilities and the dashed red lines display filtered probabilities for regime 1. The six univariate regime cycles show phases of synchronization and phases of independence. The US and Germany in particular demonstrate very

similar cycle dynamics. The next sections will assess the similarity of these cycles with different synchronization measures.

4.3.2 Multivariate Regime-Switching Results

This section focuses on multivariate benchmark regime-switching models. Table 4.5 illustrates parameter estimates for the extended sample with non-switching means.⁴ The moments of the US, the UK, and Germany in the core sample analysis turned out to be very similar to those in Table 4.5 and were therefore omitted.

The similarity between results of the core and the extended sample indicates the stability of multivariate models. Figure 4.2 illustrates this behavior for the smoothed probabilities of the core (solid blue line) and the extended sample (dashed red line). As for the moment parameters, both samples show similar dynamics in their regime processes. This observation is surprising given the different sample sizes and the diverging univariate dynamics. However, given a dominant group of assets, individual regime dynamics might vanish. Occasionally, however, the two cycles in Figure 4.2 differ strongly. In these instances, a different formation might determine the underlying dynamics (in the two samples).

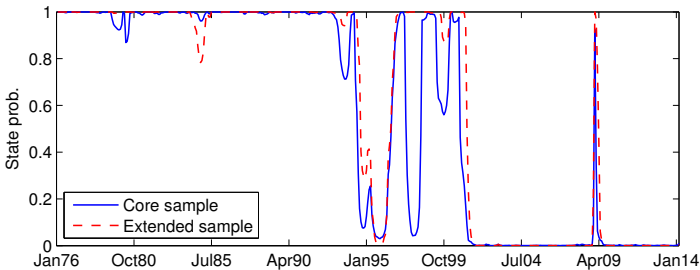


Figure 4.2: Smoothed regime probabilities for the core and the extended sample based on multivariate regime-switching inference. The figure indicates the probability of staying in a bull market regime (regime 1). The period under study is from December 1975 to February 2014.

⁴Undisclosed results for the multivariate model with switching means revealed the same picture as for the univariate model. In regime 2, means have also turned out to be insignificant.

The results clearly indicate that multivariate models might ignore the differing switching dynamics of individual assets. Simply extending the model to a larger sample therefore does not sufficiently account for individual regime dynamics. Consequently, a trade-off between model accuracy and simplicity exists. The impact of these differences across cycles will be assessed throughout this chapter.

Table 4.5 illustrates the parameter estimates for the multivariate model. In the first regime, standard deviations range from an annualized 15% in the US to 24% in the Pacific area. In the second regime, they range from 17% in the US to 25% in Germany. Compared to the univariate results, standard deviations across regimes are more similar. The regime-dependent correlations can partially explain this similarity. The correlations absorb some of the regime-dependence inherent in the covariance matrix. Pelletier (2006) gives further details about this behavior.

Table 4.5: *Parameter estimates for the six sample markets based on a multivariate regime-switching model, covering monthly returns from December 1975 to February 2014. Means are treated as regime-independent.*

	μ	σ_1	σ_2	\mathbf{P}	s_1	s_2
US	0.11*** (0.023)	0.15*** (0.007)	0.17*** (0.001)	s_1	0.99	0.02
				s_2	0.01	0.98
UK	0.12*** (0.029)	0.20*** (0.008)	0.21*** (0.001)	Duration	76.3	63.8
Germany	0.11*** (0.033)	0.20*** (0.008)	0.25*** (0.001)	Log Lik		4931.4
Japan	0.08** (0.033)	0.21*** (0.009)	0.21*** (0.001)	dim(θ)		50
Pacific ex Japan	0.13*** (0.035)	0.24*** (0.010)	0.24*** (0.001)	AIC		-9762.9
				BIC		-9556.5
				HQ		-9681.6
Switzerland	0.13*** (0.027)	0.18*** (0.008)	0.18*** (0.001)			

Values are shown on an annualized basis; standard errors are depicted in parentheses; the regime duration (in months) is measured by $\text{Duration} = \frac{1}{1-p_{ii}}$; \mathbf{P} is the transition probability matrix and s_1 and s_2 are the corresponding states; Log Lik is the log-likelihood; $\text{dim}(\theta)$ gives number of model parameters; and AIC, BIC, and HQ are the Akaike, Bayes-Schwarz, and the Hannan-Quinn information criterion.

*, **, and *** indicate the significance of the parameter estimates (different from zero) at the 90%, 95%, and 99% level.

Table 4.6 further presents the correlations across bull and bear market regimes (regimes 1 and 2). The lower triangular values represent the correlations in regime 1 and the upper triangular values those in regime 2. The bull market regime mostly shows lower correlations than the bear market regime. These observations are in line with common regime-switching literature. Contagion thereby causes equity markets to react more similarly in negative market environments (Schwendener, 2010).

Table 4.6: *Regime-dependent correlation matrices for the six sample markets based on a multivariate regime-switching model, covering monthly returns from December 1975 to February 2014. Lower triangular matrix: bull market state; upper triangular matrix: bear market state.*

	US	UK	GE	JP	PA	SW
US		0.90	0.90	0.66	0.85	0.82
UK	0.52		0.92	0.72	0.90	0.89
Germany	0.40	0.46		0.64	0.87	0.87
Japan	0.27	0.37	0.33		0.69	0.69
Pacific ex Japan	0.48	0.50	0.38	0.35		0.82
Switzerland	0.45	0.55	0.73	0.41	0.41	

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

In fact, most stylized facts of financial data can be observed within this regime-switching framework. Schwendener (2010) provides a good overview of the stylized facts inherent in financial data. These include asymmetry of returns, correlation breakdown, volatility clustering, volatility co-movement, contagion, decoupling, and safe haven. Schwendener shows that regime-switching models can cover these features more efficiently than linear market models. The previous results have confirmed most of the features that apply to equity markets.

Figure 4.3 compares the smoothed bull market regime probabilities of the multivariate and the univariate models. The multivariate regime cycles remove noise and show a very distinct structure. However, during the first half of the sample, the multivariate regime cycles show a markedly different behavior than the univariate regime cycles. Even if the univariate regime processes did not reveal any clear interrelationship so

far, their behavior can at least be differentiated from that of a common multivariate cycle. These observations further support the assumption that standard multivariate regime-switching models insufficiently cover the switching dynamics of individual assets.

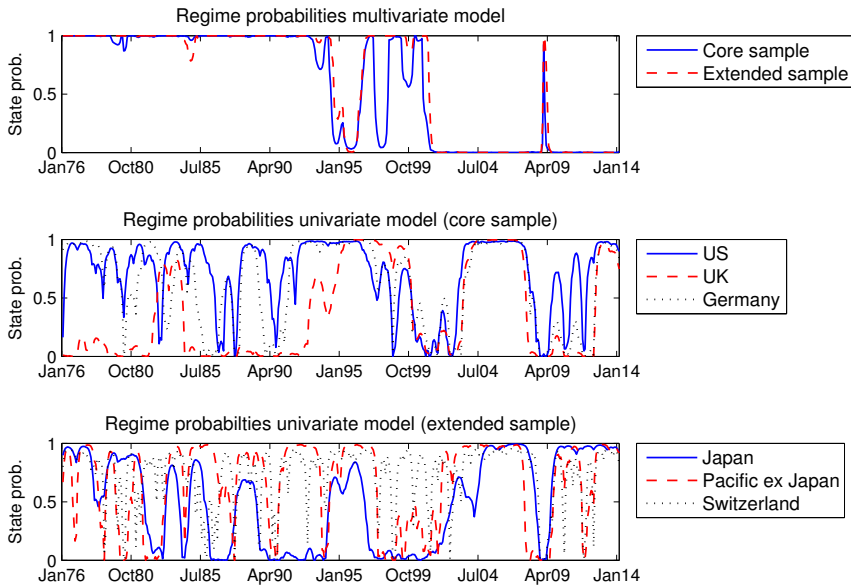


Figure 4.3: Comparison of smoothed regime probabilities inferred from the univariate and the multivariate regime-switching models (probability of residing in a bull market regime). The period of analysis is from January 1976 to February 2014. Results are depicted for the core and the extended sample.

4.3.3 Likelihood Ratio Tests

Chapter 2 specified the benchmark models with switching mean, switching variance, and switching correlation parameters. This specification allows for flexible model structures. However, each switching variable comes at the price of an extended parameter set. At the same time, the switching of these variables might be unneces-

sary. The previous section, for example, demonstrated that switching means might not contribute significant information.

A restriction of certain parameters can simplify the model estimation. Restricting means to be non-switching can thereby be of particular value. Switching means might be the source of unconditional density functions with multiple local modes. Multimodality causes the MLE optimization to stop at a local instead of the global maximum.⁵ These problems are circumvented by restricting means to being regime-independent. The optimization of restricted models can consequently result in more feasible solutions. However, such a simplification must not significantly influence the model likelihood.

The likelihood ratio test (LR-test) evaluates unrestricted and restricted models with respect to their differences in likelihood. The test assumes the twofold difference of these likelihoods to be approximately Chi-square distributed

$$2[\log L(\hat{\theta}) - \log L(\tilde{\theta})] \approx \chi_m^2,$$

where the degree of freedom m equals the number of parameter restrictions, where $L(\hat{\theta})$ is the likelihood of the unrestricted model, and where $L(\tilde{\theta})$ is the likelihood of the restricted model (Hamilton, 1994). For $m = 1$, critical values to reject the null hypothesis of equal likelihoods are 2.70, 3.84, and 6.64 for the 10%, 5%, and 1% levels of significance, respectively.

Note that the LR-test requires the restricted model to be a nested version of the unrestricted model. Accordingly, models with different regime numbers cannot be compared due to nuisance parameters. For example, the null hypothesis of only one regime (non-switching model) would not satisfy the regularity conditions of a standard regime-switching model, as some parameters would remain unidentified (Hamilton, 2005).

Table 4.7 illustrates different specifications of the multivariate benchmark models.

⁵The Gibbs sampling algorithm partially prevents this problem. The algorithm applies Markov chain Monte Carlo (MCMC) simulation instead of a gradient descent method (commonly used in MLE optimization). The former approach runs through a predefined number of simulations and can jump out of local maxima. The latter approach stops as soon as a feasible optimum is found, which, however, is not necessarily the global maximum.

The fully specified model allows for switching means, switching standard deviations, and switching correlations. Moreover, this specification defines the initial state probabilities $\xi_{1|0}$ as additional parameters to be estimated.⁶ The basic specification, on the other hand, considers the case of non-switching means and fixed initial probabilities. The last four columns in Table 4.7 refer to cases where some parameters are restricted to being regime-independent or exogenously defined.

Table 4.7: *Likelihood ratio statistics for the core and the extended sample. The period of analysis is from December 1975 to February 2014.*

	Full specification	Basic specification	Parameter restrictions			
			Mean	Std.	Correlation	Initial probs.
Core sample	2449.7 <i>21</i>	2449.0 <i>17</i>	2449.0 <i>18</i>	2444.2** <i>18</i>	2401.8*** <i>18</i>	2449.7 <i>20</i>
Extended sample	4933.0 <i>57</i>	4931.4 <i>50</i>	4931.4 <i>51</i>	4922.9*** <i>51</i>	4917.6*** <i>42</i>	4932.9 <i>56</i>

The table shows the likelihood values for the core and the extended sample. Italic values indicate the corresponding number of estimated parameters. Likelihood values are depicted for the fully specified benchmark regime-switching model and for the basic specification with restricted means and initially set regime probabilities (0.5). The last four columns indicate individual parameter restrictions. These four specifications have been tested for differences in model fit compared to the fully specified model setup.

*, **, and *** indicate statistically significant differences in model fit at the 10%, 5%, and 1% level.

The likelihood ratio test is defined by $2 \left[\log L(\hat{\theta}) - \log L(\tilde{\theta}) \right] \approx \chi_m^2$.

Table 4.7 depicts the likelihood values for the core (3 markets) and the extended sample (6 markets). In addition, values in italics depict the number of model parameters. For the LR-test, the fully specified model defines the unrestricted likelihood $L(\hat{\theta})$. For the case of restricted means, the likelihood differences turn out to be insignificant. This observation confirms the previous conditioning of the benchmark models. In contrast, the restriction of standard deviations or correlations causes significant differences in model fit for both samples. Finally, predetermined initial probabilities cause only insignificant differences. This is due to the relatively long data series, which reduces the influence of steady-state probabilities on the likelihood function. These insights substantiate the decision to subsequently restrict the means of the benchmark models. This restriction will improve model convergence and reduce

⁶In the case of fixed initial state probabilities, these are exogenously defined as 0.5 in both states.

the parameter set. The assumption of non-switching means is especially useful for the convergence of the flexible model in Chapter 5. Further, it supports the asset allocation problem in Chapter 7. Portfolio optimization is generally very sensitive to return expectations. Switching means might therefore cause more extreme portfolio allocations and turnovers. In what follows, these considerations justify the use of non-switching means.

On the other hand, the results clearly confirm the importance of switching standard deviation and correlation terms. These parameters significantly contribute to the model likelihood.

4.4 Regime Synchronization Analysis

The previous analysis revealed initial discrepancies between the univariate regime cycles of individual assets. Further, the analysis of multivariate models has shown that the aggregation of data in a joint cycle does not necessarily deliver more accurate or more stable results.

Figure 4.4 extends these analyses. The six quadrants present scatterplots for pairwise combinations of the univariate regime cycles. Each quadrant depicts the US regime probabilities (x-axis) against the regime probabilities of one of the remaining markets (y-axis).⁷ In addition, the top-left quadrant plots the US probabilities against themselves. This graph visualizes perfect synchronization. The opposite case of fully independent regime cycles would result in evenly distributed plot points within the unit field.

Most country pairs show clusters in the lower-left or upper-right corner. In these cases, both processes jointly reside in regime 2 or in regime 1. However, some country pairs reveal additional clusters. For example, the US-UK graph depicts a cluster in the lower-right corner. In this case, the US resides in a bull regime whereas the UK returns are driven by a bear regime. The US-UK, the US-Germany, and the US-Japan plots further show sparse upper-left corners. This behavior indicates that

⁷Each cross in the six quadrants (Figure 4.4) depicts the observation of two countries' univariate regime probabilities at a particular point in time.

the UK, the German, and the Japanese markets are unlikely to reside in a bull regime while the US market is in a bear regime. However, the scatter plots cannot assess which market leads or lags the other. For example, the plot for the Japanese market might imply that the US is rarely in a bear regime while Japan is still in a bull regime. However, it might also imply that Japan is rarely in a bull regime while the US still resides in a bear regime. The following results on the synchronization of regime cycles will analyze these dependence structures in more detail.

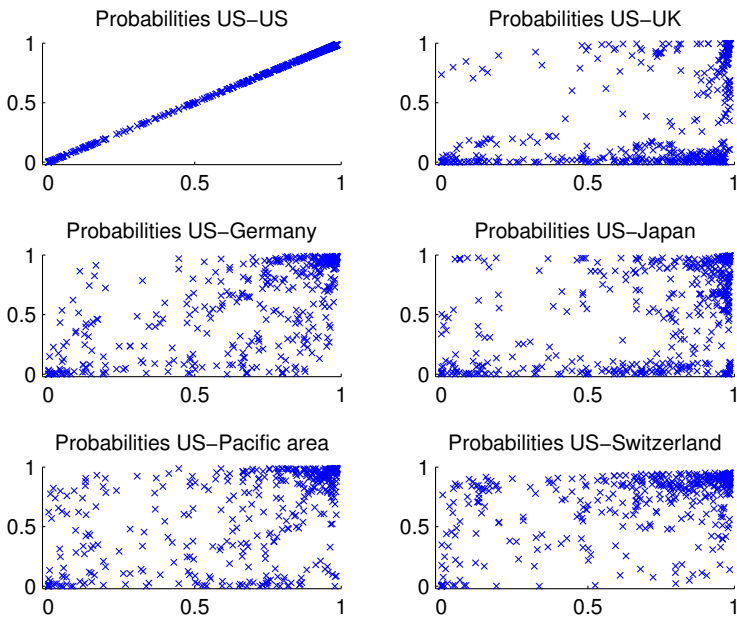


Figure 4.4: Scatterplots of pairwise univariate regime probabilities. The figure plots jointly occurring regime probabilities of the US and a corresponding sample market (to stay in a bull market). The x-axis depicts the probability of the US and the y-axis the probability of the other market. The period under study is from January 1976 to February 2014.

4.4.1 Analysis of In-Phase Synchronization

Pairwise Correlation and Concordance

Table 4.8 presents initial results for the pairwise synchronization of regime cycles. The lower triangular values depict the correlations between two regime processes. The upper triangular values illustrate the concordance between these processes.

In terms of correlation, the in-phase synchronization among individual regime cycles is generally low. Only Germany shows a moderate degree of regime correlation with the US and Switzerland. In accordance with previous results, the regime correlations of Japan, Pacific ex Japan, and the UK are rather low.

For the concordance values, a similar picture emerges with respect to the relative concordance between individual markets. Due to the transformation into binary probabilities, concordance values generally turn out higher. Except for Japan, all countries show concordance values above 0.70. Note, however, that these values might be biased. Due to the serial correlation of the Markov processes, uncorrelated cycles tend to indicate higher concordance values than expected. The analysis of SPPS and SNS will therefore provide more reliable information.

Table 4.8: *Correlation parameters (lower triangular values) and concordance indices (upper triangular values) between univariate regime processes, covering the period from December 1975 to February 2014.*

	US	UK	GE	JP	PA	SW
US		0.54	0.79	0.67	0.75	0.79
UK	0.40		0.68	0.53	0.52	0.51
Germany	0.69	0.54		0.60	0.64	0.74
Japan	0.43	0.12	0.21		0.71	0.61
Pacific ex Japan	0.56	0.23	0.32	0.41		0.71
Switzerland	0.51	0.40	0.61	0.24	0.37	

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

SPPS and SNS

Table 4.9 presents the results for SPPS and SNS. The lower triangular values disclose the difference between the mean regime probabilities of two markets. The corresponding significance for the null hypothesis of equal means is indicated by ***, **, and * for the 1%, 5%, and 10% level of significance.

These results reject the null of equal means for most country pairs. Significant exceptions are described. However, the analysis of means presents only the first step to testing for perfect correlation. The upper triangular values show GMM estimated correlations between the regime processes. In some cases, these differ significantly from the unadjusted correlations in Table 4.8. The standard errors to test for perfect correlation are again HAC consistent estimates. Despite this HAC correction, the hypothesis of perfect correlation is rejected for each pair of regime cycles.

Table 4.9: *Tests for strong perfect positive synchronization (SPPS) and strong non-synchronization (SNS) between univariate regime processes, covering the period from December 1975 to February 2014. Lower triangular matrix: differences between the mean regime probabilities of two processes; upper triangular matrix: pairwise correlations between regime probability processes estimated with GMM.*

	US	UK	GE	JP	PA	SW
US		0.28	0.54	0.33	0.38	0.36
UK	-0.40		0.47	0.10'''	0.16'''	0.26
Germany	-0.13	0.27		0.17'''	0.21	0.41
Japan	-0.21	0.19*	-0.08***		0.40	0.19
Pacific ex Japan	-0.09**	0.31	0.04***	0.12**		0.25
Switzerland	0.07***	0.47	0.20	0.28	0.16	

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

*, **, and *** indicate equal means between two regime processes at the 1%, 5%, and 10% level of statistical significance.

''' indicates that pairwise correlations are not statistically significantly different from zero at the 1% level.

The correlation estimates in the upper triangular matrix are also used to testing for strong non-synchronization. The 10%, 5%, and 1% significance levels for the null

of non-synchronization are depicted by ', ', and '''. The results indicate that the UK regime process is uncorrelated with those of Germany and Japan. This result is consistent with previous observations, where the UK's switching behavior differed significantly from other markets.

The results of in-phase synchronization analysis confirm that the two polar assumptions of standard regime-switching models can generally be rejected. However, this synchronization analysis has two main problems: first, it detects only contemporaneous effects; and second, it averages out possible observations over the sample path. The next section will resolve the first issue by allowing for possible shifts among regime processes. Rolling-window analysis presents a possible solution to the latter issue. This technique will be introduced in Section 4.4.3.

4.4.2 Analysis of Phase-Shifted Synchronization

Cross-Correlation

Simple correlation analysis revealed only moderate signs of dependence among individual regime processes. The analysis of cross-correlations should partially relax this problem. Table 4.10 shows the cross-correlations between regime cycles. The lag period is defined as a maximum of six months. Upper triangular values depict the optimal number of leads or lags between two cycles. The optimal lag is thereby determined by the highest observed correlation over all lags. Corresponding correlation values are illustrated in the lower triangular matrix.

Despite this additional flexibility, the correlation only gradually increases. Most country pairs indicate that zero-lag correlations return the best results. Only the regime processes of Japan and the Pacific area show shifted dependence with the US, the UK, and Germany. However, the correlations increase only insignificantly with these additional lags. These results illustrate that a symmetric shift in cycles can generally be rejected.

Table 4.10: Analysis of cross-correlation between two univariate regime processes, covering the period from December 1975 to February 2014. Upper triangular values depict the lag with the highest correlation (in months) and lower triangular values return the corresponding correlation value (maximum lag: 6 months).

	US	UK	GE	JP	PA	SW
US		0	-1	2	1	0
UK	0.40		1	2	1	0
Germany	0.70	0.56		4	4	0
Japan	0.45	0.14	0.23		0	0
Pacific ex Japan	0.57	0.23	0.37	0.41		0
Switzerland	0.51	0.41	0.61	0.24	0.37	

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

Dynamic Time Warping

Previous techniques mainly focused on the two polar cases of synchronization. However, it would be interesting to see if non-linear transformations result in intermediate degrees of synchronization. Table 4.11 therefore presents results for the dynamic time warping of regime cycles. The depicted values represent the relative distance between two cycles.⁸ Unity indicates perfect negative dependence and implies that the cycles always reside in opposite states. In contrast, a value of zero indicates perfect – possibly time-shifted – synchronization among cycles.

The lower triangular matrix depicts the results for a maximum lag of six months and the upper triangular matrix shows the results without any lag. Lower DTW values thereby imply better matching cycles. The potential synchronization among two cycles significantly improves with the lag size. Most country pairs were able to markedly reduce their standardized DTW values with increasing lags.

⁸This distance is measured as the total cycle distance divided by the sample length $DTW(T, T)/T$.

Table 4.11: Analysis of dynamic time warping between two univariate regime processes for a maximal lag length of zero (upper triangular values) and six months (lower triangular values). The period under study is from December 1975 to February 2014.

	US	UK	GE	JP	PA	SW
US		0.25	0.19	0.32	0.23	0.20
UK	0.21		0.35	0.45	0.26	0.27
Germany	0.09	0.31		0.34	0.30	0.22
Japan	0.24	0.41	0.26		0.32	0.33
Pacific ex Japan	0.15	0.19	0.20	0.24		0.26
Switzerland	0.14	0.24	0.15	0.27	0.20	

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

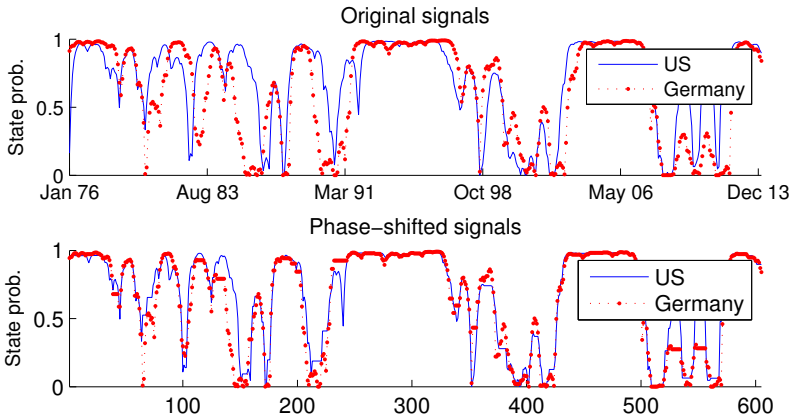


Figure 4.5: Dynamic time warping of the US and the German regime process. a) Original regime processes; b) phase-shifted regime processes. The period under study is from January 1976 to December 2014. The x-axis of the lower graph indicates the length of the warping path. Due to the one-to-many point comparison, the warping paths are of different length than the original cycles.

Figure 4.5 shows how the original regime cycles of the US and Germany are warped to better fit the opposing cycle. Especially phases with minor noise can be adapted.

Moreover, the visual inspection reveals that regime-switches are occasionally shifted by some months. These results support the use of DTW to identify asynchronous regime dependencies.

Further, Figure 4.6 matches cycles of the US and Germany for the period between December 2000 and April 2009. Parallel moving red lines indicate that the two cycles are synchronized. As soon as multiple lines concentrate in one point, this might indicate a change of synchronization. During such transitional stages, the US changes its regime whereas Germany still resides in its original state. The German process cannot match the new regime probability of the US. Therefore, it is linked to an earlier point. As soon as Germany switches its state, its regime probabilities once again match those of the US, and the regime cycles are resynchronized.⁹ This behavior clearly reveals the time-varying structure of regime cycle synchronization and puts further emphasis on the analysis in Section 4.4.3.

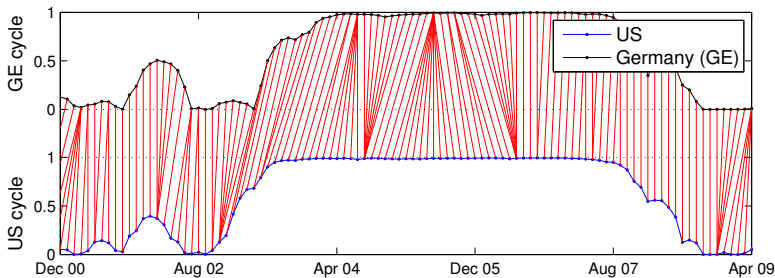


Figure 4.6: Matching of the US and the German univariate regime processes using dynamic time warping (up to six months lag). The period under study is from December 2000 to April 2009.

Longest Common Subsequence

A potential drawback of the DTW technique is that it matches each individual observation. This assumption might be too restrictive, given the possibility that two regime cycles temporarily behave independently. The LCSS approach resolves this problem and ignores phases without resemblance.

⁹The same behavior also applies vice versa.

Table 4.12: *Analysis of longest common subsequence between two univariate regime processes for a maximal warping length of zero (upper triangular values) and six months (lower triangular values). The period under study is from December 1975 to February 2014.*

	US	UK	GE	JP	PA	SW
US		0.41	0.68	0.45	0.64	0.70
UK	0.45		0.53	0.40	0.41	0.38
Germany	0.81	0.60		0.40	0.53	0.62
Japan	0.54	0.48	0.51		0.48	0.43
Pacific ex Japan	0.74	0.47	0.65	0.56		0.58
Switzerland	0.78	0.44	0.71	0.50	0.68	

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

Table 4.12 presents the results of the LCSS approach. In contrast to DTW, a unit LCSS value describes perfectly matched regime processes, whereas a zero value depicts perfect non-synchronization.¹⁰ The measure indicates the relative frequency of events when two – possibly phase-shifted – cycles show similar regime probabilities. Similarity is thereby conditioned on a threshold function. For the underlying example, the threshold value is set at 0.20. Accordingly, the difference between two regime probabilities needs to be lower than 20% for the cycles to be considered equal. Again, the upper triangular matrix represents the case of no lags. The lower triangular matrix allows for up to six months of lag between the two regime processes.

The results in Table 4.12 indicate the similarity of the US to other cycles. In the case of the six-month phase shift, the values increase markedly. Especially the Pacific area and Switzerland show much closer resemblance to the assets in the core sample.

Figure 4.7 presents the regime shifts between the US and the UK when dynamically adjusted by LCSS. These shifts take a similar form as for DTW. When accounting for the non-synchronous switches through LCSS, most distortions of the cycles can

¹⁰Note that the standardized LCSS value is not bounded by unity. However, the best possible matching is indicated by unity.

be eliminated or at least reduced. Due to the focus on similarity, periods of independence are accurately detected. The significant reduction of the cycle probabilities before the 1990s confirm this effect. The LCSS results therefore provide further input for the specification of an appropriate regime-switching model.

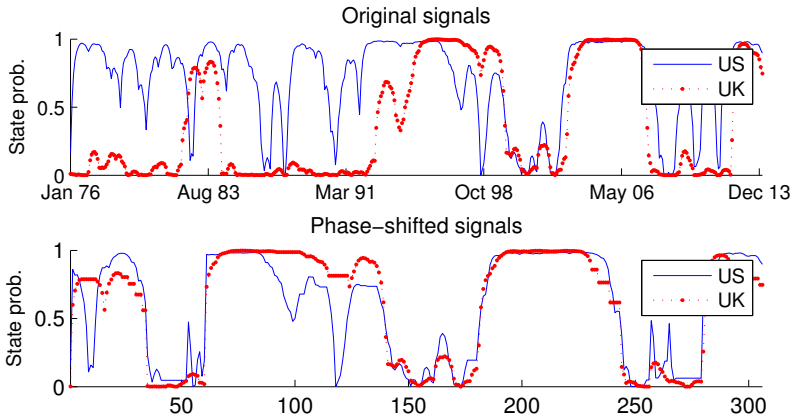


Figure 4.7: Longest common subsequence of the US and the UK regime process. a) Original regime processes; b) phase-shifted regime processes. The period under study is from January 1976 to December 2014. The x-axis of the lower graph indicates the length of the warping path. Due to the one-to-many point comparison, the warping paths are of different length than the original cycles.

Finally, Figure 4.8 presents the matching of the US and the German regime cycle. Red lines indicate these matches. Given the LCSS threshold value, non-matching phases are ignored. These phases are depicted by blank spaces and can be interpreted as times of non-synchronization. Whenever the red lines reveal a vertical structure, they imply in-phase synchronization. However, if the lines are skewed to the left or to the right, they imply some form of phase-shifted synchronization. Again, the patterns are very conclusive and persist for a limited time. This behavior further supports the hypothesis that regime processes change their degrees of synchronization across time and should therefore be analyzed for time-varying synchronization.

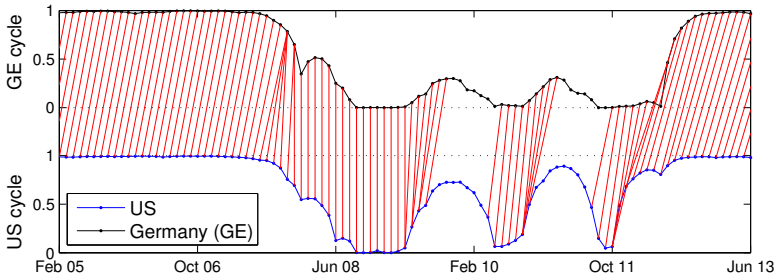


Figure 4.8: Matching of the US and the German univariate regime processes using longest common subsequence (up to six months lag). The period under study is from February 2005 to June 2013.

4.4.3 Analysis of Time-Varying Synchronization

The DTW and LCSS analysis presented initial evidence for time-varying synchronization across regime processes. At certain times, the cycles are perfectly synchronized. At other times, however, their behavior is purely idiosyncratic. Of course, there is no unique model able to capture all of these dependencies. Rolling-window analysis, however, might help to capture some of these effects.

Figure 4.9 exemplifies such an analysis, which determines the correlations over a 60-month rolling window. The correlations are thereby measured between the world regime process and a specific country's regime process. The six market cycles were separated into three groups based on visual inspection.

The correlations between the country cycles and the world cycle are mostly non-negative. Periods of perfect correlation or full independence are thereby of limited nature and strongly depend on the selected asset. Only the US, Germany, and the UK show periods of almost perfect correlation with the world regime cycle (mainly after 1998). This might be an indication for world market integration and for a possible increase in regime synchronization of individual markets. Nevertheless, most markets still show significant differences in their correlations with the world cycle. These observations further highlight the time-varying degree of synchronization and

therefore advocate the specification of a regime-switching model, which dynamically accounts for the dependence structure of individual regime cycles.

Figure 4.9 illustrates that some assets show more similar behavior than others. These assets might therefore be clustered and governed by a joint regime process. Figure 4.10 further presents the squared differences between the regime probabilities of the US and the other two assets in the core sample. It shows that the cycles of the US and Germany are very similar most of the time. The UK, on the other hand, experiences strong deviations from the US regime cycle. These results indicate that some assets might show more similar switching dynamics over time than other assets. Their joint analysis therefore presents a logical next step.

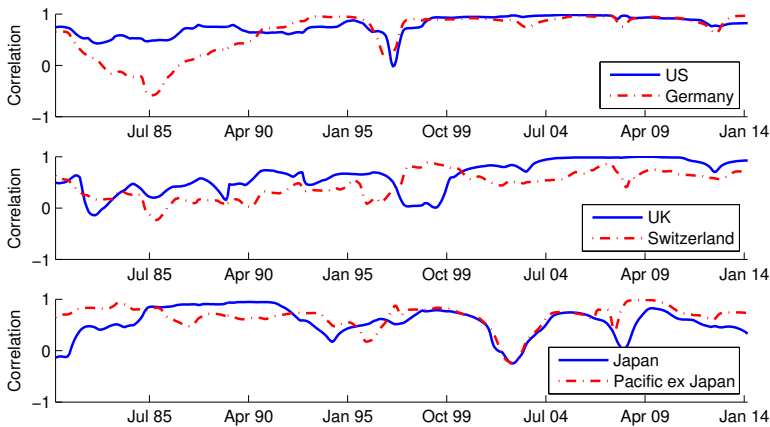


Figure 4.9: Time-varying correlation dynamics of univariate regime processes (using a 60-month rolling window). Time-varying correlation is measured between sample cycles and the world index regime cycle. Markets are grouped based on visual inspection. The period under study is from January 1976 to February 2014.

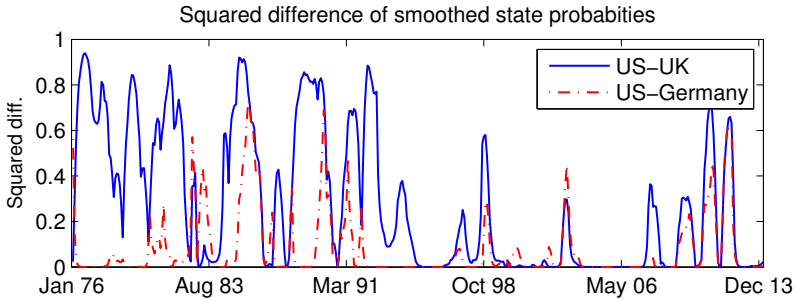


Figure 4.10: *Squared differences between the regime probabilities of the US and the other two markets in the core sample. The period under study is from January 1976 to February 2014.*

The rolling-window analysis in this section has partially explained the low dependence structures among regime cycles. Put simply, the structure is not constant over time. The model presented in Chapter 5 will incorporate these findings into the specification of a more flexible regime-switching model.

4.5 Summary

This chapter analyzed different country and regional indices for their regime-switching dynamics. Analysis revealed that neither univariate nor multivariate models appropriately cover the regime dynamics among assets.

Multivariate regime-switching models imply that all underlying assets follow the same or fully independent regime process(es). Tests for perfect synchronization and for strong non-synchronization rejected these assumptions.

Consequently, an alternative form of synchronization seems to govern inter-regime dynamics. Subsequent analysis has tested for phase-shifted and time-varying degrees of synchronization. Phase-shifted techniques were able to partially improve the degree of synchronization. However, the optimal LCSS path initially indicated that the dependence of regime cycles varies over time. Subsequent rolling-window analysis confirmed this dynamic synchronization. An appropriate model should therefore

consider not only regime processes but also the dependence structures between these processes to be time-varying.

Additional synchronization tests showed that some assets resemble a more homogeneous switching behavior than others. A regime-switching model should therefore cluster these assets. To keep the model parsimonious, the regime dynamics would subsequently be measured among these clusters.

The next chapter builds on these insights to specify a flexible regime-switching model. This model allows groups of assets to be governed by individual Markov chains, where these Markov chains underlie a time-varying degree of synchronization.

Chapter 5

Flexible Regime-Switching Models

Standard multivariate regime-switching models *a priori* assume switches in regimes to be perfectly synchronized or fully independent across assets. However, the analysis of empirical data in this study has revealed a different picture. Neither perfect synchronization nor full independence appropriately describe the switching behavior across time series. In fact, individual time series show changing regime-dependence structures.

This chapter introduces a model which resembles these regime and synchronization dynamics. The combination of the two polar dependence assumptions allows the model to replicate intermediate degrees of synchronization.

Phillips (1991) was among the first to propose the idea of non-perfectly correlated regime cycles in a joint model. Two lines of research have emerged from this idea (see Table 5.1 for a summary of the most important studies).

The first line of research assumes the existence of a common regime cycle, which is shared by all underlying assets. However, individual assets might depend on this cycle with a time lag, and shifts are either synchronous (Hamilton & Perez-Quiros, 1996; Paap et al., 2009) or asynchronous (Ang & Bekaert, 1999, 2002a; Cakmakli et al., 2011). Synchronous shifts imply identical lead or lag times irrespective of the prevailing regime. For example, asset *a* might lead switches in asset *b* by a certain time period irrespective of whether it switches from a bear to a bull state or vice versa. In contrast, asynchronous shifts depend on the prevailing regime. In regime 1,

asset a might lead switches in asset b by a different period length than in regime 2 (Cakmakli et al., 2011).

The second line of research assumes that each asset follows its own regime process. However, individual processes show some degree of synchronization. If regime cycles are assumed to be perfectly synchronized, the assets follow a joint Markov chain process. On the other hand, if cycles are assumed to behave fully independently, assets are driven by individual processes. The actual degree of synchronization emerges from the weighted linear combination of these two extreme cases of interdependence (Bengoechea et al., 2006; Camacho & Perez-Quiros, 2006). The weights of these cases are inferred from the data and can be either constant or time-varying. This approach has been applied by Camacho and Perez-Quiros (2006), Bengoechea et al. (2006), Hamilton and Owyang (2012), and Leiva-Leon (2012a, 2012b).

So far, most research on the dynamics of regime synchronization has been limited to bivariate time series (see Table 5.1). This limitation is mainly motivated by the size of the parameter set and by the model complexity. Larger time series cause a rapid increase in both dimensions, as every additional asset requires its own Markov chain process.

The present study introduces a feasible solution to partially circumvent these problems. This solution derives from the model specification of Camacho and Perez-Quiros (2006) and Leiva-Leon (2012a, 2012b). Individual regime processes thereby show a dynamic degree of dependence. However, whereas the previous models assumed that each process governs a single asset, the specification presented here applies to asset groups.

The dynamic dependence structure and the focus on asset groups are motivated by the results in Chapter 4. The analysis of pairwise regime cycles highlighted the finding that the synchronization of individual regime cycles changes over time but that some assets show more similar switching dynamics than others. These assets can therefore be clustered. Consequently, assets show relatively similar dynamics within a group but heterogeneous dynamics across groups. As a result, the proposed model does not imply an individual Markov chain for each asset. Instead, assets in the same cluster are governed by a joint Markov chain process.

Table 5.1: Literature overview of regime-switching approaches that account for synchronized regime processes.

Author	Model dimension	Switching terms			VAR	Other terms	dim(θ)		dim(θ) in standard specification	
		Mean	Std	Corr					2 assets	6 assets
Common but shifted cycles	Phillips (1991)	Yes	-	-	No	Yes	-	14-16	12-14	56-60
	Ang and Bekaert (1999, 2002a)	Yes	Yes	Yes	-	-	-	14	14	66
	Cakmakli, Paap, and Van Dijk (2011)	Yes	Yes	Yes	-	Yes	-	27	16	60
	Camacho et al. (2006)	Yes	-	-	No	Yes	-	16	15	55
Individual but synchronized cycles	Bengoechea et al. (2006)	Yes	-	-	No	Yes	-	13	16	-
	Otranto (2010)	No	No	Yes	-	-	DCC	77	32	120
		No	No	Yes	-	-	DCC	63	28	108
		No	No	No	-	-	DCC	63	29	75
	Leiva-Leon (2012)	Yes	No	No	-	-	-	15	18	64
Guidolin and Timmermann (2007)		Yes	-	-	Yes	Yes	-	48	32	120

The table shows different studies that accounted for the synchronization between individual regime-switching processes. Model dimension describes the setup of the model, switching terms indicates which parameters were defined as regime-dependent, VAR defines whether the regime-switching model followed VAR dynamics. dim(θ) gives the number of model parameters. The last two columns define the number of parameters that these models would require in case of a two-state setup with switching means, switching variances, and switching correlations.

Following the basic model setup of Camacho and Perez-Quiros (2006), these joint Markov chains, in turn, show some degree of – possibly time-varying – synchronization. In line with Billio et al. (2006) and Otranto (2010), this specification is subsequently called the flexible regime-switching model.¹ The term “flexible” refers to the instance that underlying asset groups can reside in different states, but are covered in a joint model.

The flexible model clearly competes with standard multivariate regime-switching models that allow for more than two regimes. Guidolin and Timmermann (2007) showed that the consideration of more than two regimes significantly increases the explanation power of a model. The additional regimes capture stages of transition. This feature is especially useful when analyzing different asset classes, as they might show more heterogeneous switching behaviors. A model that infers this data from just two regimes might thus return weak results.

However, a model with more than two regimes poses two potential drawbacks: first, even if regimes might cover stages of transition, all assets still follow the same common regime process; and second, the parameter set significantly increases with every additional regime. The latter argument describes a particular strength of the flexible regime-switching model. Every asset in this model follows only two regimes. The combination of these regimes already results in a higher state space. Whether asset b resides in the same regime as asset a is thereby irrelevant for the moments of asset a . As a consequence, the parameter set remains low because the moment parameters are reused in non-synchronized states. The synchronization of the Markov chains is then inferred from the data.

The next section specifies the general structure of the flexible regime-switching model. Following the framework in Chapter 2, it first describes the model’s data-generating process, followed by the Markov chain process. The subsequent sections then present corresponding state inference and parameter estimation techniques.

¹Otranto (2010) presented a flexible regime-switching model, which clustered multivariate time series. However, Otranto did not state particular synchronization dynamics. He therefore estimated a large parameterized model with a transition probability matrix that matched the size of the state space. This thesis, in contrast, specifies a model that infers corresponding synchronization dynamics across clusters. This setup significantly reduces the parameter set. Furthermore, the model provides a factorization algorithm to efficiently estimate PSD correlation matrices.

5.1 General Model Specification

5.1.1 Data-Generating Process

The multivariate DGP of Chapter 2 is readily extended to the flexible regime-switching model. To simplify the flexible model, its mean vector is assumed to be regime-independent $\mu_{s_t} = \mu$. Previous likelihood ratio tests demonstrated that regime-invariant means do not significantly reduce the likelihood of the analyzed models. Chapter 6 will confirm these results for the flexible model.

Similar to the baseline specification, the covariance matrix Σ_{s_t} is defined as being regime-dependent. This assumption is also in line with previous likelihood ratio test statistics. The covariance matrix is again decomposed into standard deviation and correlation terms

$$\begin{aligned} \mathbf{y}_t &= \mu_{s_t} + \Sigma_{s_t}^{1/2} \varepsilon_t & \varepsilon_t &\sim \mathcal{N}(0, I) \\ \Sigma_{s_t} &= D_{s_t} R_t D_{s_t} \\ R_t &= R_{s_t}, \end{aligned} \tag{5-1}$$

where D_{s_t} is a diagonal matrix of regime-dependent standard deviations, and where R_{s_t} is a regime-dependent correlation matrix.

Chapter 2 already stated different causes for decomposing the covariance matrix. The flexible model specification reveals a further reason: in this model, individual assets might reside in different states at the same time. Consequently, standard deviation parameters – and to some extent correlation parameters – can be reused in intermediate states. However, this approach does not as easily apply to covariances. The next section will illustrate this reuse of parameters and the corresponding algorithm in detail.

5.1.2 Bivariate Model with Multiple Markov Chains

The previous DGP specification has not revealed any structural differences between the flexible model and the benchmark model of Chapter 2. In fact, most differences emerge from the regime-generating process.

Similar to the benchmark specification, flexible regime-switching underlies a Markovian regime-generating process. In the multivariate benchmark model, however, all assets follow a common Markov chain. The flexible model, in contrast, assumes that assets underlie individual regime processes. For example, a simple two-state bivariate model implies that both assets reside either in regime 1 or in regime 2. The flexible model extends this state space. Table 5.2 shows how the same two regimes define four different states s_t^* . Asset a can reside in one of the two regimes irrespective of asset b 's current regime.

This specification does not imply that the two regime processes are fully independent. In fact, their dependence structure needs to be inferred from the data. For example, the two processes might be perfectly synchronized. In this case, they would reside exclusively in $s_t^* = 1$ and $s_t^* = 4$. Phillips (1991) pointed out that the true dependence structure is expected to lie somewhere between full independence and perfect synchronization. A flexible regime-switching model approximates this behavior.

s_t^*	s_t^a	s_t^b
1	1	1
2	1	2
3	2	1
4	2	2

Table 5.2: *State space of the bivariate switching model.*

The flexible setup is akin to a standard multivariate model with four instead of two regimes. However, the flexible model's specification reveals a particular strength: It depends only on two basic regimes. Accordingly, moment parameters need not be estimated for four different states. For example, whether the overall state resides in

$s_t^* = 1$ or in $s_t^* = 3$ does not influence the moment parameters of asset a . In both cases, asset a resides in regime 1, irrespective of asset b 's current regime.

States $s_t^* = 1$ and $s_t^* = 4$ describe the known regimes of the multivariate benchmark model. In these states, both assets reside in the same regime and their switches are perfectly synchronized. In the intermediate cases $s_t^* = 2, 3$, parameters can be reused from the polar states $s_t^* = 1, 4$.² Formulas 5-2 and 5-3 exemplify this approach for a bivariate two-state specification. For ease of exposition, both mean and covariance terms are defined as regime-dependent. The vectors in 5-2 combine the means of individual assets according to their prevailing regime. In formula 5-3, the same applies to standard deviations

$$\mu_1 = \begin{pmatrix} \mu_1^a \\ \mu_1^b \end{pmatrix} \quad \mu_2 = \begin{pmatrix} \mu_1^a \\ \mu_2^b \end{pmatrix} \quad \mu_3 = \begin{pmatrix} \mu_2^a \\ \mu_1^b \end{pmatrix} \quad \mu_4 = \begin{pmatrix} \mu_2^a \\ \mu_2^b \end{pmatrix} \quad (5-2)$$

$$\begin{aligned} \Sigma_1 &= \begin{pmatrix} (\sigma_1^a)^2 & \rho_1 \sigma_1^a \sigma_1^b \\ \rho_1 \sigma_1^a \sigma_1^b & (\sigma_1^b)^2 \end{pmatrix} & \Sigma_2 &= \begin{pmatrix} (\sigma_1^a)^2 & \rho_2 \sigma_1^a \sigma_2^b \\ \rho_2 \sigma_1^a \sigma_2^b & (\sigma_2^b)^2 \end{pmatrix} \\ \Sigma_3 &= \begin{pmatrix} (\sigma_2^a)^2 & \rho_3 \sigma_2^a \sigma_1^b \\ \rho_3 \sigma_2^a \sigma_1^b & (\sigma_1^b)^2 \end{pmatrix} & \Sigma_4 &= \begin{pmatrix} (\sigma_2^a)^2 & \rho_4 \sigma_2^a \sigma_2^b \\ \rho_4 \sigma_2^a \sigma_2^b & (\sigma_2^b)^2 \end{pmatrix}, \end{aligned} \quad (5-3)$$

where μ_k^i and σ_k^i depict the mean and standard deviation of asset i in state k , and where ρ_k defines the state-dependent correlation between the two assets.

However, this approach does not apply to the parameters of the correlation matrix. Correlation coefficients of states 1 and 4 cannot simply be reused in states 2 and 3. In contrast to moment parameters, correlation defines the interrelation between two time series. A correlation matrix should therefore not be specified solely with regard to the prevailing regime of asset a or asset b . States 2 and 3 thus require separately estimated correlation parameters ρ_2 and ρ_3 .

Again, different ways exist to specify such correlation dynamics. Models with a

²Of course, parameters from the intermediate cases could also be reused to define the two polar states. However, the standard regime-switching specification implies that the polar states should be used as reference cycles.

shifted regime cycle often assume the leading asset's state to determine the correlation matrix (see Cakmakli et al., 2011). In dynamic synchronization models, the covariance or correlation matrix is usually defined as being regime-independent (Camacho & Perez-Quiros, 2006; Leiva-Leon, 2012a, 2012b), as following a diagonal structure (Hamilton & Owyang, 2012), or as being fully parameterized in each state (Otranto, 2010).

However, none of these solutions seems appropriate. Either they ignore the content of information inherent in regime-dependent correlations or they over-parameterize the model. The latter point is of particular relevance for multivariate flexible models, as their parameter sets increase faster than the size of the underlying time series.

This thesis develops a novel approach to limiting this increase in correlation parameters. After introducing the multivariate flexible model, Section 5.3 will specify this algorithm.

Synchronization of the Markov Chains

The state space extension presents an efficient way of accounting for different forms of regime dependence. However, the previous setup did not reveal any information about the dependence structure of the underlying Markov chains. To complete the model specification, these synchronization dynamics need to be specified:

$$sync(s_t^a, s_t^b).$$

The conventional approach in the literature is to *a priori* define this relationship structure (Leiva-Leon, 2012b, 2012a). Regime cycles are thereby commonly assumed to be either fully synchronized or completely independent. By contrast, the model presented here proposes an intermediate degree of synchronization, located between these two polar cases (following Camacho & Perez-Quiros, 2006; Leiva-Leon, 2012b, 2012a). As this intermediate dependence is unknown, however, the model is best described in terms of the two polar cases (Phillips, 1991; Camacho & Perez-Quiros, 2006; Bengoechea et al., 2006; Leiva-Leon, 2012a, 2012b). In case

the regime processes of asset a and asset b are completely independent

$$\Pr(s_t^a = j_a, s_t^b = j_b | \delta_t = 0) = \Pr_a(s_t^a = j_a) \Pr_b(s_t^b = j_b) \quad (5-4)$$

holds, where $\delta_t = 0$ indicates full independence and where $\Pr(s_t^a = j_a, s_t^b = j_b | \delta_t = 0)$ is the conditional joint state probability.³ Further, $\Pr_a(s_t^a = j_a)$ and $\Pr_b(s_t^b = j_b)$ describe the filtered marginal regime probabilities of asset a and b .

In contrast, if the two assets are fully dependent (or synchronized), both assets are assumed to be governed by the same regime process

$$\Pr(s_t^a = j_a, s_t^b = j_b | \delta_t = 1) = \Pr_{ab}(s_t^a = s_t^b = j_a), \quad (5-5)$$

where $\delta_t = 1$ describes perfect synchronization. In both specifications, the probability law of $\Pr_a(s_t^a = j_a)$, $\Pr_b(s_t^b = j_b)$, and $\Pr_{ab}(s_t^a = s_t^b = j_a)$ is defined by individual Markov chain processes. Consequently, each process has its own transition probability matrix. This separation into individual processes reduces the parameter set and makes the model more flexible.⁴

Specifications 5-4 and 5-5 reveal the conditionality of the joint state-probabilities on δ_t . Camacho and Perez-Quiros (2006) and Bengoechea et al. (2006) found that the synchronization of regime cycles can be described by $(1 - \delta_t)$ times 5-4 and δ_t times 5-5, where $0 \leq \delta_t \leq 1$. In line with Camacho and Perez-Quiros (2006), the weight δ_t represents a measure of regime cycle synchronization.

To combine the information inherent in 5-4 and 5-5, a probability law for δ_t needs to be defined. Camacho and Perez-Quiros (2006) assumed a time-invariant dependence structure among the two cycles, that is, $\delta_t = \delta$. Consequently, they defined a fixed

³For ease of exposition, the conditionality on past information (filtration) and on the parameter set θ is omitted.

⁴Applied to the example in Table 5.2, this implies that no (4×4) matrix, but three individual (2×2) matrices need to be estimated. Two of these (2×2) matrices cover the individual regime dynamics and the third matrix covers the joint dynamics. This approach reduces the parameter set from 12 parameters (4×4 matrix) to only six parameters (2×2 matrices). Further, Chapter 7 will show that this specification offers additional advantages for the multistage optimization procedure, as the regime processes can be simulated individually.

term δ as part of the parameter set θ . The factors δ and $(1 - \delta)$ can then be expressed in terms of probabilities and serve as weights for the conditional joint probabilities. Based on these weights, an unconditional joint probability results from

$$\begin{aligned} \Pr(s_t^a = j_a, s_t^b = j_b) &= \Pr(s_t^a = j_a, s_t^b = j_b | \delta = 0) \Pr(\delta = 0) + \\ &\Pr(s_t^a = j_a, s_t^b = j_b | \delta = 1) (1 - \Pr(\delta = 0)). \end{aligned}$$

However, the results in Chapter 4.4 illustrated a dynamic dependence structure among individual regime cycles. These results contradict a time-invariant specification of δ_t . Similar to Leiva-Leon (2012a, 2012b), this thesis thus follows a dynamic approach and assumes a time-varying synchronization factor δ_t . Leiva-Leon (2012a, 2012b) proposed that this factor is driven by an additional Markov chain. The unobserved state variable s_t^δ of this process is time-varying. Similar to the other Markov chains, it is inferred from the data.

Due to this additional regime-variable, the state space in Table 5.2 is extended by another dimension. Table 5.3 presents this extended state space, where the last column indicates the regime of δ_t . In case $s_t^\delta = 1$, the regime cycles of a and b are independent. In case $s_t^\delta = 2$, asset a and b follow a common regime cycle. This information is collected in the joint state variable s_t^* . The joint probability of each case is thereby described by

$$\begin{aligned} \Pr(s_t^* = j) &= \Pr(s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta) \\ &= \Pr(s_t^a = j_a, s_t^b = j_b | s_t^\delta = j_\delta) \Pr(s_t^\delta = j_\delta), \end{aligned} \quad (5-6)$$

where $\Pr(s_t^a = j_a, s_t^b = j_b | s_t^\delta = j_\delta)$ is defined as in 5-4 and 5-5. A particular strength of this approach is its low level of complexity. The additional Markov chain does not increase the set of moment parameters. Further, it also maintains the size of the conditional density function. The conditional pdfs for states 1 to 4 resemble the same information as those for states 5 to 8. Hence, the state variable s_t^δ does not influence the conditional pdf but only the probabilities of its conditional states.

s_t^*	s_t^a	s_t^b	s_t^δ
1	1	1	1
2	1	2	1
3	2	1	1
4	2	2	1
5	1	1	2
6	1	2	2
7	2	1	2
8	2	2	2

Table 5.3: Extended state space with an additional synchronization factor.

Leiva-Leon (2012a, 2012b) showed that the results in Table 5.2 can be replicated by integrating $\Pr(s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta)$ through s_t^δ . However, the joint probability $\Pr(s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta)$ is in fact better suited to calculate the different regime processes of a , b , ab , and δ . Section 5.2 will discuss this process in more detail.

5.1.3 Multivariate Model with Multiple Markov Chains

So far, research on dynamic regime synchronization has focused mainly on bivariate time series. This emphasis clearly emerges from the argumentation presented in the previous section. First, more than two assets would require a synchronization factor δ_t for each pair of regime cycles in the sample. Second, the state space and the parameter set would significantly increase with every additional asset.

This study contributes to the solution of this problem. The proposed structure implies that multiple assets jointly define a regime cycle. Such asset groups underlie the same properties as in the previous model. However, regime processes are no longer inferred for univariate but for clustered time series. Assets within a cluster contemporaneously switch their regimes. Across clusters, the switching behavior is heterogeneous. The degree of synchronization across these clusters is thereby still measured by δ_t .

Based on this specification, the bivariate model is readily extended to multivariate time series. Mean and standard deviation parameters linearly increase with the sample size. Intermediate states $s_t^* = 2, \dots, K^{nc-1}$ reuse the moments of $s_t^* = 1, K^{nc}$. Thereby, K defines the number of states that each cluster follows (usually two) and the exponent nc accounts for the number of clusters.⁵

Given the above reasons, however, this approach does not apply to correlation matrices. An alternative specification would be to condition the correlation matrix on the regime of a particular asset or – as in the prevailing case – that of a particular cluster. The correlation matrix in equation 5-7, however, illustrates another solution. This “composite correlation matrix” is defined by two clusters a and b , which follow individual regime processes. In case $s_t^* = 1$ or $s_t^* = K^2$, both clusters follow the same regime ($s_t^a = s_t^b$). Consequently, the correlation across all assets is conditioned on either regime 1 or regime 2. In intermediate states $s_t^* = 2, 3$, the clusters reside in different regimes and the correlation matrix follows a more complex specification:

$$R_{s_t^*=2} = \begin{bmatrix} \begin{bmatrix} \rho_{11|s_t^a=1} & \rho_{12|s_t^a=1} \\ \rho_{21|s_t^a=1} & \rho_{22|s_t^a=1} \end{bmatrix} & \dots & \begin{bmatrix} \rho_{33|s_t^b=2} & \rho_{34|s_t^b=2} \\ \rho_{43|s_t^b=2} & \rho_{44|s_t^b=2} \end{bmatrix} \\ \begin{bmatrix} \rho_{31|s_t^a=1, s_t^b=2} & \rho_{32|s_t^a=1, s_t^b=2} \\ \rho_{41|s_t^a=1, s_t^b=2} & \rho_{42|s_t^a=1, s_t^b=2} \end{bmatrix} & & \end{bmatrix}. \quad (5-7)$$

Equation 5-7 exemplifies this concept for the intermediate state $s_t^* = 2$. In this state, cluster a is driven by regime 1 ($s_t^a = 1$) and cluster b is driven by regime 2 ($s_t^b = 2$).⁶ Formula 5-7 shows that the correlation within a cluster is solely conditioned on this cluster’s regime. Consequently, the regime in cluster a does not influence the correlation parameters in cluster b and vice versa.

This specification implies that the correlations within cluster a and b resemble the values of $s_t^* = 1$ and $s_t^* = 4$. Hence, as the values of $s_t^* = 1, 4$ can be reused for $s_t^* = 2, 3$, no additional parameters are required in the MLE parameter set θ so far.

A different picture emerges for the off-diagonal blocks. For ease of exposition, for-

⁵The bivariate model structure implies that $nc = 2$. Further, it is assumed that $K = 2$ for all clusters.

⁶This specification resembles state $s_t^* = 2$ in Table 5.3.

mula 5-8 abstracts the content of 5-7:

$$R_{s_t^*=2} = \begin{bmatrix} A_{s_t^a} & M'_{s_t^*} \\ M_{s_t^*} & B_{s_t^b} \end{bmatrix}, \quad (5-8)$$

where $A_{s_t^a}$ is an $(m \times m)$ regime-dependent correlation matrix for the assets in cluster a . The same structure applies to the $(n \times n)$ matrix $B_{s_t^b}$.

A different structure applies to the $(n \times m)$ matrix $M_{s_t^*}$. This matrix measures the correlations between assets of different clusters. It is therefore conditioned on the regimes of both clusters. The conditionality of this matrix is best expressed in terms of the overall state s_t^* . Again, if both clusters reside in the same regime, the correlations resemble those in $s_t^* = 1, 4$. However, in intermediate cases $s_t^* = 2, 3$, new correlation information emerges. The correlation matrices of $s_t^* = 1, 4$ do not contain these parameters. Values in $M_{s_t^*}$ thus represent additional parameters to be included in θ .

This specification leads to a slower growth of the parameter set, compared to a standard four-state multivariate regime-switching model. The latter would estimate the entire correlation matrix for each state. The current specification, however, requires only the correlation matrices of states $s_t^* = 1, K^2$ and the parameters in $M_{s_t^*}$ for $s_t^* = 2, 3$. Especially for time series with a larger number of assets, this approach significantly reduces the parameter set compared to a four-state model.⁷ At the same time, it covers the full dynamics of the data.

However, a particular problem arises with regard to the parameter estimation. Chapter 2 stated different reasons for estimating Cholesky decomposed parameters instead of correlation parameters. The main reason was to guarantee PSD matrices. The same problem applies to the current specification. Especially the correlation matrices in $s_t^* = 2$ and $s_t^* = 3$ might be subject to non-PSD results. The corresponding correlation matrices resemble information from assets that remain in different regimes,

⁷Note that a standard multivariate regime-switching model usually implies that the returns in individual regimes are normally distributed. Their weighted combination then allows one to replicate the non-normality of underlying data. Given the combination of clusters with different regimes in the same correlation matrix, however, the i.i.d. assumption should no longer hold as regime-dependent parameters are not independently observed for a particular regime. Nevertheless, the assumption still holds because the state spaces in Table 5.2 and 5.3, and not the individual regime processes, describe the final states.

which increases the chances of negative correlations. Consequently, a PSD structure is difficult to guarantee when optimizing the parameter set element-by-element.⁸

Cholesky decomposition allows one to partially circumvent these problems. However, it poses a further difficulty: Formulas 5-7 and 5-8 clearly separate the different clusters and define by which regime(s) a particular parameter is governed. This uniqueness does not apply to $L_{s_t^*}$, which defines the lower triangular Cholesky factor. The parameters in $L_{s_t^*}$ that define the block $A_{s_t^a}$ also influence the parameters in block $B_{s_t^b}$ and $M_{s_t^*}$.⁹ Consequently, the information of $L_{s_t^*}$ in $s_t^* = 1, 4$ cannot simply be reused in $s_t^* = 2, 3$. To calculate $L_{s_t^*}$ in $s_t^* = 2, 3$, an alternative approach needs to be defined. This approach must guarantee that the resulting correlation matrix $R_{s_t^*}$ is PSD and that it depicts the same block-diagonal correlations in $s_t^* = 2, 3$ as in $s_t^* = 1, 4$.

Chapter 5.3 develops an efficient algorithm to calculate these correlations in $s_t^* = 2, 3$. This algorithm presents one of the main contributions of this thesis to existing research.

Model Specification with More Than Two Clusters

The previous section illustrated a simple way of extending the bivariate synchronization model to multivariate time series. At the same time, the state space remained unaffected, as the model was still governed by only two Markov chains. The underlying assets were simply divided into two homogeneous subgroups, which were driven by the same two regime processes.

Larger samples, however, might show more distinct switching behavior. A separation into two clusters might thus be too restrictive. Clearly, increasing the number of clusters presents a trade-off between within cluster similarity and model feasibility. Additional clusters enable groups with more homogeneous switching behavior. However, they also imply a larger parameter set and a higher model complexity.

⁸Due to negative correlation values in the intermediate cases, the estimation of weakly bounded $(-1, +1)$ correlation matrices might often violate the PSD property.

⁹Note that the opposite does not hold, because the Cholesky algorithm starts from the upper-left corner of L .

Chapter 6 will show that a separation into two and three clusters provides accurate results. The previous specification with two clusters is thus extended by another cluster c . Following this specification, the processes of the three clusters are either fully independent or perfectly synchronized. In case the regime processes of the clusters are independent, specification 5-4 is extended by another marginal probability process for cluster c

$$\begin{aligned} \Pr(s_t^a = j_a, s_t^b = j_b, s_t^c = j_c | \delta_t = 0) \\ = \Pr_a(s_t^a = j_a) \Pr_b(s_t^b = j_b) \Pr_c(s_t^c = j_c). \end{aligned} \quad (5-4')$$

Contrastingly, if the individual clusters are fully dependent, all assets are governed by the joint regime process

$$\Pr(s_t^a = j_a, s_t^b = j_b, s_t^c = j_c | \delta_t = 1) = \Pr_{abc}(s_t^a = s_t^b = s_t^c = j_a). \quad (5-5')$$

Similar to the bivariate model, 5-4' and 5-5' are weighted by the probability process of δ_t . So far, the factor has determined the pairwise synchronization of two cycles by weighting their extreme forms of dependence. However, in case of more than two clusters, this feature no longer applies. The factor δ_t jointly applies to all three clusters. Consequently, it presents an average measure of synchronization between the possible cycle combinations (the tuples $a - b$, $a - c$, and $b - c$).¹⁰ The specification of pairwise synchronization measures, in contrast, would further increase the parameter set and would make the model infeasible to estimate.

Further, the correlation matrix in 5-9 shows that the additional cluster creates smaller cluster blocks and increases the number of submatrices. The matrix in 5-9 shows that an increase in clusters leads to an increase in off-diagonal block parameters. If each asset were finally defined by an individual cluster with two regimes, the number of correlation parameters would equal that of a standard four-regime model.¹¹ Note

¹⁰This procedure partially resembles the multivariate analysis of average cycle correlation applied by Candelon et al. (2008, 2009).

¹¹Every asset would represent an individual cluster. Consequently, each asset pair would be defined by two correlations when both assets reside in the same regime and by two correlations when they reside

that the current specification would still allow for more flexibility, as every asset can reside in its own regime.

$$R_{s_t^*} = \begin{bmatrix} \begin{bmatrix} \rho_{11|s_t^a=1} & \rho_{12|s_t^a=1} \\ \rho_{21|s_t^a=1} & \rho_{22|s_t^a=1} \end{bmatrix} & \dots & \dots \\ \vdots & \begin{bmatrix} \rho_{33|s_t^b=2} & \rho_{34|s_t^b=2} \\ \rho_{43|s_t^b=2} & \rho_{44|s_t^b=2} \end{bmatrix} & \vdots \\ \dots & \begin{bmatrix} \rho_{53|s_t^b=2, s_t^c=2} & \rho_{54|s_t^b=2, s_t^c=2} \end{bmatrix} & \begin{bmatrix} \rho_{55|s_t^c=2} \end{bmatrix} \end{bmatrix} \quad (5-9)$$

5.2 State Inference

The state inference procedure follows Hamilton's (1989, 1994) basic algorithm. Due to the inclusion of multiple regime cycles, however, it needs to be extended.¹² This extended algorithm underlies the same two-step procedure as in Hamilton (1989): an updating step and a forecasting step.

Step 1: The first step calculates the likelihood function and updates the state probabilities. Given the estimated parameters θ , the conditional joint densities can be calculated as the product of the state-specific density and the state probability

$$\begin{aligned} p(y_t, s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1}; \theta) = \\ f(y_t | s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta, \Omega_{t-1}; \theta) \\ \times \Pr(s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1}; \theta), \end{aligned}$$

where $f(y_t | s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta, \Omega_{t-1}; \theta)$ is the conditional density function and where $\Pr(s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1})$ is defined as in 5-6.¹³ As shown in Table 5.3, the different regime processes can be combined in the joint state s_t^* . This

in different regimes: $2N(N-1)/2 + 2N(N-1)/2$. This parameter number matches the number of correlations in a standard four-state regime-switching model: $4N(N-1)/2$.

¹²This extension follows the remarks in Camacho and Perez-Quiros (2006) and in Leiva-Leon (2012a, 2012b).

¹³Illustrated for the case with two clusters.

adjustment results in a simpler notation of the conditional joint densities

$$p(y_t, s_t^* = j | \Omega_{t-1}; \theta) = f(y_t | s_t^* = j, \Omega_{t-1}; \theta) \times \Pr(s_t^* = j | \Omega_{t-1}; \theta),$$

where $\Pr(s_t^* = j | \Omega_{t-1}; \theta)$ is specified as shown in equation 5-6.

The next step is to calculate the marginal densities of a specific state variable. This is done by summing the conditional joint densities over the remaining state variables (see also Leiva-Leon, 2012a, 2012b)

$$\begin{aligned} f(y_t, s_t^a = j_a | \Omega_{t-1}; \theta) &= \sum_{j_b=1}^2 \sum_{j_\delta=1}^2 p(y_t, s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1}; \theta) \\ f(y_t, s_t^b = j_b | \Omega_{t-1}; \theta) &= \sum_{j_a=1}^2 \sum_{j_\delta=1}^2 p(y_t, s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1}; \theta) \\ f(y_t, s_t^a = s_t^b = j_a | \Omega_{t-1}; \theta) &= \sum_{j_\delta=1}^2 p(y_t, s_t^a = j_a, s_t^b = j_a, s_t^\delta = j_\delta | \Omega_{t-1}; \theta) \\ f(y_t, s_t^\delta = j_\delta | \Omega_{t-1}; \theta) &= \sum_{j_a=1}^2 \sum_{j_b=1}^2 p(y_t, s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1}; \theta). \end{aligned} \tag{5-10}$$

Step 2: The second step of the algorithm updates and forecasts the state probabilities. The marginal distributions in 5-10 are therefore divided by their unconditional densities

$$\begin{aligned} \Pr(s_t^a = j_a | \Omega_t; \theta) &= \frac{f(y_t, s_t^a = j_a | \Omega_{t-1}; \theta)}{f(y_t | \Omega_{t-1}; \theta)} \\ \Pr(s_t^b = j_b | \Omega_t; \theta) &= \frac{f(y_t, s_t^b = j_b | \Omega_{t-1}; \theta)}{f(y_t | \Omega_{t-1}; \theta)} \\ \Pr(s_t^a = s_t^b = j_{ab} | \Omega_t; \theta) &= \frac{f(y_t, s_t^a = s_t^b = j_{ab} | \Omega_{t-1}; \theta)}{\sum_{j_m=1}^2 f(y_t, s_t^a = s_t^b = j_m | \Omega_{t-1}; \theta)} \end{aligned}$$

$$\Pr(s_t^\delta = j_\delta | \Omega_t; \theta) = \frac{f(y_t, s_t^\delta = j_\delta | \Omega_{t-1}; \theta)}{f(y_t | \Omega_{t-1}; \theta)},$$

where

$$f(y_t | \Omega_{t-1}; \theta) = \sum_{j_a=1}^2 \sum_{j_b=1}^2 \sum_{j_\delta=1}^2 p(y_t, s_t^a = j_a, s_t^b = j_b, s_t^\delta = j_\delta | \Omega_{t-1}; \theta). \quad (5-11)$$

Further, summing the logarithm of 5-11 over time returns the log-likelihood value of the time series. The resulting state probabilities are collected in the vectors $\xi_{t+1|t}^a$, $\xi_{t+1|t}^b$, $\xi_{t+1|t}^{ab}$, and $\xi_{t+1|t}^\delta$, where $\xi_{t+1|t}^i = \begin{pmatrix} \Pr(s_{t+1}^i = 1 | \Omega_t; \theta) \\ \Pr(s_{t+1}^i = 2 | \Omega_t; \theta) \end{pmatrix}$ for $i = a, b, ab$, and δ . Due to the independence of the different Markov chains, each process is updated individually

$$\begin{aligned} \xi_{t+1|t}^a &= \mathbf{P}_a \xi_{t|t}^a \\ \xi_{t+1|t}^b &= \mathbf{P}_b \xi_{t|t}^b \\ \xi_{t+1|t}^{ab} &= \mathbf{P}_{ab} \xi_{t|t}^{ab} \\ \xi_{t+1|t}^\delta &= \mathbf{P}_\delta \xi_{t|t}^\delta, \end{aligned}$$

where \mathbf{P}_i represents the corresponding transition probability matrix for $i = a, b, ab$, or δ . Following 5-6, the updated state probabilities are again combined to build the joint state probabilities.

These two steps are subsequently repeated for $t = 1, \dots, T$.

5.3 Parameter Estimation

Section 5.1 highlighted central aspects of the flexible regime-switching model. Three features are thereby of particular relevance for parameter estimation: the factorization of correlation matrices (composite correlation matrices), the clustering of under-

lying assets, and the definition of appropriate starting and boundary values. These features are subsequently analyzed in greater detail.

5.3.1 Factorization of Correlation Matrices

The previous section presented an intuitive approach to reusing certain correlation parameters in different states of the flexible model. In this model, the two states $s_t^* = 1, K^{nc}$ match those of common regime-switching models, where all assets reside in the same regime (see Table 5.3).¹⁴ However, the flexible model can also follow intermediate states $s_t^* = 2, \dots, K^{nc}-1$, where individual asset clusters reside in different regimes. Clearly, the correlations within an asset cluster resemble those in $s_t^* = 1$ or $s_t^* = K^{nc}$. Only the correlations across groups contain new information. Step 1 in Figure 5.1 illustrates this situation for a two-regime ($K = 2$) two-cluster ($nc = 2$) model.

Below, this thesis develops an algorithm that incorporates this reuse of correlation parameters for all intermediate states $s_t^* = 2, \dots, K^{nc}-1$. In general, the combination of these correlation submatrices is intuitive. However, the simple exchange of certain submatrices – as shown in step 1 – does not guarantee PSD correlation matrices. Consequently, the algorithm has to rely on the previously presented Cholesky decomposition.

In Chapter 2, this decomposition was simple. For each regime, the model defined independent Cholesky factors L , which were then reassembled within the MLE procedure. However, this simplicity does not apply to the intermediate states of the flexible model. The Cholesky factors in these states are subject to an additional constraint. The reassembled correlation matrix $R_{s_t^*} = L_{s_t^*} L_{s_t^*}'$ in $s_t^* = 2, \dots, K^{nc}-1$ must contain the same block-diagonal correlations as in $s_t^* = 1$ or $s_t^* = K^{nc}$ (see step 1 in Figure 5.1). As shown in Section 5.1.3, the conditionality of the correlations in $R_{s_t^*}$ is clearly observable (either conditioned on the regime of one cluster: block-

¹⁴Note that the state space effectively follows $2K^{nc}$ states due to the synchronization factor δ_t , which defines an additional dimension (see also Table 5.3). However, this additional dimension does not influence the number of conditional pdfs, but solely their weights. Consequently, the moment and correlation parameters for states 1 to K^{nc} depict the same information as those for states $K^{nc} + 1$ to $2K^{nc}$.

diagonal; or on the regimes of both clusters: off-block-diagonal). However, this conditionality is not as easily observable for the parameters in $L_{s_t^*}$. Consequently, for $s_t^* = 2, \dots, K^{nc}-1$ the parameters in $L_{s_t^*}$ must be recursively adjusted to guarantee that the block-diagonal correlations resemble those in $s_t^* = 1$ or $s_t^* = K^{nc}$.

Recursive hereby refers to the initial specification of the correlation matrices in $s_t^* = 1, K^{nc}$ and to the subsequent adjustment of the Cholesky factors $L_{s_t^*}$ in $s_t^* = 2, \dots, K^{nc}-1$. Steps 2 and 3 in Figure 5.1 depict this recursive procedure (for two clusters):

In step 2, the correlation matrices of the two polar states $s_t^* = 1, 4$ are Cholesky decomposed. In contrast, the Cholesky factors of intermediate states $s_t^* = 2, 3$ extend the parameter set only by their off-block-diagonal elements $\check{M}_{s_t^*}: \theta = \{\dots, L_{s_t^*=1}, L_{s_t^*=K^{nc}}, \check{M}_{s_t^*=2}, \dots, \check{M}_{s_t^*=K^{nc}-1}, \dots\}$.¹⁵ The matrices in 5-12 exemplify the elements of this block structure for a two-cluster model

$$R_{s_t^*} = \begin{bmatrix} A_{s_t^*} & M'_{s_t^*} \\ M_{s_t^*} & B_{s_t^*} \end{bmatrix} \quad L_{s_t^*} = \begin{bmatrix} \check{A}_{s_t^*} & 0 \\ \check{M}_{s_t^*} & \check{B}_{s_t^*} \end{bmatrix}, \quad (5-12)$$

where the elements $\{A, B, M\}$ and $\{\check{A}, \check{B}, \check{M}\}$ define the cross-cluster parameters for the correlation matrix and for the Cholesky decomposed matrix, respectively. The parameter set θ is subsequently fed into the MLE optimization routine. The algorithm in Figure 5.2 (*part I*) illustrates the above procedure in technical terms.

Step 3 in Figure 5.1 describes the process within the optimization routine. In a first step, the correlation matrices for $s_t^* = 1, 4$ are reassembled. The Cholesky factors $L_{s_t^*}$ for $s_t^* = 2, 3$ are then modified by the factorization algorithm. This algorithm adjusts the elements in $L_{s_t^*}$, which match the block-diagonal clusters (indicated in gray in Figure 5.1). In contrast, the off-block-diagonal input parameters $\check{M}_{s_t^*}$ enter into the matrix $L_{s_t^*}$ unadjusted.

¹⁵To initialize the MLE optimization, $\check{M}_{s_t^*}$ for $s_t^* = 2, \dots, K^{nc} - 1$ is, for example, defined by \check{M}_1 or $\check{M}_{K^{nc}}$. Alternatively, the parameters in $\check{M}_{s_t^*}$ for $s_t^* = 2, \dots, K^{nc} - 1$ could also be set to zero. The solution applied in this thesis will be described in Section 5.3.3.

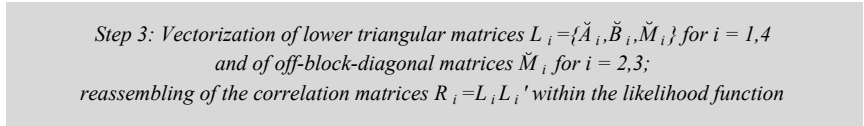
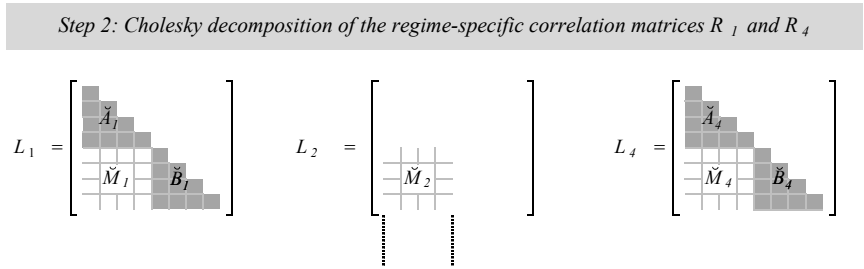
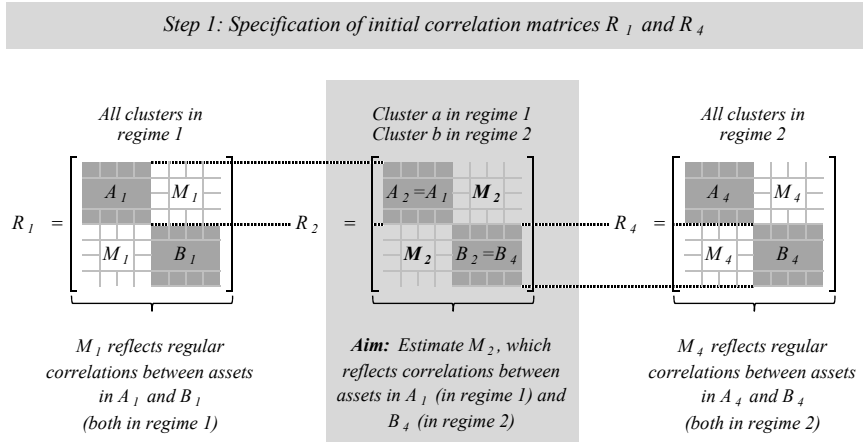
Again, the factorization algorithm in Figure 5.2 (*part 2*) describes this procedure in technical terms. It is conducted for $s_t^* = 2, \dots, K^{nc} - 1$ and follows the order of the Cholesky-Banachiewicz algorithm. The procedure therefore starts from the upper-left corner and proceeds row-by-row.

The basic Cholesky algorithm applies to diagonal elements in $L_{s_t^*}$ ($i = j$; see also formula 2-18). In addition, the test for $L_{s_t^*;jj} \notin \mathbb{R}^+$ after each full iteration of i ensures that these diagonal elements in $L_{s_t^*}$ are non-negative and therefore that the solution is unique.

In case element (i, j) in $R_{s_t^*}$ belongs to one of the block-diagonal clusters $(R_{s_t^*;ij} \in \zeta_{s_t^*}$, where $\zeta_{s_t^*} = \{A_{s_t^*}^a, B_{s_t^*}^b\}$), the corresponding element $L_{s_t^*;ij}$ is adjusted (see second line in the algorithm; *part 2*). Hereby, $\zeta_{s_t^*} = \{A_{s_t^*}^a, B_{s_t^*}^b\}$ defines the set of within-cluster correlations. It is defined by the prevailing intermediate state $s_t^* = 2$ or $s_t^* = 3$. Consequently, cluster a and cluster b do not reside in the same regime, and their corresponding submatrices A and B depict correlations from different regimes. For example, if $s_t^* = 2$, then Table 5.3 implies that clusters a and b are in regimes 1 and 2, respectively. As a result, $\zeta_{s_t^*=2}$ would be defined by $\{A_{s_t^*=1}^a, B_{s_t^*=2}^b\}$. In case (i, j) in $R_{s_t^*}$ represents an off-block-diagonal element $(R_{s_t^*;ij} \notin \zeta_{s_t^*})$, the corresponding element in $\check{M}_{s_t^*}$ enters into $L_{s_t^*}$.

Finally, the resulting Cholesky factors $L_{s_t^*}$ for $s_t^* = 2, \dots, K^{nc} - 1$ are reassembled to build the correlation matrices.¹⁶ The above-mentioned test for $L_{s_t^*;jj} \notin \mathbb{R}^+$ guarantees that these correlation matrices are PSD.

¹⁶An alternative solution would be to define the Cholesky factors $L_{s_t^*}$ for $s_t^* = 1, K^{nc}$, to re-assemble them inside the optimization routine, and to assign the corresponding correlation terms to $s_t^* = 2, \dots, K^{nc} - 1$. The off-diagonal blocks of correlations would then present part of the parameter set. This solution would guarantee PSD results for $s_t^* = 1, K^{nc}$, but not for the intermediate states $s_t^* = 2, \dots, K^{nc} - 1$.



The block-diagonal elements of L_2 and L_3 (states where the clusters reside in different regimes) are specified so that they resemble the block-diagonal correlation values of R_1 and R_4 . Only the values of clusters \check{M}_2 and \check{M}_3 are estimated in the optimization routine (for L_2 and L_3).

Figure 5.1: Illustration of the factorization algorithm to calculate the composite correlation matrices for the flexible regime-switching models.

Part 1: Specification of initial parameter set (outside optimization routine):

1. Cholesky decomposition of $R_{s_t^*}$: $L_{s_t^*} = \text{chol}(R_{s_t^*})$ for $s_t^* = 1, K^2$
2. Specify naïve input for the off-block-diagonal elements $\check{M}_{s_t^*}$ for $s_t^* \neq 1, K^2$, where

$$L_{s_t^*} = \begin{bmatrix} \check{A}_{s_t^*} & 0 \\ \check{M}_{s_t^*} & \check{B}_{s_t^*} \end{bmatrix},$$

and where $\check{A}_{s_t^*} \in \mathbb{R}^{m \times m}$, $\check{B}_{s_t^*} \in \mathbb{R}^{n \times n}$, $\check{M}_{s_t^*} \in \mathbb{R}^{n \times m}$

3. Stack parameters in

$$\theta = \left\{ \dots, \text{vech}(L_1), \text{vech}(L_{K^2}), \text{vec}(\check{M}_2), \dots, \text{vec}(\check{M}_{K^2-1}), \dots \right\},$$

where $\text{vech}(\cdot)$ is a half-vectorization and where $\text{vec}(\cdot)$ defines a vectorization.

Figure 5.2: Factorization algorithm to define composite correlation matrices.

Part 2: Factorization algorithm (within optimization routine):

1. Reassemble $L_{s_t^*}$ and reverse the Cholesky factorization:

$$R_{s_t^*} = L_{s_t^*}' L_{s_t^*}' \quad \text{for } s_t^* = 1, K^{nc}$$

2. Build the lower triangular matrices $L_{s_t^*}$ for $s_t^* = 2, \dots, K^{nc}-1$ according to the following algorithm:

for $i = 1$ to N **do**

for $j = 1$ to N **do**

$$L_{s_t^*,ij} = \begin{cases} \sqrt{1 - \sum_{m=1}^{i-1} L_{s_t^*,im}^2}, & \text{if } i = j \\ \frac{1}{L_{s_t^*,jj}} \left(R_{s_t^*,ij} - \sum_{m=1}^{j-1} L_{s_t^*,im} L_{s_t^*,jm} \right), & \text{if } i > j \cap R_{s_t^*,ij} \in \zeta_{s_t^*} \\ \check{M}_{s_t^*,i-m,j}, & \text{if } i > j \cap R_{s_t^*,ij} \notin \zeta_{s_t^*} \\ 0, & \text{otherwise,} \end{cases}$$

where $\zeta_{s_t^*} = \{A_{s_t^*}^a, B_{s_t^*}^b\}$,

and where m is the size of cluster a .

end for

Stop: if $L_{s_t^*,jj} \notin \mathbb{R}^+$

end for

Figure 5.2 (cont.): Factorization algorithm to define composite correlation matrices.

5.3.2 Clustering of Assets

The developed factorization algorithm presents an efficient way of reusing correlation parameters in different states. However, the specification of this algorithm strongly depends on the structure of the underlying clusters. An appropriate clustering technique therefore presents an integral part of the multivariate flexible regime-switching model.

This thesis applies two different data-clustering techniques: k -means clustering and hierarchical clustering. For a detailed description of k -means and hierarchical clustering, one is referred to Everitt, Landau, Leese and Stahl (2011) and to Ritter (2014). The current section focuses on the adaptation of these techniques to regime processes.

k -means clustering aims to partition the underlying univariate regime processes of the N assets, $S = \{S^1, \dots, S^N\}$, into K mutually exclusive clusters so that the within-cluster sum of distance is minimized

$$\min_{\{m_k\}_{k=1}^K} \sum_{k=1}^K \sum_{S \in C_k} d(S, m_k),$$

where m_k is the centroid of cluster C_k , K is the number of clusters, and d computes the distance between regime processes and the corresponding centroid m_k (Wu, 2012; Aggarwal & Reddy, 2014). This centroid is typically defined by the mean of the within-cluster observations μ_k .¹⁷ Moreover, the distance function d is most widely represented by the squared Euclidean distance $\|S - m\|^2$ (Wu, 2012; Aggarwal & Reddy, 2014).

Note, that a regime process moves between zero and unity. However, its mean does not reveal any information about the behavior of the regime-switching process over time. For example, the mean of a regime process with the balanced occurrence of bull and bear states is close to 0.5. This number gives no further information about the underlying regime dynamics and is therefore uninformative.¹⁸ The same

¹⁷ k -means specifically refers to the K different means $\{\mu_k\}_{k=1}^K$.

¹⁸The only exception is a process that resides mostly in one regime.

applies to the mean of a cluster $m_k = \mu_k$. For this reason, regime cycles need to be standardized relative to a reference cycle. This standardization determines the degree of deviation between two processes and returns appropriate input data for k -means clustering.

The univariate regime cycle of the Datastream World Index builds the reference cycle in this thesis. To standardize individual cycles, their absolute distance to the world regime cycle is calculated

$$d(S_t^a - S_t^w) = |S_t^a - S_t^w|, \quad \forall t$$

where S_t^a and S_t^w describe particular observations of the regime cycles of asset a and of the world index.

Clearly, this approach is biased by the assumption that the world index resembles the true regime process. A further problem arises when the world regime probability is close to 0.5. In this case, assets in bull states (S_t^i close to unity) and in bear states (S_t^i close to zero) depict the same distance to the world index even though the cycles are totally different.

Agglomerative hierarchical clustering represents a valuable alternative. This technique initially treats each univariate regime process $S = \{S^1, \dots, S^N\}$ as an individual cluster. The most similar pair of clusters is then merged. The algorithm continues with this new set and merges clusters until the intended number of clusters is reached (Aggarwal & Reddy, 2014; Ritter, 2014). Whereas k -means relies on actual observations, hierarchical clustering bases on dissimilarity measures (distance metrics). These measures adopt the previous standardization of regime cycles and therefore make this step redundant.

Hierarchical clustering follows a two-step approach. In a first step, it calculates the direct distance between two individual regime processes. This distance is determined for each combination of regime cycles and is typically defined by Euclidean distance. In a second step, the linkage criteria compares distance information of step one to determine the proximity of clusters. This thesis uses single linkage selection, where

distance is measured as the shortest distance among the observations of the two clusters

$$\min_{S^a \in C_i, S^b \in C_j} d(S^a, S^b),$$

where d describes the distance measure defined in step one (Aggarwal & Reddy, 2014).

Hierarchical clustering clearly circumvents the specification of a reference cycle. Nevertheless, the structure of the underlying univariate regime cycles still represents a weakness. The univariate regime-switching model ignores any dependence structure across individual assets. According to Camacho and Perez-Quiros (2006), perfect synchronization therefore tends to be underestimated. This shortcoming is of particular relevance for the clustering of cycles, as perfectly synchronized cycles are harder to detect. Nevertheless, the clustering results still present a valuable source of information to segregate the sample into individual groups.

Clustering based on the full sample history poses a further problem. These clusters rely on information that was not available at the time of their formation. Theoretically, the clusters should be specified at time t , with respect to univariate smoothed regime probabilities until that date: $\{\xi_{i|t}\}_{i=0}^t$. Based on these clusters, the flexible model's regime probabilities would then be estimated until a later date T : $\{\xi_{i|i}\}_{i=t}^T$. The analysis of subperiods, however, has shown that clusters remain relatively stable over time. Therefore, the overlapping of cluster formation and regime calculation periods does not significantly distort the structure of the clusters.

5.3.3 Use of Prior Information

This thesis estimates the underlying regime-switching models using maximum likelihood estimation (MLE). In general, this estimation technique presents an efficient way of maximizing a likelihood function. However, the mixture distribution of a regime-switching model may show many local maxima. As a result, the optimization of the corresponding likelihood surface might stop at a local instead of the global

maximum. Nevertheless, this local optimum might still present a valuable solution if it resides closely to the global maximum.

To ensure this closeness, appropriate starting and boundary values for the underlying parameters need to be defined. The specification of these values requires reliable prior information, which will be defined below.

The underlying assumption of regime-independent means partially mitigates the previously mentioned problem. Due to non-switching means, the density function becomes unimodal and the optimization of the likelihood function always stops at the global maximum. Despite this simplification, starting and boundary parameters, if set appropriately, might still support the estimation procedure for three reasons: First, given appropriate starting values, the likelihood function shows a faster convergence rate. Second, some parameters still require boundary conditions (for example, transition probabilities, standard deviations, or correlations). Finally, without already approaching the global solution, the search method might be unable to optimize likelihood functions with large parameter sets.

Optimal Starting Values

Defining appropriate starting values for the flexible model seems quite complex at first. However, its structure presents a particular advantage. If the underlying assets reside in a regime of perfect synchronization ($s_t^\delta = 2$), the sample mimics the properties of the multivariate benchmark model. In contrast, if the underlying assets reside in a regime of full independence ($s_t^\delta = 1$), the sample can be divided into individual clusters. In turn, each of these clusters once again exhibits the properties of the benchmark model.¹⁹

The resulting parameter estimates of the benchmark models therefore provide good start-up parameters for the flexible setup. This approach is very supportive, given the fact that standard benchmark models are simpler to estimate than the flexible regime-switching models. Clearly, the DGP setup needs to be equal in the flexible and

¹⁹Depending on the size of the cluster, this is either a univariate or a multivariate benchmark regime-switching model.

the benchmark model. This restriction also applies to the definition of the regime-dependent variables. Moreover, the estimation in the benchmark models should rely on the same observation period as in the flexible model.²⁰

Given these assumptions, benchmark models provide good start-up values for most parameters of the flexible setup. In the case of non-synchronized states ($s_t^\delta = 1$), for example, benchmark models return starting values for the sample moments and for the transition probabilities of each cluster.

In the case of perfectly synchronized states ($s_t^\delta = 2$), all assets reside in the same regime. Consequently, the multivariate benchmark model provides starting values for the joint transition probability matrix \mathbf{P}_{ab} or \mathbf{P}_{abc} as well as for the correlation matrices. The latter are additionally Cholesky-decomposed to receive the final input factors $L_{s_t^*}$.

The above specification provides a very intuitive approach to defining the starting values of the flexible model. However, this setup leaves unanswered how the benchmark models initially set their starting values. Hereby, a different approach applies. Schwendener (2010) remarked that the starting values should already possess the observed stylized facts of financial markets. This specification increases the chances that the estimated likelihood will end up close to the global maximum on the likelihood surface. In this respect, standard deviation well describes the properties of the different regimes (at least for two regimes; see Schwendener, 2010; Spremann, 2007; Spremann & Gantenbein, 2007). Consequently, the sample observations are separated into two pools, based on the standard deviation of a reference cycle. For this purpose, the Datastream World Index serves as a reference cycle. It separates the two pseudo-regimes based on the median standard deviation of its returns (24-month rolling window). Observations at time t that show a lower standard deviation than the median belong to regime 1 ($\sigma_{med}^{world} > \sigma_t^{world}$), and vice versa. Moment and correlation parameters are then estimated for the observations of the corresponding regime. These parameters serve as starting values for the benchmark model. Similarly, starting values for the transition probabilities can be estimated from the relative frequency of changes between ($\sigma_{med}^{world} > \sigma_t^{world}$) and ($\sigma_{med}^{world} \leq \sigma_t^{world}$).

²⁰This approach does not bias the results of the flexible model. The starting values depend only on information that was available at the time of the flexible model's specification (no posterior information).

Thus far, two parameters of the flexible model have not been analyzed. These include the transition probabilities of the synchronization factor s_t^δ and the correlation matrices of intermediate states. The former are initially set at 0.95 for p_{11}^δ and p_{22}^δ . Analysis of the dynamic synchronization measures in Chapter 4.4.3 revealed similar numbers.

The correlation matrices in intermediate states follow the same reference-cycle-approach as presented above. However, the current setup is slightly different, as the sample is governed by more than one regime process. Consequently, standard deviation is not defined globally by the Datastream World Index but by selected assets in the corresponding clusters. For example, the joint event $(\sigma_{med}^a > \sigma_t^a) \cap (\sigma_{med}^b \leq \sigma_t^b)$ defines the observations in the intermediate state $s_t^* = 2$, where cluster a is in regime 1 and cluster b is in regime 2. Each cluster further needs to define a particular asset as its volatility indicator. The correlations can then be calculated for the observations in the corresponding regime pool. Finally, Cholesky decomposition is applied and the off-block-diagonal elements in the lower triangular matrix are used as starting values $\check{M}_{s_t^*}$.

Optimal Parameter Bounds

Appropriate parameter bounds are defined much easier than the previously defined start-up values. In fact, the latter even support the definition of appropriate parameter bounds. This definition applies to both the flexible and the benchmark models.

Boundaries for means can be set very unrestrictedly. Basically, they can range from $-\infty$ to $+\infty$. However, to improve the convergence of the search algorithm, boundaries can be set more restrictively. Given well-defined starting values, means are limited by lower and upper bounds of -2.0% and +2.0% around the starting value. These bounds result in an annualized return bandwidth of 48.0%. In case of regime-dependent means, the corresponding boundaries are set relative to the mean of the respective regime.

For standard deviations, boundaries need to be set more restrictively. However, as illustrated in connection with covariance factorization (see Chapter 2), standard deviations are only weakly bounded. To avoid singularities in the likelihood function,

standard deviations only need a fixed lower limit. This is set at 1.0% annualized volatility. To improve likelihood convergence, the upper limit was set at 60.0% annualized volatility. However, neither limit has ever been reached.

For the correlation parameters and the Cholesky factors, the limits are easily set at -1 and $+1$ and need no further specification.

Finally, the specification of the transition probability matrices is already restricted by the properties of probabilities. Consequently, the probability values need to lie between zero and unity. However, given the implied assumptions about Markov chains, these bounds need to be additionally restricted (see Section 2.2.2). The probabilities of residing in the respective regime (p_{11} and p_{22}) are limited by lower and upper bounds of 0.75 and 0.99, respectively. The lower limit guarantees stable regime dynamics, whereas the upper limit assures ergodicity and irreducibility. Finally, the lower bounds for the transition probabilities of the flexible model (p_{11} and p_{22}) were additionally restricted by $\max \{0.75, (1 - 0.2) \cdot \text{starting value}\}$. For the upper bounds, similar limits were set by $\min \{0.99, (1 + 0.2) \cdot \text{starting value}\}$.

5.4 Summary

This chapter has introduced a multivariate regime-switching model, which allows individual groups of assets to reside in different regimes. Previous research mainly analyzed this model structure in a bivariate context. Particular reasons for this limitation are the increasing model complexity and the structure of the covariance matrix. This study presents a robust solution to these problems. Assets with similar regime processes can be clustered to reduce the number of underlying Markov chains. Further, different clustering techniques have been presented to group these cycles. Moreover, the use of composite correlation matrices reduces the increase in the number of parameters to estimate. The corresponding factorization algorithm is a particular contribution of this thesis and allows the model to incorporate more flexible correlation structures.

Chapter 6

Empirical Analysis

The empirical analysis of flexible regime-switching focuses on two particular model features: first, the synchronization dynamics between asset clusters; and second, the specification of the composite correlation matrix and its impact on model likelihood. These two aspects represent major contributions of the present work to research in this field and will receive special attention throughout the current chapter.

The following section analyzes these features by comparing different model setups. Section 6.2 then presents detailed results for the flexible model with two and three clusters. These models are applied to the core and the extended sample, which will highlight the importance of appropriate sample clustering.¹ Moreover, it will intuitively demonstrate the shortcomings of standard multivariate regime-switching models. Section 6.3 evaluates the capacity of the flexible model to capture the true degree of regime cycle synchronization. This analysis is conducted by means of a kernel density estimation. Section 6.4 then compares the forecasting power of the different models. Finally, the main findings of this chapter are summarized in Section 6.5.

6.1 Model Comparison

Chapter 4 clearly illustrated the synchronization dynamics between the regime processes of individual assets. Results revealed the necessity of considering time-varying

¹The underlying dataset is the same as in Chapter 4.

synchronization in multivariate regime-switching models. This claim, in turn, led to a more complex model setup (see Chapter 5). The flexible model allows for a dynamic switching behavior, but requires additional parameters. To assess this trade-off between model accuracy and complexity, the flexible model is evaluated in two ways.

First, the DGP specification of the flexible model is tested for possible restrictions of the composite correlation matrix. In theory, this correlation matrix and the corresponding factorization algorithm seem very efficient. However, it needs to be evaluated whether the additional flexibility contributes significantly to the model fit. Due to the nested structure of the DGP, likelihood ratio tests apply to this comparison.

Second, the flexible model is compared to alternative regime-switching and non-switching models. This comparison determines the value of synchronization dynamics in the flexible setup. The different models are, however, non-nested and can therefore not be compared directly. Information criteria thus help to evaluate the relative quality of these non-nested models through penalized likelihood functions.

Table 6.1 illustrates the first part of the model comparison. It depicts the log-likelihood values and the size of the parameter set for different models. Panel A presents the results for the core sample and panel B for the extended sample.

The columns in Table 6.1 describe different model specifications. The full and basic specifications match those in Chapter 4. Again, full specification implies no switching restrictions and resembles the DGP in formulas 2-16 and 5-1. In contrast, the basic specification assumes regime-independent means and fixed initial regime probabilities (steady-state probabilities). Further, columns four and five depict modified versions of the basic specification. Column four restricts the correlation matrix to be regime-independent. The specification in column five, in contrast, assumes that the correlation matrix is regime-dependent, but relies solely on the state of the US cluster.² Similar to Cakmakli et al. (2011), the regime of a specific underlying asset defines the state of the entire correlation matrix. This restriction helps to evaluate the information content of the proposed composite correlation matrix. The restriction applies only to the flexible model, as all other regime-switching models already

²For ease of exposition, this asset group is called the US cluster, although the US might only be one of many equally important assets in that cluster.

imply that all assets reside in the same regime. Columns four and five depict the corresponding LR-test statistics next to the likelihood values (indicated by stars). The basic model specification thereby defines the unrestricted model for the LR-test. Statistically significant differences in model fit at the 1%, 5%, and 10% level are indicated by ***, **, and *, respectively.

The rows in Table 6.1 describe five different models, which belong to three categories: First, the linear model (LIN) presents a simple non-switching one-regime model. Its corresponding parameter set consists solely of moment parameters and of a single correlation matrix.

Second, the two multivariate models (MV_2 and MV_4) belong to the category of standard regime-switching and follow the baseline specification in Chapter 2. In line with Guidolin and Timmermann (2007), one of these models follows a four-regime process (MV_4). Due to its increased state space, this model should partially resemble the properties of flexible regime-switching.

Finally, the last two models implement the structure from Chapter 5. These models belong to the category of flexible regime-switching. The two specifications differ with regard to the number of underlying clusters (two clusters: FLEX_2C; three clusters: FLEX_3C).³

Table 6.1 provides five main insights: First, the likelihood values of the full specification and the basic specification differ only marginally. Undisclosed LR-tests have shown no significant differences in model fit between the two specifications. This observation holds true both across models and across sample sizes. These results further support the flexible model's DGP setup, where means are defined as non-switching (corresponds to the basic specification).

Second, non-switching correlations significantly reduce the likelihood value of all models (see column 4). These results confirm previous observations that switching correlations significantly contribute to the model fit. Moreover, they support Pelletier's observation (2006) that standard deviations as well as correlations need to be regime-dependent.⁴

³The next section defines the corresponding clusters of the sample.

⁴LR-test results for the variances are omitted. However, these are at a level of significance similar to that of the correlation results.

Table 6.1: Comparison of different models and specifications in terms of marginal log-likelihood. The period of analysis is from December 1975 to February 2014.

	Full specification	Basic specification	Restriction of correlation matrix	
			Non-switching	Driven by the US
Panel A: Core sample				
Linear model	2360.0 9			
Multivariate model (2 regimes)	2449.7 21	2449.0 17	2360.0*** 14	
Multivariate model (4 regimes)	2508.2 51	2484.8 39	2360.0*** 30	
Flexible model 2 clusters	2482.5 34	2478.9 27	2456.2*** 20	2460.6*** 23
Flexible model 3 clusters	2503.9 39	2500.3 31	2453.4*** 22	2480.8*** 25
Panel B: Extended sample				
Linear model	4770.1 27			
Multivariate model (2 regimes)	4933.0 57	4931.4 50	4770.1*** 35	
Multivariate model (4 regimes)	5024.7 123	5004.6 102	4770.1*** 57	
Flexible model 2 clusters	4963.6 76	4956.5 66	4925.8*** 41	4951.6 56
Flexible model 3 clusters	5000.9 87	4991.9 76	4938.0*** 43	4973.3*** 58

The table shows the likelihood values for the core and the extended sample. Italic values indicate the corresponding number of estimated parameters. Likelihood values are depicted for the fully specified benchmark regime-switching model and for the basic specification with restricted means and initially set regime probabilities (0.5). The last two columns indicate restrictions of the correlation matrix. Column five assumes the cluster in which the US resides to define the overall state of the correlation matrix. The last two specifications have been tested for differences in model fit compared to the basic specification.

*** indicate statistically significant differences in model fit at the 1% level.

The likelihood ratio test is defined by $2 [\log L(\hat{\theta}) - \log L(\bar{\theta})] \approx \chi_m^2$.

Consequently, non-switching correlations present no alternative to the composite correlation matrix in the flexible model. However, the rejection of non-switching correlations does not yet support the use of composite correlation matrices. In line with Cakmakli et al. (2011), the correlation matrix can alternatively be conditioned on the regime process of a particular asset. Even large state spaces would subse-

quently be governed by only K different correlation matrices. This restriction would result in a smaller parameter set and would simplify the flexible model.

Column five presents the results of this restriction. Thereby, the cluster with the US index conditions the entire correlation matrix. The LR-test statistics indicate significant differences in model fit compared to the basic specification.⁵ The composite correlation structure thus adds statistically significant value to the model likelihood. Third, insights one and two apply to both sample sizes. Despite its larger parameter set, the flexible model is still competitive in the extended sample. FLEX_2C represents the only exception. Its parameter set would justify a restriction of the correlation matrix to depend solely on the US regime (column 5). Section 6.2.3 will show that this is due to the model's low number of clusters in relation to the size of its sample. Despite this exception, the likelihood values of both flexible models still compete with alternative setups.

Fourth, flexible regime-switching models more closely resemble the likelihood value of the multivariate benchmark model with four instead of two regimes (MV_4). At the same time, however, the pace of parameter extension in the flexible models more closely resembles MV_2.

Finally, the likelihood values of the flexible models are very competitive considering the size of their parameter sets. Unfortunately, as these models are non-nested, their likelihood values cannot be compared directly. Due to the presence of nuisance parameters, the LR-test statistic would no longer follow a Chi-squared distribution. However, information criteria present an alternative way of comparing such non-nested models.

Table 6.2 proposes three information criteria for comparing the quality of different models: the Akaike information criterion (AIC), the Bayes-Schwarz information criterion (BIC), and the Hannan-Quinn information criterion (HQ). These measures of fit are very intuitive, as they simply penalize a model's likelihood value according to the dimension of its parameter set. This trade-off between in-sample fit and model parsimony supports the comparison of different models (Guidolin, 2013). However, information criteria give no indication as to the absolute model quality, but solely

⁵Note that the latter relies on composite correlation matrices (as introduced in Chapter 5).

as to its relative strength compared to other models. Further, the structure of the penalizing function depends on the chosen information criterion. For example, the Akaike information criterion is specified as

$$AIC = 2\dim(\theta) - 2\log(L(\theta)),$$

where $\dim(\theta)$ measures the size of the parameter set and where $\log(L(\theta))$ depicts the log-likelihood value. Guidolin (2013) demonstrated that this criterion has optimal asymptotic properties, but tends to select too-large non-linear models in small samples.

Similarly, the Bayes-Schwarz information criterion can be specified as

$$BIC = \dim(\theta) \log(T) - 2\log(L(\theta)),$$

where T measures the length of the observation period. This specification reveals that BIC penalizes the parameter size more strongly than AIC. Consequently, BIC tends to select simpler model specifications. Finally, the Hannan-Quinn information criterion is defined by

$$HQ = 2\dim(\theta) \log(\log(T)) - 2\log(L(\theta)).$$

According to Guidolin (2013), this criterion performs strongly in both small samples and non-linear models. Further, the HQ criterion tends to return values between those of AIC and BIC.

Table 6.2 presents the sample results for these measures. The model with the lowest information criterion thereby presents the best relative quality. For this purpose, Table 6.2 illustrates the two lowest values for each criterion in bold numbers. In addition, it also depicts the saturation ratio, measured as total observations (NT) divided by the number of model parameters $\dim(\theta)$: $NT/\dim(\theta)$, where N is the number of underlying assets. Guidolin (2013) showed that values around 35 and above present

a good trade-off between sample size and parameter number. However, the absolute size of the parameter set might still be burdensome for model estimation. Table 6.2 reveals that all models show comparably high saturation values, with MV_4 being the only exception. However, the extended sample limits the validity of this statement, as some models already exhibit high absolute parameter numbers. For example, FLEX_3C and MV_4 require 76 and 102 parameters.

Table 6.2: Comparison of different models with respect to their information criteria. The period of analysis is from December 1975 to February 2014.

	dim(θ)	Sat. ratio	Lik	AIC	BIC	HQ
Panel A: Core sample						
Linear model	9	164.7	2360.0	-4702.1	-4664.9	-4687.4
Multivariate model (2 regimes)	17	87.2	2449.0	-4863.9	-4793.8	-4836.3
Multivariate model (4 regimes)	39	38.0	2484.8	-4891.5	-4730.6	-4828.1
Flexible model 2 clusters	27	54.9	2478.9	-4903.9	-4792.4	-4860.0
Flexible model 3 clusters	31	47.8	2500.3	-4938.5	-4810.6	-4888.1
Panel B: Extended sample						
Linear model	27	109.8	4770.1	-9486.3	-9374.8	-9442.4
Multivariate model (2 regimes)	50	59.3	4931.4	-9762.9	-9556.5	-9681.6
Multivariate model (4 regimes)	102	29.1	5004.6	-9805.1	-9384.2	-9639.4
Flexible model 2 clusters	66	44.9	4956.5	-9780.9	-9508.5	-9673.6
Flexible model 3 clusters	76	39.0	4991.9	-9831.9	-9518.2	-9708.3

Saturation ratio is defined as $NT/\text{dim}(\theta)$, where N is the number of parameters, T the time series length, and $\text{dim}(\theta)$ is the size of the parameter set. Lik is the log-likelihood. AIC depicts the Akaike information criterion, BIC the Bayes-Schwarz information criterion, and HQ the Hannan-Quinn information criterion.

Despite these objections, flexible models generally show the lowest information criteria. For the core sample, these results are very distinct. Except for the BIC value of FLEX_2C, the flexible models return the lowest values in the sample. FLEX_3C even performs slightly better than FLEX_2C. This behavior is due to the small difference in parameter size in relation to the additional model flexibility. For the extended sample, FLEX_3C still presents the best overall results. However, these results are no longer as distinct. The convergence of information criteria across

models might be due to the overly increased parameter set of the two flexible models. Further, FLEX_2C performs poorly due to its low number of clusters. Section 6.2.3 will show that two clusters underrepresent the extended sample. Consequently, the model fit of FLEX_2C is expected to be comparably weak for the extended sample. Due to the BIC's preference of simpler models, MV_2 shows the lowest BIC value again. Overall, however, the results still favor flexible regime-switching (FLEX_3C).

6.2 Empirical Results of the Flexible Model

The previous observations have demonstrated the quality of flexible regime-switching models. The following subsections analyze these models in greater detail, focusing on their parameter sets and switching dynamics.

To further emphasize the power of the flexible setup, its results are again compared to benchmark models by considering the univariate and the multivariate model with two regimes (UV_2 and MV_2). The choice of these models is motivated by their widespread use. Moreover, they enable direct comparison with the flexible model, as the latter relies only on two basic regimes (despite its larger state space).

Below, the parameters of the flexible models will be analyzed for the core and the extended sample. The insights of this analysis prove to be of particular value in at least three ways: first, they highlight the importance of appropriately specified clusters; second, they show the influence of sample size on the cluster number; and third, they reveal the potential bias of standard multivariate regime-switching models. Analysis will show that even a single asset with diverging switching dynamics might bias the entire parameter set of the multivariate benchmark model. Before analyzing the flexible models, however, their clusters will be defined for the two samples.

6.2.1 Clustering of Data

The clustering of assets with similar switching dynamics forms a vital part of flexible regime-switching. Chapter 5 presented two different techniques for clustering

regime cycles: k -means and hierarchical clustering. Interestingly, the resulting clusters turned out similarly across these techniques. The following analysis therefore focuses on the results of k -means clustering.

Table 6.3: *Grouping of sample markets using k -means clustering. a) Core sample; b) extended sample. The period under study is from December 1976 to February 2014, based on univariate regime processes.*

	<i>Core sample</i>			<i>Extended sample</i>					
	US	UK	GE	US	UK	GE	JP	PA	SW
2 Clusters	<i>a</i>	<i>b</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>a</i>
3 Clusters	<i>a</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>c</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>a</i>

Legend: GE = Germany; JP = Japan; PA = Pacific ex Japan; SW = Switzerland.

Table 6.3 highlights the k -means results for two and three clusters. The comparison of the core and the extended sample indicates that the US and Germany belong to the same group, irrespective of the sample size. As already observed in Chapter 4, the UK shows more individual cycle dynamics.

Of course, the assumption of three clusters in the core sample results in a separate group for each country. This resembles the structure of a multivariate model with independent regime dynamics. In comparison, however, the flexible model still accounts for additional synchronization dynamics.

The extended sample further reveals the heterogeneous behavior of Japan. Its allocation to a separate group applies to the model with two and three clusters. To assess the stability of the groups, a 60-month rolling-window analysis for the period from January 1986 to February 2014 was performed. The resulting asset groups remained relatively constant, irrespective of the number of assets and clusters.

Furthermore, the sample was tested for the optimal number of clusters. Figure 6.1 shows the silhouette values for the k -means clustering approach of the extended sample. Very low or negative values in a particular group indicate a weak clustering pattern. However, the silhouette values for two and three clusters in Figure 6.1 are stable. An increase in the number of clusters causes a lot of groups that cover

only a single asset. Consequently, two and three clusters are an appealing choice for the current sample. More than three clusters would not further improve results. Hamilton and Owyang (2012) observed similar numbers for their multivariate sample of the 50 US states. Even for this large sample they found that two and three clusters return good results.

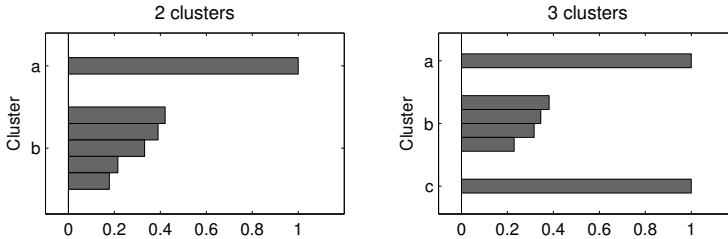


Figure 6.1: *Silhouette values of k -means clustering (x-axis) using different cluster numbers: a) Two clusters, b) three clusters.*

6.2.2 Analysis of the Core Sample

This section applies the two flexible models to the core sample. The resulting parameters and regime processes will help to evaluate the models in a relatively stable environment. In Section 6.2.3, these initial findings will be used to further analyze the results of the extended sample. Analysis will demonstrate the power of the flexible model and the importance of a sufficient amount of clusters. In fact, analysis will show that standard multivariate regime-switching models suffer from the same problem. Theoretically, their sample belongs to a single cluster and can therefore not appropriately cover the switching dynamics of individual assets.

Flexible Model with Two Clusters

Table 6.4 shows the parameter estimates for the core sample. For better visual inspection, underlying markets were sorted with respect to their clusters. The US and Germany form cluster *a* (white background) and the UK forms cluster *b* (gray background).

The moment parameters in Table 6.4 closely resemble those of UV_2 (see Table 4.4). This similarity is explained by the structure of the flexible model. Due to the multiple regime processes and well-chosen asset clusters, the flexible model resembles the dynamics of individual assets much closer than the multivariate model. Such assets are no longer forced to match a single overall regime cycle.

At the same time, the flexible model accounts not only for correlation across underlying assets but also for synchronization among individual regime processes. These features, in turn, do not apply to univariate models. Clearly, the flexible setup presents an optimal mixture of the features of both univariate and multivariate regime-switching models.

The previous statement is further confirmed as the moment parameters in Table 6.4 strongly differ from those in MV_2 (see Table 4.5).⁶ In the latter, means are comparably low and standard deviations for the two regimes lie close together. This indicates a potential bias in the switching dynamics of MV_2. As stated throughout this thesis, the aggregation of all information into a common regime process presents a particular drawback of standard multivariate regime-switching models. This aggregation is also responsible for the poorly diverging regime properties of MV_2.⁷ The flexible model, on the other hand, depicts higher means and shows statistically significant differences between the standard deviations of regime 1 and 2.

⁶Note that Table 4.5 presents only the results for the extended sample. However, as remarked in Chapter 4, the moment parameters of the core sample differ only insignificantly.

⁷The simplest way of inferring more individual dynamics in MV_2 would be to downsize the underlying sample. The sample would have to be reduced until it consists solely of homogeneously switching assets. This approach would put more emphasis on the switching dynamics of individual assets as these more closely resemble the joint regime process.

Table 6.4: Flexible regime-switching results for the core sample with two clusters. The period under study is from December 1975 to February 2014.

	μ	σ_1	σ_2		
US	0.13*** (0.022)	0.10*** (0.007)	0.21*** (0.001)	Log Lik	2478.9
				dim(θ)	27
Germany	0.14*** (0.029)	0.14*** (0.008)	0.28*** (0.001)	AIC	-4903.9
				BIC	-4792.4
UK	0.13*** (0.023)	0.11*** (0.007)	0.23*** (0.001)	HQ	-4860.0

White background: cluster a ; gray background: cluster b . Values are shown on an annualized basis; standard errors are depicted in parentheses.

Log Lik is the log-likelihood. $\text{dim}(\theta)$ is the size of the parameter set. AIC depicts the Akaike information criterion, BIC the Bayes-Schwarz information criterion, and HQ the Hannan-Quinn information criterion.

*** indicates the significance of the parameter estimates (different from zero) at the 99% level.

Formula 6-1 further highlights the transition probabilities of the flexible model. Cluster a 's probabilities to remain in regime 1 or 2 are 0.86 and 0.80. These are lower than in the univariate models of the US and Germany.⁸

$$\begin{aligned}
 \mathbf{P}_a &= \begin{bmatrix} 0.86 & 0.20 \\ 0.14 & 0.80 \end{bmatrix} & \mathbf{P}_b &= \begin{bmatrix} 0.80 & 0.02 \\ 0.20 & 0.98 \end{bmatrix} \\
 \mathbf{P}_{ab} &= \begin{bmatrix} 0.90 & 0.19 \\ 0.10 & 0.81 \end{bmatrix} & \mathbf{P}_\delta &= \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}
 \end{aligned} \tag{6-1}$$

Analysis has shown that the probabilities of staying in the prevailing regime (p_{11} and p_{22}) generally turn out lower in the flexible model than in the benchmark models. This effect is due to the synchronization dynamics of the flexible model. For example, \mathbf{P}_a describes the transition probabilities in case cluster a switches its regime independently ($\delta = 0$). In contrast, \mathbf{P}_{ab} depicts the transition probabilities if the regimes of both clusters are synchronized ($\delta = 1$). This separation enables each

⁸The results of a bivariate benchmark model for the US and Germany would be even more meaningful, as these would best mimic cluster a 's behavior. Undisclosed results confirmed this assumption.

asset to be driven by two individual regime processes, where each process covers particular subperiods of the sample.⁹ Thus, the model covers more dynamics of the individual assets, as the data is not aggregated in a single joint Markov chain. The resulting regime processes tend to be more reactive and therefore generally show lower probabilities of remaining in the preceding regime.

When UV_2 and MV_2 are applied to a selected subperiod, a similar effect can be observed. This analysis has been conducted for cluster *a*, cluster *b*, and the entire core sample. The subperiods were selected according to the start and end points of $\Pr(s_t^\delta = 1) > 0.5$ in the flexible model. Transition probabilities in MV_2 (cluster *a* or *b*) have turned out similar to those in \mathbf{P}_a and \mathbf{P}_b .¹⁰ This analysis was also conducted for the opposite case of full synchronization. Likewise, the transition probabilities of MV_2 (core sample) have resembled the data in \mathbf{P}_{ab} of the flexible model.

Despite their lower values in cluster *a*, probabilities p_{11} and p_{22} still indicate that assets reside longer in bull market states (regime 1) than in bear market states (regime 2; see Formula 6-1). This behavior is generally observed for financial markets (see Ang & Bekaert, 1999, 2002a).

A different picture emerges for cluster *b*, which contains only the UK index. Cluster *b* shows higher transition probabilities for the bear market than for the bull market. However, this behavior is in line with the UK's univariate benchmark results. Chapter 4 showed that the UK has a higher transition probability in regime 2 (0.98) than in regime 1 (0.96). The bear market probability in Formula 6-1 matches this observation. However, cluster *b*'s bull market probability turns out to be much lower than in the univariate case. Again, this can be explained by the additional synchronization dynamics. The UK jointly underlies most bull market events with the rest of the sample. Hence, transitions are governed by \mathbf{P}_{ab} . In contrast, the remaining (idiosyncratic) bull markets of the UK describe only temporary events of much shorter duration.

Matrix \mathbf{P}_{ab} depicts the joint transition probabilities, in case both clusters reside in the same regime. These probabilities should therefore match the results in MV_2 (0.98

⁹Not accounting for the additional regime process of the synchronization factor δ .

¹⁰Note that cluster *b* consists solely of the UK and would therefore be described by UV_2.

in both regimes; see Table 4.5). The high outcome of the values in MV_2 thereby describes a typical shortcoming of standard multivariate regime-switching models. The regime process is generally governed by the best sample fit. Consequently, the dynamics of individual assets only marginally influence the underlying regime process. This leads to higher regime durations, because only switches across all assets can significantly influence the regime cycle in MV_2 .

The flexible model, on the other hand, accounts for individual dynamics across clusters. P_{ab} measures only the transition probabilities when all assets truly reside in the same regime. Consequently, it presents a more realistic picture of the true switching dynamics. This results in lower probabilities of staying in regime 1 or 2 (0.90 and 0.81). Irrespective of their absolute values, these observations again confirm the longer relative duration of bull market regimes.

Finally, the transition probability matrix of the synchronization factor P_δ indicates that the stages of regime synchronization and desynchronization are very persistent (both 0.99). However, this observation depends strongly on cluster definition. An appropriate grouping of the underlying assets causes stable synchronization dynamics.¹¹ On the other hand, the synchronization factor might undergo more volatile dynamics either in case of asset misclassifications or in case of an insufficient number of clusters.

Table 6.5: *Correlation matrices of the flexible model with two clusters (core sample). The period under study is from December 1975 to February 2014.*

State 1			State 2			State 3			State 4		
US	GE	UK	US	GE	UK	US	GE	UK	US	GE	UK
US	1		US	1		US	1		US	1	
GE	0.56	1	GE	0.56	1	GE	0.61	1	GE	0.61	1
UK	0.53	0.82	1	UK	0.43	0.39	1	UK	0.79	0.74	1

Legend: GE = Germany; light-gray shaded areas: regime 1; dark-gray shaded areas: regime 2.

¹¹Of course, the grouping also depends on the structure of the underlying sample. For example, a more heterogeneous group of assets would require more clusters than a homogeneous group. The analysis of the extended sample will further illustrate this point.

Table 6.5 presents the correlation matrices for the four states. These states refer to the extended state space s_t^* presented in Table 5.3.¹² States 1 and 4 resemble the known (unmixed) correlation matrices of the benchmark model. In contrast, intermediate states 2 and 3 define mixed correlation matrices (composite matrices estimated with the factorization algorithm). The gray-marked clusters in these two matrices reuse correlation parameters of state $s_t^* = 1$ or $s_t^* = 4$. Light-gray areas thereby indicate that the correlations belong to regime 1. Dark-gray areas depict the same information for regime 2.

Due to this structure, only inter-cluster correlations had to be estimated separately for states 2 and 3 (white areas). Interestingly, the inter-cluster correlations turned out negative in state 3.¹³ These results are very appealing, given that so far none of the benchmark models has revealed negative correlations (not even MV_4).

Figure 6.2 shows that state 3 only occurred during the first half of the sample period. The provision for such events might be of particular interest for asset allocation decisions.

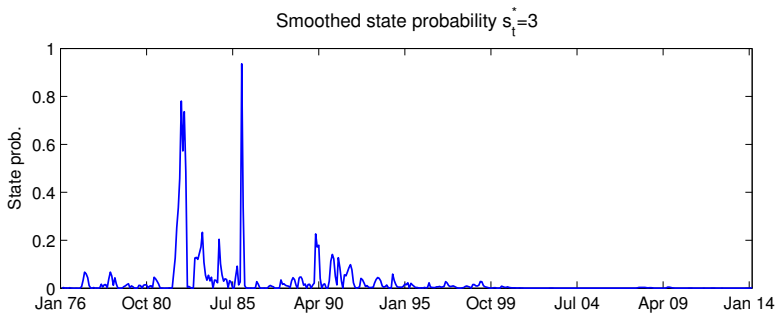


Figure 6.2: *Smoothed regime probabilities of state 3 in the flexible regime-switching model with two clusters.*

Figure 6.3 further depicts the smoothed state probabilities of clusters a and b (probabilities to stay in a bull market). In addition, it presents the corresponding synchronization process (probability of synchronized cycles, $s_t^\delta = 2$).

¹²Even though the state space was extended by the synchronization dimension, this has no influence on the conditional pdfs.

¹³In this state, cluster a resides in a bear regime and cluster b in a bull regime.

The regime processes in the top graph exhibit a volatile behavior. This is due to the low probabilities of cluster a and b for staying in their prevailing regime (p_{11} and p_{22}). Nevertheless, phases of synchronization and independence are clearly evident. The top graph shows how the Markov chains of the two clusters become synchronized after 1993.

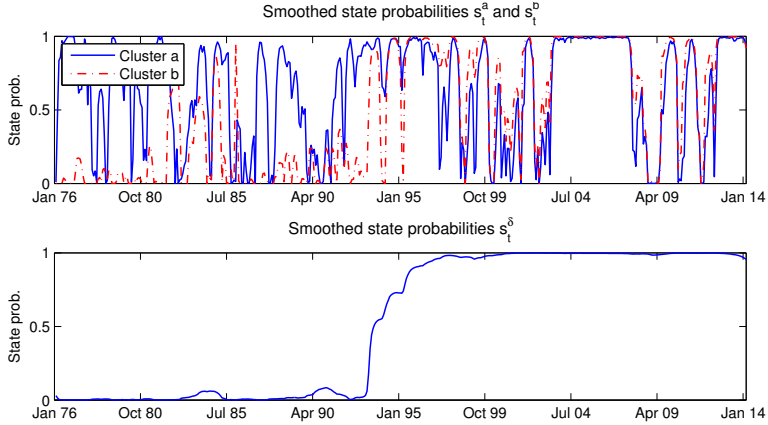


Figure 6.3: *Smoothed state probabilities for the flexible model with two clusters (core sample). a) Probability process for clusters a and b ; b) probability process for regime synchronization s_t^δ .*

The synchronization process, in contrast, is very stable and shows one particular jump (bottom graph). Between March 1993 and November 1995, the cycle switches from near zero to near unity. This behavior illustrates a change from fully independent Markov chains to perfectly synchronized processes. The end of the recessions in the early 1990s can partially explain this behavior. The US, the German, and the UK financial markets jointly switched to a bull regime and adapted their future switching behavior. Due to this adaption of world-wide equity market behavior, the probability for synchronized regimes remained high after the switch in 1993.

Flexible Model with Three Clusters

In FLEX_3C, the US, the UK, and Germany each represent a separate cluster. The previous analysis has revealed the similarity between the moment parameters of the flexible model, on the one hand, and the corresponding univariate benchmark models, on the other. This similarity also applies to FLEX_3C. The estimated moment parameters of this model are therefore omitted.

$$\begin{aligned}
 \mathbf{P}_a &= \begin{bmatrix} 0.83 & 0.20 \\ 0.17 & 0.80 \end{bmatrix} & \mathbf{P}_b &= \begin{bmatrix} 0.80 & 0.04 \\ 0.20 & 0.96 \end{bmatrix} & \mathbf{P}_c &= \begin{bmatrix} 0.83 & 0.20 \\ 0.17 & 0.80 \end{bmatrix} \\
 \mathbf{P}_{abc} &= \begin{bmatrix} 0.92 & 0.11 \\ 0.08 & 0.89 \end{bmatrix} & \mathbf{P}_\delta &= \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}
 \end{aligned} \tag{6-2}$$

Formula 6-2 presents the transition probabilities for the three cluster model. In FLEX_2C, the US and Germany were aggregated in cluster *a*. The current specification divides the two countries into separate clusters (*a* and *c*). However, as expected, the transition probabilities for the two clusters turn out to be similar. This observation supports the decision in FLEX_2C to group the two underlying assets. Likewise, the transition probabilities of cluster *b* match those of the UK in FLEX_2C. However, the joint transition probability \mathbf{P}_{abc} reveals significant differences. The probability of staying in a bear market regime is markedly higher (0.89) than in the two cluster model (0.81). Consequently, joint bear market regimes last almost twice as long in the three cluster model (9.1 months vs. 5.3 months). This increased duration is due to the specification of the synchronization factor. The individual regime processes tend to be more volatile. However, the joint regime process covers only those events where all assets reside in the same regime. These events are usually described by a certain stability.

The synchronization process behaves similarly as in FLEX_2C (see also Figure 6.4). Except for a short trough between December 1994 and December 1997, the smoothed synchronization processes of FLEX_2C and FLEX_3C closely match each other. The Mexican peso crisis, which started in December 1994, presents a potential reason for this trough. At first glance, this emerging market crisis should not have

influenced developed market regimes. However, the US intervened in the Mexican peso crisis, and thus its market was more affected by this crisis. Because synchronization dynamics were averaged across clusters, the US market therefore influenced synchronization across all markets.

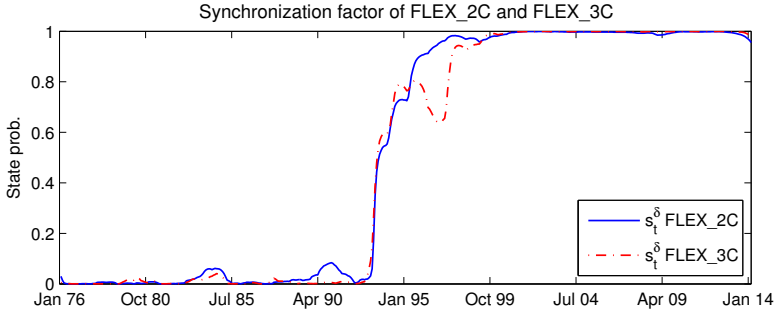


Figure 6.4: Smoothed synchronization process for the flexible models with two and three clusters (core sample).

6.2.3 Analysis of the Extended Sample

Flexible Model with Two Clusters

Table 6.6 presents results for the flexible model as applied to the extended sample. The US, the UK, Germany, Switzerland, and the Pacific area belong to cluster *a* (white background) while Japan defines cluster *b* (gray background).

In Table 6.6, the moment parameters of cluster *a* show similar values compared to the benchmark model MV_2 (see Table 4.5). Means are comparably low and standard deviations do not show large differences between regimes 1 and 2. However, the moment parameters in cluster *a* strongly differ from those in UV_2 (see Table 4.4). The univariate model displays significantly higher means for each asset. Further, its standard deviations differ across the two regimes. The diverging parameter values in the flexible model and the univariate models indicate a possible misspecification of cluster *a*. Further, the flexible model's close resemblance with the parameters in MV_2 is a clear sign for the missing flexibility in cluster *a*.

In general, moment parameters strongly depend on the grouping of the assets in the sample. The assets in cluster *a* underlie heterogeneous switching dynamics and thus should not belong to the same group. However, cluster analysis revealed even stronger switching heterogeneity between the assets of cluster *a* and the Japanese market. Therefore, the current groups present the best possible solution in case of two clusters. However, the above results generally question whether two clusters appropriately cover the extended sample.

Table 6.6: Flexible regime-switching results for the extended sample with two clusters. The period under study is from December 1975 to February 2014.

	μ	σ_1	σ_2		
US	0.11*** (0.025)	0.15*** (0.007)	0.17*** (0.001)	Log Lik	4956.5
				dim(θ)	66
UK	0.12*** (0.034)	0.20*** (0.009)	0.20*** (0.001)	AIC	-9780.9
				BIC	-9508.5
Germany	0.11*** (0.034)	0.19*** (0.009)	0.25*** (0.001)	HQ	-9673.6
Switzerland	0.13*** (0.028)	0.18*** (0.009)	0.18*** (0.001)		
Pacific ex Japan	0.12*** (0.035)	0.24*** (0.010)	0.24*** (0.001)		
Japan	0.07* (0.040)	0.17*** (0.011)	0.27*** (0.001)		

White background: cluster *a*; gray background: cluster *b*. Values are shown on an annualized basis; standard errors are depicted in parentheses.
Log Lik is the log-likelihood. dim(θ) is the size of the parameter set. AIC depicts the Akaike information criterion, BIC the Bayes-Schwarz information criterion, and HQ the Hannan-Quinn information criterion.
***, **, and * indicate the significance of the parameter estimates (different from zero) at the 99%, 95% and 90% level.

A different picture emerges for cluster *b*. The standard deviation of Japan is more than 50% higher in regime 2 than in regime 1. Cluster *b* thereby closely resembles Japan’s results in UV_2 (see Table 4.4). This behavior is in line with the previous argumentation: Japan shows a strong switching heterogeneity compared to the other

assets in cluster a . It therefore needs to follow an individual regime process, which, in turn, accounts for these dynamics.

Interestingly, the misclassification in cluster a does not influence the moment parameters in cluster b . This observation is clearly due to the stability of the flexible model. In standard multivariate regime-switching models, a single asset might already bias all parameters in the sample. In the flexible model, in contrast, this problem is mainly limited to a particular cluster.

Formula 6-3 further presents the corresponding transition probabilities. As expected, cluster a shows very high probabilities of residing in the underlying regime and implies a stable process. The high probabilities most likely are due to the misspecification of cluster a . Its regime process governs only switches that apply jointly to a multitude of assets. Individual dynamics are averaged out, which makes the process highly consistent.

$$\begin{aligned} \mathbf{P}_a &= \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix} & \mathbf{P}_b &= \begin{bmatrix} 0.95 & 0.09 \\ 0.05 & 0.91 \end{bmatrix} \\ \mathbf{P}_{ab} &= \begin{bmatrix} 0.80 & 0.20 \\ 0.20 & 0.80 \end{bmatrix} & \mathbf{P}_\delta &= \begin{bmatrix} 0.99 & 0.10 \\ 0.01 & 0.90 \end{bmatrix} \end{aligned} \quad (6-3)$$

Given the insights from the core sample analysis, at least the UK should be excluded from cluster a . Because the assets in this cluster underlie too heterogeneous switching dynamics, the joint process cannot capture the assets' individual dynamics. Subsequent analysis with three clusters will shed further light on this aspect. Cluster b , in contrast, shows stable but short-lived cycle dynamics.

Formula 6-3 also depicts the transition probability matrix of the synchronization process. The probability of staying in the regime of perfect synchronization is comparably low (0.90). Consequently, synchronization phases only last a few months. Given the heterogeneous sample in cluster a , this result has only limited force. FLEX_3C in the extended sample will further clarify whether this behavior describes the true structure of the underlying assets.

Figure 6.5 confirms the previous observations for the synchronization factor in generally depicting low synchronization between the Markov chains of the two clusters.

The processes only show some degree of synchronization after the end of the East Asian and the Russian crises in 1998, and during the start of the dot-com bubble. However, with the burst of the dot-com bubble, the cycles became fully independent again. The second synchronization event occurred between August 2008 and April 2009. This event describes the collapse of Lehman Brothers and the subsequent beginning of the worldwide financial crisis. Except for these two events, however, the regime cycles behaved fully independently.

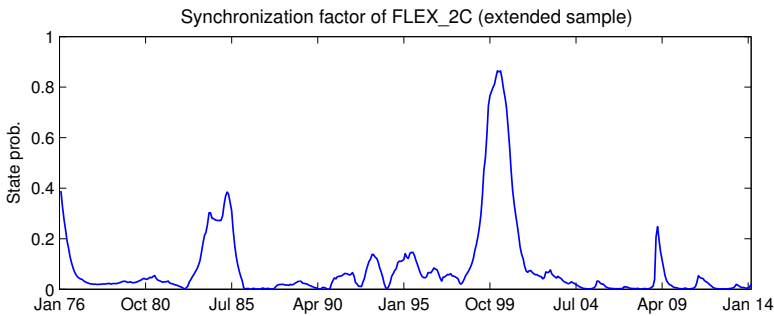


Figure 6.5: *Smoothed state probabilities for the flexible model with two clusters (extended sample).*

Flexible Model with Three Clusters

The application of FLEX_2C proved to be efficient in small samples. However, the previous analysis has shown that the extended sample cannot be appropriately segregated by only two clusters. FLEX_3C therefore presents a possible model extension. Note, however, that the synchronization factor no longer measures the pairwise dependence between two clusters. Instead, it measures the average synchronization among all three clusters.

Cluster a in FLEX_2C consisted of the US, the UK, Germany, Switzerland, and Pacific ex Japan. These markets were clustered because the Japanese market revealed even more heterogeneous switching dynamics. The additional cluster in FLEX_3C is expected to improve this classification and to return more reliable results. The three clusters in FLEX_3C are thereby defined by the US, Germany, Switzerland, and the Pacific area (cluster a), by Japan (cluster b), and by the UK (cluster c).

Due to the additional cluster, the model is able to more closely mimic the switching dynamics of its underlying assets. The estimated moment parameters resemble those in UV_2 (see Table 4.4) and were therefore omitted. However, formula 6-4 presents the corresponding transition probabilities of FLEX_3C. The transition probabilities of cluster a thereby turn out similar to those of cluster a in FLEX_2C (core sample). This is due to the homogeneous switching dynamics of the underlying assets. The transition probabilities of cluster b and c resemble those of Japan and the UK in UV_2 (see Table 4.4), respectively.

In contrast, the transition probabilities of the synchronization process turn out lower than before (see formula 6-4). The duration of the synchronization regime is still very high, but allows for more switches in cycle synchronization. In fact, this behavior became already evident from the synchronization factor in the core sample (FLEX_3C) and is now even more distinct (see the Mexican peso crisis). Potential reasons for this observation are the increased sample size and the setup of the synchronization factor for describing average dependence across all three clusters.

$$\begin{aligned}
 \mathbf{P}_a &= \begin{bmatrix} 0.95 & 0.19 \\ 0.05 & 0.81 \end{bmatrix} & \mathbf{P}_b &= \begin{bmatrix} 0.92 & 0.11 \\ 0.08 & 0.89 \end{bmatrix} & \mathbf{P}_c &= \begin{bmatrix} 0.86 & 0.02 \\ 0.14 & 0.98 \end{bmatrix} \\
 \mathbf{P}_{abc} &= \begin{bmatrix} 0.92 & 0.12 \\ 0.08 & 0.88 \end{bmatrix} & \mathbf{P}_\delta &= \begin{bmatrix} 0.94 & 0.04 \\ 0.06 & 0.96 \end{bmatrix}
 \end{aligned} \tag{6-4}$$

Table 6.7 presents the composite correlation matrices for the eight states s_t^* . Again, the gray-shaded areas present the three diagonal cluster-blocks. The reuse of correlation parameters in states 2 to 7 is clearly observable. At first glance, this effect mainly concerns cluster a , as cluster b and c consist solely of one asset.

However, visual inspection of the correlation parameters shows that the reuse of parameters also applies to off-block-diagonal elements. The factorization algorithm ensures that elements are reused when the regimes of the corresponding assets remain the same across two states s_t^* . For example, the correlation between Japan and the US is the same in $s_t^* = 1$ and $s_t^* = 2$ because the regimes of their clusters remain unchanged in these states. In contrast, the correlations of these assets with the UK are affected by a regime-change in cluster c .

Table 6.7: Correlation matrices of the flexible model with three clusters (extended sample). The period under study is from December 1975 to February 2014.

State 1							State 2						
	US	GE	SW	PA	JP	UK		US	GE	SW	PA	JP	UK
US	1	0.66	0.57	0.57	0.42	0.74	US	1	0.66	0.57	0.57	0.42	0.42
GE	0.66	1	0.73	0.59	0.39	0.80	GE	0.66	1	0.73	0.59	0.39	0.45
SW	0.57	0.73	1	0.54	0.49	0.76	SW	0.57	0.73	1	0.54	0.49	0.50
PA	0.57	0.59	0.54	1	0.46	0.70	PA	0.57	0.59	0.54	1	0.46	0.45
JP	0.42	0.39	0.49	0.46	1	0.51	JP	0.42	0.39	0.49	0.46	1	0.45
UK	0.74	0.80	0.76	0.70	0.51	1	UK	0.42	0.45	0.50	0.45	0.45	1

State 3							State 4						
	US	GE	SW	PA	JP	UK		US	GE	SW	PA	JP	UK
US	1	0.66	0.57	0.57	0.05	0.74	US	1	0.66	0.57	0.57	0.05	0.42
GE	0.66	1	0.73	0.59	0.29	0.80	GE	0.66	1	0.73	0.59	0.29	0.45
SW	0.57	0.73	1	0.54	0.39	0.76	SW	0.57	0.73	1	0.54	0.39	0.50
PA	0.57	0.59	0.54	1	0.19	0.70	PA	0.57	0.59	0.54	1	0.19	0.45
JP	0.05	0.29	0.39	0.19	1	0.61	JP	0.05	0.29	0.39	0.19	1	0.45
UK	0.74	0.80	0.76	0.70	0.61	1	UK	0.42	0.45	0.50	0.45	0.45	1

State 5							State 6						
	US	GE	SW	PA	JP	UK		US	GE	SW	PA	JP	UK
US	1	0.54	0.55	0.58	0.00	0.17	US	1	0.54	0.55	0.58	0.00	0.68
GE	0.54	1	0.77	0.50	0.47	0.57	GE	0.54	1	0.77	0.50	0.47	0.59
SW	0.55	0.77	1	0.46	0.11	0.48	SW	0.55	0.77	1	0.46	0.11	0.65
PA	0.58	0.50	0.46	1	0.55	0.01	PA	0.58	0.50	0.46	1	0.55	0.62
JP	0.00	0.47	0.11	0.55	1	0.51	JP	0.00	0.47	0.11	0.55	1	0.45
UK	0.17	0.57	0.48	0.01	0.51	1	UK	0.68	0.59	0.65	0.62	0.45	1

State 7							State 8						
	US	GE	SW	PA	JP	UK		US	GE	SW	PA	JP	UK
US	1	0.54	0.55	0.58	0.44	0.17	US	1	0.54	0.55	0.58	0.44	0.68
GE	0.54	1	0.77	0.50	0.37	0.57	GE	0.54	1	0.77	0.50	0.37	0.59
SW	0.55	0.77	1	0.46	0.51	0.48	SW	0.55	0.77	1	0.46	0.51	0.65
PA	0.58	0.50	0.46	1	0.47	0.01	PA	0.58	0.50	0.46	1	0.47	0.62
JP	0.44	0.37	0.51	0.47	1	0.61	JP	0.44	0.37	0.51	0.47	1	0.45
UK	0.17	0.57	0.48	0.01	0.61	1	UK	0.68	0.59	0.65	0.62	0.45	1

Legend: GE = Germany, SW = Switzerland, PA = Pacific ex Japan, JP = Japan; light-gray shaded areas: regime 1; dark-gray shaded areas: regime 2.

In addition, the estimated correlations reveal two interesting features: first, contagion does not emerge as pronounced for $s_t^* = 1$ and $s_t^* = 8$, as correlations turn out relatively similar across these states; and second, in the intermediate states 2 to 7, Japan (cluster *b*) and the UK (cluster *c*) show very low correlations with some assets (close to zero). These features can be very appealing for portfolio allocation and risk management.

Finally, Figure 6.6 depicts the regime process of the synchronization factor. The process indicates a more volatile synchronization behavior, confirming the initial assumptions of time-varying regime dependence. Further, even after the switch in 1993, the process shows some volatile tendencies. These are clearly caused by the different financial and economic crises in the 1990s and 2000s.

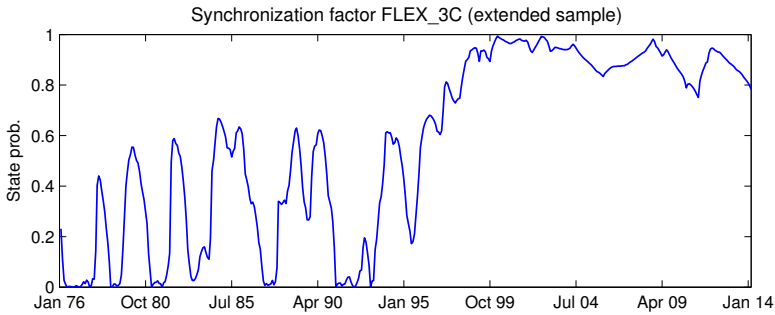


Figure 6.6: *Smoothed synchronization process for the flexible model with three clusters (extended sample).*

The previous results clearly illustrate the power of FLEX_3C to appropriately detect corresponding asset dynamics in large samples. This information is useful in many respects and allows one not only to detect the switching dynamics of individual assets but, as the next section will show, to produce even more stable results.

6.2.4 Model Robustness

The robustness of the flexible models is evaluated in two ways: first, with respect to the stability of the synchronization factor in individual subperiods; and second, with respect to the impact of misclassified assets on model stability (assignment to wrong cluster).

Figure 6.7 illustrates the synchronization factor for the core sample. The flexible models with two clusters (FLEX_2C) has been analyzed pre- and post-September 1993. This date marks the switch of the synchronization factor from fully independent to perfectly synchronized cycles. Consequently, it presents the optimal date for separating the sample. A well-defined model is thereby expected to return similar synchronization dynamics for respective subperiods as for the full sample.

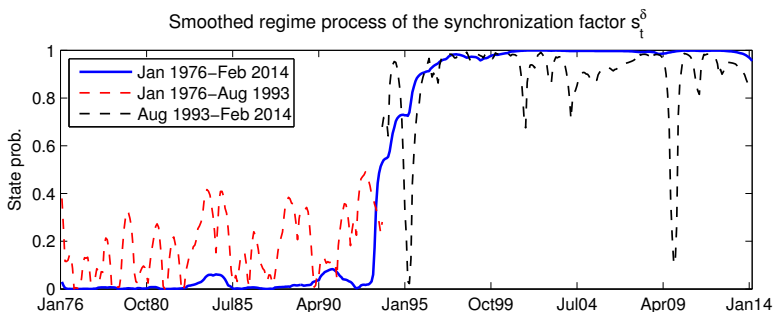


Figure 6.7: *Smoothed regime process for the synchronization factor: Analysis of subperiods (core sample).*

The results in Figure 6.7 confirm this claim. The pre-1993 sample returns only very moderate probabilities for $s_t^\delta = 2$ (perfect synchronization). In contrast, the regime process of the sample after 1993 shows mainly probabilities very close to unity. Figure 6.7 shows that the two subperiods closely match the process of the full sample. Similar dynamics have also been detected for the extended sample with FLEX_3C (not shown). Furthermore, both models have returned very stable and similar moment and correlation parameters both in the subsamples and in the full sample. The flexible setup clearly outperforms standard regime-switching models in terms of sta-

bility. This is very appealing given its comparably low increase in the parameter set compared to MV_2.

The second test for robustness analyzed the appropriate clustering of individual assets. Previous results have already revealed the sensitivity of flexible models towards the appropriate number of clusters. This problem is evident and can only be circumvented through detailed investigation of the sample. However, analysis has shown that this problem also affects standard multivariate regime-switching models.

The current analysis focuses specifically on the misclassification of individual assets. Therefore, the clusters of the extended sample were adjusted to the US, Germany, and Switzerland (cluster *a*), the UK (cluster *b*), and Japan and the Pacific area (cluster *c*). These clusters were formed based on a second-best clustering solution. Consequently, the groups do not represent completely arbitrary clustering choices.

$$\begin{aligned} \mathbf{P}_a &= \begin{bmatrix} 0.85 & 0.25 \\ 0.15 & 0.75 \end{bmatrix} & \mathbf{P}_b &= \begin{bmatrix} 0.77 & 0.07 \\ 0.23 & 0.93 \end{bmatrix} & \mathbf{P}_c &= \begin{bmatrix} 0.82 & 0.03 \\ 0.18 & 0.97 \end{bmatrix} \\ \mathbf{P}_{abc} &= \begin{bmatrix} 0.92 & 0.25 \\ 0.08 & 0.75 \end{bmatrix} & \mathbf{P}_\delta &= \begin{bmatrix} 0.98 & 0.01 \\ 0.02 & 0.99 \end{bmatrix} \end{aligned} \quad (6-5)$$

Formula 6-5 presents corresponding transition probabilities for the alternative model based on FLEX_3C. The original FLEX_3C results and those in formula 6-5 show very similar switching dynamics. Visual inspection of their corresponding synchronization processes in Figure 6.8 further confirms this assumption.¹⁴ In addition, a completely arbitrary model with the US, the UK, Germany, and the Pacific area (cluster *a*), Japan (cluster *b*), and Switzerland (cluster *c*) was formed. These wrong clusters illustrate a completely different picture in Figure 6.8. The tests have been conducted for both flexible models and both samples sizes and have shown similar results.

¹⁴The moment and correlation parameters of the two specifications have also turned out to be similar and were therefore omitted.

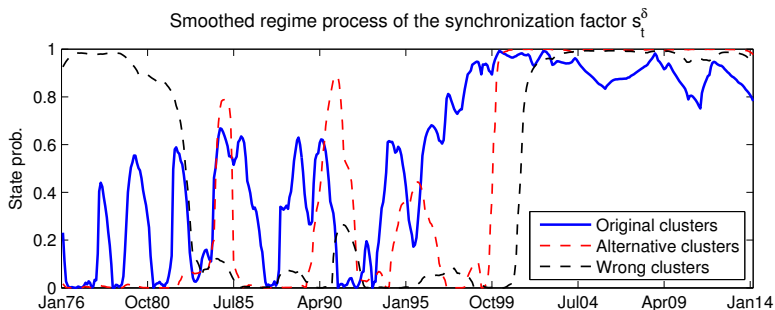


Figure 6.8: *Smoothed regime process for the synchronization factor: Analysis of alternative cluster specifications (extended sample).*

Consequently, the results of flexible models are stable with regard to the clustering of individual assets. Note, however, that assets have only been assigned to another cluster if their switching dynamics also matched with that cluster. More obvious misspecification clearly lead to severe changes in model dynamics.

6.3 Accuracy of the Synchronization Factor

Chapter 4 presented the synchronization dynamics for pairwise univariate regime cycles. It showed that this synchronization analysis returns reliable numbers when the processes are independent. However, Camacho and Perez-Quiros (2006) demonstrated that analysis is ineffective when the regime processes underlie joint dynamics. Similar to the specification of Camacho and Perez-Quiros (2006) (see Chapter 5), the flexible regime-switching model should not underlie this restriction.

A simulation approach will therefore help to assess the ability of the flexible model to detect perfectly synchronized cycles. Results will be compared to the simple analysis of correlations between univariate regime processes.

This approach follows the remarks of Camacho and Perez-Quiros (2006) and consists of three steps:

First, a regular bivariate regime-switching model for the US and the UK sample is estimated. This model contains the moment parameters, correlations, transition

probabilities, and regime probabilities for the two countries. It implies perfect dependence between the two countries' regime dynamics. Following Ang and Bekaert (2002b), a second parameter set is estimated for a bivariate model with independent cycles.

Second, these input parameters are used to simulate 100 pairs of return series (Camacho & Perez-Quiros, 2006). In the dependent case, both simulated return series are based on the same regime process. In the independent case, the return series rest on their individual regime processes.

Third, a univariate regime model is applied to both scenarios. For each return series, two regime probability processes emerge (one for each asset). Following Harding and Pagan (2006), the two regime processes are then tested with respect to their pairwise correlation. Finally, the resulting 100 correlations for each scenario serve as input for a kernel density estimation.

The same procedure is applied to the flexible model. However, the flexible model estimates a synchronization process for each return series. This process returns a time-varying dependence series. Hence, the median value of each process needs to be calculated for the kernel density estimation. This assumption should not influence the results, given that the underlying processes are either fully independent or perfectly synchronized over the entire sample. Consequently, the information content of the synchronization factor is not biased by using its median.

Figure 6.9 illustrates the kernel density estimates for the two techniques. The results for the independent case turn out relatively similar. This is in accordance with expectations, as independent cycles imply no further joint dynamics and can be easily detected. In the dependent case, however, the flexible model better resembles the correlation between regimes than the univariate model. These results are in line with Camacho and Perez-Quiros (2006).

The flexible model is able to infer corresponding dynamics despite the serial correlation of the underlying Markov chains. The correlation of univariate models, in contrast, returns much lower values. These observations clearly confirm the power of the flexible setup to appropriately detect the synchronization dynamics among individual regime cycles. Similar observations are also expected for intermediate

degrees of synchronization. Due to the dynamic structure of the model, it should therefore also appropriately cover time-varying degrees of synchronization.

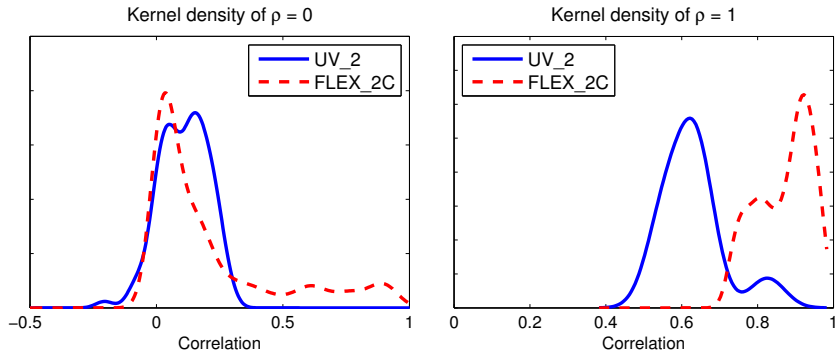


Figure 6.9: Kernel density estimates for the correlations between individual and synchronized regime cycles for UV_2 and FLEX_2C. a) Perfectly independent Markov chain processes ($\rho = 0$); b) perfectly synchronized Markov chains ($\rho = 1$).

6.4 Model Forecasting Power

To further evaluate the flexible regime-switching models, out-of-sample tests were conducted. Two different tests were applied to the different regime-switching models: a forecast of standard deviations and a test for turning-point prediction. These tests provide interesting insights about the forecasting power of the different models. The one-period-ahead variance forecast is specified as

$$Var_t[y_{t+1}] = \sigma_{t+1|t}^2 = \sigma_1^2 \xi_{1,t+1|t} + \sigma_2^2 \xi_{2,t+1|t} + \xi_{1,t+1|t} \xi_{2,t+1|t} (\mu_1 - \mu_2)^2,$$

where σ_k^2 describes the variance in state k and where $\xi_{k,t+1|t} = \Pr(s_{t+1}^2 = k | \Omega_t)$.¹⁵ In the flexible models, the state probabilities are represented by the probabilities of

¹⁵Note that $Var_t[y_{t+1}]$ is calculated for each asset in the sample. This implies that each asset i is defined by its corresponding mean and variance parameters $\mu_{i,k}$ and $\sigma_{i,k}^2$. Due to the structure of the

the corresponding cluster. The aggregation of these probabilities is illustrated in formula 5-6.

Due to the regime-independent means, the one-period-ahead standard deviation simplifies to

$$\sigma_{t+1|t} = \sqrt{\sigma_1^2 \xi_{1,t+1|t} + \sigma_2^2 \xi_{2,t+1|t}}.$$

In addition to the flexible models, this approach was also applied to the multivariate benchmark model, where all assets share a common regime process. Figure 6.10 presents the observed 12-month rolling-window standard deviation (annualized) for the three countries in the core sample (Observed). Furthermore, it shows the forecasting results of each country for the multivariate (MV_2) and the two flexible models (FLEX_2C and FLEX_3C).

Figure 6.10 illustrates the similarity between the two flexible approaches. Further, it shows that the flexible models more closely resemble the true standard deviations than the standard multivariate regime-switching model. The standard model tends to underestimate market volatility. Due to its structure, the underlying regime process is very stable and therefore adjusts its expectations only slowly. The flexible model, in contrast, underlies a very flexible structure and can react swiftly to market movements. This is also confirmed by its lower probabilities of staying in the prevailing regime.

flexible regime-switching model, its assets are further conditioned on the corresponding regime process $\xi_{k,t+1|t}^i$ for $i = a, b, ab, \delta$.

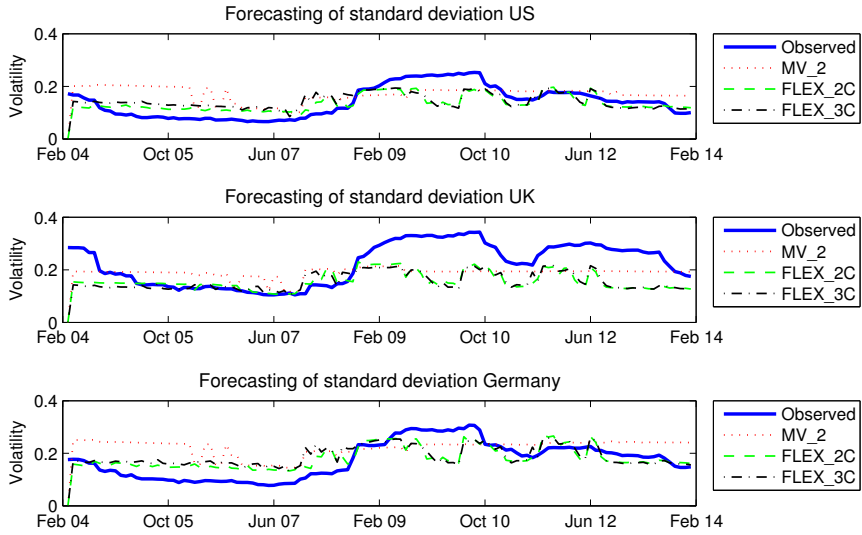


Figure 6.10: *Rolling-window forecasts of standard deviations for the US, the UK, and Germany. Legend: All values describe annualized standard deviations. Observed is the 24-month rolling-window observable standard deviation of the sample. The period of analysis is from March 2004 to February 2014.*

The test for turning-point prediction (TP) was specified as

$$TP = T^{-1} \sum_{t=1}^T \left(d_t - \hat{d}_t \right)^2,$$

where

$$d_t = \mathbb{I}_{\xi_{1,t}|T \geq 0.5} = \mathbb{I}_{\Pr(s_t=1|\Omega_T; \theta) \geq 0.5}$$

$$\hat{d}_t = \xi_{1,t|t-1} = \Pr(s_t = 1|\Omega_{t-1}; \theta),$$

and where \mathbb{I} is the indicator function.¹⁶ TP measures the average difference between binary smoothed regime probabilities estimated at date T (d_t) and forecasted prob-

¹⁶This procedure follows Hamilton and Perez-Quiros (1996) and is slightly modified to match the underlying model specification.

abilities at time t (\hat{d}_t), where $T \geq t$. To avoid the inclusion of future information in \hat{d}_t , estimates were carried out on a rolling-window basis. Again, for the flexible model, TP was carried out individually for each cluster following the aggregation in formula 5-6.

Table 6.8: *Turning-point prediction of different regime-switching models. The period under study is from December 1976 to February 2014.*

	MV_2	FLEX_2C		FLEX_3C		
		Cluster a	Cluster b	Cluster a	Cluster b	Cluster c
Core sample	0.148	0.080	0.103	0.049	0.070	0.048
Extended sample	0.558	0.502	0.205	0.023	0.036	0.053

Table 6.8 shows the results for the core and the extended sample. Values close to zero indicate a good prediction of future regimes, whereas values close to unity describe weak forecasting power. The multivariate model provides very weak turning-point estimates in both samples. In contrast, FLEX_2C and FLEX_3C show comparably good estimates. However, the results of FLEX_2C in the extended sample turned out about as high as those of MV_2. This supports the previous claim that two clusters insufficiently represent the extended sample. Finally, FLEX_3C reveals better average results than FLEX_2C in both samples. The regime expectations of flexible models prove to closely match future observations.

6.5 Summary

The empirical application of flexible regime-switching has demonstrated the quality of this approach. Flexible regime-switching models clearly compete with standard multivariate regime-switching setups. In a well-specified flexible model, individual clusters closely resemble the moment parameters of univariate regime-switching models. In contrast to univariate models, however, flexible models also account for correlations across assets and for regime cycle synchronization across clusters. These features, in turn, more closely resemble the properties of multivariate regime-

switching models. Flexible models therefore combine the features of both univariate and multivariate regime-switching.

The analysis of flexible models has revealed two important model prerequisites: a careful selection of the number of clusters, and the appropriate clustering of the underlying assets. Thorough model analysis can easily ensure these requirements. Analysis has shown that standard regime-switching models theoretically suffer from the same problems. Due to their non-flexible structure, however, they cannot circumvent these problems as easily as the flexible model.

Analysis has further detected three typical problems of standard multivariate regime-switching models: first, that parameter and regime estimates might change dramatically over the sample period; second, that a sample with heterogeneously switching assets might show only moderate parameter differences between regimes; and third, that the probabilities of staying in the prevailing regime generally turn out very high, as only changes in dynamics across a majority of assets lead to a joint regime switch. Even a single asset with diverging switching dynamics might thereby bias the entire parameter set of a multivariate regime-switching model.

Due to their dynamic structure, flexible models can easily deal with these problems and can therefore avoid potentially biased estimates. Given an appropriate number of clusters, the model can also handle samples with very distinct switching dynamics. The forecasting results have further confirmed that these model properties lead to very efficient and accurate volatility and turning-point predictions. Whether these results also contribute to an investor's asset allocation problem will be analyzed in the next chapter. Analysis will also shed further light on the out-of-sample performance and the optimal real-world application of the flexible strategies.

Chapter 7

Asset Allocation under Regime-Switching

The previous chapters have revealed the synchronization and switching dynamics of flexible regime-switching models. These observations differ clearly from the multivariate benchmark results of Chapter 4. Therefore, a next step is to compare these models with respect to their asset allocation. The analysis of an investor's portfolio choice problem will help to evaluate the portfolio stability and the out-of-sample performance of these models.

This chapter first specifies the investor's portfolio choice problem before applying it to the regime-switching models. The specification starts with a simple buy-and-hold strategy. Thereafter, portfolio rebalancing is introduced. Rebalancing allows the investor to adjust his portfolio allocation with the arrival of new information. This enhancement can be of particular value considering the time-varying investment opportunity set of regime-switching models and, in particular, of the flexible setup.

7.1 Asset Allocation Problem

The asset allocation problem can be stated as follows. An investor faces a T -month investment horizon at time t and aims to maximize his expected utility over terminal wealth $U(W_{t+T})$

$$\max_{\alpha_t, \dots, \alpha_{t+T-1}} E_t[U(W_{t+T})] \quad \text{s.t. } \alpha'_i \mathbf{1} \leq 1, \quad (7-1)$$

where W_{t+T} is the end of period wealth, $\alpha_t, \dots, \alpha_{t+T-1}$ are the portfolio weights for the respective periods, and $i = t, \dots, t + T - 1$. The investor's preference is thereby defined as (CRRA) power utility over terminal wealth

$$U(W_{t+T}) = \frac{W_{t+T}^{1-\gamma}}{1-\gamma},$$

where γ is the coefficient of relative risk aversion and where $\gamma > 1$. CRRA utility is chosen because of its widespread use and comparability to other findings (Verhofen, 2006). At time t , the investor determines his allocation to the N risky assets $\alpha_t \equiv [\alpha_t^1, \dots, \alpha_t^N]$, whereas $1 - (\alpha'_t \mathbf{1}_N)$ is allocated to the risk-free investment r^f . This risk-free asset is described by the monthly T-bill yield.¹

7.1.1 Buy-and-Hold Allocation Strategy

In case the investor follows a buy-and-hold strategy, portfolio weights depend only on the single vector α_t and the corresponding risk-free allocation. The investor solves the asset allocation problem at time t and no rebalancing applies. Further, transaction costs are ignored and short-selling is restricted. Consequently, the asset

¹In line with Ang and Bekaert (2002a) and Guidolin and Timmermann (2005b, 2007), this thesis implies a partial equilibrium framework where an asset's return process is exogenous. The risk-free rate is assumed to be both known and constant. It equals the average 1-month T-bill yield over the sample period and amounts to annualized 2.5% (see also Guidolin & Timmermann, 2007).

allocation problem in 7-1 is simplified to

$$\begin{aligned} & \max_{\alpha_t} E_t \left[\frac{W_{t+T}^{1-\gamma}}{1-\gamma} \right], \\ \text{s.t.} \quad & W_{t+T} = W_t \left\{ (1 - (\alpha'_t \mathbf{1}_N)) \exp(r^f T) + \alpha'_t \exp \left(\sum_{i=1}^T y_{t+i} \right) \right\}, \\ & e'_j \alpha_t \in [0, 1] \text{ for } j = 1, \dots, N, \\ & \alpha'_t \mathbf{1}_N \leq 1, \end{aligned}$$

where W_t is set to unity, y_{t+1} is the $(N \times 1)$ vector of continuously compounded returns², and e_j describes a $(N \times 1)$ vector of zeros with unity in row j . Finally, the conditions in lines three and four restrict the short-selling of individual assets and of the overall portfolio, respectively.

Because the Markov chain process is hidden, the investor does not know the regime at time t , but is only provided with the corresponding state probabilities. In general, he is confronted with uncertainty about both future regimes and the prevailing regime (Morger, 2006).

In contrast, Ang and Bekaert (2002a) introduced an asset allocation problem in which the investor knows the prevailing regime with certainty and in which only future regimes remain uncertain. Assuming the current regime as known describes the most extreme scenario. Uncertainty about the prevailing regime, in contrast, causes asset allocation to deviate less from the regime-independent i.i.d. solution of a myopic portfolio.³

The present study treats the current regime as latent. This assumption is more realistic as the regimes are non-observable in reality. Further, this assumption requires an investor to account for future revisions of state probabilities when initially defining the portfolio weights.

²The returns are simulated based on the predictive distribution of the underlying regime-switching model.

³At least this holds for a one-period horizon.

Following Guidolin and Timmermann (2007), this thesis approximates the integral of expected utility using Monte Carlo methods. For the simple case of a buy-and-hold investor, Barberis (2000), Honda (2003), Guidolin and Timmermann (2007) demonstrated that the expected utility can be approximated by

$$\max_{\alpha_t} H^{-1} \sum_{h=1}^H \left\{ \frac{\left[(1 - (\alpha'_t \mathbf{1}_N)) \exp(r^f T) + \alpha'_t \exp\left(\sum_{i=1}^T y_{t+i,h}\right) \right]^{1-\gamma}}{1 - \gamma} \right\},$$

where H defines the number of Monte Carlo simulations and where $\alpha'_t \exp\left(\sum_{i=1}^T y_{t+i,h}\right)$ is the return of the equity portfolio in simulation h . Each simulation path relies on an individual draw from the predictive distribution of the underlying regime-switching model. For further details on this simulation approach, see Guidolin and Timmermann (2005b, 2007).

7.1.2 Asset Allocation with Rebalancing

If the investor allows for portfolio rebalancing, the weights α in 7-1 are adjusted every $\varphi = T/RB$ months (Guidolin & Timmermann, 2007). Hereby, RB defines the number of equally spaced rebalancing points $t_{rb} = t, t + \varphi, \dots, t + (RB - 1)\varphi$.⁴ In order to simplify the notation in α and W , rebalancing points are subsequently indicated by the index $rb = 0, 1, \dots, RB - 1$.

The observations of the portfolio weights in 7-1 ($\alpha_t, \dots, \alpha_{t+T-1}$) match these rebalancing points. The same applies to the weights of the risk-free investment $1 - (\alpha'_{rb} \mathbf{1}_N)$. Similar to the buy-and-hold strategy, transaction costs are ignored and short-selling is restricted. The asset allocation problem then follows a modified version of 7-1

⁴Consequently, if the investor follows a buy-and-hold strategy, $RB = 1$ and $\varphi = T$.

$$\max_{\alpha_t, \dots, \alpha_{RB-1}} E_t \left[\frac{W_{t+T}^{1-\gamma}}{1-\gamma} \right],$$

s.t.

$$W_{rb+1} = W_{rb} \left\{ (1 - (\alpha'_{rb} \mathbf{1}_N)) \exp(\varphi^{rf}) + \alpha'_{rb} \exp \left(\sum_{k=1}^{\varphi} y_{t_{rb}+k} \right) \right\}, \quad (7-2)$$

$$e'_j \alpha_{rb} \in [0, 1] \text{ for } j = 1, \dots, N,$$

$$\alpha'_{rb} \mathbf{1}_N \leq 1 \quad \forall \text{ } rb.$$

Given this specification, the derived utility of wealth is defined by

$$J(W_{rb}, y_{rb}, \theta_{rb}, \xi_{rb}, t_{rb}) \equiv \max_{\alpha_{rb}, \dots, \alpha_{RB-1}} E_{t_{rb}} \left[\frac{W_{RB}^{1-\gamma}}{1-\gamma} \right],$$

where θ_{rb} is the vector of regime-switching parameters and where ξ_{rb} is the $(K \times 1)$ vector of filtered state probabilities conditional on the available information at time t_{rb} . Given the assumption of power utility, the Bellman equation is simplified to

$$J(W_{rb}, y_{rb}, \theta_{rb}, \xi_{rb}, t_{rb}) = Q(y_{rb}, \theta_{rb}, \xi_{rb}, t_{rb}) \frac{W_{rb}^{1-\gamma}}{1-\gamma},$$

where $\gamma \neq 1$ (Guidolin & Timmermann, 2005b). Further, the updating equation introduced in 2-12 can be adjusted to incorporate investors' learning

$$\xi_{rb+1|t} = \frac{\hat{\mathbf{P}}_t^\varphi \xi_{rb|t} \odot \eta_{rb}(y_{rb+1}; \hat{\theta}_t)}{\mathbf{1}' \left(\hat{\mathbf{P}}_t^\varphi \xi_{rb|t} \odot \eta_{rb}(y_{rb+1}; \hat{\theta}_t) \right)}, \quad (7-3)$$

where the hat operator describes an estimation and where $\hat{\mathbf{P}}_t^\varphi$ is the φ -period ahead transition probability matrix. Moreover, $\xi_{rb|t}$ is the filtered state probability at time t_{rb} , given the parameter set $\hat{\theta}_t$ estimated at the initial period $t_{rb=0}$. Finally, $\eta_{rb}(y_{rb+1}; \hat{\theta}_t)$ is the $K \times 1$ vector of the conditional densities of y_{rb+1} given the

parameter set $\hat{\theta}_t$. Guidolin and Timmermann (2007) showed that formula 7-3 helps the investor to optimally revise his perception about the underlying state.

Because wealth at time t_{rb} is known, the recursion of $Q(\cdot)$ can be expressed as

$$Q(y_{rb}, \xi_{rb}, t_{rb}) = \max_{\alpha_{rb}} E_{t_{rb}} \left[\left(\frac{W_{rb+1}}{W_{rb}} \right)^{1-\gamma} Q(y_{rb+1}, \xi_{rb+1}, t_{rb+1}) \right]. \quad (7-4)$$

In contrast to the buy-and-hold strategy, the optimal weights reflect not only the demand for hedging against stochastic shifts in the investment opportunity set, but also for hedging against the revision of beliefs about future state probabilities $\xi_{rb+i|t}$, where $i = 1, \dots, T$ (Guidolin & Timmermann, 2005b). For the implementation of the rebalancing strategy, it is referred to Guidolin and Timmermann (2007).

The Bellman equation is solved numerically using backward induction methods. Further, the interval $[0, 1]$, which defines the domain of each regime probability ξ_{rb} , is discretized on G points. Guidolin and Timmermann (2007) assumed that $Q(y_{rb+1}, \xi_{rb+1}, t_{rb+1})$ is known at the points $\xi_{rb+1} = \xi_{rb+1}^j$, where $j = 1, \dots, G^{K-1}$. To obtain $Q(y_{rb}, \xi_{rb}, t_{rb})$ in 7-4,

$$E_{t_{rb}} \left[\left\{ (1 - (\alpha'_{rb} \mathbf{1})) \exp(\varphi^f) + \alpha'_{rb} \exp(R_{rb+1,h}(s_{rb})) \right\}^{1-\gamma} Q(\xi_{rb+1}^j, t_{rb+1}) \right] \quad (7-5)$$

can be maximized with respect to α_{rb} , where $R_{rb+1,h}(s_{rb}) = \sum_{k=1}^{\varphi} y_{t_{rb}+k,h}(s_{rb})$. This multiple integral is again calculated using Monte Carlo methods. For each grid-point, H draws of $R_{rb+1,h}(s_{rb})$ are generated by the regime-switching model.⁵ The grid-points are thereby defined by $\xi_{rb} = \xi_{rb}^j$, where $j = 1, \dots, G^{K-1}$.

For a standard multivariate regime-switching model with two states, this discretization grid has only a single dimension. For example, a model with six grid-points is evaluated at $\xi_{1,rb}^{(j)} = \{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ (regime 1) and consequently also at $\xi_{2,rb}^{(j)} = 1 - \xi_{1,rb}^{(j)}$ (regime 2). In case of more than two regimes, the dimension of grid-

⁵Note that each of these draws consists of φ individual observations, which are separately simulated by the regime-switching model.

points G^{K-1} grows exponentially, and the probabilities need to be further restricted: $\sum_{i=1}^K \xi_{i,rb}^{(j)} = 1, \forall j$. Fortunately, this restriction does not apply to the flexible model. Given its cluster-wise specification, the individual Markov chains still follow only two regimes $\xi_{rb}^{(j,m)}$ for $m = a, b, ab$, and δ (two clusters) or $m = a, b, c, abc$, and δ (three clusters). However, each underlying Markov chain is defined by its own discretization grid G_m^{K-1} . In case of the two-cluster model, this implies a separate grid for clusters a and b , for the joint process ab , and for the synchronization process δ . The probabilities of these grids can be combined simply, just as for the specification of the joint probability in 5-6. This results in a discretization grid for the higher-dimensional state s_t^* , which forms the basis of subsequent return simulations. Finally, the expectation in 7-5 is approximated by

$$H^{-1} \sum_{h=1}^H \left[\left\{ (1 - (\alpha'_{rb} \mathbf{1}_N)) \exp(\varphi^f) + \alpha'_{rb} \exp(R_{rb+1,h}(s_{rb})) \right\}^{1-\gamma} Q\left(\xi_{rb+1,h}^{(j)}, t_{rb+1}\right) \right], \quad (7-6)$$

where $\xi_{rb+1,h}^{(j,m)}$ for $m = a, b, ab, \delta$ equals the grid-element $\xi_{rb+1}^{(j,m)}$, which is closest to

$$\xi_{rb+1,h}(\hat{\theta}_t) = \frac{\hat{\mathbf{P}}_t^\varphi \xi_{rb|t} \odot \eta_{rb}(y_{rb+1,h}; \hat{\theta}_t)}{\mathbf{1}' \left(\hat{\mathbf{P}}_t^\varphi \xi_{rb|t} \odot \eta_{rb}(y_{rb+1,h}; \hat{\theta}_t) \right)}, \quad (7-7)$$

measured by the distance function $\sum_{i=1}^{K-1} |\xi_{rb+1}^{(j,m)} e_i - \xi_{rb+1,h} e_i|$ (Guidolin & Timmermann, 2005b). This algorithm is iterated backwards over all RB rebalancing points and until $\alpha_{rb=0}$ is reached. Appendix B in Guidolin and Timmermann (2005b) and Section 4.3.2 in Morger (2006) provide further details about this algorithm.

7.2 Asset Allocation Results

In line with the empirical analysis of Chapter 6, this section focuses on two aspects of the flexible regime-switching model: first, it analyzes the impact of flexible regime-switching on asset allocation; and second, it examines whether the proposed composite correlation matrix influences model performance.

Analysis concentrates on four basic models: the linear non-switching model (LIN), the multivariate benchmark regime-switching model (MV_2), the two-cluster flexible model (FLEX_2C), and the three-cluster flexible model (FLEX_3C). In addition, a modified version of FLEX_2C and FLEX_3C will be analyzed. To evaluate the power of the composite correlation matrix, this version restricts the correlation matrix to follow cluster a 's regime. All models thereby underlie regime-independent means.

The dataset remains the same as in Chapter 6. The estimation period for the parameter set θ ranges from December 1975 to January 2004. This allows for a non-overlapping holding period of up to 120 months (until February 2014).

7.2.1 Portfolio Weights

The analysis of portfolio weights starts with a simple buy-and-hold investment problem. This allows one to evaluate the different models without considering any revision in expectations. The current section concentrates on two dimensions of the investment problem: the holding period T and the risk aversion γ . Asset weights are initially compared with respect to these dimensions. The number of Monte Carlo simulations H was set to 30'000.

For better visual inspection, portfolio weights are chiefly analyzed for the core sample, as results for the extended sample have revealed a similar allocation behavior.

Investment Horizon

Analysis of the investment horizon is conducted for periods between 1 and 120 months. Further, a risk aversion of $\gamma = 5$ was assumed, as this corresponds to a typically used value in financial literature. The comparison of portfolio weights with respect to the investment horizon is thereby useful for determining initial model differences.

For LIN, Figure 7.1 presents a very stable allocation along the investment horizon. As shown by Samuelson (1969), if the returns y_t are i.i.d. across time, the weights for CRRA utility are constant. Consequently, the T -month investment problem delivers the same weights as a myopic, one-period problem (see also Ang & Bekaert, 1999). The linear model in Figure 7.1 confirms this observation.

Interestingly, a similar picture emerges for MV_2. Portfolio weights are also relatively stable for different investment horizons. Moreover, a majority of funds is allocated towards the US index. This trend even increases slightly over time. The reason for this allocation is the relative strength of the US. Regime-switching results in Chapter 4 have demonstrated that the US shows higher returns and lower volatilities than the UK and Germany (in both regimes). Given also the relatively high correlations between these countries, the potential of diversification is limited.

The flexible model reveals a different picture. Portfolio weights of FLEX_2C approximate those of LIN. The US/Germany cluster and the UK cluster can switch regimes independently. Due to this temporary independence, the UK might reside in a bull regime while the US and Germany reside in a bear regime. In this case, the lower bull market volatility of the UK clearly competes with the risk-return profile of the US. Moreover, due to the composite correlation matrix, intermediate states $s_t^* = 2$ and $s_t^* = 3$ are described by low or even negative correlations. Consequently, the diversification among assets is more pronounced compared to MV_2. These results highlight the power of flexible regime-switching and the significant influence of the composite correlation matrices on portfolio allocation.

Visual inspection also shows that the US allocation in FLEX_2C even decreases slightly along the investment horizon (in favor of the UK). This behavior is explained

by the rather short durations of the underlying regime processes (a, b, ab, δ). Consequently, joint events, where the US and the UK reside in different states, take some time to occur. For longer investment horizons, however, the likelihood of such events increases and leads to a change in portfolio exposure.

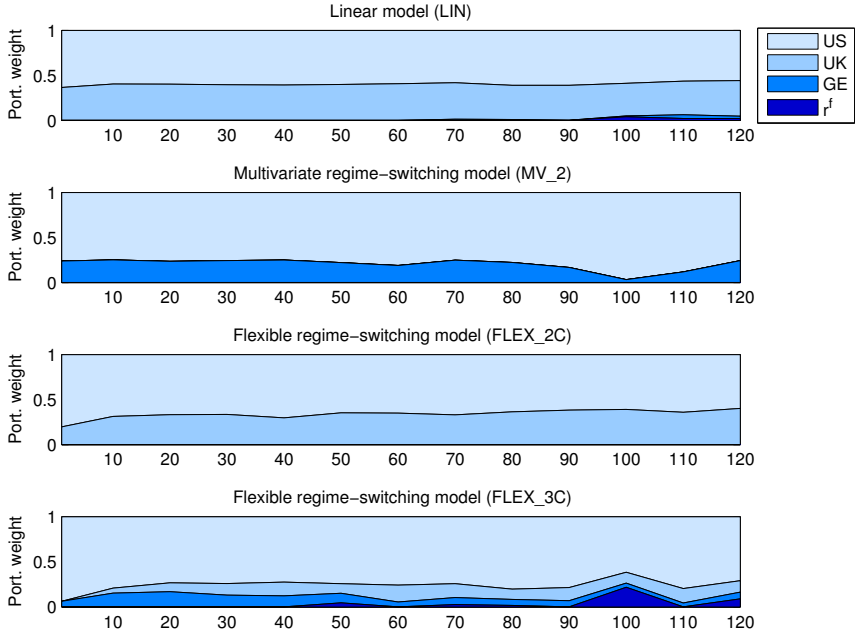


Figure 7.1: *Change of portfolio weights with increasing holding periods, covering the period from January 2004 to February 2014. Results are depicted for the core sample. Risk aversion γ is set to 5. The x-axis depicts the length of the holding period HP (in months).*

Despite these good properties of flexible regime-switching models, portfolio diversification might not be optimally reached with a two-cluster model. Similar to LIN, no funds are allocated to Germany. As both the US and Germany belong to cluster a , they always follow the same regime process. Consequently, Germany does not provide any useful diversification, given its high correlation with the US and given its comparably weak risk-return features. Separating the US and Germany might therefore improve results. The last graph in Figure 7.1 indeed shows an increase in the

allocation to Germany and even to the risk-free asset. However, this exposure is low and comes at the price of an overly reduced allocation towards the UK. The US, in contrast, shows a high and stable exposure (mostly above 70%), which is due to the setup of FLEX_3C. Synchronization in this model is jointly measured between all three clusters. Hence, individual dynamics across two clusters cannot be determined as accurately. Consequently, the occurrence of intermediate states across two clusters decreases, which leads to a closer approximation of the results in MV_2.

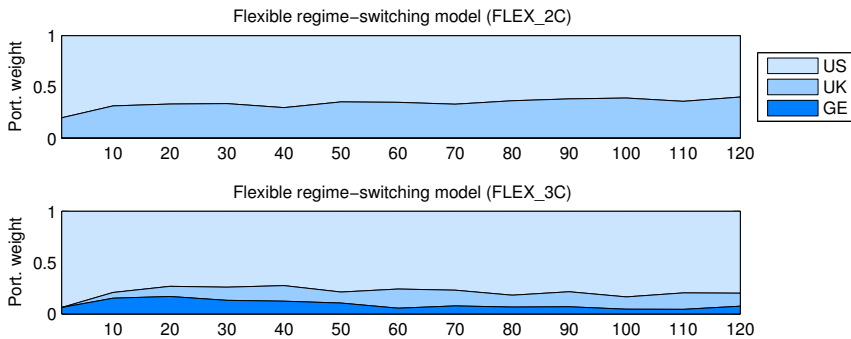


Figure 7.2: *Change of portfolio weights with increasing holding periods, covering the period from January 2004 to February 2014. Results are depicted for the core sample. Risk aversion γ is set to 5. The x-axis depicts the length of the holding period HP (in months). The investment in the risk-free assets is restricted (equity-only portfolio).*

Figure 7.2 further presents the results for FLEX_2C and FLEX_3C without allocation to the risk-free asset. As expected, allocation changed only marginally and most weights remained unaffected. FLEX_3C presented the only exception, where the weight of the risk-free asset was reallocated to the UK. This observation highlights the favorable behavior of the UK in intermediate states. Due to the previously mentioned averaging of the synchronization factor, however, the dynamics of the UK are less pronounced in FLEX_3C.

Risk Aversion

The second dimension of portfolio allocation considers the risk aversion γ . Portfolio weights were generated for values of risk aversion between 2 and 10. The holding period was set to 60 months.

Figure 7.3 provides two main insights: First, portfolios become more equally weighted with increasing γ . FLEX_2C presents the only exception. Its allocation to the US strongly increases with γ . Only at very high levels of risk aversion does the US show slightly decreasing weights. This behavior can be explained by the correlation dynamics in intermediate states. FLEX_2C accounts strongest for these effects and therefore contributes influential information that leads to more distinct allocations even at high levels of risk.

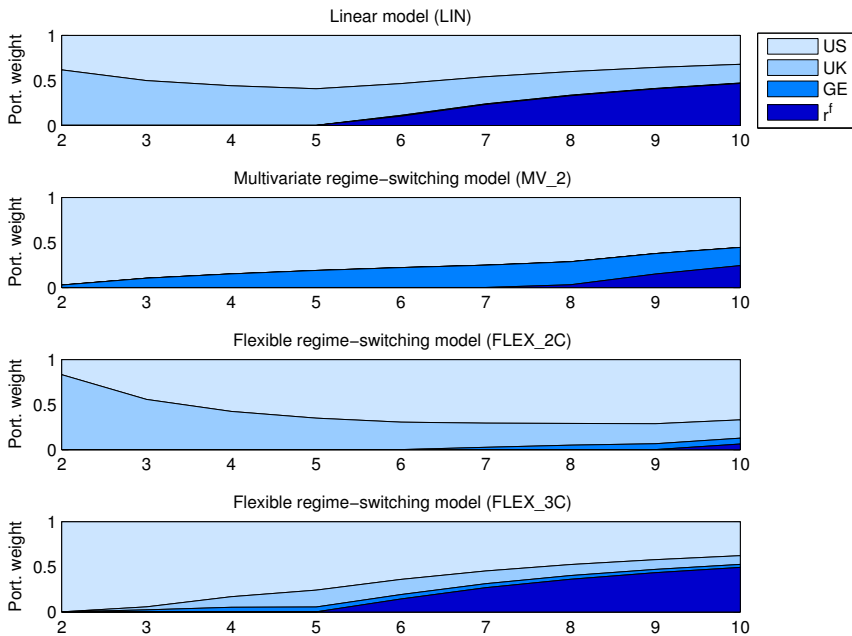


Figure 7.3: Change of portfolio weights with increasing risk aversion, covering the period from January 2004 to February 2014. Results are depicted for the core sample. The holding period HP is set to 60 months. The x-axis depicts the level of risk aversion γ .

Second, as expected, risk-free allocation increases with γ . However, even at high levels of risk aversion, MV_2 and FLEX_2C show comparably low exposures to the risk-free asset. In contrast, LIN and FLEX_3C show very similar dynamics and end up with risk-free exposures of almost 50% for $\gamma = 10$.

The initial buy-and-hold results have highlighted model dynamics with respect to different holding periods and risk aversions. FLEX_2C has thereby revealed similar allocations as the linear model. However, analysis has further shown that this similarity strongly depends on the investor's degree of risk aversion. In general, FLEX_2C exhibits very good properties and is therefore expected to perform well in the core sample.

FLEX_3C also presented appealing allocation results. However, the empirical analysis in Chapter 6 has demonstrated that the power of FLEX_3C is best observed in larger samples. The analysis of the extended sample has thereby revealed similar insights into allocation dynamics as the core sample. The next section will compare these models in terms of out-of-sample performance.

7.2.2 Performance of the Strategies

This section compares the different models with regard to their out-of-sample performance. As stated above, two particular features are thereby of interest: the relative performance of the flexible regime-switching models and the impact of composite correlation on portfolio returns.

To best capture the regime-switching dynamics in portfolio analysis, this section accounts for the buy-and-hold as well as for the rebalancing strategy. The latter follows the same dimensions as previously defined (holding period and risk aversion). Moreover, different rebalancing periods of 3, 6, and 12 months have been tested. Because results across rebalancing periods have turned out similarly, for this analysis it was set to 6 months.

As presented in Section 7.1, the model estimation relies on a Bayesian portfolio choice problem. To estimate the rebalancing strategy, appropriate grid points therefore need to be defined. For MV_2, six equally-spaced grid points were set on the

domain $[0, 1]$: $\xi_{1,rb}^j = \{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. Because of the multiple Markov chains in the flexible models, their number of grid points was slightly reduced. In FLEX_2C it was set to four: $\xi_{1,rb}^{(j,m)} = \{0, 0.33, 0.67, 1.0\}$ for $m = a, b, ab, \delta$. Likewise, in FLEX_3C it was set to three: $\xi_{1,rb}^{(j,m)} = \{0, 0.5, 1.0\}$ for $m = a, b, c, abc, \delta$. Due to the multi-period focus of the rebalancing strategy, the number of Monte Carlo simulations H was also reduced to 10'000.

Results are presented for both the core (Tables 7.1 and 7.2) and the extended sample (Tables 7.3 and 7.4). The parallel analysis of the two samples will further clarify the influence of sample size on the number of asset clusters. Moreover, it will present the optimal area of application of the different models.

Buy-and-Hold Versus Rebalancing

Tables 7.1 and 7.2 display the results for the core sample. Table 7.1 depicts performance figures for the buy-and-hold strategy, whereas Table 7.2 shows those for the rebalancing strategy. Performance is measured in terms of annualized means, annualized standard deviations, and Sharpe ratios of the corresponding portfolio returns. The underlying models remain the same as in the previous section (LIN, MV_2, FLEX_2C, and FLEX_3C).

Tables 7.1 and 7.2 point out three main findings for the core sample: First, the linear model and FLEX_2C perform best in terms of average return and in terms of volatility. Thereby, FLEX_2C demonstrates comparably good return figures for all holding periods. This outperformance is likely due to the fast implementation of the flexible regime dynamics in the asset allocation model. The previous section has shown that FLEX_2C underlies shorter regime cycle durations. Consequently, its switching dynamics should take effect in short- to medium-term investment periods (12 to 60 months). Table 7.1 illustrates this effect. However, the higher portfolio returns come at the price of higher volatilities. As a result, FLEX_2C only moderately outperforms LIN in terms of Sharpe ratio. Nevertheless, FLEX_2C presents better results than MV_2. This outperformance further confirms the power of flexible compared to standard regime-switching models.

Second, MV_2 and FLEX_3C present good risk-return-structures for very long-term investment horizons. MV_2 shows comparably high returns and low volatilities for this holding period. Due to the long duration of its underlying regime process, the regime-switching dynamics only take effect when longer holding periods are assumed. In this case, the optimization model appropriately accounts for possible switches and reallocates the underlying funds. Despite the positive risk-return dynamics of MV_2, it is outperformed by FLEX_3C. The favorable return dynamics of the FLEX_3C in longer periods are due to the structure of the synchronization factor. The regime process becomes more stable (longer durations) as synchronization is measured across all three clusters. Therefore it takes longer until the asset allocation model can account for intermediate states. Nevertheless, the observations clearly support the selection of flexible models at the account of standard multivariate regime-switching models.

Third, rebalancing has only moderate effects on the risk-adjusted performance of regime-switching models. The comparison of Tables 7.1 and 7.2 shows that the results remain stable between the buy-and-hold and the rebalancing strategies. Only MV_2 and FLEX_2C show slightly lower volatilities when rebalancing is accounted for. This might be due to the revision of expectations in case of rebalancing. Overall, however, the impact of rebalancing on risk-adjusted portfolio performance is limited in the core sample.

Figure 7.4 displays the performance of the different models over the 10-year holding period. It clearly establishes that rebalancing strategies generally outperform buy-and-hold strategies in terms of risk-unadjusted return. Further, the flexible model with rebalancing exhibits the best overall performance.

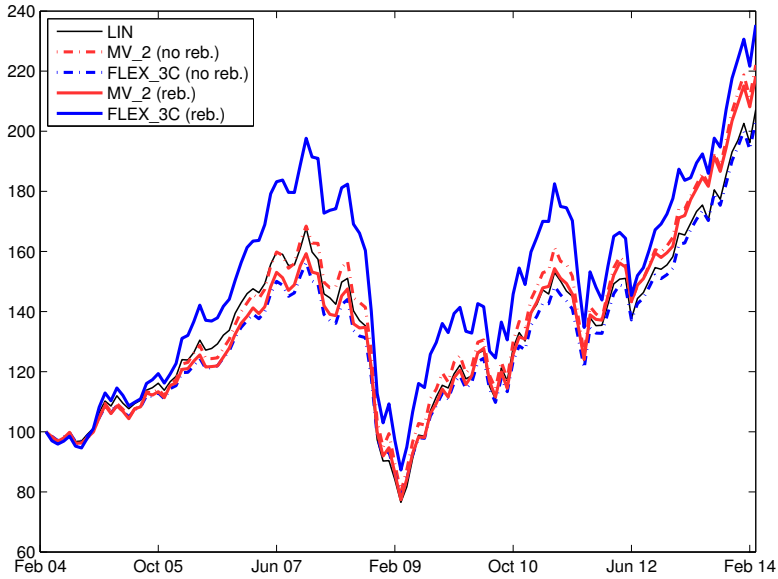


Figure 7.4: Performance comparison of the switching and non-switching models using buy-and-hold and rebalancing investment strategies. Results are depicted for the core sample. The holding period HP is set to 120 months, the risk aversion γ to 5, and the rebalancing period RB to 6 months. The period under study is from January 2004 to February 2014.

Table 7.1: Performance of the different switching and non-switching models using a buy-and-hold investment strategy (core sample). The period of analysis is from January 2004 to February 2014.

Panel A - Mean portfolio return																				
Holding per.		Gamma				Gamma				Gamma				Gamma						
		2	3	5	10	2	3	5	10	2	3	5	10	2	3	5	10			
6 months		-6.1%	-6.1%	-6.1%	-3.0%	-6.2%	-7.5%	-8.6%	-5.9%	-6.0%	-6.1%	-6.2%	-5.3%	-6.3%	-7.4%	-8.4%	-6.8%			
12 months		13.5%	12.3%	11.3%	6.1%	7.2%	8.2%	9.0%	6.0%	14.9%	12.4%	10.4%	8.3%	7.0%	7.8%	9.0%	6.5%			
36 months		14.0%	13.2%	12.5%	7.2%	10.2%	11.1%	12.0%	9.3%	15.6%	13.7%	12.2%	9.2%	9.5%	10.2%	11.7%	6.8%			
60 months		-5.4%	-5.6%	-5.6%	-2.8%	-5.9%	-5.4%	-4.8%	-3.2%	-5.2%	-5.5%	-5.7%	-5.0%	-6.1%	-5.9%	-5.5%	-2.5%			
120 months		7.3%	7.3%	7.3%	4.3%	7.6%	7.7%	8.0%	8.4%	7.2%	7.3%	7.3%	7.3%	7.4%	7.4%	7.1%	4.1%			
Panel B - Standard deviation of portfolio returns																				
Holding per.		Gamma				Gamma				Gamma				Gamma						
		2	3	5	10	2	3	5	10	2	3	5	10	2	3	5	10			
6 months		5.8%	6.0%	6.1%	3.1%	7.1%	7.2%	7.3%	4.8%	5.7%	6.0%	6.3%	5.7%	7.1%	7.1%	7.2%	5.6%			
12 months		8.8%	8.6%	8.4%	4.5%	8.3%	9.0%	9.5%	6.1%	9.0%	8.6%	8.3%	7.4%	8.2%	8.7%	9.1%	6.0%			
36 months		7.6%	7.3%	7.2%	4.0%	7.4%	7.7%	8.1%	5.9%	8.2%	7.5%	7.1%	5.5%	7.2%	7.4%	7.6%	4.2%			
60 months		16.3%	15.9%	15.6%	8.9%	15.1%	15.5%	16.1%	13.0%	17.1%	16.1%	15.5%	14.8%	14.9%	15.1%	15.5%	8.4%			
120 months		17.0%	16.5%	16.0%	8.8%	15.5%	16.0%	16.9%	18.5%	18.0%	16.8%	16.1%	16.3%	15.1%	15.2%	14.7%	7.7%			
Panel C - Sharpe ratio																				
Holding per.		Gamma				Gamma				Gamma				Gamma						
		2	3	5	10	2	3	5	10	2	3	5	10	2	3	5	10			
6 months		-1.05	-1.03	-1.00	-0.96	-0.88	-1.05	-1.19	-1.22	-1.05	-1.03	-0.98	-0.93	-0.88	-1.04	-1.16	-1.21			
12 months		1.54	1.44	1.34	1.38	0.87	0.92	0.95	0.98	1.65	1.44	1.24	1.12	0.85	0.90	0.99	1.08			
36 months		1.85	1.81	1.75	1.79	1.37	1.43	1.49	1.56	1.89	1.84	1.71	1.66	1.31	1.37	1.55	1.61			
60 months		-0.33	-0.35	-0.36	-0.31	-0.39	-0.35	-0.30	-0.24	-0.31	-0.34	-0.37	-0.34	-0.41	-0.39	-0.36	-0.30			
120 months		0.43	0.44	0.45	0.49	0.49	0.48	0.47	0.46	0.40	0.43	0.45	0.45	0.49	0.49	0.48	0.52			

Table 7.2: Performance of the different switching and non-switching models using an investment strategy with rebalancing (core sample). The period of analysis is from January 2004 to February 2014. Rebalancing frequency RB is set to 6 months.

MV_2		FLEX_2C										FLEX_3C									
		Panel A - Mean portfolio return																			
		Gamma					Gamma					Gamma									
Holding per:		2	3	5	10		2	3	5	10		2	3	5	10						
6 months		-6.2%	-7.5%	-8.6%	-5.9%		-6.0%	-6.1%	-6.2%	-5.3%		-6.3%	-7.4%	-8.4%	-6.8%						
12 months		7.0%	7.1%	9.1%	7.5%		10.6%	9.0%	7.7%	7.0%		7.0%	7.0%	7.0%	7.0%						
36 months		9.5%	16.7%	13.1%	6.5%		14.4%	14.9%	12.3%	13.6%		9.5%	0.4%	9.5%	12.1%						
60 months		-6.1%	-5.3%	-5.0%	-2.2%		-6.0%	-5.1%	-6.0%	-5.2%		-6.1%	0.4%	-4.5%	-1.2%						
120 months		7.4%	7.4%	7.8%	6.2%		7.2%	7.2%	7.2%	7.2%		8.1%	0.4%	8.6%	8.3%						
Panel B - Standard deviation of portfolio returns																					
		Gamma					Gamma					Gamma									
Holding per:		2	3	5	10		2	3	5	10		2	3	5	10						
6 months		7.1%	7.2%	7.3%	4.8%		5.7%	6.0%	6.3%	5.7%		7.1%	7.1%	7.2%	5.6%						
12 months		8.2%	8.2%	9.6%	7.4%		8.4%	8.2%	8.2%	8.2%		8.2%	8.2%	8.2%	8.2%						
36 months		7.2%	8.9%	8.6%	4.8%		7.7%	7.9%	7.1%	7.4%		7.2%	0.3%	7.2%	7.1%						
60 months		14.9%	16.6%	15.7%	10.3%		15.0%	17.8%	15.0%	16.6%		14.9%	0.3%	16.2%	20.2%						
120 months		15.1%	15.1%	15.9%	10.9%		18.8%	18.8%	18.8%	18.8%		19.7%	0.3%	19.2%	20.0%						
Panel C - Sharpe ratio																					
		Gamma					Gamma					Gamma									
Holding per:		2	3	5	10		2	3	5	10		2	3	5	10						
6 months		-0.88	-1.05	-1.19	-1.22		-1.05	-1.03	-0.98	-0.93		-0.88	-1.04	-1.16	-1.21						
12 months		0.85	0.86	0.95	1.01		1.26	1.09	0.94	0.85		0.85	0.85	0.85	0.85						
36 months		1.31	1.89	1.53	1.35		1.87	1.89	1.72	1.84		1.31	1.53	1.31	1.70						
60 months		-0.41	-0.32	-0.32	-0.22		-0.40	-0.28	-0.40	-0.32		-0.41	1.54	-0.28	-0.06						
120 months		0.49	0.49	0.49	0.57		0.38	0.38	0.38	0.38		0.41	1.54	0.45	0.41						

Tables 7.3 and 7.4 present the results for the extended sample. For the analysis of this sample, only LIN, MV_2, and FLEX_3C were considered. The results in Chapter 6 have demonstrated that two clusters would insufficiently cover this sample.

Whereas FLEX_2C and LIN showed the best mean returns in the core sample, this is no longer the case in the extended sample. FLEX_3C performs very strongly for short-, medium, and long-term holding periods and shows a similar behavior to FLEX_2C in the core sample. However, the good performance of FLEX_3C is due to its riskier investment strategies. FLEX_3C shows a higher portfolio risk and comparably low Sharpe ratios. LIN, in contrast, still presents very low portfolio risk numbers. This also results in higher Sharpe ratios.

The analysis of the rebalancing strategy shows that the risk of FLEX_3C can be reduced (see Table 7.4). However, this is mainly due to extreme allocations towards the risk-free asset and might therefore not present a valuable solution in practice. Tables 7.3 and 7.4 further confirm that the flexible model again performs best during medium- and longer-term periods. In contrast, MV_2 performs comparably low in terms of both risk and return. Only for very short periods does it present favorable results due to extreme allocations to portfolios with high volatility exposures.

Figure 7.5 charts the performance of the different strategies over the 10-year holding period. Again, the flexible model with rebalancing performed best during that time. The flexible model with the buy-and-hold investment strategy also performed comparably well. In contrast, MV_2 shows the weakest performance and ranks even behind LIN.

The previous analysis of the core and the extended sample presents two main findings for flexible regime-switching models: First, the flexible models perform best during mid- and long-term holding periods. This focus is easily explained by the structure and durations of the underlying Markov chains. Second, for shorter holding periods, the flexible models still show values comparable to MV_2. This observation demonstrates the comparative advantage of accounting for flexible instead of *a priori* defined switching dynamics.

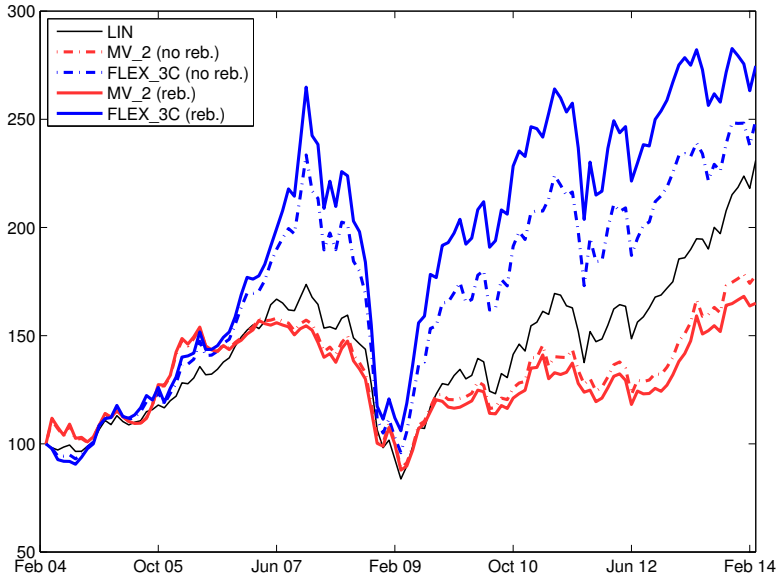


Figure 7.5: Performance comparison of the switching and non-switching models using buy-and-hold and rebalancing investment strategies. Results are depicted for the extended sample. The holding period HP is set to 120 months, the risk aversion γ to 5, and the rebalancing period RB to 6 months. The period under study is from January 2004 to February 2014.

Table 7.3: Performance of the different switching and non-switching models using a buy-and-hold investment strategy (extended sample). The period of analysis is from January 2004 to February 2014.

LIN		MV_2								FLEX_3C							
Panel A - Mean portfolio return																	
Holding per. 6 months 12 months 36 months 60 months 120 months	Gamma				Gamma				Gamma								
	2	3	5	10	2	3	5	10	2	3	5	10					
	-6.5%	-6.8%	-6.8%	-3.7%	7.7%	6.6%	5.8%	3.3%	-12.9%	-12.9%	-11.2%	-6.9%					
	13.3%	12.4%	11.6%	7.2%	13.9%	14.0%	14.0%	12.0%	16.3%	16.3%	15.4%	10.1%					
	14.7%	14.2%	13.8%	8.7%	15.2%	15.2%	15.2%	14.4%	19.2%	19.1%	17.3%	11.9%					
	-4.1%	-3.8%	-3.6%	-2.1%	-2.6%	-2.5%	-2.5%	-2.8%	1.2%	0.8%	-1.3%	-1.6%					
8.1%	8.2%	8.4%	5.5%	5.7%	5.8%	5.8%	5.8%	10.1%	9.9%	9.2%	8.7%						
Panel B - Standard deviation of portfolio returns																	
Holding per. 6 months 12 months 36 months 60 months 120 months	Gamma				Gamma				Gamma								
	2	3	5	10	2	3	5	10	2	3	5	10					
	5.8%	5.9%	6.0%	3.7%	25.4%	23.8%	22.5%	18.7%	9.9%	9.9%	8.2%	4.8%					
	9.0%	9.0%	9.0%	5.6%	17.6%	17.0%	16.5%	13.4%	12.6%	12.6%	10.9%	6.9%					
	7.8%	7.7%	7.6%	4.7%	12.5%	12.5%	12.4%	10.9%	12.4%	12.4%	10.3%	6.7%					
	15.8%	15.4%	15.1%	9.9%	16.4%	16.3%	16.3%	16.1%	23.1%	22.7%	20.0%	13.8%					
16.4%	16.0%	15.7%	9.8%	15.0%	15.0%	15.0%	14.9%	21.7%	21.4%	19.9%	17.8%						
Panel C - Sharpe ratio																	
Holding per. 6 months 12 months 36 months 60 months 120 months	Gamma				Gamma				Gamma								
	2	3	5	10	2	3	5	10	2	3	5	10					
	-1.13	-1.15	-1.13	-1.02	0.30	0.28	0.26	0.18	-1.30	-1.30	-1.36	-1.44					
	1.48	1.37	1.28	1.30	0.79	0.82	0.85	0.90	1.29	1.29	1.41	1.47					
	1.89	1.86	1.81	1.85	1.21	1.22	1.23	1.33	1.54	1.55	1.68	1.76					
	-0.26	-0.25	-0.24	-0.21	-0.16	-0.16	-0.15	-0.18	0.05	0.03	-0.07	-0.12					
0.49	0.52	0.53	0.57	0.38	0.38	0.38	0.39	0.47	0.46	0.46	0.49						

Table 7.4: Performance of the different switching and non-switching models using an investment strategy with rebalancing (extended sample). The period of analysis is from January 2004 to February 2014. Rebalancing frequency R_B is set to 6 months.

	MV_2				FLEX_3C			
Panel A - Mean portfolio return								
	Gamma				Gamma			
Holding per.	2	3	5	10	2	3	5	10
6 months	7.7%	6.6%	5.8%	3.3%	-12.9%	-12.9%	-11.2%	-6.9%
12 months	13.8%	13.9%	14.1%	13.6%	16.3%	16.3%	16.3%	16.3%
36 months	14.3%	14.9%	15.0%	14.6%	19.2%	0.4%	19.2%	19.2%
60 months	-3.2%	-2.7%	-2.6%	-2.7%	1.2%	-0.2%	1.2%	1.2%
120 months	3.9%	4.7%	5.0%	5.8%	10.1%	0.4%	10.1%	10.1%
Panel B - Standard deviation of portfolio returns								
	Gamma				Gamma			
Holding per.	2	3	5	10	2	3	5	10
6 months	25.4%	23.8%	22.5%	18.7%	9.9%	9.9%	8.2%	4.8%
12 months	19.6%	18.3%	17.2%	15.1%	12.6%	12.6%	12.6%	12.4%
36 months	14.7%	13.4%	13.1%	11.3%	12.4%	0.3%	12.4%	12.4%
60 months	17.2%	16.6%	16.4%	15.5%	23.1%	21.7%	23.1%	23.1%
120 months	15.8%	15.3%	15.2%	14.5%	21.7%	0.3%	21.7%	21.7%
Panel C - Sharpe ratio								
	Gamma				Gamma			
Holding per.	2	3	5	10	2	3	5	10
6 months	0.30	0.28	0.26	0.18	-1.30	-1.30	-1.36	-1.44
12 months	0.70	0.76	0.82	0.90	1.29	1.29	1.29	1.32
36 months	0.98	1.11	1.15	1.29	1.54	1.53	1.54	1.54
60 months	-0.19	-0.16	-0.16	-0.18	0.05	-0.01	0.05	0.05
120 months	0.25	0.31	0.33	0.40	0.47	1.54	0.47	0.47

Composite Correlation

The previous results have highlighted the out-of-sample performance of the flexible models and of their underlying investment strategies. However, these results have not disclosed whether the composite correlation structure markedly contributes to this performance.

Table 7.5 therefore presents the performance and risk figures for this restricted specification of the flexible model (core sample). Similar to the previous analysis, the correlation matrix was conditioned on the regime of cluster a . Return and risk figures thereby reveal an underperformance of this restricted model. Figure 7.6 further illustrates these observations and shows the comparative disadvantage of a restricted correlation structure. The basic specification with composite correlation matrices clearly outperforms the restricted strategies during the entire observation period. This holds for the buy-and-hold (red lines) as well as for the rebalancing strategy (blue lines). These results again highlight the content of information inherent in composite correlation matrices and their significant contribution to flexible regime-switching models.

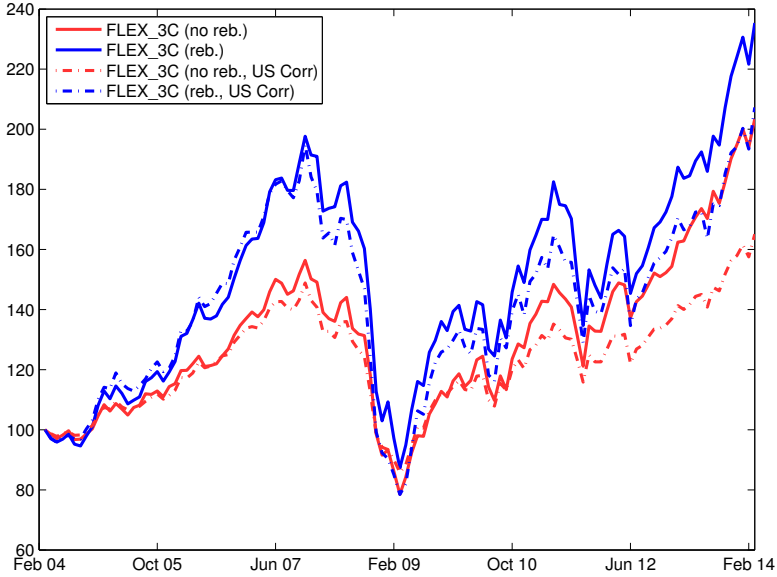


Figure 7.6: Performance comparison of the flexible regime-switching model with composite correlation matrices and with matrices that are conditioned on the regime of the US (cluster a). Results are depicted for the core sample. The holding period HP is set to 120 months, the risk aversion γ to 5, and the rebalancing period RB to 6 months. The period under study is from January 2004 to February 2014.

Table 7.5: Performance of the different flexible regime-switching models with correlation matrices that are conditioned on the regime of the US in cluster α (core sample). The period of analysis is from January 2004 to February 2014. Rebalancing frequency RB is set to 6 months.

FLEX_2C (buy-and-hold)					FLEX_3C (buy-and-hold)					FLEX_2C (rebalancing)					FLEX_3C (rebalancing)				
Panel A - Mean portfolio return																			
Holding per.		Gamma				Gamma				Gamma				Gamma					
		2	3	5	10	2	3	5	10	2	3	5	10	2	3	5	10		
6 months		-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.3%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.3%
12 months		1.3%	1.0%	0.9%	0.8%	0.9%	1.0%	1.0%	0.7%	13.4%	11.6%	10.4%	9.1%	13.4%	11.6%	10.4%	9.1%	7.0%	8.9%
36 months		1.3%	1.1%	1.0%	1.0%	1.1%	1.1%	1.1%	0.6%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	16.7%	15.1%	19.5%
60 months		-0.4%	-0.5%	-0.5%	-0.5%	-0.5%	-0.5%	-0.4%	-0.2%	-5.2%	-5.1%	-5.1%	-5.3%	-5.1%	-5.3%	-5.1%	-4.7%	-0.6%	
120 months		0.6%	0.6%	0.6%	0.6%	0.6%	0.6%	0.4%	0.2%	7.2%	7.2%	7.2%	7.2%	7.2%	7.2%	7.2%	9.5%	7.3%	
Panel B - Standard deviation of portfolio returns																			
Holding per.		Gamma				Gamma				Gamma				Gamma					
		2	3	5	10	2	3	5	10	2	3	5	10	2	3	5	10		
6 months		1.7%	1.7%	1.8%	1.9%	1.9%	1.8%	1.8%	1.1%	1.7%	1.7%	1.8%	1.9%	1.7%	1.7%	1.8%	1.8%	1.8%	1.1%
12 months		2.6%	2.5%	2.4%	2.4%	2.4%	2.5%	2.5%	1.6%	8.7%	8.5%	8.3%	8.2%	8.7%	8.5%	8.3%	8.2%	8.2%	8.2%
36 months		2.3%	2.1%	2.1%	2.0%	2.1%	2.1%	2.2%	1.2%	8.9%	8.9%	8.9%	8.9%	8.9%	8.9%	8.9%	9.6%	0.3%	12.2%
60 months		4.8%	4.6%	4.4%	4.4%	4.7%	4.7%	3.9%	2.1%	16.9%	17.8%	16.5%	17.8%	16.9%	17.8%	16.5%	17.8%	21.7%	21.1%
120 months		5.1%	4.8%	4.6%	4.7%	5.0%	5.0%	3.3%	1.7%	18.8%	18.8%	18.8%	18.8%	18.8%	18.8%	18.8%	22.8%	18.8%	18.8%
Panel C - Sharpe ratio																			
Holding per.		Gamma				Gamma				Gamma				Gamma					
		2	3	5	10	2	3	5	10	2	3	5	10	2	3	5	10		
6 months		-0.30	-0.30	-0.28	-0.27	-0.28	-0.29	-0.29	-0.29	-0.30	-0.30	-0.28	-0.27	-0.30	-0.30	-0.28	-0.27	-0.28	-0.29
12 months		0.48	0.42	0.37	0.32	0.38	0.40	0.41	0.42	1.53	1.37	1.24	1.11	1.53	1.37	1.24	1.11	0.85	0.85
36 months		0.55	0.53	0.49	0.49	0.52	0.52	0.53	0.54	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.58	1.53
60 months		-0.09	-0.10	-0.11	-0.11	-0.10	-0.10	-0.09	-0.08	-0.31	-0.28	-0.32	-0.28	-0.31	-0.28	-0.32	-0.28	-0.12	-0.01
120 months		0.12	0.13	0.13	0.14	0.12	0.12	0.13	0.13	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.42

Chapter 8

Conclusion

Multivariate regime-switching presents an efficient way of modeling the dynamics of financial markets. In standard models, however, the relationship between these dynamics is determined *a priori* and is usually assumed to be perfectly synchronized or fully independent. This thesis has developed an alternative setup. The flexible regime-switching model allows subgroups of assets to be driven by individual Markov chains. At the same time, these Markov chains underlie a dynamically changing degree of synchronization, which is inferred from the data.

In a first step, the present study introduced different regime synchronization measures. Their analysis revealed two important findings: first, the regime dynamics of individual assets show a time-varying degree of synchronization; and second, some assets show very similar dynamics and might therefore be driven by a joint process. These findings describe the key properties of the flexible regime-switching model and were also jointly responsible for its initial specification.

The flexible model efficiently covers the regime dynamics of individual assets. Its moment parameters, for example, more closely resemble the parameters of univariate models than those of multivariate models. Similarly, Chapter 7 has shown that the flexible setup provides more diversified portfolio allocations than standard regime-switching models. These observations are due to the intermediate states of the flexible model. In these states, the risk-return-structure of weakly performing assets becomes more attractive, which in turn increases their exposure. Especially medium-term performance was significantly influenced by the results of the flexible model. This period matched the duration of the assets' underlying regime processes.

Throughout this study, the empirical results have revealed a strong dependence on cluster definition. Both the number and the composition of each cluster can thereby significantly influence the parameter set of the flexible model. The clustering procedure has therefore proved to form an integral part of the model specification.

The observed dependence on cluster specification paves the way for future research. The clustering process in this thesis is not part of the model algorithm. Instead, the clusters are determined in an upstream process. The resulting cluster groups then exogenously enter the parameter set of the flexible model.

Hamilton and Owyang (2012), for example, introduced an alternative approach. In their dynamic regime-switching model, each asset is assigned to one or multiple clusters, based on a logistic variable. Unfortunately, the employed algorithm requires the covariance matrix to follow a diagonal structure. The results in the present study, however, have shown the importance of fully specified, regime-dependent covariance dynamics.

Implementing logistic variables in the flexible regime-switching model would therefore present a valuable resource for further analysis. Similarly, applying the correlation factorization algorithm to existing regime-switching models with non-synchronized regime processes also presents a promising field of future research.

Finally, given the stable dynamics of the individual clusters in the flexible model, this approach could be applied to even larger samples in future research. First, rudimentary clusters would be built, which replicate the most basic features of the extended universe. The flexible model could then be applied to each of these rudimentary clusters. These would be further divided into individual subclusters until the desired granularity is reached. As the results of the individual clusters are only weakly inter-related, additional segregation would allow one to define individual subprocesses for the refined clusters, limiting potential biases from prior estimations of parent clusters. At the same time, the whole spectrum of information in parent and subclusters could be used for the analysis of financial markets.

References

- Aggarwal, C. C. & Reddy, C. K. (2014). *Data Clustering: Algorithms and Applications*. Boca Raton, FL: CRC Press.
- Albert, J. H. & Chib, S. (1993). Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. *Journal of Business & Economic Statistics*, 11(1), 1-15.
- Ammann, M. & Verhofen, M. (2006). The Effect of Market Regimes on Style Allocation. *Financial Markets and Portfolio Management*, 20(3), 309-337.
- Andrews, D. W. (2001). Testing When a Parameter Is on the Boundary of the Maintained Hypothesis. *Econometrica*, 69(3), 683-734.
- Ang, A. & Bekaert, G. (1999). *International Asset Allocation with Time-Varying Correlations* (Working Paper No. 7056). NBER.
- Ang, A. & Bekaert, G. (2002a). International Asset Allocation With Regime Shifts. *Review of Financial Studies*, 15(4), 1137-1187.
- Ang, A. & Bekaert, G. (2002b). Regime Switches in Interest Rates. *Journal of Business & Economic Statistics*, 20(2), 163-182.
- Ang, A. & Bekaert, G. (2002c). Short Rate Nonlinearities and Regime Switches. *Journal of Economic Dynamics and Control*, 26(7-8), 1243-1274.
- Ang, A. & Bekaert, G. (2004). How Regimes Affect Asset Allocation. *Financial Analysts Journal*, 60(2), 86-99.
- Ang, A. & Chen, J. (2002). Asymmetric Correlations of Equity Portfolios. *Journal of Financial Economics*, 63(3), 443-494.
- Angeloni, A. & Sverrisson, S. (2012). *Asset Allocation in the Presence of Regime Switching in Asset Returns – Model Comparison* (Unpublished Master's thesis). Copenhagen Business School.
- Artis, M., Marcellino, M. & Proietti, T. (2004). Dating Business Cycles: A Methodological Contribution with an Application to the Euro Area. *Oxford Bulletin*

- of Economics and Statistics*, 66(4), 537-565.
- Barberis, N. (2000). Investing for the Long Run when Returns Are Predictable. *The Journal of Finance*, 55(1), 225-264.
- Barnard, J., McCulloch, R. & Meng, X.-L. (2000). Modeling Covariance Matrices in Terms of Standard Deviations and Correlations, with Application to Shrinkage. *Statistica Sinica*, 10(4), 1281-1311.
- Bekaert, G., Harvey, C. R. & Ng, A. (2005). Market Integration and Contagion. *The Journal of Business*, 78(1), 39-70.
- Bengoechea, P., Camacho, M. & Perez-Quiros, G. (2006). A Useful Tool for Forecasting the Euro-Area Business Cycle Phases. *International Journal of Forecasting*, 22(4), 735-749.
- Billio, M., Caporin, M. & Gobbo, M. (2006). Flexible Dynamic Conditional Correlation Multivariate GARCH Models for Asset Allocation. *Applied Financial Economics Letters*, 2(2), 123-130.
- Bollerslev, T. (1990). Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *The Review of Economics and Statistics*, 72(3), 498-505.
- Bordo, M. D. & Helbling, T. (2003). *Have National Business Cycles Become More Synchronized?* (Working Paper No. 10130). NBER.
- Cakmakli, C., Paap, R. & van Dijk, D. J. (2011). *Modeling and Estimation of Synchronization in Multistate Markov-Switching Models* (Tinbergen Institute Discussion Papers No. 11-002/4). Tinbergen Institute.
- Camacho, M. & Perez-Quiros, G. (2006). A New Framework to Analyze Business Cycle Synchronization. In C. Milas, P. Rothman & D. van Dijk (Eds.), *Nonlinear Time Series Analysis of Business Cycles* (Vol. 276, p. 133 - 149). Amsterdam: Elsevier.
- Candelon, B., Piplack, J. & Straetmans, S. (2008). On Measuring Synchronization of Bulls and Bears: The Case of East Asia. *Journal of Banking & Finance*, 32(6), 1022-1035.
- Candelon, B., Piplack, J. & Straetmans, S. (2009). Multivariate Business Cycle Synchronization in Small Samples. *Oxford Bulletin of Economics and Statistics*, 71(5), 715-737.

- Canova, F., Ciccarelli, M. & Ortega, E. (2007). Similarities and Convergence in G-7 Cycles. *Journal of Monetary Economics*, 54(3), 850-878.
- Cassisi, C., Montalto, P., Aliotta, M., Cannata, A. & Pulvirenti, A. (2012). Similarity Measures and Dimensionality Reduction Techniques for Time Series Data Mining. In A. Karahoca (Ed.), *Advances in Data Mining Knowledge Discovery and Applications* (p. 71-96). Rijeka: InTech.
- Chant, D. (1974). On Asymptotic Tests of Composite Hypotheses in Nonstandard Conditions. *Biometrika*, 61(2), 291-298.
- Chib, S. (1996). Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models. *Journal of Econometrics*, 75(1), 79-97.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1), 1-38.
- Diebold, F. X., Lee, J.-H. & Weibach, G. C. (1994). Regime Switching with Time-Varying Transition Probabilities. In C. Hargreaves (Ed.), *Nonstationary Time Series Analysis and Cointegration* (p. 283-302). Oxford: Oxford University Press.
- Dueker, M. J. & Sola, M. (2008). *Multivariate Markov Switching with Weighted Regime Determination: Giving France More Weight Than Finland* (Working Papers No. 2008-001). Federal Reserve Bank of St. Louis.
- Edwards, S., Biscarri, J. G. & Perez de Gracia, F. (2003). Stock Market Cycles, Financial Liberalization and Volatility. *Journal of International Money and Finance*, 22(7), 925-955.
- Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics*, 20(3), 339-350.
- Everitt, B. S., Landau, S., Leese, M. & Stahl, D. (2011). *Cluster Analysis*. Chichester: Wiley.
- Frauendorfer, K., Jacoby, U. & Schwendener, A. (2007). Regime Switching Based Portfolio Selection for Pension Funds. *Journal of Banking & Finance*, 31(8), 2265-2280.
- Garcia, R. & Perron, P. (1996). An Analysis of the Real Interest Rate Under Regime

- Shifts. *The Review of Economics and Statistics*, 78(1), 111-125.
- Geman, S. & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721-741.
- Geweke, J. & Amisano, G. (2003). *Compound Markov Mixture Models with Application in Finance* (Mimeo). NBER/NSF Time Series Conference.
- Goldfeld, S. M. & Quandt, R. E. (1973). A Markov Model for Switching Regressions. *Journal of Econometrics*, 1(1), 3-15.
- Gray, S. F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process. *Journal of Financial Economics*, 42(1), 27-62.
- Grinstead, C. M. & Snell, J. L. (2012). *Introduction to Probability*. Providence, RI: American Mathematical Society.
- Guidolin, M. (2013). *Modelling, Estimating and Forecasting Financial Data under Regime (Markov) Switching* (Working Paper). Bocconi University.
- Guidolin, M. & Timmermann, A. (2005a). *An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns* (Working Papers No. 2005-003). Federal Reserve Bank of St. Louis.
- Guidolin, M. & Timmermann, A. (2005b). *Strategic Asset Allocation and Consumption Decisions under Multivariate Regime Switching* (Australasian Meetings No. 349). Econometric Society.
- Guidolin, M. & Timmermann, A. (2006). An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns. *Journal of Applied Econometrics*, 21(1), 1-22.
- Guidolin, M. & Timmermann, A. (2007). Asset Allocation Under Multivariate Regime Switching. *Journal of Economic Dynamics and Control*, 31(11), 3503-3544.
- Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 57(2), 357-384.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Hamilton, J. D. (2005). Regime-Switching Models. In S. N. Durlauf & L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Basingtoke, Hampshire:

Palgrave Macmillan.

- Hamilton, J. D. & Owyang, M. T. (2012). The Propagation of Regional Recessions. *The Review of Economics and Statistics*, 94(4), 935-947.
- Hamilton, J. D. & Perez-Quiros, G. (1996). What Do the Leading Indicators Lead? *The Journal of Business*, 69(1), 27-49.
- Hamilton, J. D. & Susmel, R. (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime. *Journal of Econometrics*, 64(1-2), 307-333.
- Harding, D. & Pagan, A. (2002). Dissecting the Cycle: a Methodological Investigation. *Journal of Monetary Economics*, 49(2), 365-381.
- Harding, D. & Pagan, A. (2006). Synchronization of Cycles. *Journal of Econometrics*, 132(1), 59-79.
- Honda, T. (2003). Optimal Portfolio Choice for Unobservable and Regime-Switching Mean Returns. *Journal of Economic Dynamics and Control*, 28(1), 45-78.
- Kim, C.-J. (1994). Dynamic Linear Models with Markov-Switching. *Journal of Econometrics*, 60(1-2), 1-22.
- Kim, C.-J. & Nelson, C. R. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge, MA: MIT Press.
- Kose, M. A., Otrok, C. & Whiteman, C. H. (2003). International Business Cycles: World, Region, and Country-Specific Factors. *The American Economic Review*, 93(4), 1216-1239.
- Kose, M. A., Otrok, C. & Whiteman, C. H. (2008). Understanding the Evolution of World Business Cycles. *Journal of International Economics*, 75(1), 110-130.
- Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*. Berlin Heidelberg: Springer Verlag.
- Krolzig, H.-M. (2001). Markov-Switching Procedures for Dating the Euro-Zone Business Cycle. *Vierteljahrshefte zur Wirtschaftsforschung / Quarterly Journal of Economic Research*, 70(3), 339-351.
- Leippold, M. & Morger, F. (2007). International Stock Portfolios and Optimal Currency Hedging with Regime Switching. In G. N. Gregoriou (Ed.), *Asset*

- Allocation and International Investments* (p. 16-41). Basingstoke, Hampshire: Palgrave Macmillan.
- Leiva-Leon, D. (2012a). *Dynamic Synchronization of Cycles: A Markov-Switching Network Approach* (Working Paper). Bank of Canada.
- Leiva-Leon, D. (2012b). *Monitoring Synchronization of Regional Recessions: A Markov-Switching Network Approach* (Working Paper). University of Alicante.
- McCulloch, R. E. & Tsay, R. S. (1994). Bayesian Inference of Trend and Difference-Stationarity. *Econometric Theory*, 10(3-4), 596-608.
- Morger, F. (2006). *International Asset Allocation and Hidden Regime Switching* (Dissertation). University of Zürich.
- Morse, M. D. & Patel, J. M. (2007). An Efficient and Accurate Method for Evaluating Time Series Similarity. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data* (p. 569-580). New York, NY: ACM.
- Newey, W. K. (1984). A Method of Moments Interpretation of Sequential Estimators. *Economics Letters*, 14(2-3), 201-206.
- Newey, W. K. & West, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, 61(4), 631-653.
- Oates, T., Firoiu, L. & Cohen, P. R. (1999). Clustering Time Series with Hidden Markov Models and Dynamic Time Warping. In C. L. Giles & R. Sun (Eds.), *Proceedings of the IJCAI-99 Workshop on Neural, Symbolic and Reinforcement Learning Methods for Sequence Learning* (p. 17-21).
- Otranto, E. (2010). Asset Allocation Using Flexible Dynamic Correlation Models with Regime Switching. *Quantitative Finance*, 10(3), 325-338.
- Owyang, M. T., Piger, J. & Wall, H. J. (2005). Business Cycle Phases in U.S. States. *The Review of Economics and Statistics*, 87(4), 604-616.
- Paap, R., Segers, R. & van Dijk, D. (2009). Do Leading Indicators Lead Peaks More Than Troughs? *Journal of Business & Economic Statistics*, 27(4), 528-543.
- Pelletier, D. (2006). Regime Switching for Dynamic Correlations. *Journal of Econometrics*, 131(1-2), 445-473.
- Phillips, K. L. (1991). A Two-Country Model of Stochastic Output with Changes in

- Regime. *Journal of International Economics*, 31(1-2), 121-142.
- Quandt, R. E. (1958). The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes. *Journal of the American Statistical Association*, 53(284), 873-880.
- Quandt, R. E. (1972). A New Approach to Estimating Switching Regressions. *Journal of the American Statistical Association*, 67(338), 306-310.
- Ramchand, L. & Susmel, R. (1998a). Variances and Covariances of International Stock Returns: The International Capital Asset Pricing Model Revisited. *Journal of International Financial Markets, Institutions and Money*, 8(1), 39-57.
- Ramchand, L. & Susmel, R. (1998b). Volatility and Cross Correlation Across Major Stock Markets. *Journal of Empirical Finance*, 5(4), 397-416.
- Ritter, G. (2014). *Robust Cluster Analysis and Variable Selection*. Boca Raton, FL: Chapman & Hall CRC.
- Samuelson, P. A. (1969). Lifetime Portfolio Selection by Dynamic Stochastic Programming. *The Review of Economics and Statistics*, 51(3), 239-246.
- Schwendener, A. (2010). *The Estimation of Financial Markets by Means of a Regime-Switching Model* (Dissertation, No. 3795). University of St. Gallen, Bamberg: Difo-Druck.
- Smith, P. A. & Summers, P. M. (2005). How Well Do Markov Switching Models Describe Actual Business Cycles? The Case of Synchronization. *Journal of Applied Econometrics*, 20(2), 253-274.
- Spremann, K. (2007). *Finance*. München: Oldenbourg.
- Spremann, K. & Gantenbein, P. (2007). *Zinsen, Anleihen, Kredite*. München: Oldenbourg.
- Tse, Y. K. & Tsui, A. K. C. (2002). A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations. *Journal of Business & Economic Statistics*, 20(3), 351-362.
- Verhofen, M. (2006). *Bayesian Inference in Empirical Finance* (Dissertation, No. 3246). University of St. Gallen, Bamberg: Difo-Druck.
- Wu, J. (2012). *Advances in k-means Clustering: A Data Mining Thinking*. Berlin Heidelberg: Springer Verlag.

Curriculum Vitae

Raphael Vial, born on August 12, 1985, in Zug, Switzerland

Education

2010 - 2014	University of St. Gallen , Switzerland Doctoral Student in Finance
2007 - 2010	University of St. Gallen , Switzerland Master of Arts in Banking and Finance (M.A. HSG)
2004 - 2007	University of St. Gallen , Switzerland Bachelor of Arts in Business Administration (B.A. HSG)
2000 - 2004	Gymnasium Immensee , Switzerland Grammar School

Work Experience

2007 -	RCV Consulting AG , Küssnacht, Switzerland CEO
2013	EFG Capital Advisors , Miami FL, USA Portfolio Manager, Investment Division
2010 - 2013	Notenstein Private Bank , St. Gallen, Switzerland <i>former Wegelin & Co. Private Bankers</i> Quantitative Analyst, Quantitative Strategies and Research
2008	Credit Suisse Asset Management , Zürich, Switzerland Internship, Quantitative Strategies Group
2007 - 2009	Audita Treuhand & Wirtschaftsprüfung AG , St. Gallen, Switzerland