



Exploiting Common Features  
in  
Macroeconomic and Financial Data



ISBN: 978 90 361 0281 0

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **526** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

# Exploiting Common Features in Macroeconomic and Financial Data

Gebruik van gemeenschappelijke kenmerken in macroeconomische en financiële gegevens.

Thesis

to obtain the degree of Doctor  
from Erasmus University Rotterdam  
by command of the rector magnificus

Prof.dr. H.G. Schmidt

and in accordance with the decision of Doctoral Board

The public defense shall be held  
Thursday 26 January 2012 at 11:30 hours  
by  
Cem Çakmaklı  
born in Istanbul, Turkey.



## Doctoral Committee

Promoters:            Prof.dr. D.J.C. van Dijk  
                             Prof.dr. R. Paap

Other members:      Prof.dr. H.P. Boswijk  
                             Prof.dr. D. Giannone  
                             Prof.dr. S.J. Koopman  
                             Prof.dr. J.M. Maheu  
                             Prof.dr. H.K. van Dijk  
                             Prof.dr. M. Verbeek



*To My Father*





# Acknowledgements

Dear Sir/Madam

In this letter, I would like to elucidate my motivation and reasons behind my gratitude to various people.

DO NOT PANICK. This preface will not continue in such a serious mood. Rather, I would like to deliver my sincere thanks to several people for their limitless support.

Dick (van Dijk) and Richard (Paap), thank you for your guidance, and cooperation throughout my research. I am thankful to you for bringing me into the realm of Financial Econometrics and Bayesian Econometrics. I have greatly enjoyed and benefited from our conversations on research but also related to life in general. I am also indebted to Herman van Dijk for his support and guidance for my research and academic career. I am indebted to Ümit Şenesen for his longstanding support and friendship.

I am grateful to inner committee members Peter Boswijk, John Maheu and Herman van Dijk for reviewing the manuscript and providing their valuable comments and suggestions. I also would like to thank to Domenico Giannone, Siem Jan Koopman and Marno Verbeek for agreeing being in the large committee. Especially, I have greatly benefited from the conversations with John Maheu and Mark Jensen during various conferences. I have received many useful comments and suggestions from excellent colleagues that help me a lot to improve my thesis. These include Karim Abadir, Gianni Amisano, Paulo Giordani, Jim Griffin, Maria Kalli, Frank Klaessen, Gary Koop, Hashem Pesaran, Simon Potter, Rodney Strachan. I am especially indebted to Gary Koop for kindly providing me his support for Rimini Center for Economic Analysis fellowship and to Luc Bauwens for kindly providing his support for my NWO funding application. I would like to thank to Paul de Boer and Jan Brinkhuis for their friendly conversations.

The Econometric Institute and the Tinbergen Institute have provided an excellent environment for doing high-quality research. I am grateful to all of the staff members at the Econometric and Tinbergen Institutes. Special thanks go to Carien de Ruiter and Elli Hoek van Dijke (former staff member) for being so friendly and helpful.

Many thanks to my friends at EUR and TI. I think all my roommates; Ruben de Blik, Felix Lamp, Pengfei Sun, and Dennie van Dolder deserve special thanks. I apologize from all of those, especially from Felix, I locked inside the room and I forgive you for locking me inside. Pengfei I still insist that my pronunciation of English is quite understandable and Ruben thanks for the nice and deep conversations, I enjoyed a lot your accompany guys. I hope you felt the same way. Murat, thanks a lot for your cheerful jokes and conversations. Nufer, it was nice to beat (and be beaten to) you in basketball. Many thanks to Amit, Asli, Kareem, Kyle, Marcel, Nuno, Rui, Sjoerd, Umut (2), Viorel, and Vitalie for bringing more fun to my PhD. Nalan, many thanks for being there many times for listening to me. It is nice to have a friend around who is as obsessed with tiny matters as me :). Simon, in the first year of our desperate MPhil, you called me and woke me up to come to TI and to vote for you for the student council. Of course I voted for you but you were not selected :). That was a terrific start of our friendship. Please tell my greetings to ‘Gene’ and ‘Ken’. I’ll see you in Stockholm during the ceremony.

Needless to say, my deepest gratitude goes to my family. My mother Hanife, my father Mustafa, my brother Cenk and my grandma Hayriye, words are definitely not sufficient to express my gratitude to all of you. My mother, you are my role model in life, thank you for providing me your continuous trust and support. My grandma, thank you for still showing me the same care as I was two years old. My brother, life could have been much easier for you (and most probably for me) if I were to choose another path. Thanks for this unconditional, limitless support. My dear father, I hope you see me from somewhere above. You are always residing in your place reserved for you in my heart. This thesis is dedicated to you. (*Sevgili ailem, bana duyduğunuz sonsuz güven, sevgi ve desteğiniz için ne kadar teşekkür etsem azdır. İyi ki varsınız.*)

My dear wife Anil, we started to this journey with you. You have always been my engine and my source of happiness. Thank you for making me feel very lucky to have you by my side. Your endless support is in every single line of this thesis.

From Rotterdam with love

Cem Çakmaklı,

December 2011

# Contents

|          |                                                                                                      |           |
|----------|------------------------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction and Outline</b>                                                                      | <b>1</b>  |
| 1.1      | Introduction . . . . .                                                                               | 1         |
| 1.2      | Methodology . . . . .                                                                                | 4         |
| 1.3      | Outline and Summary . . . . .                                                                        | 5         |
| <b>2</b> | <b>Getting the Most Out of Macroeconomic Information for Predicting Stock Returns and Volatility</b> | <b>9</b>  |
| 2.1      | Introduction . . . . .                                                                               | 9         |
| 2.2      | Factor-augmented predictive regressions . . . . .                                                    | 13        |
| 2.3      | Data, implementation details, and benchmark models . . . . .                                         | 16        |
| 2.4      | Market timing and volatility timing . . . . .                                                        | 22        |
| 2.5      | Economic value . . . . .                                                                             | 34        |
| 2.6      | Conclusions . . . . .                                                                                | 47        |
| 2.A      | Appendices . . . . .                                                                                 | 49        |
| <b>3</b> | <b>Modeling and Estimation of Synchronization in Multistate Markov-Switching Models</b>              | <b>55</b> |
| 3.1      | Introduction . . . . .                                                                               | 55        |
| 3.2      | Model specification . . . . .                                                                        | 58        |
| 3.3      | Estimation . . . . .                                                                                 | 63        |
| 3.4      | Lead-lag effects in size-based portfolio returns . . . . .                                           | 68        |
| 3.5      | Conclusion . . . . .                                                                                 | 78        |
| 3.A      | Appendices . . . . .                                                                                 | 79        |
| <b>4</b> | <b>Measuring and Predicting Heterogeneous Recessions</b>                                             | <b>83</b> |
| 4.1      | Introduction . . . . .                                                                               | 83        |
| 4.2      | Stylized facts of US recessions . . . . .                                                            | 87        |
| 4.3      | Univariate analysis . . . . .                                                                        | 89        |

|          |                                                                         |            |
|----------|-------------------------------------------------------------------------|------------|
| 4.4      | Multivariate analysis . . . . .                                         | 102        |
| 4.5      | Severe recessions and the theory of the financial accelerator . . . . . | 111        |
| 4.6      | Conclusion . . . . .                                                    | 114        |
| 4.A      | Appendices . . . . .                                                    | 116        |
| <b>5</b> | <b>Bayesian Semiparametric Dynamic Nelson-Siegel Model</b>              | <b>125</b> |
| 5.1      | Introduction . . . . .                                                  | 125        |
| 5.2      | Model specification . . . . .                                           | 129        |
| 5.3      | Bayesian inference . . . . .                                            | 136        |
| 5.4      | Empirical results using US yield data . . . . .                         | 144        |
| 5.5      | Conclusion . . . . .                                                    | 158        |
| 5.A      | Appendices . . . . .                                                    | 160        |
|          | <b>Bibliography</b>                                                     | <b>180</b> |

# Chapter 1

## Introduction and Outline

### 1.1 Introduction

Living in the age of information, economic agents have access to a substantial amount of publicly available information on which they can base their decisions. For example, central banks require timely projections of market conditions, financial investors diversify their portfolios using their risk and return expectations, and policy makers need precise measures of current economic activity. Unfortunately, neither market conditions and financial expectations nor economic activity is precisely observed. Instead, statistical measures and predictions of those must be derived from the observed macroeconomic and financial variables. Obviously, better estimates and predictions translate into more accurate decisions. Typically, the datasets available for making predictions include several key variables postulated by economic theory. The main problem is, however, the uncertainty about which information to use, how to model it and finally, how to use it. Departing from this, modeling these datasets in a most appropriate and statistically optimal way to extract the required information and unobserved characteristics in these datasets is the central focus of this thesis.

A popular approach in practice is to employ unobserved components models to extract the required information from the observed data. These models assume an underlying process of the *unobserved* variables, denoted as states, to capture the evolution of the *observed* data that depend on these latent states, see Harvey (1990) and Durbin and Koopman (2001). Not surprisingly, these models provide a very convenient and accurate tool for extracting the unobserved key information embedded in observed macroeconomic and financial data.

One type of such key information that constitutes the major focus of economic agents is the economic and financial market conditions. Typically, economic conditions are mea-

sured using many economic and financial variables by constructing a summary measure such as a weighted average of some key variables or using statistical tools for constructing an index. In the first approach, the Conference Board in US, for example, publishes monthly coincident index for many countries based on a few key monthly indicators such as industrial production and unemployment rate, see The Conference Board (2001). In the second approach, dynamic factor models serve as a major workhorse to extract the common information in large macroeconomic and financial datasets. The Federal Reserve Banks of Chicago and Philadelphia, for example, publish such indices of economic conditions obtained from dynamic factor models. These models literally make use of tens of economic or financial variables. However, instead of using all the variables explicitly, the underlying sources of behavior of these variables are extracted in a statistically consistent way. Consequently, if the co-movement among the variables is high as is the case for many macroeconomic or financial datasets then only a few factors are sufficient for efficiently compiling the information in the dataset, see Stock and Watson (2002a) and Bai (2003).

Intuitively, as the target is to extract the economic conditions in a single index or using a few key indices, the assumption of underlying factors that lead to the observed data is a plausible one. On the other hand, economic and financial theory often postulates a factor structure on the observed data. Theoretical models of the term structure of interest rates, for example, suggests a few underlying key factors that explain the common variation in large datasets of treasury bond yields, see Duffie and Kan (1996) and Diebold and Li (2006) for example. For stock markets, the Capital Asset Pricing Model (CAPM) asserts the market portfolio index as the driving factor of stock returns.

Another type of such key information concerns the categorization of economic and financial conditions. Economic agents are often interested in finding answers of questions such as whether the economy has entered a recession or an expansion phase or whether the financial markets are in a bear regime (persistent decline in stock prices) or bull regime (persistent increase in stock prices). The answers to these questions are of crucial importance in making economic and financial decisions. One approach to obtain inference about the categorical classification of the economy and of financial markets is to use ad hoc decision rules that are based on the experts' judgmental opinions and/or definitions of the underlying phases such as recessions and expansions or bull and bear markets. The business cycle dating committee of the National Bureau of Economic Research (NBER), for example, publishes monthly recession indicators for US and a similar committee of the Center for Economic Policy Research (CEPR) publishes recession indicators for Europe. Another approach for categorization of the economic and financial data is to employ statistical data-based classification techniques, most notably, 'mixture models'. These models

enable the researcher to obtain a probabilistic inference about the state of the economy or financial markets independent of specific definitions providing a general framework for data-based clustering.

As mixture models provide a categorization of the underlying data, the resulting estimates are multicategorical variables indicating the dependence of the observation to the corresponding cluster. Accordingly, it provides a probabilistic approach for converting a continuous variable into a categorical variable. This implies that measuring the relation between the categorical variables can be imposed into mixture models to obtain inference about the unobserved states as well as the synchronization of the variables in terms of these states they embed. Indeed, a substantial part of this thesis concerns the estimation of the relation between multicategorical variables in the context of the mixture models as well as in terms of testing the potential relation between the observed multicategorical variables.

From a practical point of view, following the seminal paper of Hamilton (1989), dynamic mixture models (also called Markov switching models) constitute the major workhorse in empirical modeling of the business cycle using macroeconomic variables and in modeling the cyclical variation in financial variables. On the one hand, discrete changes observed in economic and financial data can be modeled quite accurately using these models. On the other hand, these models also exploit the dynamic evolution of the phases of the economy or financial markets; hence they convey crucial information about future economic and financial conditions.

From an econometric point of view, mixture models provide a very convenient tool for modeling non-normalities observed in economic and financial data. The fact that non-normal distributions can be modeled by means of a mixture of underlying normal distributions facilitates the estimation substantially. Closely related, mixture models provide the framework of density estimation in the sense that any distribution can be approximated by means of a number of underlying normal distributions.

This thesis consists of three parts. In the first part, we use dynamic factor models to extract the common information and indicators of macroeconomic conditions. We use these factors to make accurate (categorical) predictions of future stock returns and their volatilities. In this part, we employ a recursive classical estimation strategy of the factor-augmented models. In the second part, we use dynamic mixture models to extract the categorical information in some key macroeconomic and financial variables. Together with this we also model the information diffusion among different variables using these categorical representations. This leads us to a general framework where we can estimate the unobserved phases of macroeconomic and financial variables jointly with the lead/lag

structure between these phases of different variables in the system. In the third part of the thesis, we combine both approaches from the previous parts. We propose a specific class of dynamic factor models for modeling the dynamic evolution of a large dataset of bond yields where we estimate the density of the factors using Bayesian semiparametric techniques. These techniques involve the use of mixture models where the number of underlying mixture components and the mixture parameters are estimated together with other parameters in the model.

The remainder of this introductory chapter is as follows. Section 1.2 briefly summarizes the methods employed in this thesis. Section 1.3 provides a detailed outline of the individual chapters, together with a summary of the main theoretical and empirical findings and the main contributions of the chapters.

## 1.2 Methodology

We employ different methodologies in each chapter of this thesis depending on the research question. However, a common feature of each chapter is that these methodologies are either related to dynamic factor models or to mixture models.

A dynamic factor model is a statistical approach for modeling the common variation in a large number of time series. Once the common variation can be summarized in a few key variables, denoted as factors, then the problem of high-dimensionality reduces to the dimension of only a few key factors thereby greatly facilitating econometric inference. While these types of models date back to Sargent and Sims (1977) and Geweke and Singleton (1981), until the last decade these models could only be applied to datasets with moderate dimensions. With the increasing availability of large datasets as well as with the recent advances in computing technology there has been a renewed interest in dynamic factor models. Recent developments in theory and estimation procedures have made it possible to apply dynamic factor models also to high dimensional models. Consequently, these models have become a very popular technique for modeling the common variation in high dimensional time series, see Otrok and Whiteman (1998), Forni *et al.* (2000), Stock and Watson (2002a), Jungbacker and Koopman (2008), Mönch *et al.* (2009), among others, for example.

From a classical perspective, Stock and Watson (2002b) and Bai (2003) among others show that under mild regularity conditions related to the correlation structure, it is fairly straightforward to extract the unobserved factors in dynamic factor models. Typically, principal components extracted from the covariance or correlation matrix of the data are consistent estimates of the unobserved factors, and thus, estimation of the factors is a

matter of minutes if not seconds even in case of datasets with hundreds of variables. Even for general cases, Jungbacker and Koopman (2008) show how to obtain efficient inference using the techniques of unobserved components models.

From a Bayesian perspective, standard estimation procedures involving Markov-Chain Monte Carlo (MCMC) techniques related to unobserved components models also encompass dynamic factor models. Moreover, the modular feature of the MCMC techniques enables us to enrich the model depending on the research question. Another advantage of the Bayesian approach is that in the estimation of the factors parameter uncertainty is also taken into account. This is especially important when the model involves many parameters to estimate.

The second type of unobserved components models that is considered in this thesis are mixture models. On the one hand, mixture models provide a probabilistic approach to the classification of the data in contrast to the rule-based approaches. On the other hand, mixture models provide a very convenient and tractable way to model the non-normalities observed in macroeconomic and financial data, see, for example, Hamilton (1989) for an economic application on US business cycle and Guidolin and Timmermann (2006) for a financial application on US stock markets. When the underlying mixture components follow a normal distribution, then mixture models provide a natural approach to model the non-normal distributions by still preserving the desired normal structure. Consequently, mixture models serve as a major workhorse in the density estimation. Accordingly, Dirichlet process mixture or infinite mixture models provide a very convenient way to estimate the *density* of the variables with unknown distributions, see, for example, Jensen and Maheu (2010) for a recent financial application.

## 1.3 Outline and Summary

This thesis consists of four self-contained chapters on macroeconomic and financial applications of dynamic factor models and mixture models. While some of the chapters develop novel models to answer the research questions, some make use of existing models in answering key questions in the macro-finance interface.

The first chapter is based on Cakmakli and van Dijk (2010). This chapter documents that macroeconomic factors bear useful information for predicting monthly US excess stock returns and their volatility over the period 1980-2005. The factors are extracted from a macroeconomic dataset of 106 variables using classical inference techniques, specifically principal component analysis. These factors together with conventional financial variables such as valuation ratios are used to predict returns on a US stock market index and its

volatility using recursive predictive regressions. Additionally, before extracting factors we employ classical variable selection techniques such as Least Angle Regression (LARS), see Efron *et al.* (2004), for selection of those macroeconomic variables that are useful for return and volatility predictions. Hence, we use several types of factors that depend on the macroeconomic dataset used for factor estimation.

The performance of the resulting predictive regressions is examined using both statistical and economic evaluations. Statistical significance of the resulting predictions is analyzed using state-of-the-art statistical tests that investigate directional accuracy. Directional accuracy concerns the market and volatility timing ability of the models, i.e. whether these regressions can successfully predict the sign of the future returns and the level of the future volatility (classified as high or low). We examine the economic significance of the methodology by simulating a mean-variance investor that construct a portfolio of risky and risk-free assets using the predictive regressions. A consistent estimate of the ex-post utility perceived by the investor enables us to evaluate the economic significance of the methodology.

Results show that factor-augmented predictive regression models improve upon both benchmark models that only include valuation ratios and interest rate related variables, and possibly individual macro variables, as well as the historical average excess return. The improvements in out-of-sample forecast accuracy are both statistically and economically significant. The factor-augmented predictive regressions have superior market timing ability and volatility timing ability, while a mean-variance investor would be willing to pay an annual performance fee of several hundreds of basis points to switch from the predictions offered by the benchmark models to those of the factor-augmented models. An important reason for the superior performance of the factor-augmented predictive regressions is the stability of their forecast accuracy, whereas the benchmark models suffer from a forecast breakdown during the 1990s.

The second chapter of the thesis is based on Cakmakli *et al.* (2011). This chapter develops a Markov-Switching vector autoregressive (MS-VAR) model that allows for imperfect synchronization of cyclical regimes in multiple variables, due to phase shifts of a single common cycle. The model structure draws from the ideas in Hamilton and Perez-Quiros (1996) and Paap *et al.* (2009) and provides a general yet parsimonious framework where each of the previous models as well as the standard MS-VAR model are encompassed as special cases. This is accomplished by modeling the relation between multicategorical variables and by blending it into the MS-VAR framework. The joint one-step estimation of the model is obtained using Bayesian inference techniques.

The model has three key features: (i) the amount of phase shift can be different across regimes (as well as across variables), (ii) it allows the cycle to consist of any number of regimes, and (iii) it allows for regime-dependent volatilities and correlations. The first feature is to allow for sufficient heterogeneity among the phase shifts. This is of key importance in many applications of macroeconomics and finance. Indeed, Paap *et al.* (2009), for example, show that there is a considerable heterogeneity in the prediction ability of the US leading indicators in leading US business cycle turning points. The second feature is especially important if the number of underlying regimes in a cycle embedded in a variable is uncertain. However, allowing for any number of regimes requires modeling the relation between the multicategorical variables. Switching from binary to multicategorical variables causes some technical problems which we tackle in this chapter. The third feature is especially important for financial time series where volatility clustering as well as regime-dependent volatility together with nonlinear dependence structure are evident for many decades.

In an empirical application to monthly returns on size-based stock portfolios, a three regime model with asymmetric phase shifts and regime-dependent heteroscedasticity is found to characterize the joint distribution of returns most adequately. We show that imposing regime-dependent heteroscedasticity, allowing for more than only two regimes and allowing for asymmetric phase shifts are key ingredients of an appropriate model for this financial data. Our results on asymmetric phase shifts show that while large- and small-cap portfolios switch contemporaneously into boom and crash regimes, the large-cap portfolio leads the small-cap portfolio for switches to a moderate regime by a month.

The third chapter of thesis is constituted by a macroeconomic application using a variant of the model developed in the second chapter. Departing from the evidence in the 2007-9 recession and the ongoing mild recovery thereafter, this chapter provides empirical evidence for distinct phases of US recessions using Markov-switching models. First, using univariate analysis of monthly coincident and leading economic indicators (LEI) we show that a three regime model with a single structural break in the variance is more appropriate than other competing models for many of these variables. Interestingly, the third regime captures severe recessions contrasting the conventional view of a recovery phase as the additional third regime. Second, we conduct Markov-switching vector autoregressive models that allow for phase shifts between the cyclical regimes of LEI and industrial production index (IP). Results indicate that a three regime model with a severe recession phase explains the data better than the two regime models with only recession and expansion regimes and three regime model with an additional recovery phase. The timing of the third regime mostly corresponds with the times of substantial credit squeezes

and dramatic increases in the spreads of corporate bonds as in the recent recession of 2007-2009. These findings provide further empirical evidence for the theory of financial accelerator coined by Bernanke *et al.* (1996). There is also a substantial heterogeneity in lead times of the LEI depending on the type of recessions. The severe recession regime of the LEI leads that of IP by 6.5 months whereas for moderate recessions this lead time increases to one year.

The final chapter of the thesis combines factor models used in chapter 2 and mixture models used in chapters 3 and 4. This chapter proposes a Bayesian semiparametric dynamic factor model for modeling the evolution of US bond yields where the error distributions of the factors are modeled according to a Dirichlet process mixture. While the factor models analyzed in the first chapter concern with the classical estimation of the factors using principal components, this chapter focuses on the Bayesian estimation of the dynamic factor models cast in a general unobserved components model framework (state space models). Imposing the Dirichlet process mixture structure into the state equation enables us to estimate the factors together with the factor distribution. This semiparametric structure of the factors leads to modeling of various forms of non-normalities, such as excess kurtosis, skewness and nonlinear dependence between factors using a unified approach. An efficient and computationally tractable algorithm is implemented to obtain Bayesian inference.

We use the Bayesian semiparametric dynamic factor model structure for estimating the dynamic Nelson-Siegel model which is the leading example of the factor models for modeling the term structure of the treasury bond yields. Semiparametric structure of the factors facilitates the estimation of the non-normalities observed in large datasets of bond yields using only the factors instead of bond yields. Results show that our framework can identify two different periods with distinct characteristics. This framework is also able to capture the nonlinear dependence structure between the factors. While the relatively stable years of late 1980s and 1990s comprise the first period, the second period captures the years of severe recessions including the recessions of 1970s and 1980s and the recent recession of 2007-9 together with highly volatile periods of Federal Reserve's monetary policy experiments in the first half of 1980s.

## Chapter 2

# Getting the Most Out of Macroeconomic Information for Predicting Stock Returns and Volatility

### 2.1 Introduction

Stock return predictability remains an issue of intense debate. On the one hand, a substantial number of papers reports positive forecasting results, with the most successful predictor variables being valuation ratios such as the dividend yield, the price-earnings ratio, and the book-to-market ratio, or other financial variables such as the short-term interest rate, the yield spread, or the credit spread, see Keim and Stambaugh (1986); Campbell and Shiller (1988); Fama and French (1988); Ferson and Harvey (1991); Ang and Bekaert (2007); Campbell and Thompson (2008); Wachter and Warusawitharana (2009), among many others. On the other hand, studies such as Welch and Goyal (2008) contest this view, arguing that over long time spans, no single predictor variable can outperform the historical mean return in terms of forecast accuracy, both in-sample and out-of-sample.

Interestingly, although stock returns are assumed to be linked to business conditions, macroeconomic variables such as output growth and inflation do not seem to add any predictive power for stock returns beyond the above-mentioned valuation ratios and interest rate related variables. The few exceptions to this rule include the consumption-wealth ratio (Lettau and Ludvigson, 2001), the expenditure share on housing Piazzesi *et al.* (2007),

expenditure on durables as a fraction of its stock (Gomes *et al.*, 2009), and survey-based measures of expected business conditions (Campbell and Diebold, 2009).

Several complicating issues plague research on the predictive content of macroeconomic information for stock returns. First and foremost, numerous macroeconomic variables are available, but typically only a small number of variables is considered as possible predictors in a return regression. Presumably, the information embedded in the selected variables is more limited than the information that is actually available to an investor, who might consider a much larger set of macro variables.

Obviously, there are sound statistical reasons for limiting the number of (macro) variables in a predictive regression model. Expanding the set of predictors would exacerbate the issue of parameter estimation uncertainty, which translates into additional uncertainty in the resulting return forecasts. At the same time, with the number of predictors being restricted it becomes crucial to select the most informative variables for consideration. This essentially leads to the issue of model uncertainty, as it is not clear *a priori* which macroeconomic variables bear the most relevant information for stock returns.

Second, the relations between excess stock returns and individual predictors appear to be highly unstable, see Bossaerts and Hillion (1999); Pesaran and Timmermann (2002); Paye and Timmermann (2006); Rapach and Wohar (2006); Lettau and Van Nieuwerburgh (2008); Ravazzolo *et al.* (2008); Pettenuzzo and Timmermann (forthcoming), among others. Hence, if a given macroeconomic variable is found to be useful for forecasting stock returns over a certain period, this might very well break down at some point. Put differently, the predictive ability of individual variables strongly fluctuates over time. Note that this aggravates the model uncertainty problem in the sense that the set of variables to include in a predictive regression model is unstable.

In this paper, we apply dynamic factor models to jointly handle the issues of model uncertainty, parameter estimation uncertainty, and structural instability. Specifically, we use principal component analysis to construct a small number of factors from a large set of macro variables. These factors are then included in a predictive regression model for excess stock returns. The motivation for considering this approach is derived from macroeconomic forecasting, where various successful applications of factor models have appeared in recent years for prediction variables such as output growth and inflation, see Stock and Watson (2002b); Forni *et al.* (2003); Mönch (2008), among others. The same approach has been adopted by Ludvigson and Ng (2007) for examining the relationship between expected stock returns and volatility and by Ludvigson and Ng (2009) for analyzing the effects of macroeconomic conditions on bond risk premia. Independent,

concurrent research by Bai (2010) also demonstrates that macro factors can be exploited for improving excess stock return predictions.<sup>1</sup>

While the issue of return predictability is still disputed, there appears to be more consensus that the level of volatility is predictable to some extent. Several approaches have been developed for this purpose, ranging from GARCH and stochastic volatility models to realized volatility measures based on high-frequency returns, see Andersen *et al.* (2006) for a recent review. In addition, several studies document a relation between stock market volatility and macroeconomic conditions, with volatility being substantially higher during business cycle recessions, see Schwert (1989) and Hamilton and Lin (1996), among others. In this paper we also consider the use of macro factors extracted for predicting the volatility of monthly stock returns.

When using the factor-based approach for forecasting excess stock returns and volatility, several choices need to be made. Among others, we have to decide upon which individual variables are used in the factor construction and how many factors are included in the predictive regression model. While we adopt a standard approach for the latter issue (following previous studies in the macro forecasting literature), we examine the former issue in detail. Specifically, we consider the added value of several approaches for pre-selection of the macro variables that enter the factor construction stage, relative to the standard approach of simply including all available variables. This relates to the issue that in the standard approach the factors are taken to be the first few principal components of the macro variables, which are constructed without taking into account the purpose for which the factors are going to be used. In case the factors are used in predictive regressions as we do in the present paper, it need not necessarily be the case that these principal components contain the most relevant information for the variable that we aim to forecast, as shown by Boivin and Ng (2006) and Bai and Ng (2008), among others. The variable pre-selection procedures that we consider here explicitly take into account the predictive ability of the individual macro variables for the variables of interest, i.e. the excess stock return and volatility.

We conduct an empirical analysis of the usefulness of the factor-augmented regression models for out-of-sample prediction of monthly US excess stock returns and volatility over the period January 1980 - December 2005. In addition to the most commonly applied financial variables such as the dividend yield and the short-term interest rate, we construct

---

<sup>1</sup>We only became aware of this project while completing the current work. Several features distinguish our paper from Bai (2010). First, Bai (2010) only considers predictive regressions for excess return, while we also use macro factors for volatility. Second, we examine the possibility of pre-selection of macro variables that are used in the factor construction, while Bai (2010) does not. (She instead examines the possibility to select the most informative factors after their construction.)

factors based on a comprehensive set of 106 individual macroeconomic variables. The added value of the macro factors in the predictive regressions is assessed both in statistical as well as in economic terms. On the one hand, we consider the directional accuracy of the forecasts. For excess returns, we examine the ability of the factor-augmented predictive regressions to correctly predict its sign (market timing). For volatility, we consider the sign relative to the historical median level (volatility timing). Hit ratios, defined as the proportion of forecasts for which the actual and predicted signs match, are used to evaluate the directional accuracy of the forecasts. On the other hand, we assess the economic value of the return and volatility forecasts by using them in active mean-variance investment strategies. In addition to standard performance measures such as the Sharpe Ratio, we use a utility based metric to evaluate how much an investor would be willing to pay to use the predictions from the factor-augmented models rather than those from a benchmark model.

Our results provide convincing evidence that the factor-based approach is successful in predicting excess returns and volatility along both dimensions. First, the factor-based predictive regressions have superior market timing and volatility timing ability compared to several benchmark models that include only valuation ratios and interest rate related variables, and possibly individual macro variables. In addition, the historical average return forecast is beaten as well. Second, using the excess return and volatility forecasts in active investment strategies, we find considerable economic value of the factor-based predictive regression models. A mean-variance investor would be willing to pay an annual performance fee of several hundreds of basis points to switch from the predictions offered by the benchmark models to those of the factor-based models. The active investment strategies also outperform passive buy-and-hold benchmarks. An important finding of our analysis is that the performance of the factor-based predictive regression models is stable over time. Whereas the benchmark models, especially the model including financial variables only, suffer from a substantial deterioration in performance during the 1990s, the predictive accuracy of the factor-based models remains stable or even improves.

Several alternative approaches are available to handle the issue of model uncertainty. First, we may consider variable selection, trying to identify the most relevant individual predictors at each point in time, as in Pesaran and Timmermann (1995). This strategy has several drawbacks though. For example, possibly valuable information in the variables that are not selected for inclusion in the predictive regression is discarded completely. Additionally, selection is necessarily based on past performance, which is no guarantee for future performance, especially in light of the well-documented instability of predictive relations between returns and financial and macro variables. Second, we may apply

(Bayesian) model averaging (Avramov, 2002; Cremers, 2002), or forecast combination (Aiolfi and Favero, 2005; Timmermann, 2008; Huang and Lee, 2010; Rapach *et al.*, 2010). In principle this is an attractive way to account for model uncertainty, but still only a limited number of variables can be considered in this approach.

The paper is structured as follows. Section 2.2 develops the specification of the factor-augmented predictive regression models. Two procedures for preselecting the macro variables that are included in the factors are also discussed. Section 2.3 provides details on the data set used in the empirical analysis, various implementation issues, and the benchmark models that are used to judge the performance of the factor-augmented predictive regressions. Sections 2.4 and 2.5 present the empirical results, focusing on market timing and volatility timing ability of the predictive regressions, and on their economic value in investment strategies, respectively. Section 2.6 draws conclusions.

## 2.2 Factor-augmented predictive regressions

We consider one-period ahead forecasts of excess stock returns and volatility based on linear regression models. In this section, the model specification is described in detail in terms of predictive regressions for excess returns. Exactly the same set-up is used for (log) volatility.

In the conventional approach, the predictive regression model for excess returns takes the form

$$r_{e,t+1} = \beta_0 + \beta'_x x_t + \beta'_z z_t + \varepsilon_{t+1}, \quad (2.1)$$

where  $r_{e,t+1}$  is the excess stock return in period  $t + 1$ ,  $x_t = (x_{1t}, \dots, x_{kt})'$  is a  $(k \times 1)$  vector of macroeconomic predictor variables, and  $z_t = (z_{1t}, \dots, z_{lt})'$  is an  $(l \times 1)$  vector of financial predictor variables. Throughout it is assumed that the values of  $x_t$  and  $z_t$  are available at the end of period  $t$ .

The key issue in the empirical specification of (2.1) is the content of the vectors  $x_t$  and  $z_t$ . Multiple candidates typically are available, especially for the macroeconomic variables. Including all of them is not recommended in that case, as the resulting parameter estimation uncertainty would adversely affect forecast accuracy. As discussed in the introduction, one possibility to avoid this issue would be to apply a variable selection procedure. This typically boils down to selecting those variables that have had the strongest predictive ability for excess returns during the recent past. Alternatively, model averaging or forecast combination may be employed, constructing a weighted average of forecasts from a range of models of the form (2.1) with different variables in  $x_t$  and  $z_t$ .

In both approaches, the number of variables that can be considered is still fairly limited. Furthermore, variable selection procedures eventually completely discard the information embedded in those variables not selected for inclusion in the model.

The approach we consider in this paper rests on the assumption that the  $k$  available macroeconomic variables obey a factor structure of the form

$$x_t = \Lambda f_t + e_t, \quad (2.2)$$

where  $f_t$  is an  $(r \times 1)$  vector of common factors, with  $r$  being much smaller than  $k$ . The factors  $f_t$  are assumed to be mutually orthogonal and in decreasing order in terms of their ability to explain the variance of the macroeconomic variables  $x_t$ . The factors are unobserved but can be estimated consistently by means of principal component analysis, as discussed in Stock and Watson (2002a, 2002b), among others. Given such estimates of the factors, we consider a factor-augmented predictive regression for excess returns of the form

$$r_{e,t+1} = \beta_0 + \beta'_f f_t + \beta'_z z_t + \varepsilon_{t+1}. \quad (2.3)$$

The main advantage of this factor-based approach is that the number of macro variables  $k$  can be rather large, while still the information in all variables is exploited.

A possible drawback of the above procedure is that the factors  $f_t$  are constructed regardless of the relation between the individual predictors  $x_{it}$ ,  $i = 1, \dots, k$  and the target variable  $r_{e,t+1}$ . Intuitively, while the first few principal components capture the bulk of the variation in the macroeconomic variables  $x_t$ , it need not necessarily be the case that this is the information that is most relevant for forecasting excess returns. More accurate forecasts may be obtained by taking into account the specific target when constructing the factors. Recent studies such as Bair *et al.* (2006); Boivin and Ng (2006) and Bai and Ng (2008) show that when the variables used for factor construction are pre-selected based on the predictive ability of  $x_{it}$  for the target variable, the forecasting performance of the factor-augmented models can indeed be improved in comparison with the models that simply use all the available macro variables.

Existing procedures for constructing factors that take into account the target variable typically boil down to reducing the initial set of macroeconomic variables  $\{x_{it}\}_{i=1}^k$  by applying some sort of variable selection procedure. Here we consider the ‘thresholding’ rules advocated by Bai and Ng (2008). These are described intuitively in the remainder of this section, with full details being provided in 2.A.1.

### 2.2.1 Pre-selecting macroeconomic variables

The first variable selection procedure we consider, called ‘hard’ thresholding, is based on the individual predictive content of  $x_{it}$  for  $r_{e,t+1}$ . Specifically, for each predictor  $x_{it}$ ,  $i = 1, 2, \dots, k$ , we regress  $r_{e,t+1}$  onto  $x_{it}$  and  $z_t$ , that is, we estimate the model

$$r_{e,t+1} = \beta_0 + \beta_{x,i}x_{it} + \beta'_{z,i}z_t + \varepsilon_{i,t+1}. \quad (2.4)$$

We then construct the factors  $f_t$  by principal components as before, but only from those predictors for which the coefficient  $\beta_{x,i}$  in this regression is significantly different from zero at a given significance level  $\alpha$ .

The hard thresholding procedure evaluates each predictor in isolation, ignoring the (possibly correlated) information in other predictors. This may lead to the selection of a subset of  $x_t$  consisting of very similar macroeconomic predictors. To overcome this problem, Bai and Ng (2008) suggest a ‘soft’ thresholding rule, which does take other candidate variables into account when selecting the predictors. Specifically, the soft thresholding procedure makes use of the Least Angle Regression (LARS) methodology, originating from Efron et al. (2004), to select a limited number of informative variables out of a large group. The main idea is to add predictors one at a time, starting with the predictor that is most correlated with the target variable. This predictor’s coefficient is increased from its starting value of zero, up to the point where the residual is equally correlated with the predictor chosen initially and a second predictor. This second predictor is then added to the so-called ‘active set’, and the coefficients on both predictors in this set are now simultaneously increased, in such a way that the residual is equally correlated with these two predictors and with a third predictor, which then enters the active set. This process can be repeated until the set of candidate predictors is exhausted (provided the number of candidates  $k$  is less than the number of observations  $T$ ). In this way, a complete ranking of the candidate predictors is obtained. Alternatively, LARS can be used as a variable selection method by stopping the algorithm after a prespecified number of predictors has been selected. This procedure closely approximates the more well-known Lasso method proposed by Tibshirani (1996); see Efron et al. (2004) and Bai and Ng (2008) for a discussion of this similarity. It should be noted that the soft thresholding procedure as described above is not computationally intensive due to the availability of an efficient algorithm for LARS proposed in Efron *et al.* (2004). In fact, the soft thresholding procedure has a computational cost of the same order as ordinary least squares.

Intuitively, soft thresholding is similar to a ‘specific-to-general’ model selection procedure or ‘forward stepwise regression’, where predictors are added to the model one at

a time. The predictor that is added in a given iteration is the variable that has the maximum correlation with the residual vector from a regression of the target variable to the predictors already selected previously. The main difference is that soft thresholding performs variable selection and shrinkage of regression coefficients simultaneously. This follows from the fact that LARS effectively optimizes an objective function that can be written in terms of a penalized regression. This shrinkage property of the LARS has the potential of providing an additional advantage when the objective is to forecast excess stock returns. As shown in Campbell and Thompson (2008), shrinkage of the regression coefficients (either using theoretical considerations or econometric techniques) improves the accuracy of predictive regressions for excess returns.

## 2.3 Data, implementation details, and benchmark models

### 2.3.1 Data

We aim to predict monthly excess returns and volatility on the Standard & Poor's (S&P) 500 index. The sample period runs from January 1970 until December 2005. As discussed below, we use a ten-year moving window for model specification and estimation, such that one-month ahead forecasts are made for January 1980 onwards. The risk-free rate is proxied by the 1-month T-bill rate, for which we use the Fama-Bliss rate from CRSP. Unfortunately, this series was discontinued at the end of 2001. For the remaining years in our sample period we use the 4-week T-Bill rate measured at close on the last trading day of each month. For constructing a measure of monthly volatility, we follow the recent literature that indicates that highly accurate measures of volatility for a certain period can be obtained from higher frequency returns (see Fleming *et al.*, 2003; Andersen *et al.*, 2006, among others). This motivates us to opt for the 'realized volatility' based on daily returns. Specifically, following French *et al.* (1987) and Akgiray (1989), we measure the monthly realized volatility by

$$\sigma_t^2 = \sum_{i=1}^{n_t} (r_{i,t} - \bar{r}_t)^2 \left[ 1 + \frac{2}{n_t} \sum_{j=1}^{n_t} (n_t - j) \hat{\phi}_t^j \right], \quad (2.5)$$

where  $r_{i,t}$  denotes the daily return in day  $i$  of the month  $t$ .  $n_t$  denotes the number of trading days in month  $t$ , and  $\bar{r}$  and  $\hat{\phi}_t$  denote the mean and the first-order autocorrelation, respectively, of the daily returns in month  $t$ . The correction to the sum of squared daily

returns in (2.5) aims to account for the non-zero first-order autocorrelation in daily index returns due to non-synchronous trading of securities and other market microstructure effects (see French *et al.*, 1987; Lo and MacKinlay, 1990b, for example). In the predictive regression models, we use the natural logarithm of the volatility measure,  $\log(\sigma_t^2)$ , as dependent variable to assure positive predictions and compute the one-period ahead expectation of volatility assuming a lognormal distribution for the realized volatility measure, see Andersen *et al.* (2006).

As discussed in the previous section, we use two types of predictors, namely macroeconomic and financial variables, in the predictive regression models for excess returns and its volatility. Our selection of financial predictors  $z_t$  consists of commonly accepted variables, namely the dividend yield, the price earnings ratio, the risk free rate and its first lag and the default spread (defined as difference between the yields on Moody's Baa and Aaa rates). Valuation ratios that combine accounting measures of companies with their market valuation have a long tradition in stock return predictability and have been found to be positively correlated with future excess returns, see Fama and French (1988) and Campbell and Thompson (2008), among others. We include the risk free rate as it is also used in the construction of the excess returns and we use the default spread as a proxy for market risk. In the predictive regression model for volatility, we also include the first lag of the natural logarithm of the volatility measure to account for the strong persistence in stock return volatility.

For the macroeconomic variables  $x_t$  that are used for extracting the factors  $f_t$ , we use an updated version of the dataset of Stock and Watson (2005). The dataset consists of 127 macroeconomic variables.<sup>2</sup> We exclude four variables related to stock prices and corresponding financial ratios as well as all interest rate related variables, as we intend to separate the information in financial indicators and macroeconomic conditions. Put differently, we aim for the financial predictors and the macroeconomic factors to embed genuine, complementary information for predicting the excess returns and volatility. The remaining dataset consists of 106 macro variables, classified in the following twelve categories: output and income (17), employment and hours (29), sales (2), consumption (1), housing starts and sales (10), inventories (3), orders (5), exchange rates (5), money and credit quantity aggregates (10), price indexes (20), average hourly earnings (3) and miscellaneous (1), with the number of series in the category given in parentheses. We transform the series, whenever appropriate, to ensure stationarity by using log-levels, annual differences and annual log differences. Outliers are defined as the observations that

---

<sup>2</sup>The original dataset of Stock and Watson (2005) includes 132 variables but due to the data availability when updating the variables from their original sources we had to exclude 5 variables.

deviate from the median value of the sample by more than six interquantile ranges and in each individual series these are replaced by the median value of the previous five observations (see Stock and Watson, 2005, for details). To avoid any look-ahead bias, the macroeconomic variables are lagged by one month. By this we mean that  $x_t$  contains the values of the macro variables concerning month  $t - 1$ , which typically are released at some point during month  $t$ . Detailed description of the individual macroeconomic variables is provided in 2.A.2.

### 2.3.2 Implementation details

The empirical application of the factor-augmented predictive regression (2.3) with (2.2) requires several decisions to be made.

First, we emphasize that we use a fully recursive forecasting approach, in the sense that for predicting the excess return and volatility in period  $t + 1$ , we only use historically available information up to and including period  $t$  for all parts of the model specification, including the pre-selection of macro variables, the construction of the factors with principal component analysis, and the estimation of the coefficients in the predictive regressions. Given the existing evidence for the presence of structural breaks in the relation between excess returns and financial and macro predictors (see the references in the introduction), we use a moving window approach, with the window length fixed at ten years (120 monthly observations).<sup>3</sup> As demonstrated by Pesaran and Timmermann (2004) ignoring structural breaks can be especially costly when predicting the sign of excess returns, which is one of the aspects we focus on in our empirical analysis.

Second, a crucial issue concerns the number of factors  $r$  to be included in the predictive regression (2.3). Following Stock and Watson (2002b), among others, we use the Bayesian Information Criterion (BIC) for this purpose, where at least one and at most six factors are included. Furthermore, we generalize the model in (2.3) allowing for up to two lags of each factor to be included. To keep the computational burden at a feasible level, we do not perform an exhaustive search across all possible combinations of the first six principal components and lag structures. Instead, we assume that factors are included

---

<sup>3</sup>An alternative approach would be to consider a dynamic ‘optimal’ window length as in Pesaran and Timmermann (2002, 2007), based on statistical tests for structural breaks.

sequentially in order of importance,<sup>4</sup> while the number of lags is assumed to be the same for all included factors.<sup>5</sup>

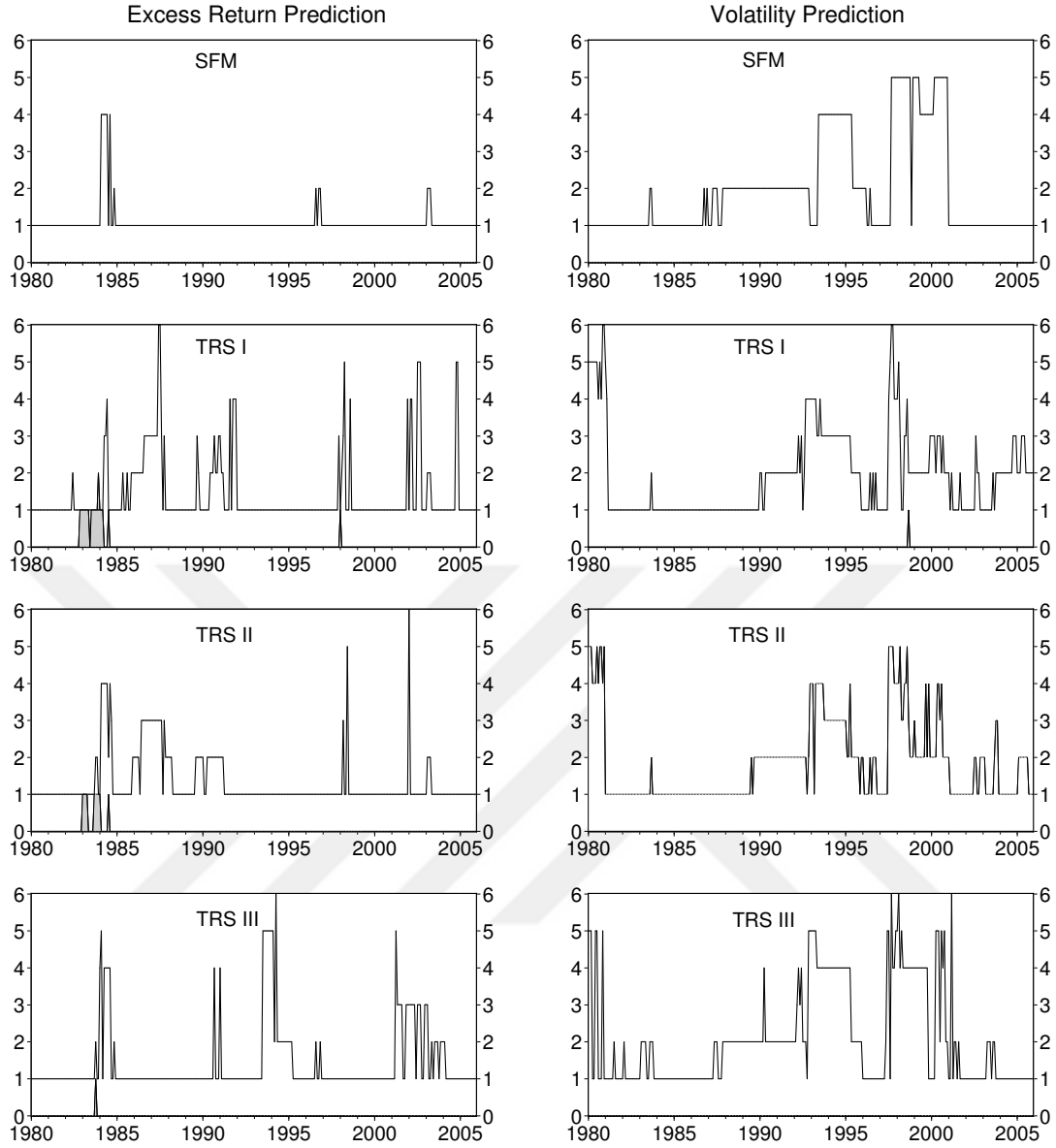
Figure 2.1 summarizes the results of this model selection procedure. The graphs on the left show that BIC selects parsimonious models for the excess return regressions. For most windows only the first principal component is included, especially when the factors are constructed from all macro variables. Multiple factors are included more often when preselection of the macro variables is applied. At first sight this finding may seem puzzling, in the sense that it intuitively less factors should be required to capture the information in a smaller set of variables. However, it should be kept in mind that the thresholding rules aim to select those macro variables that have relevant predictive information for excess returns. This should then also apply to the principal components of the selected subset of variables, such that more factors may be included. The same observations apply to the predictive regressions for volatility, although multiple factors also are included quite frequently when no preselection of the macro variables is done. Interestingly, for both returns and volatility lags of the factors are almost never included, except for the excess return regressions during a short period centered around 1980. This may be interpreted as showing that any relevant information contained in these macro variables is incorporated in stock prices and volatility within the next month.

Third, we need to specify the significance level in the hard thresholding procedure. Given the limited predictability of excess returns, we employ fairly liberal (one-sided) significance levels of 0.10 and 0.15, corresponding with threshold levels  $\alpha$  for the  $t$ -statistic of  $\beta_{x,i}$  of 1.28 and 1.04 (in absolute value), respectively, in (2.4). For these significance levels the average number of variables selected into the predictive regression for excess returns is equal to 35 and 45. For the volatility models, the corresponding numbers are 48 and 58. Figure 2.2 shows that the number of selected variables varies substantially over time. For example, for the excess return regressions and a significance level of 0.10, it ranges between 6 (for windows ending in 1982) and 79 (for some windows ending in 2001). It is also interesting to note that for the predictive regressions for volatility the number of selected variables seems to gradually decline over time. There seems to be no relation between the number of individual variables selected for the factor construction

<sup>4</sup>By this we mean that when the  $j$ -th principal component is included in the predictive regression for  $j = 2, \dots, 6$ , the first  $j - 1$  principal components are as well.

<sup>5</sup>We examine the added value of selecting the number of factors (and their lags) by examining the performance of a predictive regression model with a fixed number of factors. For this purpose, we consider the factor-augmented model that only includes the first principal component (and no lags), as this is supposed to capture the bulk of the variation in the macroeconomic variables. Results for this single-factor model are qualitatively similar to the results reported here, and are available upon request.

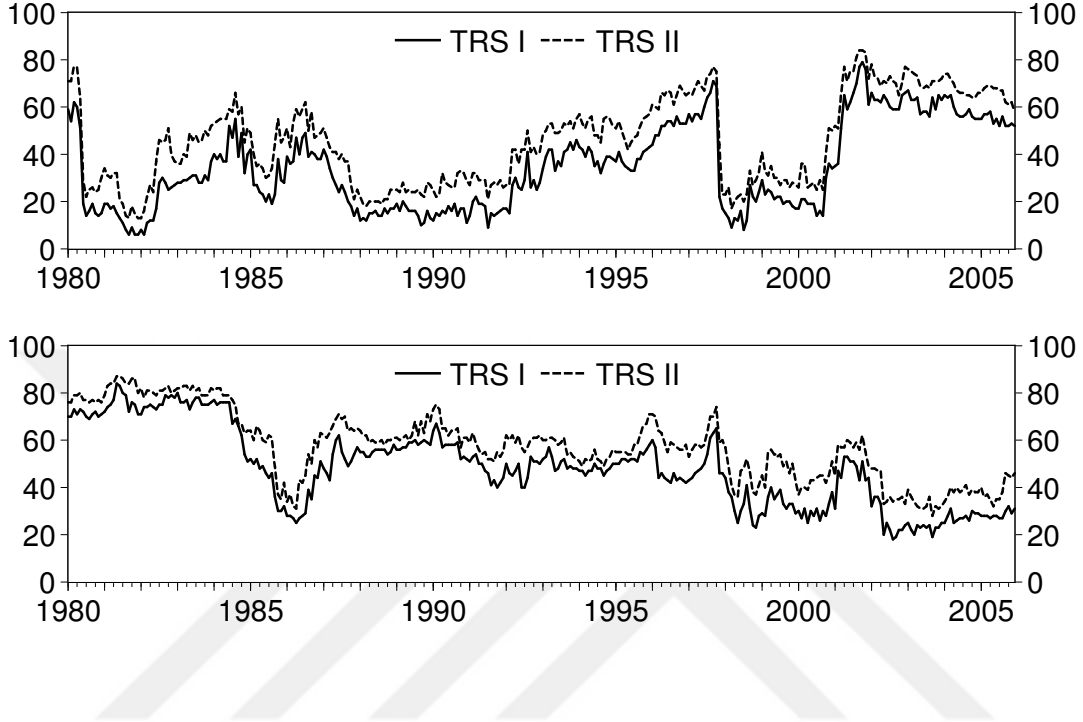
Figure 2.1: Number of factors and their lags used in each period



*Note:* The graphs show the number of factors (solid line) and the number of lags (shaded area) included in the predictive regression models for excess returns (left) and volatility (right), for the moving window ending at the date on the horizontal axis. The selection of the number of factors and their lags is made with BIC, where at least one and at most six factors are included and up to two lags of each factor. SFM denotes the ‘standard’ factor-augmented predictive regressions where factors are constructed using all available macroeconomic variables. TRS I and TRS II stand for the factor-augmented predictive regressions where factors are constructed after employing hard thresholding rules with thresholds  $|t| = 1.28$  and  $1.04$ , respectively. TRS III stands for the factor-augmented predictive regression where the factors are constructed using the first 60 predictors included in the soft thresholding procedure.

and the number of factors included in the predictive regressions, in the sense that the correlation between these two characteristics is very close to zero in all cases.

Figure 2.2: Number of factors and their lags used in each period



*Note:* The graphs show the number of macroeconomic variables that are selected using the hard thresholding approach with significance levels 0.10 (TRS I) and 0.15 (TRS II) for constructing factors that enter the predictive regressions for excess returns (top panel) and for volatility (bottom panel).

Fourth, we need to decide upon a decision rule for the soft thresholding approach for determining the subset of predictors to be included in the construction of the factors  $f_t$ . The LARS algorithm provides a complete ordering of the available macroeconomic variables according to their predictive ability. Hence, in principle it can be combined with any decision rule. An obvious approach would be to use information criteria, but in practice this results in the selection of a rather small number of predictors. Considering the limited predictability of excess returns, we select the first 60 variables that are included in the ‘active set’ in the LARS algorithm.

### 2.3.3 Benchmark models

We consider three benchmark models to evaluate the value added of the factor-augmented predictive regression models. The first benchmark model discards the information in the macroeconomic variables altogether and only includes the financial variables  $z_t$  that are

also used in the factor-augmented regressions. This benchmark enables us to observe if the macro factors indeed bear additional information for predicting the excess stock returns and volatility. The second benchmark model uses individual macroeconomic variables instead of factors. Specifically, we include the three variables with the largest  $t$ -values (in absolute value) for their coefficient in the regression (2.4) in the hard thresholding procedure.<sup>6</sup> The aim of this benchmark is to examine the benefits of using factors that contain information from a large number of macroeconomic variables relative to using the most promising individual macroeconomic predictors. The third benchmark model revisits the issue of predictive (in)ability of financial and macroeconomic predictors by using the historical average as the prediction of the excess return, see Welch and Goyal (2008). Given its strong persistence, for volatility we use a random walk forecast in this case.

## 2.4 Market timing and volatility timing

We evaluate the performance of the factor-augmented predictive regressions in two different ways. In the next section, we assess their economic value when the excess return and volatility forecasts are used in active investment strategies. In this section, we consider the quality of the forecasts in statistical terms by evaluating their directional accuracy. For excess returns, we examine the ability of the factor-augmented predictive regressions to correctly predict its sign, which we refer to as market timing. For volatility, it seems crucial to be able to predict high and low volatility regimes and therefore, following Marquering and Verbeek (2004a), we consider the sign relative to the median over the relevant 10-year moving window that is used for factor construction and model estimation, which we label volatility timing. We focus on the directional accuracy of the return and volatility predictions, as this generally is quite closely related to their economic value in terms of the performance of investment strategies based on the forecasts, see Leitch and Tanner (1991).

### 2.4.1 Methodology

We assess the market (volatility) timing ability by means of the hit ratio, defined as the proportion of months for which the sign of the excess return (volatility) is predicted

---

<sup>6</sup>In an earlier version of the paper, we used annual inflation, annual growth of industrial production and the annual growth rate of money supply, as in Pesaran and Timmermann (1995). This approach is obviously not as flexible as the current approach as it uses same variables each period. Results for this alternative specification are available upon request.

correctly. We formally test the null hypothesis of no market timing ability by means of the statistic proposed in Pesaran and Timmermann (1992). This generalization of the Henriksson-Merton (HM) statistic tests whether the empirical hit ratio is significantly higher than the expected hit ratio under independence of the signs of the actual value and the forecast. For excess returns, the statistic is defined as

$$PT = \frac{\hat{P} - \hat{P}_*}{\sqrt{\hat{V}(\hat{P}) - \hat{V}(\hat{P}_*)}} \quad (2.6)$$

where  $\hat{P}$  is the actual hit ratio  $\hat{P} = \sum_{t=0}^{n-1} I_{\{r_{e,t+1}\hat{r}_{e,t+1}\}}/n$  with  $I_{\{A\}}$  being an indicator function that takes the value 1 if the argument  $A$  (in this case the product of the actual excess return  $r_{e,t+1}$  and the forecast  $\hat{r}_{e,t+1}$ ) is positive and 0 otherwise, and  $n$  is the number of months in the forecast period. Similarly  $\hat{P}_*$  is the expected hit ratio under independence, computed as  $P_* = \hat{P}_r \hat{P}_{\hat{r}} + (1 - \hat{P}_r)(1 - \hat{P}_{\hat{r}})$  where  $\hat{P}_r = \sum_{t=1}^n I_{\{r_{e,t+1}\}}/n$  and  $\hat{P}_{\hat{r}} = \sum_{t=1}^n I_{\{\hat{r}_{e,t+1}\}}/n$  are the proportion of months for which the actual returns and the predicted returns, respectively, are positive. Finally,  $V(\hat{P}_*)$  and  $V(\hat{P})$  are estimates of the variances of  $\hat{P}_*$  and  $\hat{P}$ . Under the null hypothesis of no market timing ability, the PT statistic is standard normally distributed asymptotically.

The PT test relies upon the assumption that the indicator functions  $I_{\{r_{e,t+1}\}}$  and  $I_{\{\hat{r}_{e,t+1}\}}$  do not exhibit serial dependence. For the monthly excess returns this assumption may hold, but given the extensive evidence on volatility clustering it is likely to be violated for volatility. We assess this issue in the following way. The *PT* statistic as defined in (2.6) is asymptotically equivalent to a (one-sided) test of the significance of the slope coefficient in the regression

$$I_{\{r_{e,t+1}\}} = a + bI_{\{\hat{r}_{e,t+1}\}} + \eta_t, \quad (2.7)$$

see Breen *et al.* (1989). We examine the significance of the slope coefficient  $b$  by means of its  $t$ -statistic based on heteroscedasticity and autocorrelation consistent (HAC) standard errors to address the relevance of serial correlation in the indicator functions. The resulting test is denoted as BGJ-HAC.

In addition to the separate evaluation of the directional accuracy of the excess return and volatility forecasts, we also consider the market and volatility timing ability of the predictive regression models jointly. For this purpose, we employ a generalization of the nonparametric HM test proposed by Marquering and Verbeek (2004b) allowing for four categories, namely combinations of positive or negative excess return with high or low volatility. More explicitly, the test is based on the following contingency table, where  $\xi_t$

denotes the median level of volatility during the 10-year moving window ending in month  $t$ :

|                                                           | $r_{e,t+1} \leq 0$<br>$\sigma_{t+1}^2 \leq \xi_t$ | $r_{e,t+1} \leq 0$<br>$\sigma_{t+1}^2 > \xi_t$ | $r_{e,t+1} > 0$<br>$\sigma_{t+1}^2 \leq \xi_t$ | $r_{e,t+1} > 0$<br>$\sigma_{t+1}^2 > \xi_t$ | Total     |
|-----------------------------------------------------------|---------------------------------------------------|------------------------------------------------|------------------------------------------------|---------------------------------------------|-----------|
| $\hat{r}_{e,t+1} \leq 0, \hat{\sigma}_{t+1}^2 \leq \xi_t$ | $n_{1,1}$                                         | $n_{1,2}$                                      | $n_{1,3}$                                      | $n_{1,4}$                                   | $n_{1,0}$ |
| $\hat{r}_{e,t+1} \leq 0, \hat{\sigma}_{t+1}^2 > \xi_t$    | $n_{2,1}$                                         | $n_{2,2}$                                      | $n_{2,3}$                                      | $n_{2,4}$                                   | $n_{2,0}$ |
| $\hat{r}_{e,t+1} > 0, \hat{\sigma}_{t+1}^2 \leq \xi_t$    | $n_{3,1}$                                         | $n_{3,2}$                                      | $n_{3,3}$                                      | $n_{3,4}$                                   | $n_{3,0}$ |
| $\hat{r}_{e,t+1} > 0, \hat{\sigma}_{t+1}^2 > \xi_t$       | $n_{4,1}$                                         | $n_{4,2}$                                      | $n_{4,3}$                                      | $n_{4,4}$                                   | $n_{4,0}$ |
| Total                                                     | $n_{0,1}$                                         | $n_{0,2}$                                      | $n_{0,3}$                                      | $n_{0,4}$                                   | $n$       |

Given the above contingency table the HM test statistic takes the form

$$HM = \sum_{i,j=1}^4 \left[ \frac{\left( n_{i,j} - \frac{n_{i,0}n_{0,j}}{n} \right)^2}{\frac{n_{i,0}n_{0,j}}{n}} \right], \quad (2.8)$$

Under the null hypothesis of no joint market timing ability, the test statistic follows a  $\chi^2$  distribution with nine degrees of freedom.

Like the PT test in (2.6), the HM statistic in (2.8) assumes the absence of serial correlation in the signs of the actual and predicted excess returns and volatility. As discussed before, this is a rather unrealistic assumption for volatility in particular, invalidating the HM test. We address this issue by means of the testing framework recently proposed by Pesaran and Timmermann (2009). Pesaran and Timmermann (2009) develop tests of independence for multicategorical variables based on canonical correlations from dynamically augmented reduced rank regressions, which are robust in the presence of serial dependencies. The crucial ingredient of this approach is to define  $(3 \times 1)$  vectors  $y_{t+1} = (y_{1,t+1}, y_{2,t+1}, y_{3,t+1})'$  and  $\hat{y}_{t+1} = (\hat{y}_{1,t+1}, \hat{y}_{2,t+1}, \hat{y}_{3,t+1})'$  representing the occurrence of the different categories. Specifically, in line with the contingency table given above, define  $y_{i,t+1}$ ,  $i = 1, 2, 3$  as

$$\begin{aligned} y_{1,t+1} &= 1 \text{ if } r_{e,t+1} \leq 0 \text{ and } \sigma_{t+1}^2 \leq \xi_t, \text{ and } =0 \text{ otherwise,} \\ y_{2,t+1} &= 1 \text{ if } r_{e,t+1} \leq 0 \text{ and } \sigma_{t+1}^2 > \xi_t, \text{ and } =0 \text{ otherwise,} \\ y_{3,t+1} &= 1 \text{ if } r_{e,t+1} > 0 \text{ and } \sigma_{t+1}^2 \leq \xi_t, \text{ and } =0 \text{ otherwise,} \end{aligned}$$

and  $\hat{y}_{i,t+1}$ ,  $i = 1, 2, 3$ , similarly but based on the predictions  $\hat{r}_{e,t+1}$  and  $\hat{\sigma}_{t+1}^2$ . As shown by Pesaran and Timmermann (2009), in case the multicategorical variables are serially independent, a test of independence can be based on the canonical correlation coefficients

between  $y_{t+1}$  and  $\hat{y}_{t+1}$ . We may then consider both a maximum canonical correlation test and a trace test based on the average canonical correlation. The latter in fact is identical to the standard Pearson chi-square contingency table test of independence, as given in (2.8). In the case of serially dependent outcomes, a valid test of independence can be constructed using canonical correlations of suitably filtered versions of  $y_{t+1}$  and  $\hat{y}_{t+1}$ .

The tests of Pesaran and Timmermann (2009) can also be obtained within a regression framework. Specifically, in the absence of serial dependence a test of independence between the actual and the predicted signs of the excess returns and volatility corresponds with testing  $\gamma = 0$  in the regression

$$\theta' y_{t+1} = c + \gamma' \hat{y}_{t+1} + \eta_t, \quad (2.9)$$

where  $\theta$  is a vector of weights for the different categories of the actual signs. Note that in case of two categories, this reduces to the regression in (2.7). In the multicategory case, the null hypothesis  $\gamma = 0$  may be tested straightforwardly by means of an  $F$ -test, but the result obviously depends on the values of the ‘nuisance parameters’  $\theta$ . A standard approach in this case is to take the maximum  $F$ -test over all possible choices of  $\theta$ , which is identical to testing the null hypothesis that the largest canonical correlation between  $y_{t+1}$  and  $\hat{y}_{t+1}$  is equal to zero. Note that the largest canonical correlation between  $y_{t+1}$  and  $\hat{y}_{t+1}$  being equal to zero when the actual and predicted signs of the excess returns and volatility are independent implies that in fact all three canonical correlations between  $y_{t+1}$  and  $\hat{y}_{t+1}$  are zero in this case. This leads to the idea of the trace statistic, which exactly tests this hypothesis. The trace statistic also corresponds with testing the null hypothesis that the rank of the matrix  $\Pi$  is equal to 0 in the reduced rank regression

$$y_{t+1} = a + \Pi \hat{y}_{t+1} + \eta_t. \quad (2.10)$$

As shown by Pesaran and Timmermann (2009), in the case of serial dependence in the actual and predicted signs, robust tests of independence may be obtained by augmenting the regressions (2.9) and (2.10) by an appropriate number of lagged values of  $y$  and  $\hat{y}$ . Alternatively, the serial correlation may be accounted for by using a HAC covariance matrix. In our empirical analysis, we employ the trace statistic based on (2.10) with a HAC covariance matrix with Bartlett weights, using the Andrews (1991) procedure for automatic lag length selection. As discussed in Pesaran and Timmermann (2009) this test statistics is asymptotically chi-squared distributed with nine degrees of freedom, but this may be a poor approximation to the finite-sample distribution given the discrete nature

of  $y_{t+1}$  and  $\hat{y}_{t+1}$  in (2.10). For this reason, we employ simulation to obtain the appropriate critical values for the sample size in our empirical application.

## 2.4.2 Empirical results

Table 2.1 displays hit ratios for the excess return and volatility forecasts for all models considered. We show results for the complete out-of-sample period from January 1980 until December 2005, as well as for subperiods spanning the first and second half. Panels A and B of the table show hit ratios for excess returns and volatility individually. The bottom panel displays hit ratios for joint market timing and volatility timing, i.e. the numbers shown are the proportion of months for which the signs of excess returns and volatility both are predicted correctly. The corresponding test statistics for evaluating the significance of the market and volatility timing are shown in Table 2.2.

Table 2.1: Market timing and volatility timing

|                                                           | SFM   | TRS I | TRS II | TRS III | BM I  | BM II | BM III |
|-----------------------------------------------------------|-------|-------|--------|---------|-------|-------|--------|
| <u>Panel A: Market timing</u>                             |       |       |        |         |       |       |        |
| Jan 1980 - Dec 2005                                       | 0.625 | 0.587 | 0.574  | 0.596   | 0.580 | 0.526 | 0.592  |
| Jan 1980 - Dec 1992                                       | 0.622 | 0.558 | 0.558  | 0.590   | 0.628 | 0.532 | 0.577  |
| Jan 1993 - Dec 2005                                       | 0.628 | 0.615 | 0.590  | 0.603   | 0.532 | 0.519 | 0.609  |
| <u>Panel B: Volatility timing</u>                         |       |       |        |         |       |       |        |
| Jan 1980 - Dec 2005                                       | 0.721 | 0.724 | 0.728  | 0.728   | 0.712 | 0.718 | 0.731  |
| Jan 1980 - Dec 1992                                       | 0.622 | 0.635 | 0.641  | 0.615   | 0.628 | 0.647 | 0.628  |
| Jan 1993 - Dec 2005                                       | 0.821 | 0.814 | 0.814  | 0.840   | 0.795 | 0.788 | 0.833  |
| <u>Panel C: Joint market timing and volatility timing</u> |       |       |        |         |       |       |        |
| Jan 1980 - Dec 2005                                       | 0.455 | 0.429 | 0.420  | 0.446   | 0.407 | 0.385 | 0.442  |
| Jan 1980 - Dec 1992                                       | 0.397 | 0.346 | 0.359  | 0.391   | 0.397 | 0.346 | 0.359  |
| Jan 1993 - Dec 2005                                       | 0.513 | 0.513 | 0.481  | 0.500   | 0.417 | 0.423 | 0.526  |

*Note:* Panel A (market timing) shows hit ratios for the sign of the monthly excess stock return. Panel B (volatility timing) shows hit ratios for the sign of monthly volatility relative to the median volatility level during the 10-year rolling window that is used for model specification and estimation. Panel C (joint market timing and volatility timing) shows hit ratios for the signs of the monthly excess stock return and volatility. SFM denotes the ‘standard’ factor-augmented predictive regressions where factors are constructed using all available macroeconomic variables. TRS I and TRS II stand for the factor-augmented predictive regressions where factors are constructed after employing hard thresholding rules with thresholds  $|t| = 1.28$  and  $1.04$ , respectively. TRS III stands for the factor-augmented predictive regression where the factors are constructed using the first 60 predictors included in the soft thresholding procedure. BM I denotes the benchmark model that uses only financial variables  $z_t$ , BM II denotes the model that uses the three individual macroeconomic predictors with the highest  $t$  statistics in the regression (2.4). BM III denotes the benchmark model where the historical average is used as the excess return forecast and a random walk forecast is used for volatility.

Panel A of Table 2.1 reveals several interesting results concerning the market timing ability of the factor-augmented predictive regressions. First, over the complete out-of-sample period the factor-augmented models achieve hit ratios between 57.4% and 62.5%, which all are significantly higher than the expected hit ratio under independence according to the PT test and the HAC  $t$ -statistic of the slope coefficient in (2.7), see Table 2.2. It is also interesting to note that in most cases the values of the PT test and the BGJ-HAC statistic are close, which corresponds to the lack of serial dependence in the signs of the actual and predicted excess returns.

The hit ratios achieved by the factor-augmented models exceed those of benchmark models I and II, with the single exception that the model using the strictest hard thresholding rule (TRS II) has a slightly lower hit ratio than the first benchmark model. This finding suggests that macroeconomic variables do bear relevant additional information for market timing, relative to the valuation ratios and interest rate related variables considered here. Moreover, it is better to extract the information from a large panel of macroeconomic variables by means of factors than to include a small number of specific macro variables as in the second benchmark model. In fact, the low hit ratio of this model (also compared to the first benchmark model) suggests that including individual macroeconomic variable actually deteriorates market timing ability. Relative to the third benchmark model, the performance of the factor-augmented predictive regressions seems less convincing, with only the standard factor model achieving a clearly higher hit ratio. However, to a large extent the market timing ability of the third benchmark is spurious, in the sense that only four percent of the excess return predictions from this model are negative. Hence, an investor using this approach would almost always invest in stocks without trying to time the market. This is confirmed by the test results in Table 2.2, which shows that the third benchmark model does not have significant market timing ability. In fact, for the second sub-period the test statistics cannot be computed due to the fact that after 1992 all excess return predictions from this model are positive.

Second, comparing the predictions from the different factor-augmented models among each other, it is seen that the ‘standard’ approach of including all available macroeconomic variables for extracting the factors at each point in time performs best. Apparently, market timing ability is not enhanced by pre-selecting the (presumably) most informative predictors which the other three methods aim to achieve.

Third, the market timing ability of the factor-augmented predictive regressions seems to be fairly stable over time, in the sense that the hit ratios for the first and second half of the out-of-sample period do not differ dramatically. In fact, the hit ratios are higher for the period 1993-2005 for all four approaches. The PT and BGJ-HAC results in Table 2.2

Table 2.2: Tests for market timing and volatility timing

|                                             | SFM      | TRS I    | TRS II   | TRS III  | BM I     | BM II    | BM III   |
|---------------------------------------------|----------|----------|----------|----------|----------|----------|----------|
| Panel A: Market timing                      |          |          |          |          |          |          |          |
| <u>PT</u>                                   |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 3.46***  | 2.73***  | 2.14**   | 2.66***  | 2.00**   | 0.63     | 0.12     |
| Jan 1980 - Dec 1992                         | 2.50***  | 1.62*    | 1.49*    | 1.86**   | 2.73***  | 0.91     | 0.05     |
| Jan 1993 - Dec 2005                         | 2.38***  | 2.24**   | 1.50*    | 1.89**   | 0.11     | -0.08    | NA       |
| <u>BGJ-HAC</u>                              |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 3.12***  | 2.84***  | 2.08**   | 2.70***  | 1.82**   | 0.62     | 0.10     |
| Jan 1980 - Dec 1992                         | 2.47***  | 1.74**   | 1.55*    | 1.94**   | 2.42***  | 0.88     | 0.00     |
| Jan 1993 - Dec 2005                         | 2.05**   | 2.17**   | 1.30*    | 1.84**   | 0.10     | 0.10     | NA       |
| <u>BH-HAC</u>                               |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 2.60***  | 2.40***  | 2.24**   | 3.24***  | 2.02**   | 1.70**   | -0.81    |
| Jan 1980 - Dec 1992                         | 2.76***  | 2.27**   | 2.05**   | 2.88***  | 3.07***  | 1.99**   | -0.56    |
| Jan 1993 - Dec 2005                         | 0.60     | 0.95     | 1.00     | 1.47*    | -0.76    | 0.12     | NA       |
| Panel B: Volatility timing                  |          |          |          |          |          |          |          |
| <u>PT</u>                                   |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 8.11***  | 8.16***  | 8.32***  | 8.32***  | 7.80***  | 7.98***  | 8.17***  |
| Jan 1980 - Dec 1992                         | 3.40***  | 3.71***  | 3.98***  | 3.25***  | 3.54***  | 3.95***  | 3.19***  |
| Jan 1993 - Dec 2005                         | 8.11***  | 7.83***  | 7.91***  | 8.56***  | 7.48***  | 7.29***  | 8.34***  |
| <u>BGJ-HAC</u>                              |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 4.65***  | 4.67***  | 4.70***  | 4.65***  | 4.48***  | 4.67***  | 4.79***  |
| Jan 1980 - Dec 1992                         | 2.53***  | 2.72***  | 2.82***  | 2.39***  | 2.87***  | 2.87***  | 2.24**   |
| Jan 1993 - Dec 2005                         | 4.17***  | 4.15***  | 4.10***  | 4.33***  | 3.94***  | 3.95***  | 4.61***  |
| <u>BH-HAC</u>                               |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 3.47***  | 3.22***  | 3.27***  | 3.38***  | 4.13***  | 3.07***  | 2.34***  |
| Jan 1980 - Dec 1992                         | 1.69**   | 1.59*    | 1.50*    | 1.56*    | 1.71**   | 1.62*    | 4.80***  |
| Jan 1993 - Dec 2005                         | 5.64***  | 5.26***  | 5.48***  | 5.69***  | 4.60***  | 4.62***  | 11.84*** |
| Panel C: Joint market and volatility timing |          |          |          |          |          |          |          |
| <u>HM</u>                                   |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 81.91*** | 75.75*** | 76.97*** | 81.36*** | 70.46*** | 66.53*** | 84.43*** |
| Jan 1980 - Dec 1992                         | 27.08*** | 20.75**  | 24.00*** | 27.78*** | 37.71*** | 22.87*** | 28.20*** |
| Jan 1993 - Dec 2005                         | 75.43*** | 70.36*** | 68.19*** | 77.75*** | 60.61*** | 62.45*** | NA       |
| <u>Trace CC</u>                             |          |          |          |          |          |          |          |
| Jan 1980 - Dec 2005                         | 30.28*** | 26.64*** | 26.77*** | 27.63*** | 25.48*** | 23.05*** | 33.77*** |
| Jan 1980 - Dec 1992                         | 18.04**  | 11.31    | 14.09    | 22.74*** | 21.77*** | 11.72    | NA       |
| Jan 1993 - Dec 2005                         | 21.16**  | 22.06*** | 19.16**  | 19.82**  | 17.32**  | 20.26**  | NA       |

*Note:* PT is the test statistic for market (volatility) timing ability as defined in (2.6). BGJ-HAC is the HAC  $t$ -statistic of the slope coefficient in (2.7). BH-HAC is the HAC  $t$ -statistic of the slope coefficient in (2.11). HM is the generalized Henriksson-Merton statistic for joint market and volatility timing in (2.8). The Trace CC statistic of Pesaran and Timmermann (2009) uses the sum of canonical correlations for testing market and volatility timing jointly. NA indicates that the test could not be computed. \*, \*\*, \*\*\* indicate significance at the 10%, 5% and 1% level, respectively. Model abbreviations are as in Table 2.1.

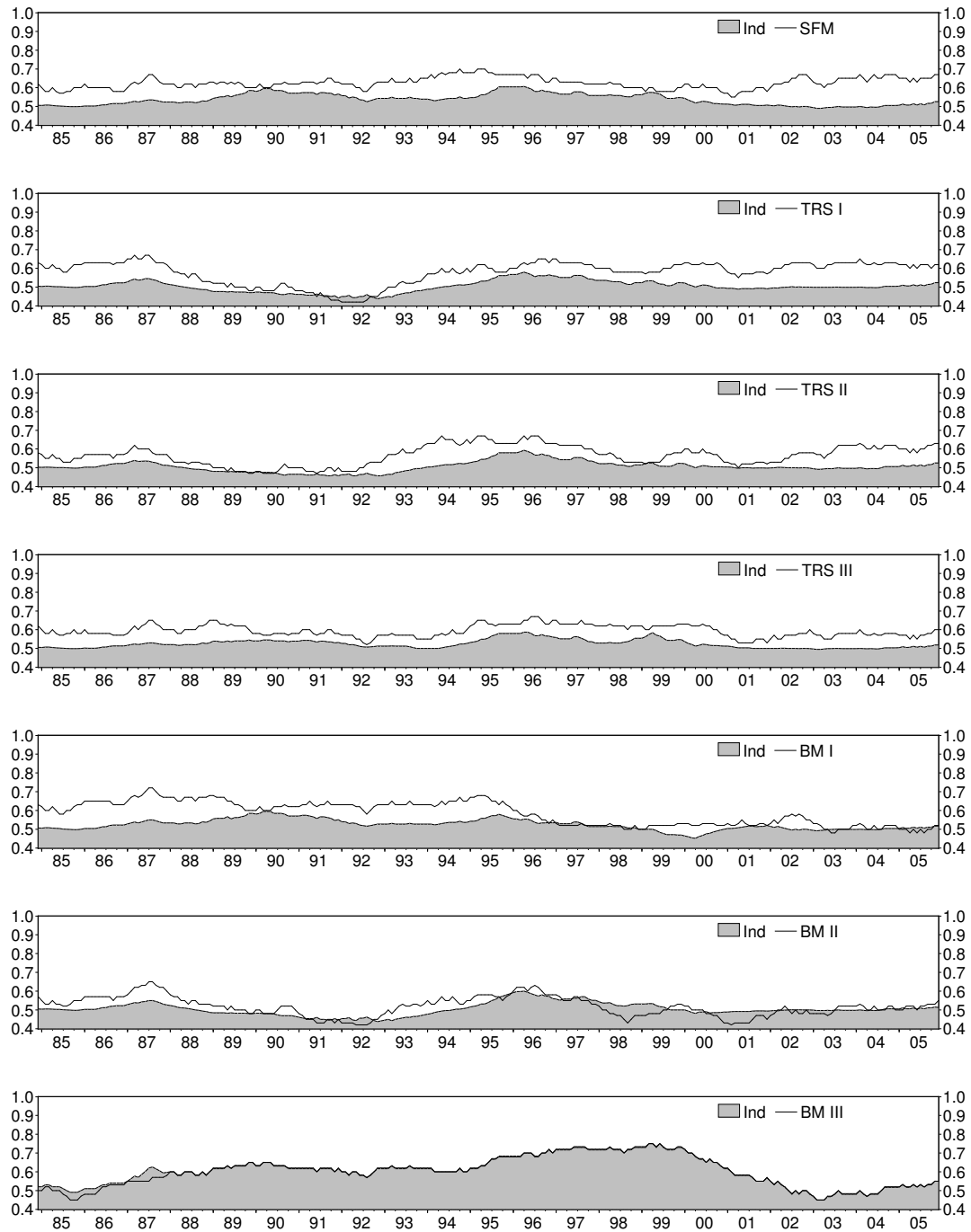
also show that the significance of the market timing ability does not differ across the two sub-periods. This finding is in sharp contrast to existing evidence, which suggests that the market timing ability of typical predictive regressions has reduced considerably since the early 1990s. This is indeed confirmed by our results for the first benchmark model, which achieves the highest hit ratio of almost 63% during the period from 1980-1992, but then suffers a dramatic decline by ten percentage points to only 53% in the period 1993-2005. The corresponding PT and BGJ-HAC statistics in Table 2.2 decline from highly significant values of 2.73 and 2.42 for the first sub-period to insignificant values close to zero for the second subperiod.

We examine the (in)stability of the market timing ability of the different models in more detail by computing hit ratios for rolling windows of 60 months. These are displayed in Figure 2.3 together with the expected hit ratios under independence of the signs of the excess return forecasts and realizations. The fifth panel of Figure 2.3 clearly reveals the large decline of the market timing ability of predictive regressions that include only financial variables between 1995-1997. After 1995 the hit ratios for the second benchmark model mostly are even below the expected hit ratios under independence, confirming the limited ability of this model with individual macro variables to predict the sign of the excess return correctly. In contrast, the performance of the ‘standard’ factor-augmented predictive regression model is much more stable and consistently above the expected hit ratio under independence. The hit ratios of the factor models with pre-selection of macro variables show considerably more fluctuation, although they do exceed the hit ratios expected under independence most of the time. For example, for the first hard-thresholding model, the hit ratio varies between 41% for windows around 1990 and approximately 65% for windows centered around 1985, 1995 and after the year 2000. Finally, the bottom panel of Figure 2.3 clearly demonstrates the lack of market timing ability of the historical mean return, in the sense that the actual hit ratio generally coincides with the expected hit ratio under independence.

The market timing results suggest that the factors extracted from macroeconomic variables may have useful information for predicting the sign of the excess stock returns. Ultimately, accurately predicting the magnitude of the excess returns may be at least as important as correct sign predictions. We assess this aspect here in two different ways. First, we examine the significance of the slope coefficient in a regression of the actual excess returns on the corresponding forecasts,

$$r_{e,t+1} = a + b\hat{r}_{e,t+1} + \eta_t, \quad (2.11)$$

Figure 2.3: Hit ratio for excess returns during five-year moving windows



*Note:* The solid line is the hit ratio for the sign of the monthly excess stock return, during the five-year moving window ending at the date displayed on the horizontal axis. The shaded area indicates the expected hit ratio under the assumption of independence between the actual and predicted signs. Model abbreviations are as in Table 2.1.

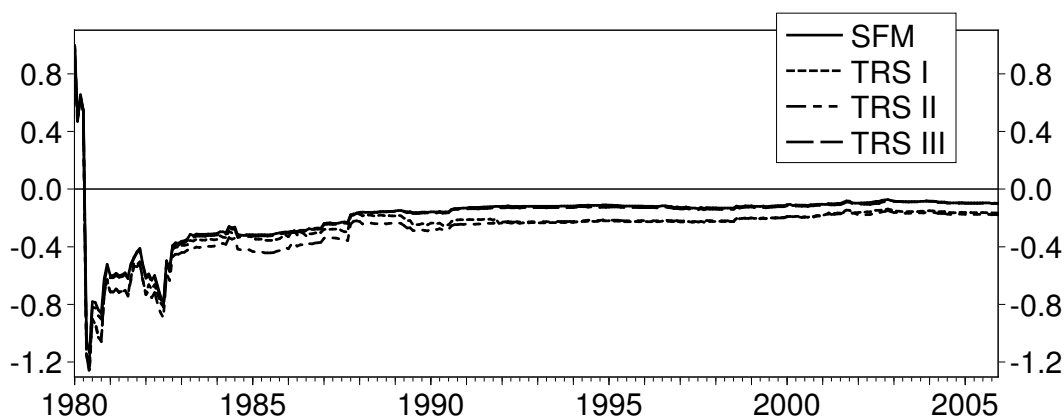
where again we use HAC standard errors to account for serial correlation. This is a robust version of the test used in Bossaerts and Hillion (1999). The results in Table 2.2 show that the performance of all models declines in the second half of the out-of-sample period. While the slope coefficient is significantly different from zero (and positive) at the 5% level or better for the subperiod 1980-1992, it becomes insignificant for the period 1993-2005. For the first benchmark model, the estimate of  $b$  in (2.11) even turns negative. The only exception is that the factor-augmented model based on soft thresholding appears to retain some of its predictive ability, with the HAC  $t$ -test being significant at the 10% level for the latter period.

Second, we evaluate the mean squared prediction error (MSE) of the factor-augmented predictive regressions relative to the MSE of the historical mean prediction used in the third benchmark model. Specifically, following Welch and Goyal (2008) and Campbell and Thompson (2008), among others, we consider the out-of-sample  $R^2$  defined as

$$R_{\text{OoS}}^2 = 1 - \frac{\sum_{t=0}^{n-1} (r_{e,t+1} - \hat{r}_{e,t+1})^2}{\sum_{t=0}^{n-1} (r_{e,t+1} - \bar{r}_{e,t+1})^2}, \quad (2.12)$$

where  $\hat{r}_{e,t+1}$  is the return prediction obtained from one of the factor-augmented regression models and  $\bar{r}_{e,t+1}$  is the historical average return computed using the ten-year moving window ending in period  $t$ , that is, the excess return prediction used in the third benchmark model. Positive (negative) values of  $R_{\text{OoS}}^2$  indicate that the factor-augmented model achieves a smaller (larger) MSE than the historical average return. By inspecting the path of  $R_{\text{OoS}}^2$  for increasing values of  $n$ , we can also assess the stability of the relative predictive accuracy of the two competing models. When the  $R_{\text{OoS}}^2$  increases, it means that the factor-augmented model outperforms the historical average, and vice versa when the  $R_{\text{OoS}}^2$  decreases. Figure 2.4 displays the path of the  $R_{\text{OoS}}^2$  during the complete out-of-sample period for the four factor-augmented predictive regressions. Over the complete out-of-sample period, the  $R_{\text{OoS}}^2$  values are negative at around  $-0.10$ , such that the historical average return achieves a smaller MSE than the elaborate factor-augmented predictive regressions. Interestingly, this appears to be due completely to a small number of erratic forecasts obtained from the factor-augmented models at the beginning of the sample period. This results in large fluctuations in the  $R_{\text{OoS}}^2$  values, such that they drop to levels below  $-1$  in July 1980. Following this unfortunate start the  $R_{\text{OoS}}^2$  generally shows an increasing pattern, indicating that over the remainder of the forecast period the factor-augmented regressions actually produce considerably more accurate excess return predictions than the historical average return.

Figure 2.4: Out-of-sample  $R^2$  of factor-augmented predictive regression models for excess returns

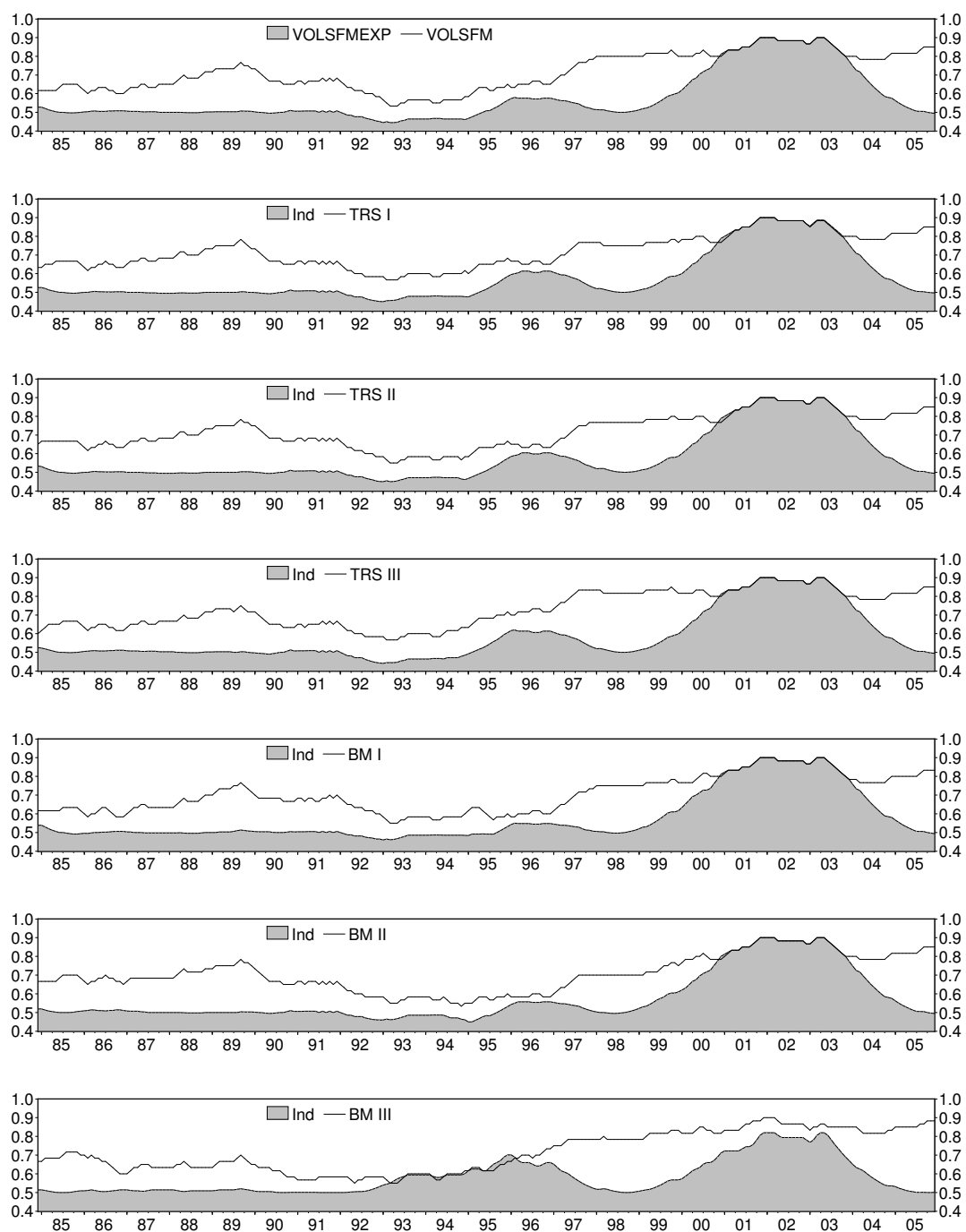


*Note:* The graph shows the time path of  $R^2_{\text{OoS}}$ , defined in (2.12), that is, one minus the ratio of the cumulative squared prediction errors for the factor-augmented regression relative to the cumulative squared prediction errors for the prevailing historical mean excess return. Model abbreviations are as in Table 2.1.

Next, we turn to the volatility timing results as shown in panel B of Table 2.1. As volatility exhibits much more persistence than excess returns, predicting the state of volatility one month ahead may be expected to be less difficult than predicting the sign of excess returns. This is indeed reflected by the high values of the hit ratios for the complete out-of-sample period as well as both subperiods. All models have hit ratios above 60%, with a substantial increase in the second half of the out-of-sample period. For the complete out-of-sample period 1980-2005 the factor-augmented models outperform the first and second benchmark only marginally. For the period after 1992, we find larger differences in hit ratios up to 5% between the factor-augmented model based on soft thresholding and the second benchmark model that includes individual macro variables. The statistics in Table 2.2 show that all models have significant volatility timing ability, both for the complete out-of-sample period as well as the two subperiods. Note that in this case the value of the BGJ-HAC statistic is considerably smaller than the PT statistic, due to the presence of positive serial correlation in the signs of the actual and predicted volatility.

Figure 2.5 shows the hit ratios for volatility forecasts over 5-year rolling windows, together with the expected hit ratios under independence of the sign forecasts and the actual signs. We observe that all models exhibit a similar pattern in their hit ratio, showing a gradual increase after 1995. Note that this does not necessarily imply that the

Figure 2.5: Hit ratio for volatility during five-year moving windows



*Note:* The solid line is the hit ratio for the sign of the monthly volatility (relative to the median volatility level during the 10-year rolling window that is used for model specification and estimation), during the five-year moving window ending at the date displayed on the horizontal axis. The shaded area indicates the expected hit ratio under the assumption of independence between the actual and predicted signs. Model abbreviations are as in Table 2.1.

volatility timing ability of the predictive regressions has improved over time. For example, the actual hit ratios are at the highest level for moving windows ending between 2000 and 2003, close to 90% for most models. This also is the case, however, for the expected hit ratios under independence, suggesting that the impressive values of the hit ratios are in fact not due to superior volatility timing ability but due to volatility clustering. Apart from this exceptional period, though, we do observe that the actual hit ratios are well above their expected values under independence, suggesting that in general volatility timing is possible with predictive regression models.

Finally, we consider the joint market and volatility timing ability, which is crucial for a mean-variance investor who would like to make investment decisions based on predictions of both excess returns and volatility. From Table 2.1, we observe that the hit ratios of joint timing take lower values compared to those of market and volatility timing individually. This shows that predicting the true state of the excess return and volatility jointly is more problematic than predicting them separately. The difference in the performance of the factor-augmented models and benchmark models in market timing ability is reflected in the joint timing results as this difference becomes larger during the second subsample. As in the market timing case many of the test statistics cannot be computed for the third benchmark model. Again this is due to the fact that some of the categories are never realized for the third benchmark due to the lack of market timing ability.

## 2.5 Economic value

In the previous section we analyzed the market timing and volatility timing ability of the factor-augmented predictive regression models. As mentioned before, this generally provides some indication of the performance of active trading strategies based on the excess return and volatility predictions. In this section, we assess this economic value of the factor-augmented predictive regression models explicitly. For this purpose, we consider a mean-variance investor with a one-month horizon, who decides upon the fraction of her wealth to be allocated to stocks using forecasts of excess returns and volatility. We use a utility based metric to evaluate how much the investor would be willing to pay to use the predictions from the factor-augmented models rather than those from the benchmarks. Specifically, we employ the procedure proposed in West *et al.* (1993), which is frequently used for measuring the economic value of dynamic investment strategies, see Fleming *et al.* (2001, 2003), de Pooter *et al.* (2008), and Della Corte *et al.* (2009, 2010), among others.

### 2.5.1 Methodology

We consider a risk-averse investor with mean-variance preferences, managing a portfolio consisting of stocks and risk-free T-bills. Each month the investor decides upon the fraction of wealth to be invested in stocks using the objective function

$$\max_{w_{t+1}} \quad \mathbb{E}_t(r_{p,t+1}) - \frac{1}{2}\gamma\text{Var}_t(r_{p,t+1}), \quad (2.13)$$

where  $\mathbb{E}_t(r_{p,t+1})$  and  $\text{Var}_t(r_{p,t+1})$  are the expected value and the variance of the portfolio return in period  $t + 1$  conditional on the information in period  $t$ , and  $\gamma$  represents the degree of relative risk aversion (RRA). The portfolio return is given by

$$r_{p,t+1} = r_{f,t+1} + w_{t+1}r_{e,t+1},$$

where  $r_{f,t+1}$  denotes the risk-free return, and  $w_{t+1}$  is the fraction of wealth allocated to stocks. We consider a myopic investor with a horizon of one month, such that the investor solves the maximization problem in (2.13) each time period in order to determine the portfolio weights for the next month. The optimal portfolio weight for stocks is given by

$$w_{t+1}^* = \frac{\mathbb{E}_t(r_{e,t+1})}{\delta\text{Var}_t(r_{p,t+1})}. \quad (2.14)$$

As (2.14) indicates, the implementation of the dynamic mean-variance strategy requires predictions of the excess return and volatility for computing the optimal portfolio weight for stocks. We assume that our investor employs the recursive predictions of the expected excess returns and volatility obtained from the factor-augmented regression models and the benchmark models for this purpose.

In order to mitigate the effects of forecast uncertainty the investor imposes restrictions on the portfolio weight. Specifically, we consider the following two scenarios. First, the investor uses the restriction  $w_{t+1}^* \in [-1, 2]$ , thus allowing for limited short-sales and leveraging of the portfolio. In the second scenario, short-sales and leveraging are not allowed at all, thus  $w_{t+1}^* \in [0, 1]$ .

Using the new return and volatility predictions that become available each month, the mean-variance investor adjusts her portfolio weights if necessary. Hence, transaction costs are an inevitable factor that must be taken into account. At the start of month  $t + 1$ , the investor changes the allocation to stocks from  $w_t$  to  $w_{t+1}$ . We assume that transaction costs are a fixed proportion  $c$  of the wealth invested, such that the portfolio return is

reduced by

$$c_{t+1} = 2c|w_{t+1} - w_t|, \quad (2.15)$$

where the multiplication by 2 follows from the fact that the investor rebalances her positions in both stocks and the riskfree T-bills. We consider no, medium and high transaction costs by setting  $c$  equal to 0%, 0.1% and 0.3%, respectively.

Assuming quadratic utility, we can use the average realized utility to consistently estimate the expected utility obtained from a given level of initial wealth  $W$ , see West *et al.* (1993), that is,

$$\bar{U} = \frac{W}{n} \sum_{t=0}^{n-1} \left( R_{p,t+1} - \frac{1}{2} \frac{\gamma}{(1+\gamma)} R_{p,t+1}^2 \right), \quad (2.16)$$

where  $R_{p,t+1} = 1 + r_{p,t+1} - c_{t+1}$  is the gross return on the portfolio net of transaction costs.

We assess the economic value of the factor-augmented predictive regression models by computing a ‘performance fee’, denoted as  $\Delta$ , that an investor should be willing to pay for switching from one investment strategy to another. For example, suppose that holding a dynamic portfolio constructed using one of the benchmark models yields the same average utility as holding a dynamic portfolio constructed using a factor-augmented predictive model that is subject to annual expenses,  $\Delta$ , expressed as a fraction of wealth invested. Because the investor would be indifferent between the two strategies we can interpret  $\Delta$  as the maximum fee an investor should be willing to pay to switch from the benchmark strategy to the strategy based on the factor-augmented models. The magnitude of this fee is given by the value of  $\Delta$  that satisfies

$$\sum_{t=0}^{n-1} \left\{ (R_{p,t+1}^a - \Delta) - \frac{1}{2} \frac{\gamma}{(1+\gamma)} (R_{p,t+1}^a - \Delta)^2 \right\} = \sum_{t=0}^{n-1} \left\{ R_{p,t+1}^b - \frac{1}{2} \frac{\gamma}{(1+\gamma)} (R_{p,t+1}^b)^2 \right\} \quad (2.17)$$

where the superscripts  $a$  and  $b$  denote the alternative strategies.

We compare the economic value of the factor-augmented predictive models relative to the three benchmark models. Additionally, we also consider three buy-and-hold strategies. These strategies assume that the investor invests a constant fraction of 100% (0%), 50% (50%), or 0% (100%) of her wealth in stocks (risk-free T-bills), respectively. Throughout the analysis, we set the relative risk aversion  $\gamma$  equal to 6.<sup>7</sup> Finally, for conventional comparison of the portfolio performances we also compute average returns, standard deviations, and Sharpe Ratios for each portfolio.

<sup>7</sup>In order to examine the sensitivity of the results to the level of risk aversion, we also consider risk aversion levels of 2 and 10. Results are shown in 2.A.3.

## 2.5.2 Empirical results

The results in Table 2.3 indicate that over the complete out-of-sample period from January 1980 until December 2005, investment strategies based on the factor-augmented predictive regressions outperform those based on the benchmark models, with a few exceptions. This holds irrespective of whether we consider Sharpe Ratios, as shown in the third column, or the performance fees for switching from the benchmark models to the factor-augmented predictive models in the three rightmost columns. While the magnitude of the economic value of the macro factors is affected by the level of transaction costs and the restrictions imposed on the portfolio weights (as discussed in more detail below), it remains positive in all scenarios.

As an example, consider the performance of the factor-augmented models relative to the first benchmark model in the scenario where limited short sales and leverage are allowed ( $w_{t+1}^* \in [-1, 2]$ ) and transaction costs are moderate (0.1%). In that case the portfolios based on the factor-augmented regressions achieve annualized Sharpe Ratios in the range 0.45-0.55, compared to 0.35 for the first benchmark model. This difference is mostly due to a higher average return achieved by the factor-augmented models ranging between 12% and 13.5% compared to 10.5% for the benchmark. The volatilities of the portfolio returns are comparable, at around 15%. The mean-variance investor is willing to pay a performance fee ranging from 131 to 326 basis points annually for switching from the first benchmark model to one of the factor-augmented models. This confirms the results of the market timing analysis in the previous section, in that the macroeconomic factors add relevant information concerning the expected returns and volatility to the information embedded in the financial variables in this first benchmark model. The possibilities to exploit this additional information become somewhat more limited when short sales and leverage are abandoned completely. The performance fees become smaller in this case, although they remain substantial ranging between 109 and 142 basis points.

The outperformance of the factor-augmented predictive regression models relative to the second benchmark model is particularly impressive. Again considering the scenario with medium transaction costs and limited short sales and leverage, the Sharpe Ratio of this benchmark is a disappointing 0.10, mostly due to a low average return of 7.2%. The performance fees in this scenario range between 622 and 816 basis points. Put differently, the investor is willing to pay between 6-8% on an annual basis to replace individual macroeconomic variables by factors based on a large database in the predictive regression models. Again, the performance fees are reduced, to about 3-3.5%, when the portfolio weights are restricted to be between 0 and 1.

Table 2.3: Performance of active trading strategies: Jan 1980 - Dec 2005, RRA=6

|                                                 | $\mu$ | $\sigma$ | SR   | $\Delta_{100}$ | $\Delta_{50}$ | $\Delta_0$ | $\Delta_{BM\ I}$ | $\Delta_{BM\ II}$ | $\Delta_{BM\ III}$ |
|-------------------------------------------------|-------|----------|------|----------------|---------------|------------|------------------|-------------------|--------------------|
| <u>Passive Strategies</u>                       |       |          |      |                |               |            |                  |                   |                    |
| I. 100% Market                                  | 13.56 | 15.10    | 0.52 |                |               |            |                  |                   |                    |
| II. 50% Market                                  | 9.60  | 7.55     | 0.52 |                |               |            |                  |                   |                    |
| III. 0% Market                                  | 5.64  | 0.90     | —    |                |               |            |                  |                   |                    |
| <u>Active Strategies</u>                        |       |          |      |                |               |            |                  |                   |                    |
| <u>Panel A: No transaction costs</u>            |       |          |      |                |               |            |                  |                   |                    |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 13.68 | 14.79    | 0.55 | 41             | −109          | 120        | 200              | 642               | −72                |
| TRS I                                           | 13.92 | 15.38    | 0.55 | 12             | −138          | 91         | 171              | 615               | −101               |
| TRS II                                          | 13.32 | 15.00    | 0.52 | −15            | −164          | 65         | 145              | 587               | −127               |
| TRS III                                         | 14.88 | 14.48    | 0.62 | 191            | 41            | 270        | 349              | 791               | 78                 |
| BM I                                            | 11.64 | 14.65    | 0.42 | −159           | −308          | −80        |                  |                   |                    |
| BM II                                           | 8.76  | 16.32    | 0.21 | −606           | −757          | −526       |                  |                   |                    |
| BM III                                          | 13.08 | 13.30    | 0.55 | 112            | −37           | 190        |                  |                   |                    |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 11.52 | 8.69     | 0.69 | 280            | 134           | 358        | 119              | 314               | 136                |
| TRS I                                           | 11.16 | 7.38     | 0.76 | 311            | 165           | 388        | 149              | 344               | 167                |
| TRS II                                          | 10.92 | 7.48     | 0.69 | 277            | 132           | 354        | 116              | 311               | 133                |
| TRS III                                         | 11.40 | 8.31     | 0.69 | 287            | 141           | 364        | 125              | 320               | 142                |
| BM I                                            | 10.44 | 8.83     | 0.55 | 162            | 16            | 239        |                  |                   |                    |
| BM II                                           | 8.64  | 9.18     | 0.35 | −33            | −180          | 44         |                  |                   |                    |
| BM III                                          | 10.92 | 9.98     | 0.52 | 145            | −1            | 223        |                  |                   |                    |
| <u>Panel B: Medium transaction costs (0.1%)</u> |       |          |      |                |               |            |                  |                   |                    |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 12.48 | 14.79    | 0.45 | −81            | −231          | −2         | 182              | 673               | −94                |
| TRS I                                           | 12.72 | 15.38    | 0.45 | −108           | −258          | −28        | 156              | 649               | −121               |
| TRS II                                          | 12.12 | 14.96    | 0.45 | −133           | −282          | −53        | 131              | 622               | −145               |
| TRS III                                         | 13.56 | 14.45    | 0.55 | 63             | −86           | 142        | 326              | 816               | 50                 |
| BM I                                            | 10.56 | 14.69    | 0.35 | −263           | −413          | −184       |                  |                   |                    |
| BM II                                           | 7.20  | 16.35    | 0.10 | −761           | −911          | −681       |                  |                   |                    |
| BM III                                          | 12.12 | 13.30    | 0.48 | 13             | −136          | 91         |                  |                   |                    |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 11.04 | 8.69     | 0.62 | 232            | 86            | 310        | 109              | 325               | 138                |
| TRS I                                           | 10.68 | 7.34     | 0.69 | 265            | 119           | 342        | 142              | 358               | 170                |
| TRS II                                          | 10.44 | 7.48     | 0.66 | 232            | 86            | 309        | 109              | 324               | 137                |
| TRS III                                         | 10.92 | 8.31     | 0.62 | 238            | 92            | 315        | 114              | 330               | 143                |
| BM I                                            | 10.08 | 8.83     | 0.48 | 124            | −22           | 201        |                  |                   |                    |
| BM II                                           | 8.04  | 9.18     | 0.28 | −93            | −239          | −16        |                  |                   |                    |
| BM III                                          | 10.44 | 9.98     | 0.48 | 95             | −51           | 173        |                  |                   |                    |
| <u>Panel C: High transaction costs (0.3%)</u>   |       |          |      |                |               |            |                  |                   |                    |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 10.08 | 14.86    | 0.31 | −327           | −477          | −248       | 147              | 739               | −138               |
| Trsy I                                          | 10.32 | 15.42    | 0.31 | −351           | −501          | −272       | 124              | 717               | −162               |
| TRS II                                          | 9.72  | 15.00    | 0.28 | −372           | −522          | −293       | 102              | 694               | −183               |
| TRS III                                         | 9.24  | 14.72    | 0.24 | −395           | −545          | −316       | 79               | 670               | −7                 |
| BM I                                            | 8.52  | 14.72    | 0.21 | −474           | −623          | −395       |                  |                   |                    |
| BM II                                           | 4.20  | 16.45    | —    | −1075          | −1226         | −995       |                  |                   |                    |
| BM III                                          | 10.08 | 13.34    | 0.35 | −188           | −336          | −109       |                  |                   |                    |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 10.08 | 8.73     | 0.52 | 136            | −10           | 213        | 89               | 349               | 178                |
| TRS I                                           | 9.84  | 7.38     | 0.55 | 173            | 27            | 250        | 126              | 385               | 141                |
| TRS II                                          | 9.48  | 7.48     | 0.52 | 140            | −5            | 217        | 93               | 352               | 145                |
| TRS III                                         | 9.96  | 8.35     | 0.52 | 139            | −7            | 216        | 92               | 352               | 144                |
| BM I                                            | 9.36  | 8.87     | 0.42 | 47             | −99           | 124        |                  |                   |                    |
| BM II                                           | 6.96  | 9.25     | 0.14 | −213           | −360          | −136       |                  |                   |                    |
| BM III                                          | 9.48  | 9.98     | 0.38 | −5             | −152          | −73        |                  |                   |                    |

*Note:* The table displays performance measures for the active mean-variance investment strategies based on monthly excess return and volatility predictions for the period January 1980 - December 2005. Columns headed  $\mu$  and  $\sigma$  contain the annualized mean and standard deviation of the portfolio returns (in percent). SR denotes the Sharpe ratio. Columns headed  $\Delta$  contain performance fees (in annualized basis points) for switching from the strategy indicated by the subscript to the strategy indicated by the corresponding row. The subscript  $x$  with  $x = 100, 50$ , and  $0$  refer to the static strategies with  $x\%$  of wealth invested in stocks. Model abbreviations are as in Table 2.1.

The third benchmark model, which uses the historical average excess return and the previous month's volatility level as predictions, appears more difficult to beat. When limited short sales and leverage are allowed the third benchmark model actually performs better than most of the factor-augmented models, with performance fees of around 100 basis points. Interestingly, all of the factor-augmented models yield higher average returns than this benchmark model. The third benchmark model achieves a considerably lower standard deviation of returns, however. Still, the factor-augmented model with soft thresholding performs slightly better than the third benchmark model by an annual performance fee of 50 basis points. Moreover, when short sales and leverage are completely restricted, all factor-augmented models have a positive performance fee as well as higher Sharpe ratios. In this case, an investor is willing to pay between 137 and 170 basis points annually to switch from the third benchmark model to the factor-augmented models. The main consequence of abandoning short sales and leverage is a substantial reduction in the standard deviation of portfolio returns based on the factor-augmented models from 15% to 8% on average, see the second column of Table 2.3. While the portfolio standard deviation is also reduced for the third benchmark model, the decline is much less pronounced (from 13.3% to 10%).

The performance fees of the factor-augmented models relative to the three benchmark models are relatively stable across the three different levels of transaction costs considered. By contrast, when comparing the dynamic strategies with static buy and hold strategies, the level of transaction costs becomes much more important in evaluating the relative portfolio performance, especially when limited short sales and leveraging are allowed. In that case, the factor-augmented models still render positive performance fees when transaction costs are not included. Medium and high transaction cost levels, however, lead to negative performance fees. Especially the mixed portfolio with a constant weight of 50% in stocks performs well, with  $\Delta$  being approximately equal to  $-500$  basis points in the scenario with high transaction costs. This picture changes dramatically when short sales and leveraging are not allowed. In this case, under medium transaction costs all factor-augmented models have a positive performance fee ranging from 86 to 119 basis points annually against the static mixed portfolio. The performance fees against the other two static portfolios are considerably higher. The investor is willing to pay around 250 basis points annually to switch from the buy-and-hold strategy in stocks to one of the factor augmented models, and more than 300 basis points to switch from the static strategy that always invests in T-bills. Interestingly, even if the transaction costs are high, factor-augmented models still perform better than or as good as the static strategies. An important finding is that the performance fees are highest when the factor models

with thresholding rules are used for portfolio construction. This indicates that a sensible pre-selection of the set of macroeconomic variables from which factors are constructed adds even more economic value than simply using all available macro variables as in the standard factor-augmented model.

The performance fees for switching from the static portfolios to the factor-augmented predictive regressions indicate the importance of imposing restrictions on the portfolio weights, to avoid excessive exposure to stocks but also to limit transaction costs. This is reflected in Table 2.4 where we display the mean and the standard deviation of the portfolio weights. For all factor-augmented models (as well as for the benchmark model), we observe that the standard deviation of the portfolio weights is reduced by 50-60% (from around 1.00 to 0.45, on average) when short sales and leverage are abandoned.

Table 2.4: Mean and standard deviation of portfolio weights

|                                                    | SFM                       | TRS I | TRS II | TRS III | BM I | BM II | BM III |
|----------------------------------------------------|---------------------------|-------|--------|---------|------|-------|--------|
| <u>Panel A: <math>w_{t+1}^* \in [-1, 2]</math></u> |                           |       |        |         |      |       |        |
|                                                    | <u>Mean</u>               |       |        |         |      |       |        |
| Jan 1980 - Dec 2005                                | 0.50                      | 0.32  | 0.34   | 0.46    | 0.51 | 0.33  | 0.90   |
| Jan 1980 - Dec 1992                                | 0.42                      | 0.08  | 0.13   | 0.34    | 0.46 | 0.22  | 0.69   |
| Jan 1993 - Dec 2005                                | 0.58                      | 0.56  | 0.56   | 0.59    | 0.56 | 0.44  | 1.11   |
|                                                    | <u>Standard Deviation</u> |       |        |         |      |       |        |
| Jan 1980 - Dec 2005                                | 0.99                      | 1.12  | 1.10   | 1.04    | 0.99 | 1.14  | 0.68   |
| Jan 1980 - Dec 1992                                | 0.98                      | 1.15  | 1.13   | 0.99    | 0.96 | 1.21  | 0.62   |
| Jan 1993 - Dec 2005                                | 1.00                      | 1.03  | 1.02   | 1.07    | 1.01 | 1.05  | 0.67   |
| <u>Panel B: <math>w_{t+1}^* \in [0, 1]</math></u>  |                           |       |        |         |      |       |        |
|                                                    | <u>Mean</u>               |       |        |         |      |       |        |
| Jan 1980 - Dec 2005                                | 0.47                      | 0.39  | 0.41   | 0.45    | 0.47 | 0.43  | 0.66   |
| Jan 1980 - Dec 1992                                | 0.45                      | 0.33  | 0.35   | 0.41    | 0.46 | 0.41  | 0.41   |
| Jan 1993 - Dec 2005                                | 0.49                      | 0.46  | 0.48   | 0.50    | 0.47 | 0.46  | 0.50   |
|                                                    | <u>Standard Deviation</u> |       |        |         |      |       |        |
| Jan 1980 - Dec 2005                                | 0.44                      | 0.44  | 0.44   | 0.44    | 0.45 | 0.45  | 0.35   |
| Jan 1980 - Dec 1992                                | 0.43                      | 0.43  | 0.43   | 0.43    | 0.44 | 0.46  | 0.43   |
| Jan 1993 - Dec 2005                                | 0.44                      | 0.44  | 0.45   | 0.46    | 0.45 | 0.44  | 0.45   |

*Note:* The table displays means and standard deviations of the portfolio weight for stocks in the active mean-variance investment strategies based on monthly predictions of excess returns and volatility. Model abbreviations are as in Table 2.1.

A crucial finding from the market timing analysis in the previous section was the stability of the performance of the factor-augmented predictive regressions, whereas the performance of the benchmark models deteriorates substantially during the second half of the out-of-sample period. This is also reflected in their (relative) economic value during the subperiods 1980-1992 and 1993-2005, as shown in Tables 2.5 and 2.6. The portfolios based on the predictions obtained from the benchmark models experience a pronounced decline in their Sharpe Ratio. For example, the first benchmark model (including financial variables only) achieves a Sharpe ratio of 0.59 over the period 1980-1992, when limited

short sales and leverage are allowed and transaction costs are medium. This drops to 0.00 for the period 1993-2005. The reduction in Sharpe ratio is less dramatic when the portfolio weights are more restricted, but still substantial from 0.66 to 0.28. For the factor-augmented models, the Sharpe ratio is approximately the same during the two sub-periods.

The performance fees reveal a similar pattern. While the dynamic strategies based on the benchmark models mostly achieve positive performance fees against the static buy-and-hold strategies during the first half of the out-of-sample period, these turn negative for the period after 1992. Strikingly, the factor-augmented models generally have higher performance fees during the second sub-period. Most notably this holds for those factor models where thresholding rules are employed for factor extraction. Specifically, models with hard thresholding with restricted portfolio weights perform better than all passive strategies even when the transaction costs are at the highest level. Allowing for short sales and leverage of the portfolio increases the performance fee for switching from the benchmark models to the factor-augmented models. However, at the same time it reduces the performance fees for switching from the static strategies, as more volatile returns cause a loss of utility. For the first half of the sample period only the first benchmark model performs marginally better than the factor-augmented models when the portfolio weights are completely restricted. When limited short sales and leverage are allowed the performance fee for switching to a factor-augmented model ranges from  $-18$  to  $-534$  basis points under medium transaction costs.

A more detailed picture of the evolution of the performance of the factor-augmented models relative to the benchmark models over the out-of-sample period is provided by Figure 2.6, showing the fees for switching from each of the three benchmark models to factor-augmented regressions for moving windows of 60 months (with medium transaction costs). An interesting observation from these graphs is that for the first benchmark model, the superior performance of the factor-augmented models already starts around the beginning of the 1990's, whereas for the third benchmark model it only starts around the year 2000. In fact, the third benchmark model is seen to perform particularly well during most of the 1990s, with negative performance fees for all four factor-augmented models. This is due to the low standard deviation of portfolio returns obtained from the third benchmark model combined with the prolonged bull market with low volatility during this period. As discussed before, the third benchmark model always predicts moderate positive excess returns after the mid 1980's but does not have any market timing ability. This strategy seems to work for the bull market of the mid 1990's when volatility was also very low. However, with the advent of more volatile periods towards the end of this

Table 2.5: Performance of active trading strategies: Jan 1980 - Dec 1992, RRA=6

|                                                 | $\mu$ | $\sigma$ | SR   | $\Delta_{100}$ | $\Delta_{50}$ | $\Delta_0$ | $\Delta_{\text{BM I}}$ | $\Delta_{\text{BM II}}$ | $\Delta_{\text{BM III}}$ |
|-------------------------------------------------|-------|----------|------|----------------|---------------|------------|------------------------|-------------------------|--------------------------|
| <u>Passive Strategies</u>                       |       |          |      |                |               |            |                        |                         |                          |
| I. 100% Market                                  | 16.20 | 15.93    | 0.52 |                |               |            |                        |                         |                          |
| II. 50% Market                                  | 11.88 | 7.93     | 0.52 |                |               |            |                        |                         |                          |
| III. 0% Market                                  | 7.68  | 0.83     | —    |                |               |            |                        |                         |                          |
| <u>Active Strategies</u>                        |       |          |      |                |               |            |                        |                         |                          |
| <u>Panel A: No transaction costs</u>            |       |          |      |                |               |            |                        |                         |                          |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                        |                         |                          |
| SFM                                             | 17.40 | 17.01    | 0.59 | 14             | −182          | 57         | −176                   | 602                     | −239                     |
| TRS I                                           | 16.32 | 18.19    | 0.48 | −246           | −444          | −203       | −437                   | 347                     | −501                     |
| TRS II                                          | 15.84 | 17.60    | 0.45 | −215           | −412          | −172       | −405                   | 376                     | −467                     |
| TRS III                                         | 17.16 | 16.28    | 0.59 | 65             | −130          | 107        | −124                   | 651                     | −187                     |
| BM I                                            | 18.84 | 16.66    | 0.66 | 189            | −6            | 232        |                        |                         |                          |
| BM II                                           | 13.20 | 18.57    | 0.31 | −594           | −792          | −551       |                        |                         |                          |
| BM III                                          | 14.76 | 11.54    | 0.62 | 246            | 55            | 288        |                        |                         |                          |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                        |                         |                          |
| SFM                                             | 14.64 | 9.94     | 0.69 | 348            | 158           | 389        | −22                    | 209                     | 152                      |
| TRS I                                           | 13.08 | 8.24     | 0.66 | 290            | 102           | 332        | −77                    | 152                     | 95                       |
| TRS II                                          | 12.84 | 8.28     | 0.62 | 268            | 79            | 309        | −100                   | 130                     | 73                       |
| TRS III                                         | 14.16 | 9.32     | 0.69 | 341            | 152           | 382        | −28                    | 202                     | 145                      |
| BM I                                            | 15.00 | 10.12    | 0.73 | 370            | 180           | 411        |                        |                         |                          |
| BM II                                           | 12.48 | 9.77     | 0.48 | 139            | −50           | 180        |                        |                         |                          |
| BM III                                          | 12.84 | 9.42     | 0.55 | 196            | 7             | 237        |                        |                         |                          |
| <u>Panel B: Medium transaction costs (0.1%)</u> |       |          |      |                |               |            |                        |                         |                          |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                        |                         |                          |
| SFM                                             | 16.20 | 17.01    | 0.48 | −112           | −308          | −69        | −165                   | 614                     | −271                     |
| TRS I                                           | 15.00 | 18.19    | 0.42 | −373           | −571          | −330       | −427                   | 358                     | −534                     |
| TRS II                                          | 14.64 | 17.56    | 0.38 | −332           | −528          | −289       | −385                   | 397                     | −492                     |
| TRS III                                         | 15.96 | 16.32    | 0.52 | −63            | −258          | −20        | −116                   | 660                     | −221                     |
| BM I                                            | 17.52 | 16.70    | 0.59 | 53             | −142          | 96         |                        |                         |                          |
| BM II                                           | 11.88 | 18.57    | 0.21 | −733           | −931          | −690       |                        |                         |                          |
| BM III                                          | 13.80 | 11.50    | 0.52 | 155            | −36           | 197        |                        |                         |                          |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                        |                         |                          |
| SFM                                             | 14.16 | 9.94     | 0.66 | 297            | 108           | 339        | −18                    | 209                     | 151                      |
| TRS I                                           | 12.60 | 8.24     | 0.59 | 244            | 55            | 285        | −70                    | 156                     | 98                       |
| TRS II                                          | 12.48 | 8.24     | 0.59 | 224            | 35            | 265        | −90                    | 136                     | 78                       |
| TRS III                                         | 13.68 | 9.35     | 0.66 | 288            | 99            | 330        | −27                    | 200                     | 142                      |
| BM I                                            | 14.52 | 10.12    | 0.66 | 316            | 126           | 357        |                        |                         |                          |
| BM II                                           | 12.00 | 9.77     | 0.45 | 88             | −101          | 130        |                        |                         |                          |
| BM III                                          | 12.36 | 9.42     | 0.48 | 146            | −43           | 188        |                        |                         |                          |
| <u>Panel C: High transaction costs (0.3%)</u>   |       |          |      |                |               |            |                        |                         |                          |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                        |                         |                          |
| SFM                                             | 13.80 | 17.11    | 0.35 | −367           | −563          | −324       | −145                   | 641                     | −337                     |
| TRS I                                           | 12.48 | 18.26    | 0.28 | −633           | −831          | −590       | −409                   | 382                     | −603                     |
| TRS II                                          | 12.24 | 17.56    | 0.28 | −570           | −767          | −527       | −347                   | 440                     | −541                     |
| TRS III                                         | 13.44 | 16.39    | 0.35 | −321           | −516          | −278       | −100                   | 683                     | −292                     |
| BM I                                            | 14.88 | 16.80    | 0.42 | −222           | −417          | −179       |                        |                         |                          |
| BM II                                           | 9.12  | 18.64    | 0.07 | −1017          | −1215         | −974       |                        |                         |                          |
| BM III                                          | 12.00 | 11.50    | 0.38 | −29            | −219          | 13         |                        |                         |                          |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                        |                         |                          |
| SFM                                             | 13.20 | 10.01    | 0.55 | 196            | 6             | 237        | −12                    | 209                     | 148                      |
| TRS I                                           | 11.76 | 8.24     | 0.48 | 150            | −38           | 192        | −56                    | 164                     | 103                      |
| TRS II                                          | 11.52 | 8.28     | 0.48 | 135            | −53           | 176        | −71                    | 149                     | 88                       |
| TRS III                                         | 12.72 | 9.39     | 0.52 | 182            | −7            | 223        | −25                    | 196                     | 135                      |
| BM I                                            | 13.44 | 10.18    | 0.55 | 207            | 18            | 249        |                        |                         |                          |
| BM II                                           | 10.92 | 9.80     | 0.35 | −14            | −203          | 28         |                        |                         |                          |
| BM III                                          | 11.28 | 9.42     | 0.38 | 47             | −142          | 89         |                        |                         |                          |

*Note:* The table displays performance measures for the active mean-variance investment strategies based on monthly excess return and volatility predictions for the period January 1980 - December 1992. See Table 2.3 for further details.

Table 2.6: Performance of active trading strategies: Jan 1993 - Dec 2005, RRA=6

|                                                 | $\mu$ | $\sigma$ | SR   | $\Delta_{100}$ | $\Delta_{50}$ | $\Delta_0$ | $\Delta_{BM\ I}$ | $\Delta_{BM\ II}$ | $\Delta_{BM\ III}$ |
|-------------------------------------------------|-------|----------|------|----------------|---------------|------------|------------------|-------------------|--------------------|
| <u>Passive Strategies</u>                       |       |          |      |                |               |            |                  |                   |                    |
| I. 100% Market                                  | 11.04 | 14.24    | 0.52 |                |               |            |                  |                   |                    |
| II. 50% Market                                  | 7.32  | 7.14     | 0.52 |                |               |            |                  |                   |                    |
| III. 0% Market                                  | 3.60  | 0.45     | —    |                |               |            |                  |                   |                    |
| <u>Active Strategies</u>                        |       |          |      |                |               |            |                  |                   |                    |
| <u>Panel A: No transaction costs</u>            |       |          |      |                |               |            |                  |                   |                    |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 9.96  | 12.12    | 0.52 | 67             | -39           | 180        | 554              | 680               | 87                 |
| TRS I                                           | 11.64 | 11.85    | 0.69 | 260            | 155           | 374        | 747              | 873               | 281                |
| TRS II                                          | 10.80 | 11.78    | 0.62 | 178            | 73            | 292        | 665              | 791               | 199                |
| TRS III                                         | 12.48 | 12.26    | 0.73 | 313            | 207           | 427        | 800              | 927               | 140                |
| BM I                                            | 4.32  | 11.95    | 0.07 | -487           | -592          | -373       |                  |                   |                    |
| BM II                                           | 4.32  | 13.61    | 0.07 | -617           | -723          | -503       |                  |                   |                    |
| BM III                                          | 11.40 | 14.86    | 0.52 | -21            | -127          | 94         |                  |                   |                    |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 8.40  | 7.14     | 0.69 | 215            | 111           | 327        | 253              | 415               | 121                |
| TRS I                                           | 9.24  | 6.34     | 0.90 | 331            | 227           | 443        | 369              | 531               | 237                |
| TRS II                                          | 8.88  | 6.55     | 0.80 | 286            | 183           | 398        | 325              | 486               | 192                |
| TRS III                                         | 8.64  | 7.07     | 0.69 | 234            | 130           | 346        | 272              | 434               | 333                |
| BM I                                            | 5.88  | 7.14     | 0.31 | -38            | -142          | 74         |                  |                   |                    |
| BM II                                           | 4.92  | 8.38     | 0.14 | -201           | -305          | -88        |                  |                   |                    |
| BM III                                          | 9.12  | 10.46    | 0.52 | -201           | -9            | 208        |                  |                   |                    |
| <u>Panel B: Medium transaction costs (0.1%)</u> |       |          |      |                |               |            |                  |                   |                    |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 8.76  | 12.09    | 0.42 | -53            | -158          | 61         | 512              | 731               | 76                 |
| TRS I                                           | 10.56 | 11.85    | 0.59 | 145            | 39            | 258        | 709              | 928               | 273                |
| TRS II                                          | 9.60  | 11.78    | 0.52 | 56             | -49           | 170        | 620              | 840               | 185                |
| TRS III                                         | 11.28 | 12.26    | 0.62 | 182            | 76            | 296        | 746              | 967               | 310                |
| BM I                                            | 3.48  | 11.95    | 0.00 | -565           | -670          | -451       |                  |                   |                    |
| BM II                                           | 2.64  | 13.65    | —    | -790           | -896          | -675       |                  |                   |                    |
| BM III                                          | 10.32 | 14.86    | 0.45 | -130           | -237          | -15        |                  |                   |                    |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 7.92  | 7.10     | 0.62 | 168            | 64            | 280        | 231              | 438               | 125                |
| TRS I                                           | 8.76  | 6.30     | 0.83 | 284            | 181           | 396        | 347              | 553               | 241                |
| TRS II                                          | 8.40  | 6.55     | 0.73 | 238            | 134           | 350        | 301              | 507               | 195                |
| TRS III                                         | 8.16  | 7.03     | 0.62 | 187            | 83            | 299        | 250              | 456               | 144                |
| BM I                                            | 5.64  | 7.14     | 0.28 | -63            | -167          | 49         |                  |                   |                    |
| BM II                                           | 4.20  | 8.42     | 0.07 | -270           | -374          | -158       |                  |                   |                    |
| BM III                                          | 8.52  | 10.46    | 0.48 | 43             | -61           | 157        |                  |                   |                    |
| <u>Panel C: High transaction costs (0.3%)</u>   |       |          |      |                |               |            |                  |                   |                    |
| <u><math>w_{t+1}^* \in [-1, 2]</math></u>       |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 6.36  | 12.09    | 0.21 | -295           | -400          | -181       | 427              | 836               | 52                 |
| TRS I                                           | 8.16  | 11.85    | 0.38 | -89            | -194          | 25         | 632              | 1040              | 257                |
| TRS II                                          | 7.20  | 11.81    | 0.31 | -191           | -296          | -77        | 530              | 938               | 156                |
| TRS III                                         | 8.64  | 12.33    | 0.42 | -84            | -189          | 30         | 638              | 1047              | 263                |
| BM I                                            | 1.92  | 11.99    | —    | -721           | -827          | -608       |                  |                   |                    |
| BM II                                           | 0.72  | 13.79    | —    | -1138          | -1244         | -1023      |                  |                   |                    |
| BM III                                          | 8.16  | 14.93    | 0.31 | -351           | -458          | -236       |                  |                   |                    |
| <u><math>w_{t+1}^* \in [0, 1]</math></u>        |       |          |      |                |               |            |                  |                   |                    |
| SFM                                             | 6.96  | 7.10     | 0.48 | 73             | -31           | 185        | 186              | 483               | 134                |
| TRS I                                           | 7.80  | 6.30     | 0.66 | 190            | 86            | 301        | 302              | 599               | 250                |
| TRS II                                          | 7.44  | 6.55     | 0.59 | 140            | 37            | 252        | 253              | 550               | 201                |
| TRS III                                         | 7.20  | 7.03     | 0.52 | 92             | -12           | 204        | 205              | 502               | 153                |
| BM I                                            | 5.16  | 7.10     | 0.21 | -113           | -217          | -1         |                  |                   |                    |
| BM II                                           | 2.88  | 8.49     | —    | -411           | -515          | -299       |                  |                   |                    |
| BM III                                          | 7.56  | 10.50    | 0.38 | -61            | -166          | 52         |                  |                   |                    |

*Note:* The table displays performance measures for the active mean-variance investment strategies based on monthly excess return and volatility predictions for the period January 1993 - December 2005. See Table 2.3 for further details.

decade, the market timing ability of the predictive models gains more importance. This explains the dramatic improvement in the performance of the factor-augmented models starting from the end of the 1990's, with the fees relative to the third benchmark model increasing to almost 1200 basis points at the end of 2003.

An important result from Table 2.6 is that the factor-augmented models that use thresholding rules for predictor selection perform better than the factor model that simply uses all available variables, especially when limited short sales and leverage are allowed for. This confirms the findings of Bai and Ng (2008) in the context of forecasting inflation. When the aim is to predict a variable with limited predictive content such as stock returns, our findings show that a reasonable set of thresholding rules is even more crucial also economically. For switching from the passive strategy of investing 100% in stocks to the first hard thresholding rule with completely restricted portfolio weights, for example, an investor would pay 284 basis points per annum compared to a fee of only 168 for the standard factor model. This difference increases when limited short sales are allowed, for which the precision of return predictions becomes more important rather than the prediction of the direction of the market. When this scenario is considered the factor model with the soft thresholding rule, which is relatively more sophisticated than others, outperforms the other factor models.

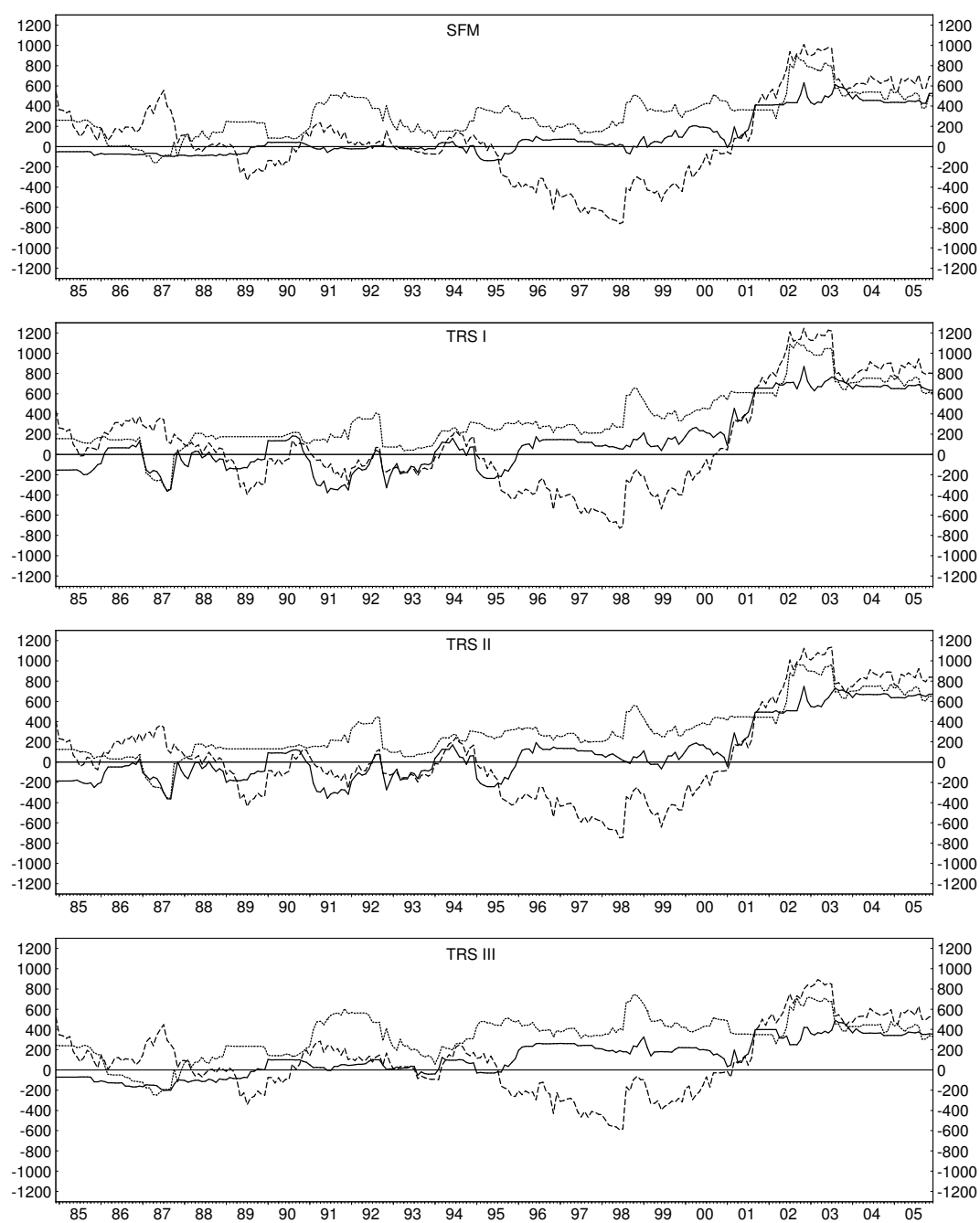
The majority of previous studies focuses on the market timing ability of predictive models for (excess) returns only. This is either because the objective of the econometric methods is directed only at return prediction or because the investment strategy is 'discrete', involving switching between stocks and risk-free investments based only on the predicted sign of the excess return (see Pesaran and Timmermann, 1995; Aiolfi and Favero, 2005, among others). From this perspective, it is also interesting to evaluate the performance of the factor-augmented models using only the information in the excess return forecasts. We implement this in our mean-variance framework by only using the return forecast in (2.14) while setting the conditional volatility equal to the median volatility value during the estimation sample.<sup>8</sup> Table 2.7 exhibits the results of these strategies for the complete out-of-sample period and the two subperiods. For the sake of brevity, we only display results for medium transaction costs (0.1%). Results for other transaction cost levels are in line with these and available upon request.

A comparison of the results in Table 2.7 with those in Panel B of Tables 2.3-2.6 reveals that ignoring the volatility predictions increases both the mean and volatility of the portfolio returns. The overall effect on the Sharpe ratio is slightly negative. The performance fees for switching from the benchmark models to factor-augmented models

---

<sup>8</sup>Recall that we also used this as benchmark value for determining "high" and "low" volatile periods.

Figure 2.6: Economic performance of factor-augmented predictive regression models during five-year moving windows



*Note:* Performance fees  $\Delta$  (in basis points per annum) computed according to (2.17) for five-year moving windows ending at the date displayed on the horizontal axis. The solid line, the dotted line and the dashed lines represent the performance fees of the indicated factor-augmented predictive regression relative to the first, second, and third benchmark models. Model abbreviations are as in Table 2.1.

Table 2.7: Performance of active trading strategies based on excess return forecasts

|                                       | $\mu$ | $\sigma$ | SR   | $\Delta_{100}$ | $\Delta_{50}$ | $\Delta_0$ | $\Delta_{\text{BM I}}$ | $\Delta_{\text{BM II}}$ | $\Delta_{\text{BM III}}$ |
|---------------------------------------|-------|----------|------|----------------|---------------|------------|------------------------|-------------------------|--------------------------|
| Panel A: January 1980 - December 2005 |       |          |      |                |               |            |                        |                         |                          |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 14.28 | 18.05    | 0.48 | -252           | -405          | -172       | 353                    | 1191                    | -45                      |
| TRS I                                 | 13.56 | 17.67    | 0.45 | -271           | -423          | -191       | 332                    | 1169                    | -64                      |
| TRS II                                | 13.08 | 17.60    | 0.42 | -310           | -462          | -230       | 293                    | 1129                    | -103                     |
| TRS III                               | 14.28 | 17.60    | 0.48 | -200           | -352          | -120       | 403                    | 1239                    | 7                        |
| BM I                                  | 10.80 | 18.15    | 0.28 | -605           | -758          | -525       |                        |                         |                          |
| BM II                                 | 4.92  | 20.30    | —    | -1464          | -1619         | -1382      |                        |                         |                          |
| BM III                                | 12.24 | 15.80    | 0.42 | -205           | -355          | -125       |                        |                         |                          |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 12.00 | 10.18    | 0.62 | 232            | 86            | 310        | 159                    | 516                     | 231                      |
| TRS I                                 | 11.52 | 8.49     | 0.69 | 287            | 141           | 364        | 213                    | 568                     | 286                      |
| TRS II                                | 10.92 | 8.38     | 0.62 | 235            | 90            | 313        | 162                    | 517                     | 235                      |
| TRS III                               | 11.52 | 9.80     | 0.59 | 245            | 98            | 323        | 403                    | 1239                    | 207                      |
| BM I                                  | 10.44 | 10.25    | 0.45 | 74             | -73           | 151        |                        |                         |                          |
| BM II                                 | 7.32  | 10.95    | 0.14 | -284           | -431          | -206       |                        |                         |                          |
| BM III                                | 10.80 | 11.81    | 0.45 | 1              | -147          | 79         |                        |                         |                          |
| Panel B: January 1980 - December 1992 |       |          |      |                |               |            |                        |                         |                          |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 18.84 | 19.85    | 0.55 | -204           | -404          | -160       | -131                   | 1044                    | -198                     |
| TRS I                                 | 16.68 | 19.71    | 0.45 | -397           | -596          | -353       | -324                   | 850                     | -391                     |
| TRS II                                | 16.68 | 19.43    | 0.45 | -355           | -554          | -311       | -282                   | 890                     | -349                     |
| TRS III                               | 17.16 | 18.71    | 0.52 | -217           | -415          | -174       | -145                   | 1021                    | -212                     |
| BM I                                  | 18.96 | 18.98    | 0.59 | -73            | -271          | -29        |                        |                         |                          |
| BM II                                 | 10.44 | 21.44    | 0.14 | -1263          | -1465         | -1218      |                        |                         |                          |
| BM III                                | 11.52 | 10.57    | 0.38 | -5             | -195          | 36         |                        |                         |                          |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 14.76 | 11.71    | 0.62 | 239            | 48            | 281        | -3                     | 291                     | 222                      |
| TRS I                                 | 13.92 | 8.90     | 0.69 | 341            | 152           | 382        | 101                    | 392                     | 325                      |
| TRS II                                | 13.20 | 8.73     | 0.62 | 279            | 90            | 320        | 39                     | 330                     | 263                      |
| TRS III                               | 14.40 | 10.81    | 0.62 | 259            | 69            | 300        | 18                     | 310                     | 243                      |
| BM I                                  | 14.52 | 11.33    | 0.59 | 242            | 51            | 283        |                        |                         |                          |
| BM II                                 | 11.64 | 11.40    | 0.35 | -52            | -242          | -10        |                        |                         |                          |
| BM III                                | 11.64 | 10.46    | 0.38 | 16             | -174          | 58         |                        |                         |                          |
| Panel C: January 1993 - December 2005 |       |          |      |                |               |            |                        |                         |                          |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 9.72  | 15.97    | 0.38 | -301           | -408          | -185       | 803                    | 1331                    | 100                      |
| TRS I                                 | 10.56 | 15.35    | 0.45 | -153           | -260          | -37        | 948                    | 1474                    | 247                      |
| TRS II                                | 9.48  | 15.42    | 0.38 | -269           | -376          | -154       | 831                    | 1358                    | 130                      |
| TRS III                               | 11.28 | 16.39    | 0.48 | -186           | -294          | -70        | 920                    | 1449                    | 173                      |
| BM I                                  | 2.64  | 16.97    | —    | -1110          | -1218         | -993       |                        |                         |                          |
| BM II                                 | -0.60 | 18.91    | —    | -1659          | -1768         | -1541      |                        |                         |                          |
| BM III                                | 12.96 | 19.68    | 0.48 | -410           | -519          | -291       |                        |                         |                          |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 9.12  | 8.35     | 0.66 | 225            | 121           | 337        | 315                    | 732                     | 240                      |
| TRS I                                 | 9.00  | 8.00     | 0.66 | 232            | 128           | 345        | 322                    | 739                     | 248                      |
| TRS II                                | 8.64  | 7.97     | 0.62 | 192            | 88            | 304        | 282                    | 698                     | 207                      |
| TRS III                               | 8.64  | 8.63     | 0.59 | 157            | 53            | 270        | 248                    | 665                     | 215                      |
| BM I                                  | 6.24  | 8.87     | 0.31 | -90            | -194          | 23         |                        |                         |                          |
| BM II                                 | 2.88  | 10.29    | —    | -510           | -614          | -397       |                        |                         |                          |
| BM III                                | 9.84  | 13.03    | 0.48 | -16            | -122          | 99         |                        |                         |                          |

*Note:* The table displays performance measures for the active mean-variance investment strategies based on monthly excess return predictions, with the volatility fixed at the median level during the 10-year rolling window that is used for model specification and estimation. Results are shown for medium transaction costs (0.1%) and with RRA=6. See Table 2.3 for further details.

increase substantially for almost all cases. Especially for the third benchmark model, which uses the historical average excess return as prediction, the performance fees almost double. This clearly shows how this benchmark model benefits from volatility timing rather than market timing in returns, confirming the results of Fleming *et al.* (2001, 2003). The benefits of volatility timing do not seem so pronounced for factor-augmented models as the increase of the performance fee for switching from static portfolios is very limited. When portfolio weights are completely restricted, in almost all cases the factor-augmented models have positive performance fees, regardless whether the competing strategies are static or dynamic but based on benchmark models.

## 2.6 Conclusions

While the basic question whether stock returns are predictable at all has not been settled conclusively, the voluminous literature on this topic seems to agree that individual macroeconomic variables do not have predictive ability for returns, at least not for prolonged periods of time. The relation between stock returns and individual macroeconomic predictors appears to be subject to relatively frequent structural breaks. Hence, when a given macroeconomic variable may appear to be useful for forecasting stock returns over a certain period, this might very well break down at some point. Put differently, studying return predictability with a limited number of individual macroeconomic variables over a long time span is unlikely to find positive results.

While the issue of return predictability is still disputed, there appears to be more consensus that the level of volatility is predictable to some extent. In addition, several studies document a relation between stock market volatility and macroeconomic conditions, with volatility being substantially higher during business cycle recessions. Nevertheless, it is unclear a priori which specific macroeconomic variables have the most relevant information for future volatility.

The point of departure for this paper is that in reality we may consider numerous macroeconomic variables, and their combined information may be useful for predicting future stock returns and volatility to some extent. Based upon an empirical analysis we document that factors extracted from a large set of macroeconomic variables do indeed have predictive ability for monthly US excess stock returns and volatility for the period 1980-2005. Factor-augmented predictive regression models improve upon both benchmark models that only include valuation ratios and interest rate related variables and possibly individual macro variables, as well as naive predictors such as the historical average excess return and previous month's volatility. The improvements in out-of-sample forecast accu-

racy are both statistically and economically significant. The factor-augmented predictive regressions have superior market timing ability and volatility timing ability, while a mean-variance investor would be willing to pay an annual performance fee of several hundreds of basis points to switch from the predictions offered by the benchmark models to those of the factor-augmented models. An important reason for the superior performance of the factor-augmented predictive regressions is the stability of their forecast accuracy, whereas the benchmark models suffer from a forecast breakdown during the 1990s.



## 2.A Appendices

### 2.A.1 Thresholding rules

This appendix provides descriptions of the two procedures for preselection of a subset of macro variables that are used for constructing the factors that enter the predictive regressions. Further details of the thresholding rules can be found in Bai and Ng (2008).

#### Hard thresholding

The ‘hard’ thresholding procedure selects macro variables  $x_{it}$  based on their individual predictive content for the excess returns  $r_{e,t+1}$ . The algorithm for hard thresholding is as follows:

1. For each macroeconomic predictor  $x_{it}$ ,  $i = 1, 2, \dots, k$ , estimate the regression

$$r_{e,t+1} = \beta_0 + \beta_{x,i}x_{it} + \beta'_{z,i}z_t + \varepsilon_{i,t+1},$$

by least squares, where  $r_{e,t+1}$  denotes the excess stock return in period  $t + 1$  and  $z_t$  is the  $(l \times 1)$  vector of valuation ratios and interest rate related variables that are always included.

2. Sort the macro variables in decreasing order of the absolute value of the  $t$ -statistic of  $\hat{\beta}_{x,i}$ , denoted  $t_i$ .
3. Choose a (one-sided) significance level  $\alpha$  and select the subset of predictors for which  $|t_i|$  exceeds the corresponding threshold.
4. Construct the factors  $f_t$  by means of the principal components for the selected subset of predictors.

#### Soft thresholding

The ‘soft’ thresholding procedure takes other candidate macro variables into account when selecting the predictors, based on the Least Angle Regression (LARS) methodology. Predictors are added to the model one at a time, starting with the variable that is most correlated with the excess return. This predictor’s coefficient is increased from its starting value of zero, up to the point where the residual is equally correlated with the predictor chosen initially and a second predictor. This second predictor is then added to the so-called ‘active set’, and the coefficients on both predictors in this set are now simultaneously

increased, in such a way that the residual is equally correlated with these two predictors and with a third predictor, which then enters the active set. This process can be repeated until the set of candidate predictors is exhausted (provided the number of candidates  $k$  is less than the number of observations  $T$ ). In this way, a complete ranking of the candidate predictors is obtained.

Technically, the  $j$ -th iteration of the LARS algorithm proceeds as follows.

1. Compute the correlations  $\hat{c}_{(j)}$  between the candidate predictor variables and the residual excess return after the  $(j - 1)$ -st iteration (thus, with  $j - 1$  candidate predictors already included in the active set). Collecting the predictors (which are assumed to be standardized to have mean zero and variance one) in the  $(T \times k)$  matrix  $X$  and denoting the estimate of the expected excess return obtained in the  $(j - 1)$ -st iteration as  $\hat{\mu}_{(j-1)}$ , we have  $\hat{c}_{(j)} = X'(r_e - \hat{\mu}_{(j-1)})$ .
2. Identify the candidate predictors with the largest absolute correlations and define  $\mathcal{S}_{(j)}$  as the set of the corresponding indices, that is

$$\hat{c}_{(j),\max} = \max_i \{|\hat{c}_{(j),i}|\} \quad \text{and} \quad \mathcal{S}_{(j)} = \{i : |\hat{c}_{(j),i}| = \hat{c}_{(j),\max}\},$$

where  $\hat{c}_{(j),i}$  denotes the  $i$ -th element of  $\hat{c}_{(j)}$ .

3. Define the active matrix  $X_{(j)}$  corresponding to  $\mathcal{S}_{(j)}$  by ‘signing’ the predictors, that is

$$X_{(j)} = (s_{(j),i} X_i)_{i \in \mathcal{J}} \quad \text{with} \quad s_{(j),i} = \text{sign}(\hat{c}_{(j),i}).$$

4. Let  $G_{(j)} = X_{(j)}' X_{(j)}$  and  $A_{(j)} = (\iota_{(j)}' G_{(j)}' \iota_{(j)})^{-1/2}$  where  $\iota_{(j)}$  is a vector of ones corresponding to the size of  $\mathcal{S}_{(j)}$ . Then define  $u_{(j)} = X_{(j)} w_{(j)}$ , where  $w_{(j)} = A_{(j)} G_{(j)}^{-1} \iota_{(j)}$ , and  $a_{(j)} \equiv X' u_{(j)}$ .

5. Update the estimate of the expected excess return as  $\hat{\mu}_{(j)} = \hat{\mu}_{(j-1)} + \hat{\gamma} u_{(j)}$ , where

$$\hat{\gamma} = \min_{i \in \mathcal{S}_{(j)}^c}^+ \left\{ \frac{\hat{c}_{(j)} - \hat{c}_{(j),i}}{\hat{A}_{(j)} - \hat{a}_{(j),i}}, \frac{\hat{c}_{(j)} + \hat{c}_{(j),i}}{\hat{A}_{(j)} + \hat{a}_{(j),i}} \right\}$$

with  $\min^+$  indicating that the minimum is taken over only positive components. The algorithm is initialized with  $\hat{\mu}_{(0)} = 0$ .

## 2.A.2 Macroeconomic variables

Table A.1: Base Set of Predictors used for Extracting Factors

| Short Name                      | Transf.      | Description                                                             |
|---------------------------------|--------------|-------------------------------------------------------------------------|
| <i>Output and Income</i>        |              |                                                                         |
| PI                              | $\Delta \ln$ | Personal Income (AR, Bil. Chain 2000\$) (TCB)                           |
| PI less transfers               | $\Delta \ln$ | Personal Income Less Transfer Payments (AR, Bil. Chain 2000\$) (TCB)    |
| IP: total                       | $\Delta \ln$ | Industrial Production Index - Total Index                               |
| IP: products                    | $\Delta \ln$ | Industrial Production Index - Products, Total                           |
| IP: final prod                  | $\Delta \ln$ | Industrial Production Index - Final Products                            |
| IP: cons gds                    | $\Delta \ln$ | Industrial Production Index - Consumer Goods                            |
| IP: cons dble                   | $\Delta \ln$ | Industrial Production Index - Durable Consumer Good                     |
| IP: cons nondble                | $\Delta \ln$ | Industrial Production Index - Nondurable Consumer Good                  |
| IP: bus eqpt                    | $\Delta \ln$ | Industrial Production Index - Business Equipment                        |
| IP: matls                       | $\Delta \ln$ | Industrial Production Index - Materials                                 |
| IP: dble mats                   | $\Delta \ln$ | Industrial Production Index - Durable Goods Materials                   |
| IP: nondble mats                | $\Delta \ln$ | Industrial Production Index - Nondurable Goods Materials                |
| IP: mfg                         | $\Delta \ln$ | Industrial Production Index - Manufacturing (Sic)                       |
| IP: res util                    | $\Delta \ln$ | Industrial Production Index - Residential Utilities es                  |
| IP: fuels                       | $\Delta \ln$ | Industrial Production Index - Fuels                                     |
| NAPM prodn                      | lv           | NAPM Production Index (Percent)                                         |
| Cap util                        | $\Delta \ln$ | Capacity Utilization (Mfg) (TCB)                                        |
| <i>Employment and hours</i>     |              |                                                                         |
| Help wanted index               | lv           | Index Of Help-Wanted Advertising In Newspapers (1967=100;Sa)            |
| Help wan/emp                    | lv           | Employment: Ratio; Help-Wanted Ads: No. Unemployed Clf                  |
| Emp CPS total                   | $\Delta \ln$ | Civilian Labor Force: Employed, Total (Thous. Sa)                       |
| Emp CPS nonag                   | $\Delta \ln$ | Civilian Labor Force: Employed, Nonagric. Industries (Thous.,Sa)        |
| U: all                          | lv           | Unemployment Rate: All Workers, 16 Years & Over (%Sa)                   |
| U: mean duration                | $\Delta \ln$ | Unemploy.By Duration: Average(Mean)Duration In Weeks (Sa)               |
| U < 5 wks                       | $\Delta \ln$ | Unemploy.By Duration: Persons Unempl.Less Than 5 Wks (Thous.,Sa)        |
| U 5-14 wks                      | $\Delta \ln$ | Unemploy.By Duration: Persons Unempl.5 To 14 Wks (Thous.,Sa)            |
| U 15+ wks                       | $\Delta \ln$ | Unemploy.By Duration: Persons Unempl.15 Wks + (Thous.,Sa)               |
| U 15-26 wks                     | $\Delta \ln$ | Unemploy.By Duration: Persons Unempl.15 To 26 Wks (Thous.,Sa)           |
| U 27+ wks                       | $\Delta \ln$ | Unemploy.By Duration: Persons Unempl.27 Wks + (Thous.,Sa)               |
| UI claims                       | $\Delta \ln$ | Average Weekly Initial Claims, Unemploy. Insurance (Thous.) (TCB)       |
| Emp: total                      | $\Delta \ln$ | Employees On Nonfarm Payrolls: Total Private                            |
| Emp: gds prod                   | $\Delta \ln$ | Employees On Nonfarm Payrolls-Goods-Producing                           |
| Emp: mining                     | $\Delta \ln$ | Employees On Nonfarm Payrolls-Mining                                    |
| Emp: const                      | $\Delta \ln$ | Employees On Nonfarm Payrolls-Construction                              |
| Emp: mfg                        | $\Delta \ln$ | Employees On Nonfarm Payrolls-Manufacturing                             |
| Emp: dble gds                   | $\Delta \ln$ | Employees On Nonfarm Payrolls-Durable Goods                             |
| Emp: nondbles                   | $\Delta \ln$ | Employees On Nonfarm Payrolls-Nondurable Goods                          |
| Emp: services                   | $\Delta \ln$ | Employees On Nonfarm Payrolls-Service-Providing                         |
| Emp: TTU                        | $\Delta \ln$ | Employees On Nonfarm Payrolls-Trade, Transportation, And Utilities      |
| Emp: wholesale                  | $\Delta \ln$ | Employees On Nonfarm Payrolls-Wholesale Trade                           |
| Emp: retail                     | $\Delta \ln$ | Employees On Nonfarm Payrolls-Retail Trade                              |
| Emp: FIRE                       | $\Delta \ln$ | Employees On Nonfarm Payrolls-Financial Activities                      |
| Emp: Govt                       | $\Delta \ln$ | Employees On Nonfarm Payrolls-Government                                |
| Avg hrs                         | lv           | Avg Weekly Hrs of Workers On Private Nonfarm Payrolls - Goods-Producing |
| Overtime: mfg                   | $\Delta \ln$ | Avg Weekly Hrs of Workers On Private Nonfarm Payrolls - Mfg Overtime    |
| Avg hrs: mfg                    | lv           | Average Weekly Hours, Mfg. (Hours) (TCB)                                |
| NAPM empl                       | lv           | Napm Employment Index (Percent)                                         |
| <i>Sales</i>                    |              |                                                                         |
| Retail sales                    | $\Delta \ln$ | Sales Of Retail Stores (Mil. Chain 2000\$) (TCB)                        |
| M&T sales                       | $\Delta \ln$ | Manufacturing And Trade Sales (Mil. Chain 1996\$) (TCB)                 |
| <i>Consumption</i>              |              |                                                                         |
| Consumption                     | $\Delta \ln$ | Real Consumption (AC) a0m224/gmdc (a0m224 is from TCB)                  |
| <i>Housing starts and sales</i> |              |                                                                         |
| HStarts: Total                  | ln           | Housing Starts:Nonfarm(1947-58);Total Farm&Nonfarm(1959-) (Thous.,Saar) |
| HStarts: NE                     | ln           | Housing Starts:Northeast (Thous.U.)S.A.                                 |
| HStarts: MW                     | ln           | Housing Starts:Midwest(Thous.U.)S.A.                                    |
| HStarts: South                  | ln           | Housing Starts:South (Thous.U.)S.A.                                     |
| HStarts: West                   | ln           | Housing Starts:West (Thous.U.)S.A.                                      |

continued on next page

*continued from previous page*

| Short Name                                  | Transf. | Description                                                                  |
|---------------------------------------------|---------|------------------------------------------------------------------------------|
| BP: Total                                   | ln      | Housing Authorized: Total New Priv Housing Units (Thous.,Saar)               |
| BP: NE                                      | ln      | Houses Authorized By Build. Permits:Northeast(Thou.U.)S.A                    |
| BP: MW                                      | ln      | Houses Authorized By Build. Permits:Midwest(Thou.U.)S.A.                     |
| BP: South                                   | ln      | Houses Authorized By Build. Permits:South(Thou.U.)S.A.                       |
| BP: West                                    | ln      | Houses Authorized By Build. Permits:West(Thou.U.)S.A.                        |
| <i>Inventories</i>                          |         |                                                                              |
| NAPM invent                                 | lv      | Napm Inventories Index (Percent)                                             |
| M&T invent                                  | Δln     | Manufacturing And Trade Inventories (Bil. Chain 2000\$) (TCB)                |
| M&T invent/sales                            | Δlv     | Ratio, Mfg. And Trade Inventories To Sales (Based On Chain 2000\$) (TCB)     |
| <i>Orders</i>                               |         |                                                                              |
| PMI                                         | lv      | Purchasing Managers' Index (Sa)                                              |
| NAPM new ordrs                              | lv      | Napm New Orders Index (Percent)                                              |
| NAPM vendor del                             | lv      | Napm Vendor Deliveries Index (Percent)                                       |
| Orders: cons gds                            | Δln     | Mfrs' New Orders, Consumer Goods And Materials (Bil. Chain 1982\$) (TCB)     |
| Orders: cap gds                             | Δln     | Mfrs' New Orders, Nondefense Capital Goods (Mil. Chain 1982\$) (TCB)         |
| <i>Exchange rates</i>                       |         |                                                                              |
| Ex rate: avg                                | Δln     | United States;Effective Exchange Rate(Merm)(Index No.)                       |
| Ex rate: Switz                              | Δln     | Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.\$)                  |
| Ex rate: Japan                              | Δln     | Foreign Exchange Rate: Japan (Yen Per U.S.\$)                                |
| Ex rate: UK                                 | Δln     | Foreign Exchange Rate: United Kingdom (Cents Per Pound)                      |
| Ex rate: Canada                             | Δln     | Foreign Exchange Rate: Canada (Canadian\$ Per U.S.\$)                        |
| <i>Money and credit quantity aggregates</i> |         |                                                                              |
| M1                                          | Δln     | Money Stock:M1(Curr,Trav.Cks,Dep,Other Ck'able Dep)(Bil\$,Sa)                |
| M2                                          | Δln     | Money Stock:M2M1+O'night Rps, Euro\$,G/P&B/D Mmmfs&Sav&Sm Time Dep(Bil\$,Sa) |
| M2 (real)                                   | Δln     | Money Supply:M2 In 1996 Dollars (Bci)                                        |
| MB                                          | Δln     | Monetary Base, Adj For Reserve Requirement Changes(Mil\$,Sa)                 |
| Reserves tot                                | Δln     | Depository Inst Reserves:Total, Adj For Reserve Req Chgs(Mil\$,Sa)           |
| Reserves nonbor                             | Δln     | Depository Inst Reserves:Nonborrowed,Adj Res Req Chgs(Mil\$,Sa)              |
| C&I loans                                   | Δln     | Commercial&Industrial Loans Outstanding In 1996 Dollars (Bci)                |
| C&I loans                                   | lv      | Wkly Rp Lg Com'l Banks:Net Change Com'l&Indus Loans(Bil\$,Saar)              |
| Cons credit                                 | Δln     | Consumer Credit Outstanding-Nonrevolving(G19)                                |
| Inst cred/PI                                | Δlv     | Ratio, Consumer Installment Credit To Personal Income (Pct.) (TCB)           |
| <i>Price indexes</i>                        |         |                                                                              |
| PPI: fin gds                                | Δln     | Producer Price Index: Finished Goods (82=100,Sa)                             |
| PPI: cons gds                               | Δln     | Producer Price Index: Finished Consumer Goods (82=100,Sa)                    |
| PPI: int mat'ls                             | Δln     | Producer Price Index: Intermed Mat.Supplies & Components(82=100,Sa)          |
| PPI: crude mat'ls                           | Δln     | Producer Price Index: Crude Materials (82=100,Sa)                            |
| Com: spot price                             | Δln     | Spot market price index: bls&crb: all commodities(1967=100)                  |
| NAPM com price                              | lv      | Napm Commodity Prices Index (Percent)                                        |
| CPI-U: all                                  | Δln     | Cpi-U: All Items (82-84=100,Sa)                                              |
| CPI-U: apparel                              | Δln     | Cpi-U: Apparel&Upkeep (82-84=100,Sa)                                         |
| CPI-U: transp                               | Δln     | Cpi-U: Transportation (82-84=100,Sa)                                         |
| CPI-U: medical                              | Δln     | Cpi-U: Medical Care (82-84=100,Sa)                                           |
| CPI-U: comm.                                | Δln     | Cpi-U: Commodities (82-84=100,Sa)                                            |
| CPI-U: dbles                                | Δln     | Cpi-U: Durables (82-84=100,Sa)                                               |
| CPI-U: services                             | Δln     | Cpi-U: Services (82-84=100,Sa)                                               |
| CPI-U: ex food                              | Δln     | Cpi-U: All Items Less Food (82-84=100,Sa)                                    |
| CPI-U: ex shelter                           | Δln     | Cpi-U: All Items Less Shelter (82-84=100,Sa)                                 |
| CPI-U: ex med                               | Δln     | Cpi-U: All Items Less Midical Care (82-84=100,Sa)                            |
| PCE defl                                    | Δln     | Pce, Impl Pr Defl:Pce (1987=100)                                             |
| PCE defl: dlbes                             | Δln     | Pce, Impl Pr Defl:Pce; Durables (1987=100)                                   |
| PCE defl: nondble                           | Δln     | Pce, Impl Pr Defl:Pce; Nondurables (1996=100)                                |
| PCE defl: services                          | Δln     | Pce, Impl Pr Defl:Pce; Services (1987=100)                                   |
| <i>Average hourly earnings</i>              |         |                                                                              |
| AHE: goods                                  | Δln     | Avg Hourly Earnings of Workers On Private Nonfarm Payrolls - Goods-Producing |
| AHE: const                                  | Δln     | Avg Hourly Earnings of Workers On Private Nonfarm Payrolls - Construction    |
| AHE: mfg                                    | Δln     | Avg Hourly Earnings of Workers On Private Nonfarm Payrolls - Manufacturing   |
| <i>Miscellaneous</i>                        |         |                                                                              |
| Consumer expect                             | Δlv     | U. Of Mich. Index Of Consumer Expectations(Bcd-83)                           |

*Note:* This table lists the individual macro series used in factor extraction. The column headed 'Transf.' indicates the transformation applied to original series: lv = no transformation (levels are used), ln = logarithm of the level, Δlv = first differences of the level, and Δln = first differences of the log level.

## 2.A.3 Alternative levels of risk aversion

Table A.1: Performance of active trading strategies for RRA=2

|                                       | $\mu$ | $\sigma$ | SR   | $\Delta_{100}$ | $\Delta_{50}$ | $\Delta_0$ | $\Delta_{BM\ I}$ | $\Delta_{BM\ II}$ | $\Delta_{BM\ III}$ |
|---------------------------------------|-------|----------|------|----------------|---------------|------------|------------------|-------------------|--------------------|
| Panel A: January 1980 - December 2005 |       |          |      |                |               |            |                  |                   |                    |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                  |                   |                    |
| SFM                                   | 17.40 | 21.41    | 0.55 | 147            | 371           | 714        | 447              | 1146              | 228                |
| TRS I                                 | 16.80 | 19.19    | 0.59 | 187            | 405           | 747        | 480              | 1178              | 262                |
| TRS II                                | 15.60 | 19.26    | 0.52 | 49             | 273           | 615        | 348              | 1046              | 129                |
| TRS III                               | 16.20 | 20.06    | 0.52 | 88             | 312           | 655        | 387              | 1086              | 169                |
| BM I                                  | 12.96 | 21.41    | 0.35 | -300           | -75           | 268        |                  |                   |                    |
| BM II                                 | 6.48  | 22.55    | 0.03 | -1000          | -775          | -432       |                  |                   |                    |
| BM III                                | 15.84 | 23.04    | 0.45 | -81            | 144           | 487        |                  |                   |                    |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                  |                   |                    |
| SFM                                   | 12.36 | 11.22    | 0.59 | -16            | 207           | 549        | 151              | 435               | 100                |
| TRS I                                 | 12.48 | 8.80     | 0.80 | 46             | 269           | 610        | 213              | 497               | 162                |
| TRS II                                | 11.76 | 9.01     | 0.69 | -37            | 186           | 527        | 130              | 414               | 79                 |
| TRS III                               | 11.52 | 9.84     | 0.59 | -78            | 145           | 486        | 89               | 373               | 38                 |
| BM I                                  | 10.92 | 11.15    | 0.48 | -167           | 56            | 398        |                  |                   |                    |
| BM II                                 | 8.04  | 11.33    | 0.21 | -451           | -228          | 113        |                  |                   |                    |
| BM III                                | 11.88 | 13.16    | 0.48 | -116           | 107           | 449        |                  |                   |                    |
| Panel B: January 1980 - December 1992 |       |          |      |                |               |            |                  |                   |                    |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                  |                   |                    |
| SFM                                   | 21.48 | 23.69    | 0.59 | 210            | 441           | 805        | -124             | 868               | 364                |
| TRS I                                 | 19.56 | 21.03    | 0.55 | 144            | 374           | 738        | -190             | 801               | 298                |
| TRS II                                | 19.08 | 20.89    | 0.55 | 101            | 332           | 695        | -232             | 758               | 255                |
| TRS III                               | 21.12 | 22.38    | 0.59 | 243            | 473           | 838        | -91              | 901               | 397                |
| BM I                                  | 22.56 | 23.52    | 0.62 | 334            | 565           | 929        |                  |                   |                    |
| BM II                                 | 12.96 | 24.04    | 0.21 | -659           | -428          | -63        |                  |                   |                    |
| BM III                                | 17.28 | 22.66    | 0.42 | -154           | 76            | 441        |                  |                   |                    |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                  |                   |                    |
| SFM                                   | 15.12 | 12.75    | 0.59 | -11            | 218           | 580        | 10               | 259               | 154                |
| TRS I                                 | 14.88 | 9.25     | 0.80 | 48             | 277           | 638        | 68               | 317               | 213                |
| TRS II                                | 14.40 | 9.35     | 0.73 | -2             | 227           | 588        | 18               | 267               | 163                |
| TRS III                               | 15.12 | 11.99    | 0.62 | 2              | 231           | 593        | 23               | 272               | 167                |
| BM I                                  | 15.00 | 12.82    | 0.59 | -20            | 209           | 571        |                  |                   |                    |
| BM II                                 | 12.24 | 11.78    | 0.38 | -270           | -41           | 321        |                  |                   |                    |
| BM III                                | 13.68 | 13.27    | 0.45 | -165           | 64            | 426        |                  |                   |                    |
| Panel C: January 1993 - December 2005 |       |          |      |                |               |            |                  |                   |                    |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                  |                   |                    |
| SFM                                   | 13.44 | 18.78    | 0.52 | 82             | 300           | 622        | 1004             | 1418              | 93                 |
| TRS I                                 | 14.04 | 17.11    | 0.62 | 214            | 432           | 754        | 1136             | 1549              | 226                |
| TRS II                                | 12.00 | 17.39    | 0.48 | -7             | 211           | 533        | 915              | 1329              | 5                  |
| TRS III                               | 12.72 | 18.46    | 0.48 | 24             | 242           | 564        | 947              | 1360              | 36                 |
| BM I                                  | 3.24  | 18.67    | 0.00 | -923           | -704          | -383       |                  |                   |                    |
| BM II                                 | 0.00  | 20.78    | -    | -1338          | -1120         | -797       |                  |                   |                    |
| BM III                                | 14.40 | 23.38    | 0.45 | -12            | 207           | 530        |                  |                   |                    |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                  |                   |                    |
| SFM                                   | 9.60  | 9.39     | 0.66 | -22            | 195           | 516        | 291              | 610               | 46                 |
| TRS I                                 | 10.08 | 8.24     | 0.80 | 43             | 260           | 581        | 356              | 675               | 112                |
| TRS II                                | 9.00  | 8.56     | 0.62 | -72            | 145           | 465        | 240              | 559               | 4                  |
| TRS III                               | 9.00  | 9.53     | 0.55 | -95            | 122           | 442        | 217              | 536               | -27                |
| BM I                                  | 6.72  | 9.04     | 0.35 | -313           | -96           | 225        |                  |                   |                    |
| BM II                                 | 3.84  | 10.77    | 0.03 | -632           | -415          | -94        |                  |                   |                    |
| BM III                                | 10.08 | 13.09    | 0.48 | -69            | 149           | 470        |                  |                   |                    |

*Note:* The table displays performance measures for the active mean-variance investment strategies based on monthly excess return and volatility predictions for the period January 1980 - December 1992 with relative risk aversion  $\gamma$  in (2.16) set equal to 2. Results are shown for medium transaction costs (0.1%). See Table 2.3 for further details.

Table A.2: Performance of active trading strategies for RRA=10

|                                       | $\mu$ | $\sigma$ | SR   | $\Delta_{100}$ | $\Delta_{50}$ | $\Delta_0$ | $\Delta_{\text{BM I}}$ | $\Delta_{\text{BM II}}$ | $\Delta_{\text{BM III}}$ |
|---------------------------------------|-------|----------|------|----------------|---------------|------------|------------------------|-------------------------|--------------------------|
| Panel A: January 1980 - December 2005 |       |          |      |                |               |            |                        |                         |                          |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 10.44 | 11.85    | 0.42 | 168            | -382          | -280       | 68                     | 458                     | -282                     |
| TRS I                                 | 11.40 | 13.20    | 0.45 | 79             | -481          | -377       | -23                    | 373                     | -379                     |
| TRS II                                | 10.56 | 12.78    | 0.38 | 61             | -495          | -392       | -40                    | 354                     | -394                     |
| TRS III                               | 11.76 | 11.99    | 0.52 | 287            | -264          | -162       | 187                    | 577                     | -164                     |
| BM I                                  | 9.24  | 11.43    | 0.31 | 99             | -448          | -346       |                        |                         |                          |
| BM II                                 | 7.56  | 13.13    | 0.14 | -294           | -853          | -750       |                        |                         |                          |
| BM III                                | 10.08 | 9.08     | 0.48 | 437            | -97           | 2          |                        |                         |                          |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 9.84  | 7.10     | 0.59 | 582            | 56            | 153        | 105                    | 243                     | 79                       |
| TRS I                                 | 9.72  | 6.65     | 0.62 | 603            | 79            | 176        | 127                    | 266                     | 101                      |
| TRS II                                | 9.48  | 6.58     | 0.59 | 588            | 64            | 160        | 112                    | 251                     | 86                       |
| TRS III                               | 9.96  | 6.86     | 0.62 | 612            | 87            | 184        | 136                    | 274                     | 110                      |
| BM I                                  | 9.00  | 7.41     | 0.45 | 478            | -49           | 48         |                        |                         |                          |
| BM II                                 | 7.80  | 7.59     | 0.28 | 339            | -188          | -91        |                        |                         |                          |
| BM III                                | 9.36  | 7.52     | 0.48 | 505            | -23           | 75         |                        |                         |                          |
| Panel B: January 1980 - December 1992 |       |          |      |                |               |            |                        |                         |                          |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 14.28 | 14.27    | 0.45 | 92             | -581          | -487       | -154                   | 498                     | -586                     |
| TRS I                                 | 13.92 | 16.28    | 0.38 | -293           | -990          | -892       | -547                   | 126                     | -995                     |
| TRS II                                | 13.44 | 15.76    | 0.38 | -243           | -934          | -837       | -495                   | 172                     | -939                     |
| TRS III                               | 14.52 | 13.99    | 0.48 | 163            | -507          | -412       | -81                    | 568                     | -511                     |
| BM I                                  | 15.12 | 13.89    | 0.55 | 244            | -425          | -331       |                        |                         |                          |
| BM II                                 | 11.16 | 15.45    | 0.24 | -414           | -1101         | -1004      |                        |                         |                          |
| BM III                                | 11.40 | 7.27     | 0.52 | 627            | 5             | 92         |                        |                         |                          |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 12.60 | 8.28     | 0.59 | 668            | 41            | 129        | -18                    | 170                     | -4                       |
| TRS I                                 | 11.64 | 7.69     | 0.52 | 617            | -7            | 81         | -66                    | 121                     | -51                      |
| TRS II                                | 11.52 | 7.59     | 0.52 | 611            | -13           | 74         | -72                    | 115                     | -58                      |
| TRS III                               | 12.12 | 7.79     | 0.59 | 658            | 33            | 121        | -26                    | 162                     | -11                      |
| BM I                                  | 12.96 | 8.42     | 0.62 | 687            | 59            | 147        |                        |                         |                          |
| BM II                                 | 11.04 | 8.31     | 0.38 | 498            | -129          | -41        |                        |                         |                          |
| BM III                                | 11.40 | 6.69     | 0.55 | 664            | 44            | 131        |                        |                         |                          |
| Panel C: January 1993 - December 2005 |       |          |      |                |               |            |                        |                         |                          |
| $w_{t+1}^* \in [-1, 2]$               |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 6.60  | 8.63     | 0.35 | 238            | -200          | -92        | 272                    | 420                     | -7                       |
| TRS I                                 | 8.76  | 9.08     | 0.59 | 420            | -19           | 89         | 454                    | 603                     | 174                      |
| TRS II                                | 7.80  | 8.80     | 0.48 | 340            | -98           | 11         | 375                    | 524                     | 95                       |
| TRS III                               | 9.00  | 9.49     | 0.55 | 400            | -41           | 68         | 434                    | 584                     | 153                      |
| BM I                                  | 3.24  | 7.93     | -    | -34            | -469          | -362       |                        |                         |                          |
| BM II                                 | 3.84  | 10.18    | 0.03 | -185           | -629          | -519       |                        |                         |                          |
| BM III                                | 8.64  | 10.57    | 0.48 | -185           | -629          | -519       |                        |                         |                          |
| $w_{t+1}^* \in [0, 1]$                |       |          |      |                |               |            |                        |                         |                          |
| SFM                                   | 6.96  | 5.54     | 0.59 | 498            | 69            | 175        | 220                    | 313                     | 156                      |
| TRS I                                 | 7.80  | 5.40     | 0.76 | 588            | 159           | 265        | 311                    | 403                     | 247                      |
| TRS II                                | 7.44  | 5.27     | 0.73 | 564            | 136           | 242        | 287                    | 380                     | 223                      |
| TRS III                               | 7.68  | 5.72     | 0.73 | 566            | 137           | 243        | 289                    | 381                     | 224                      |
| BM I                                  | 5.04  | 6.03     | 0.24 | 278            | -152          | -46        |                        |                         |                          |
| BM II                                 | 4.44  | 6.62     | 0.14 | 186            | -246          | -139       |                        |                         |                          |
| BM III                                | 7.32  | 8.24     | 0.45 | 347            | -89           | 19         |                        |                         |                          |

*Note:* The table displays performance measures for the active mean-variance investment strategies based on monthly excess return and volatility predictions for the period January 1980 - December 1992 with relative risk aversion  $\gamma$  in (2.16) set equal to 10. Results are shown for medium transaction costs (0.1%). See Table 2.3 for further details.

## Chapter 3

# Modeling and Estimation of Synchronization in Multistate Markov-Switching Models

### 3.1 Introduction

Synchronization of cyclical behavior of economic and financial time series variables has generated considerable interest in recent years. Typical examples include the synchronization of bull and bear phases in financial markets (see Edwards *et al.*, 2003; Bekaert *et al.*, 2005, among others) and the synchronization of business cycles in different countries or regions (see Bordo and Helbling, 2003; Artis *et al.*, 2004; Canova *et al.*, 2007; Kose *et al.*, 2003, 2008, among others). Much attention in this research area has been focused on the extreme cases of independence and perfect synchronization of cycles. In the first case, the cycles in different variables are purely idiosyncratic, while in the second case the variables are driven by a single common cycle such that regime shifts occur contemporaneously. A wide range of econometric techniques has been developed to examine these polar cases, ranging from (non-)parametric testing procedures (see Artis *et al.*, 1997; Harding and Pagan, 2006; Pesaran and Timmermann, 2009), to unobserved components models (see Koopman and Harvey, 1997), to Markov-Switching Vector Autoregressive (MS-VAR) models (see Krolzig, 1997; Artis *et al.*, 2004).

Perhaps not surprisingly, it is often found that neither independence nor perfect synchronization are adequate representations of the cyclical dynamics in economic and financial variables. A variety of mechanisms may lead to the intermediate case of ‘imperfect synchronization’. For example, it may be that cycles in individual variables are partly

due to a common component and partly due to idiosyncratic factors. This seems relevant when considering, for example, bull and bear phases in different stock markets (see Kose *et al.*, 2006, for example), or business cycles in different countries or regions, (see Kose *et al.*, 2003, 2008; Owyang *et al.*, 2005; Del Negro and Otrok, 2008, for example). A second possibility is that the different variables in fact share a single common cyclical component, but subject to different phase shifts. This approach is appropriate in the context of co-movement of asset returns in the sense that shifts from bull to bear phases and vice versa may occur earlier in some assets than in others due to differences in liquidity or gradual information diffusion, among other reasons. Similarly, leading indicator variables of the economy are expected to enter to different business cycle phases earlier than coincident indicator variables. Imperfect synchronization due to different phase shifts of a single common cycle has been incorporated by Koopman and Azevedo (2008) in unobserved components models, and by Hamilton and Perez-Quiros (1996) and Paap *et al.* (2009) in MS-VAR models.

In this paper, we extend the framework of MS-VAR models for describing imperfect synchronization due to different phase shifts of a single common cycle. We generalize the models of Hamilton and Perez-Quiros (1996) and Paap *et al.* (2009) in two important directions. First, we allow for multiple regimes and consider the possibility of different phase shifts for the different regimes. Second, we allow for regime-dependent volatilities and correlations, in addition to regime switching in the means. In this way we obtain a general framework where we can simultaneously estimate the degree of synchronization together with the unobserved regimes.

Allowing for multiple regimes is relevant in the context of financial markets as well as of business cycle analysis. For example, when modeling stock returns using Markov-switching models, existing evidence suggests that three or four regimes are required to characterize the dynamics adequately. Next to regimes representing bull and bear markets, additional regimes may be necessary to capture ‘crashes’ and ‘recoveries’ (Guidolin and Timmermann, 2006), or ‘bull market corrections’ and ‘bear market rallies’ (Maheu *et al.*, 2010). Similarly, Boldin (1996) and Clements and Krolzig (2003), among others, show that a three regime model provides a more accurate and robust characterization of the US business cycle. While the specification in Hamilton and Perez-Quiros (1996) can be applied to MS-VAR models with multiple regimes, their formulation allows only for ‘symmetric’ phase shifts, that is, the regime switches in one variable lead/lag transitions of the other variable by a fixed number of periods. This may be too restrictive given the empirical evidence that, for example, leading indicator variables have a longer lead time at business cycle peaks than at troughs, (see The Conference Board, 2001; Paap *et al.*,

2009, for example). Paap *et al.* (2009) in fact provide an MS-VAR specification that allows the phase shifts to be different depending on the type of transition. However, the formulation they use can only be applied to the MS-VAR model with two regimes.

Allowing for regime-specific volatilities may be of crucial importance for reliable and accurate identification of the regimes, especially in the context of asset returns. Together with the sign of the returns, the main identifying characteristic of bull and bear markets is their different levels of volatilities, see Maheu and McCurdy (2000) and Guidolin and Timmermann (2006), among others. This gives rise to additional complications in case of imperfectly synchronized cycles, however, as there is no unique way to define a proper covariance matrix. We demonstrate that this complication may be overcome by using the decomposition of the covariance matrix in volatilities and correlations. It is fairly straightforward to guarantee that the resulting covariance matrix is positive definite, for example, while still allowing for a fair amount of flexibility in specifying the regime dynamics of the volatilities and correlations.

We adopt a Bayesian approach for estimation and inference, based upon the aim to obtain the degree of synchronization by estimating the phase shifts in the cycle of one variable relative to the cycle of another variable. As these phase shifts are represented by means of discrete valued parameters in the model, frequentist estimation procedures are not feasible.

We illustrate the potential of our framework for modeling imperfect synchronization using an empirical application for size-based US stock portfolios. There is considerable evidence that returns on (portfolios of) large stocks lead returns on (portfolios of) small stocks, see Lo and MacKinlay (1990a), McQueen *et al.* (1996), Richardson and Peterson (1999), and Hou (2007), among others. Using monthly US large-cap and small-cap portfolio returns over the period 1963-2007, we find that a three-regime model with regime-dependent heteroscedasticity captures the dynamics of the portfolio returns most adequately. Allowing for different volatilities and correlations is of crucial importance for identification of the regimes. The regimes can be characterized as ‘boom’ (high average returns with moderate volatility and low correlation), ‘moderate’ (mean returns close to zero, low volatility and high correlations), and ‘bear/crash’ (negative mean return, high volatility and high correlation). We find heterogeneity in the phase shifts across different regimes, in the sense that the large-cap portfolio leads the small-cap portfolio for switches to the moderate regime by a single month, while the two portfolios switch contemporaneously into the boom and crash regimes. This suggests that during turmoil periods, with either good or bad news, information diffusion occurs rapidly and lead-lag effects are not present. News ‘travels slowly’ during more moderate times, in the sense that it is

incorporated only gradually into small stocks, such that their prices adjust with a delay compared to large stocks.

Several alternatives to MS-VAR models for analyzing synchronization of cycles are available. One possibility is to use non-parametric ‘dating algorithms’ for identifying ‘turning points’ or regime switches. Once these turning points are determined, independence and perfect synchronization may be tested in a straightforward manner (see Artis *et al.*, 1997; Harding and Pagan, 2006). However, this approach typically is limited to only two regimes (corresponding with bull and bear markets in financial markets or recession and expansion periods in business cycle analysis). Another alternative is provided by multivariate unobserved components models, where trend and (common or idiosyncratic) cycle components of the series are modeled separately, as in the similar cycle model of Koopman and Harvey (1997). Our motivation for adopting MS-VAR models is that the conventional turning points approach is intrinsically embedded in these models without defining content-specific decision rules. This obviously provides a general framework for analyzing synchronization, reconciling traditional dating algorithms with modern parametric econometric tools.

The remainder of the paper is organized as follows. In Section 3.2 we describe the general Markov-switching VAR model with different phase shifts for multiple regimes and with regime-dependent heteroskedasticity. We discuss the Bayesian estimation methodology with prior and likelihood specifications in Section 3.3, with full details being provided in the Appendix. In Section 3.4 we consider the empirical application to size-based portfolios. Finally, we conclude in Section 3.5.

## 3.2 Model specification

In this section we put forward the MS-VAR model with imperfect synchronization due to phase shifts of a single common cycle. The key feature of the model is that the amount of phase shift can be different across regimes. Furthermore, it allows the cycle to consist of any number of regimes  $J \geq 2$ , and accommodates the possibility of regime-dependent volatilities and correlations. To simplify the exposition, we describe the model for the bivariate case, but it can be extended to include multiple variables in a straightforward manner.

Let  $y_{1,t}$  and  $y_{2,t}$  denote the observations of the variables of interest in period  $t$ . We assume that the  $J$  regimes are characterized by different means of  $y_{1,t}$  and  $y_{2,t}$ . We also allow their variances and contemporaneous correlation to be regime dependent, as described in detail below, but we assume that autoregressive coefficients are identical

across regimes. In case of first-order autoregressive dynamics, our assumptions imply the model specification

$$\begin{aligned} y_{1,t} - \mu_{1,S_{1,t}} &= \phi_{1,1}(y_{1,t-1} - \mu_{1,S_{1,t-1}}) + \phi_{1,2}(y_{2,t-1} - \mu_{2,S_{2,t-1}}) + \varepsilon_{1,t}, \\ y_{2,t} - \mu_{2,S_{2,t}} &= \phi_{2,1}(y_{1,t-1} - \mu_{1,S_{1,t-1}}) + \phi_{2,2}(y_{2,t-1} - \mu_{2,S_{2,t-1}}) + \varepsilon_{2,t}, \end{aligned} \quad (3.1)$$

where  $S_{l,t} \in \{1, 2, \dots, J\}$ ,  $l = 1, 2$  are latent multinomial variables taking the value  $j$  if  $y_{l,t}$  is in regime  $j$  at time  $t$ ,  $\mu_{l,S_{l,t}} = \mathbb{E}[y_{l,t}|S_{l,t}]$  denotes the mean of  $y_{l,t}$  in the regime indicated by  $S_{l,t}$ , and

$$\begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix} \sim NID(0, \Sigma_t). \quad (3.2)$$

For now, we do not make the specification of the covariance matrix  $\Sigma_t$  explicit, but we return to this issue in detail below.

The model is completed by specifying the dynamic properties of the unobserved regime indicators  $S_{1,t}$  and  $S_{2,t}$ . We assume that  $S_{1,t}$  and  $S_{2,t}$  are first-order  $J$ -state homogenous Markov processes with constant transition probabilities

$$\Pr(S_{l,t} = j | S_{l,t-1} = i) = p_{ij,l}, \quad \text{for } i, j = 1, \dots, J \text{ and } l = 1, 2. \quad (3.3)$$

Without loss of generality, we assume that  $y_{1,t}$  is the ‘reference series’ and we define the properties of  $S_{2,t}$ , the regime indicator of  $y_{2,t}$ , relative to  $S_{1,t}$ . Different specifications of the relation between these two Markov processes imply different types of relations between the cycles of the two variables. Two extreme cases can be distinguished. First, we may assume that  $S_{1,t}$  and  $S_{2,t}$  are independent, that is, each variable has its own idiosyncratic cycle. Second, we can assume that  $S_{1,t}$  and  $S_{2,t}$  are identical, that is,

$$S_{2,t} = S_{1,t}, \quad (3.4)$$

or, put differently, there is a single cycle governing both variables and hence there is only one transition probability specification as in (3.3) where the subscript  $l$  is dropped from the notation. Following Harding and Pagan (2006), we refer to the latter case as ‘perfect synchronization’ (PS). It is useful to note that the specification with independent Markov processes  $S_{1,t}$  and  $S_{2,t}$  does not imply that the cycles in the two variables do not exhibit any synchronization at all. In fact, letting  $\lambda_{l,j}$  denote the unconditional probability that

$S_{l,t} = j$  for  $l = 1, 2$  and  $j = 1, \dots, J$ , the two cycles are in the same regime with probability

$$\Pr(S_{2,t} = S_{1,t}) = \sum_{j=1}^J \lambda_{1,j} \lambda_{2,j} > 0.$$

In practice, the degree of synchronization between cycles in different financial and macroeconomic variables may not be perfect but still higher than the expected level under independence as given above, see the empirical evidence in Harding and Pagan (2006) and Candelon *et al.* (2008, 2009), among others. Hence, neither perfect synchronization nor independence may be adequate representations of the relation between cycles. A natural and elegant approach to accommodate this possibility is to assume that the cycle in  $y_{1,t}$  leads/lags the cycle in  $y_{2,t}$  by  $\kappa$  periods, as suggested by Hamilton and Perez-Quiros (1996). In other words, there is a common cycle but it affects the different variables with a certain phase shift. We refer to this case as imperfect synchronization with ‘symmetric’ phase shifts (IS-SPS), where symmetry refers to the fact that all possible regime transitions in the two variables differ by the same number of time periods  $\kappa$ . This specification can be formulated as

$$S_{2,t-\kappa} = S_{1,t}. \tag{3.5}$$

Obviously, if  $\kappa$  equals zero, it reduces to the perfect synchronization case in (3.4).

The specification in (3.5) may still be too restrictive, in the sense that the phase shift of the cycle may not be the same for all possible regime transitions. For example, leading indicator variables typically have a considerably longer lead time at business cycle peaks than at troughs, see The Conference Board (2001). For this purpose, Paap *et al.* (2009) consider a two-regime model with possibly different phase shifts  $\kappa_{12}$  and  $\kappa_{21}$  for these two types of turning points.

In case of multiple regimes  $J > 2$ , several extensions of the specification in (3.5) are possible to allow for the possibility of such ‘asymmetric’ phase shifts. First, following Paap *et al.* (2009) we can assume a specific lead/lag structure for every possible transition from regime  $i$  to regime  $j$ . While this specification obviously provides a great deal of flexibility, it is not without problems as the number of phase shift parameters increases rapidly when the number of regimes  $J$  becomes larger. For a model with  $J$  regimes, it would result in  $J(J-1)$  phase shift parameters  $\kappa_{ij}$ ,  $i, j = 1, \dots, J$ ;  $i \neq j$ . In addition, the  $\kappa_{ij}$  parameters are identified only by specific transitions from regime  $i$  to regime  $j$ . Hence, in order to obtain accurate estimates of these parameters we need a fair number of all possible regime switches to occur.

If the regimes are ordered in a logical way (for example according to the magnitude of the mean of the reference series  $y_{1,t}$ ), another possible extension of (3.5) would be to assume identical phase shifts for all possible upward and downward regime changes. Independent of the number of regimes  $J$ , this would require estimating only two phase shifts parameters, one for transitions to regimes with a higher mean and one for transitions to regimes with a lower mean. However, as all regime changes to the same direction are assumed to be subject to identical phase shifts, the asymmetry in the lead/lag characteristics of the cycles is highly restricted.

Here we propose an alternative solution, which aims to strike a balance between flexibility and parsimony. Specifically, instead of specifying the phase shifts for each possible type of regime switch, we assume that the regime indicator  $S_{1,t}$  itself is shifted but allowing the amount of phase shift to be different across regimes. That is, we generalize (3.5) to the case of imperfect synchronization with ‘asymmetric’ phase shifts (IS-APS) as

$$S_{2,t-\kappa_{S_{1,t}}} = S_{1,t}, \quad (3.6)$$

where the subscript  $S_{1,t}$  to  $\kappa$  indicates that the regime indicator is shifted by a possibly different number of time periods for each regime. Hence, this specification involves a separate regime shift parameter  $\kappa_j$  for each regime  $j = 1, \dots, J$ . Put differently, we assume that the lead/lag time is different per regime, such that each regime in the second series starts later or earlier by  $\kappa_j$  periods. The advantage of this specification is that it provides a parsimonious solution (as the number of phase shift parameters grows only linearly with the number of regimes), while at the same time it is quite flexible and allows for considerable heterogeneity in the amount of phase shifts of the common cycle.

The specification in (3.6) is not ‘complete’, in the sense that it may lead to situations where for some time periods  $S_{2,t}$  is assigned multiple values or is not defined at all. To illustrate this problem explicitly, suppose the reference series  $y_{1,t}$  is in regime 1 up to period  $\tau - 1$  and switches to regime 2 in period  $\tau$ , where it remains in subsequent periods. Using the formulation in (3.6) for determining the regime of the second series implies a phase shift of  $\kappa_1$  periods until  $t = \tau - 1$  and  $\kappa_2$  periods from  $t = \tau$  onwards. In particular, for  $t = \tau - 1$  and  $\tau$  we have

$$\begin{aligned} S_{2,\tau-1-\kappa_1} &= S_{1,\tau-1}, \\ S_{2,\tau-\kappa_2} &= S_{1,\tau}. \end{aligned}$$

If  $\kappa_1 > \kappa_2$ , the shift at  $\tau - 1$  is greater than the shift at  $\tau$ . Hence, by applying this dating procedure we do not assign any of the regimes to the second variable for periods  $[\tau - \kappa_1, \tau - \kappa_2 - 1]$ . In the opposite case, we assign both regimes 1 and 2 simultaneously for periods  $[\tau - \kappa_2, \tau - 1 - \kappa_1]$ . This example demonstrates that we need a further decision rule to resolve these possible conflicts in determining the relevant regime for  $y_{2,t}$ . We suggest to use the rule that the regime with the larger amount of phase shift is assigned to such conflicting periods. By the regime with the larger amount of phase shift we mean the regime with the larger phase shift parameter in terms of its real value. Hence, in the example above, this rule implies assigning regime 1 to the second variable in case  $\kappa_1 > \kappa_2$  for time periods that are left undetermined by (3.6), that is, setting  $S_{2,t} = 1$  for  $t \in [\tau - \kappa_1, \tau - \kappa_2 - 1]$ . Similarly, we set  $S_{2,t} = 2$  for  $t \in [\tau - \kappa_2, \tau - 1 - \kappa_1]$  in the opposite case when  $\kappa_1 > \kappa_2$ .

The main advantage of the proposed decision rule to assign the regime with the larger amount of phase shift to conflicting periods is that it implies that the phase shift is equal to  $\kappa_j$  when entering regime  $j$ . Possible alternatives would not achieve this. For example, assigning the regime with the smaller amount of phase shift to conflicting periods would lead to a specification where the phase shift is equal to  $\kappa_j$  when leaving regime  $j$ .

In sum, the specification in (3.6) together with the decision rule to assign the regime with the larger  $\kappa_j$  parameter in case of conflicted time periods constitutes a general framework for imperfect synchronization with multiple regimes, which encompasses perfect synchronization and imperfect synchronization with symmetric phase shifts.

Finally, we return to the specification of the structure of the covariance matrix to complete the general model framework. As discussed before, in many applications assuming homoskedasticity for the error terms may not be appropriate. In fact, regime-dependent variances and co-variances may be crucial for accurate identification of the regimes. To facilitate the specification of a regime-dependent covariance matrix we decompose it into the volatilities and correlations, that is

$$\Sigma_t = D_t R_t D_t, \tag{3.7}$$

where  $D_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t})$  is a diagonal matrix with standard deviations of the error terms as diagonal elements and  $R_t$  is a matrix with ones on the diagonal and the correlation  $\rho_t$  as the off-diagonal element (see, for example, Barnard *et al.*, 2000).

In case of perfect synchronization, it is straightforward to specify a regime-dependent covariance matrix. The values of the variances and correlation just depend on the value of  $S_{1,t}$ . Imperfect synchronization (either with symmetric or asymmetric phase shifts) allows

for the possibility that the different series are in different regimes in a given period. Consequently, there are several ways of relating the variances and correlations to the regimes. The easiest solution is to maintain the same assumption as in the perfect synchronization case, namely that the variance and correlation regimes are completely determined by the regime indicator of the first series. This results in

$$\begin{aligned}\sigma_{l,t} &= \sigma_{l,S_{1,t}} \quad \text{for } l = 1, 2, \\ \rho_t &= \rho_{S_{1,t}},\end{aligned}\tag{3.8}$$

where we also allow for regime switching in the correlation parameter (see, for example, Pelletier, 2006).

Specification (3.8) implies perfect synchronization in the variance regimes of both series. In case phase shifts lead to imperfect synchronization in the means  $\mu_l$ ,  $l = 1, 2$ , this assumption may not be realistic. Instead, it seems more natural to relate the variances of two series to their individual regime indicator  $S_{1,t}$  and  $S_{2,t}$ . This leads to the specification

$$\begin{aligned}\sigma_{l,t} &= \sigma_{l,S_{l,t}} \quad \text{for } l = 1, 2, \\ \rho_t &= \rho_{S_{1,t}},\end{aligned}\tag{3.9}$$

where we use the regime indicator of the first series  $S_{1,t}$  to describe the regime of the correlation parameter.

### 3.3 Estimation

We use a Bayesian approach for estimation and inference in the MS-VAR model with imperfect synchronization due to asymmetric phase shifts using Markov Chain Monte Carlo (MCMC) techniques. Specifically, we use Gibbs sampling together with data augmentation (see Geman and Geman, 1984; Tanner and Wong, 1987) to obtain posterior results. The main reason for adopting a Bayesian approach is that we treat the lead/lag times  $\kappa_j$ ,  $j = 1, \dots, J$ , as unknown parameters to be estimated. Given that the  $\kappa_j$ 's are discrete, a frequentist approach (such as maximum likelihood combined with the EM algorithm) is infeasible. In Section 3.3.1 we derive the likelihood function of the model, while we discuss the specifications of the prior distributions in Section 3.3.2. In Section 3.3.3 we outline the Gibbs sampling algorithm for simulating from the posterior distribution. Finally, in Section 3.3.4 we discuss how to select a preferred specification from several competing models with possibly different numbers of regimes, different types of synchronization, and different specifications of the covariance matrix.

### 3.3.1 Likelihood function

We follow the common practice in the literature to treat the unobserved regimes in Markov-Switching models as parameters to be estimated. For this purpose we derive the complete data likelihood function. We do so for the bivariate MS-VAR model as given in (3.1) with (3.2), with imperfect synchronization due to asymmetric phase shifts as specified in (3.6) and covariance matrix as in (3.7) with (3.9). In vector notation we can write the model as

$$Y_t - \mathcal{M}_{S_t} = \Phi(Y_{t-1} - \mathcal{M}_{S_{t-1}}) + \mathcal{E}_t, \quad \text{with} \quad \mathcal{E}_t \sim N(0, \Sigma_t), \quad (3.10)$$

where

$$Y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}, \quad S_t = \begin{pmatrix} S_{1,t} \\ S_{2,t} \end{pmatrix}, \quad \mathcal{E}_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix},$$

and

$$\mathcal{M}_{S_t} = \begin{pmatrix} \mu_{1,S_{1,t}} \\ \mu_{2,S_{2,t}} \end{pmatrix}, \quad \text{and} \quad \Phi = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{pmatrix}.$$

Given the assumption of multivariate normality for the shocks  $\mathcal{E}_t$ , the conditional density of  $Y_t$  given the past observations  $Y^{t-1} = \{Y_1, \dots, Y_{t-1}\}$  and given the past and current states  $S^t = \{S_1, \dots, S_t\}$  is given by

$$f(Y_t | Y^{t-1}, S^t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \Phi, \kappa) = \left( \frac{1}{\sqrt{2\pi}} \right)^2 |\Sigma_t|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \mathcal{E}_t' \Sigma_t^{-1} \mathcal{E}_t \right), \quad (3.11)$$

where  $\mu_l = (\mu_{l,1}, \dots, \mu_{l,J})'$  and  $\sigma_l^2 = (\sigma_{l,1}^2, \dots, \sigma_{l,J}^2)'$  for  $l = 1, 2$ ,  $\rho = (\rho_1, \dots, \rho_J)'$  and  $\kappa = (\kappa_1, \dots, \kappa_J)'$ . The complete data likelihood function for model (3.10) conditional on the first observation equals

$$f(Y^T, S^T | Y_1, \theta) = \left( \prod_{i=1}^J \prod_{j=1}^J (p_{ij}^{T_{ij}}) \right) \prod_{t=2}^T f(Y_t | Y^{t-1}, S^t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \Phi, \kappa), \quad (3.12)$$

where  $T$  denotes the sample size,  $T_{ij}$  is the number of transitions from regime  $i$  to regime  $j$ , and  $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \text{vec}(\Phi), \kappa, \text{vec}(P))$  represents all the model parameters, with  $P = \{p_{ij}\}_{i,j=1}^J$  the matrix with transition probabilities. The likelihood function conditional only on the model parameters can be obtained by summing (3.12) over all the possible states as

$$f(Y^T | Y_1, \theta) = \sum_{S_{1,1}=1}^J \sum_{S_{1,2}=1}^J \cdots \sum_{S_{1,T}=1}^J f(Y^T, S^T | Y_1, \theta). \quad (3.13)$$

### 3.3.2 Prior distributions

To estimate the model parameters together with the unobserved regimes, we would like to obtain posterior results that are driven by the data rather than by the prior distributions. Therefore, we impose rather diffuse prior specifications for the model parameters.

The parameters of key interest are the lead/lag times or phase shifts  $\kappa_j$ ,  $j = 1, \dots, J$ . For those, we use a discrete uniform prior where we assign equal probability to each possible value of  $\kappa$  in a predefined set  $\mathcal{C}$ , that is

$$f(\kappa) \propto \begin{cases} 1 & \text{for all } \kappa \in \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.14)$$

The specification of the set  $\mathcal{C}$  determines the type of synchronization that is assumed. For example, when  $\mathcal{C} = \{\kappa \in \mathbb{Z}^J \mid -c_j \leq \kappa_j \leq c_j, j = 1, \dots, J\}$  for certain positive valued  $c_j$ ,  $j = 1, \dots, J$ , we allow for imperfect synchronization with the phase shifts being restricted only by the bounds  $c_j$ . Additional restrictions may be imposed such that, for example, the difference between the phase shifts of the distinct regimes cannot exceed a certain threshold  $d$ . In this case,  $\mathcal{C}$  can be specified as  $\mathcal{C} = \{\kappa \in \mathbb{Z}^J \mid -c_j \leq \kappa_j \leq c_j, |\kappa_i - \kappa_j| \leq d, i, j = 1, \dots, J\}$ . Note that setting  $d$  equal to zero (while  $c_j > 0$  for  $j = 1, \dots, J$ ) results in a model where phase shift parameters are identical across regimes, that is, corresponding with the specification of Hamilton and Perez-Quiros (1996) given in (3.5). Setting  $c_j = 0$  for  $j = 1, \dots, J$  we obtain an MS-VAR model with perfect synchronization as in (3.4).

For the transition probabilities we use an uninformative Dirichlet prior

$$p_{1j}, \dots, p_{Jj} \sim \text{Dir}(1, \dots, 1). \quad (3.15)$$

When this non-informative prior for the transition probabilities is used special attention must be paid to the prior specifications of the regime-dependent parameters. This follows from the fact that the value of the complete data likelihood function is the same if we switch all regime dependent parameters together with the corresponding transition probabilities. This ‘label switching problem’ complicates proper posterior analysis as the posterior distributions of the regime dependent parameters become multimodal, see Frühwirth-Schnatter (2001) and Geweke (2007) for discussion. To circumvent this problem, we define the prior for the regime dependent mean parameters  $\mu = (\mu'_1, \mu'_2)'$  in such a way that it identifies the regimes. Specifically, we set the prior specification for  $\mu$

as

$$f(\mu) \propto \begin{cases} 1 & \text{if } \mu_1 \in \{\mu_1 \in \mathbb{R}^J | \mu_{1,1} < \dots < \mu_{1,j} < \dots < \mu_{1,J}\} \\ 0 & \text{elsewhere.} \end{cases} \quad (3.16)$$

For the autoregressive coefficients we use flat priors, restricted to the stationarity region, that is,

$$f(\Phi) \propto 1, \quad (3.17)$$

when the characteristic roots of  $\Phi$  lie outside the unit circle and 0 otherwise.

Finally, for the regime-dependent variance parameters as well as for the correlation parameters we take uninformative priors

$$f(\sigma_{l,j}^2) \propto \sigma_{l,j}^{-2} \quad \text{for } l = 1, 2 \text{ and } j = 1, \dots, J, \quad (3.18)$$

$$f(\rho_j) \propto (1 - \rho_j^2)^{-\frac{3}{2}}, \quad j = 1, \dots, J. \quad (3.19)$$

These priors follow from decomposing an inverted Wishart prior where the degrees of freedom approaches zero for the covariance matrix into the distributions of the variances and the correlations.

### 3.3.3 Posterior simulation

The posterior distribution for the model parameters is proportional to the product of the likelihood function (3.13) and the joint prior  $f(\theta)$ , with the latter obtained as the product of (3.15)–(3.19). Gibbs sampling together with data augmentation leads to the following algorithm to draw from the joint posterior distribution:

1. Sample  $\mu$  from  $f(\mu | \Phi, \sigma^2, \rho, S^T, Y^T)$ .
2. Sample  $\Phi$  from  $f(\Phi | \mu, \sigma^2, \rho, S^T, Y^T)$ .
3. Sample  $\sigma_l^2$  from  $f(\sigma_l^2 | \mu, \Phi, \rho, S^T, Y^T)$  for  $l = 1, 2$ .
4. Sample  $\rho_j$  from  $f(\rho_j | \mu, \Phi, \sigma_{i,j}^2, S^T, Y^T)$  for  $j = 1, \dots, J$ .
5. Sample  $p_{ij}$  from  $f(p_{ij} | S_1^T)$ , for  $i, j = 1, \dots, J$ .
6. Sample  $\kappa_j$  from  $f(\kappa_j | \mu, \Phi, \sigma^2, \rho, S_1^T, Y^T)$  for  $j = 1, \dots, J$ .
7. Sample  $S_1^T$  from  $f(S_1^T | P, \mu, \Phi, \sigma^2, \rho, \kappa, Y^T)$ .

Details on the derivation of the conditional posterior distributions resulting from the likelihood function and prior distributions are given in Appendix 3.A.1.

### 3.3.4 Model selection

The multistate MS-VAR model with imperfect synchronization with asymmetric phase shifts provides a flexible and intuitive framework for characterizing partial synchronization of cycles in different macroeconomic and financial variables. However, moving away from more restricted but more parsimonious forms of synchronization may also result in overfitting. Therefore, we want to establish the potential of the model by comparing it with more restricted (nested) alternative models, in particular MS-VAR models with symmetric phase shifts and with perfect synchronization.

Selecting the best model is not an easy task when the competing alternatives embody regime switching dynamics. In most cases standard testing procedures apply only when the number of regimes is the same in the models being compared. In this case marginal likelihood based comparisons, e.g. Bayes factors, can be implemented. When the task is determination of the number of the regimes, however, a complication arises when improper priors are used for the regime dependent parameters of interest. In this case, Bayes factors are not properly defined and tend to select the more parsimonious model (see Gelfand and Dey, 1994, for details). An alternative way of model comparison is by means of predictive Bayes factors computed using predictive likelihoods. This approach has the advantage that it is not affected by the choice of prior distributions and overfitting, while it is directly related to the posterior model probabilities. For a given model, the predictive likelihood of the observation at  $t_0 + 1$ ,  $Y_{t_0+1}$ , conditional on the previous observations  $Y^{t_0}$ , is given by

$$f(Y_{t_0+1}|Y^{t_0}) = \int f(Y_{t_0+1}|\theta) f(\theta|Y^{t_0}) d\theta, \quad (3.20)$$

where  $p(\theta|Y^{t_0})$  is the posterior distribution of the model parameters  $\theta$  given the observations until  $t_0$ , and  $p(Y_{t_0+1}|\theta)$  is the density of the observation  $y_{t_0+1}$ , which can be written as

$$f(Y_{t_0+1}|\theta) = \sum_{j=1}^J f(Y_{t_0+1}|S_{1,t_0+1} = j, \theta) f(S_{1,t_0+1} = j|\theta, Y^{t_0}). \quad (3.21)$$

We can use the posterior simulator to obtain the distribution of the model parameters and estimate the predictive likelihood by  $G^{-1} \sum_{g=1}^G f(Y_{t_0+1}|Y^{t_0}, \theta^{(g)})$  where  $G$  is a large number of draws from the posterior distribution. This can be extended to compute the predictive likelihood of a sequence  $Y^{t_0+1, t_0+h} = (Y'_{t_0+1}, \dots, Y'_{t_0+h})'$  where

$$f(Y^{t_0+1, t_0+h}|Y^{t_0}) = \int \prod_{l=1}^h f(Y_{t_0+l}|Y^{t_0+l-1}, \theta) f(\theta|Y^{t_0}) d\theta, \quad (3.22)$$

see Frühwirth-Schnatter (2006), for details.

### 3.4 Lead-lag effects in size-based portfolio returns

In this section, we use the MS-VAR models to analyze the degree of synchronization in monthly returns of stock portfolios with different levels of market capitalization. Empirical evidence convincingly has demonstrated that asset returns are not normally distributed. Markov-switching models provide an interesting approach for describing this non-normality, by means of a mixture of underlying normal distributions, see Timmermann (2000) and Perez-Quiros and Timmermann (2001), among others. A further attractive property of Markov-switching models is that they formalize the idea of cyclical variation in financial markets, with phases corresponding with periods of bull and bear sentiment, for example, during which asset prices have quite different characteristics. Not surprisingly then, Markov-switching models have become increasingly popular in modeling and forecasting asset prices. Previous research focuses mostly on univariate modeling of asset returns. This generally leads to the finding that a model with two or three regimes captures the univariate dynamics in aggregate stock returns and identifies bull and bear markets successfully (see Turner *et al.*, 1989; Chauvet and Potter, 2000; Maheu and McCurdy, 2000; Guidolin and Timmermann, 2006, among others). Relatively fewer papers consider multivariate settings to analyze the dependence between different asset returns (see Perez-Quiros and Timmermann, 2000; Ang and Bekaert, 2002; Guidolin and Timmermann, 2006, for example). More regimes typically seem to be required for an adequate representation of the joint return distribution. For example, using excess returns of size-based stock portfolios and bonds, Guidolin and Timmermann (2006) show that a four regime model representing bear, bull, recovery and crashes performs better than competing models with fewer regimes.

Previous applications of multivariate regime-switching models to asset returns assume perfect synchronization in the cyclical dynamics of the included assets. While it does not seem unreasonable to represent the idea of regimes in financial markets by means of a single common cycle, the assumption that regime switches occur contemporaneously for different assets may not be appropriate. In particular, a substantial body of empirical research has documented the presence of lead-lag effects between different types of stocks. Determinants of the lead-lag effect include size, analyst coverage, institutional ownership and trading volume, see Lo and MacKinlay (1990a); Brennan *et al.* (1993); Badrinath *et al.* (1995); Chordia and Swaminathan (2000), among others. Focusing on size, the typical finding is that returns on (portfolios of) large stocks lead returns on (portfolios

of) small stocks, see Lo and MacKinlay (1990a); McQueen *et al.* (1996); Richardson and Peterson (1999); Hou (2007), among others. Several explanations have been put forward for this size-based lead-lag effect, including time-varying expected returns and market microstructure biases such as infrequent trading, see Boudoukh *et al.* (1994). Most empirical evidence, however, points in the direction of the tendency of small-caps to adjust more slowly to new (common) information than large caps as the underlying reason. Previous studies typically examine the lead-lag effect in size-based portfolios using cross-autocorrelations within a linear VAR framework. Interestingly, McQueen *et al.* (1996) and Hou (2007) document a ‘directional asymmetry’ in the lead-lag effect. Specifically, McQueen *et al.* (1996) find that large caps lead small caps only in response to good news, but not bad news. Hou (2007) reports the opposite effect, namely that small stocks adjust slowly to negative news but not to positive news. Here we examine whether the lead-lag effect between small caps and large caps can also be established when Markov-switching models are used for describing their joint distribution. In particular, we adopt our MS-VAR framework where the regime dynamics are governed by a single common cycle that is subject to phase shifts, which possibly are different across regimes.

We examine the presence of lead-lag effects in size-based portfolios based on all NYSE, AMEX, and NASDAQ stocks over the period July 1963 - December 2007. We consider monthly returns on equal-weighted large-cap and small-cap portfolios, consisting of the 20% largest and smallest stocks (sorted according to market capitalization), respectively.<sup>1</sup> We use the returns in excess of the risk-free rate.<sup>2</sup>

We estimate MS-VAR models with two, three and four regimes. For the relation between the regime dynamics in the two portfolio returns series, we consider the four possibilities discussed in Section 3.2, namely (i) independent cycles (IND); (ii) perfect synchronization (PS); (iii) imperfect synchronization with a single phase shift of the common cycle (IS-SPS); and (iv) imperfect synchronization with regime-specific phase shifts (IS-APS). For the specification of the covariance matrix, we examine the two characterizations discussed in the Section 3.2. On the one hand, we impose perfect synchronization in the variance dynamics, as in (3.8). Here we assume that the regime dynamics of the covariance matrix are governed by the Markov process that also drives the regime dynamics of the expected excess returns on the large-cap portfolio. Note that the large-cap portfolio

---

<sup>1</sup>The data is obtained from Kenneth French’s website data library <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/>

<sup>2</sup>For the risk-free rate we use Fama-Bliss risk free rates from CRSP (Center for Research in Security Prices). This series is discontinued in December 2001. For the remaining part of the sample period we use the 4-week T-Bill rate measured at close on the last trading day of each month, obtained from the St. Louis Federal Reserve (<http://research.stlouisfed.org/fred2/>).

resembles the market index much more closely than the small-cap portfolio and, therefore, is a more adequate measure to capture the dynamics of volatility and correlation in the stock market. This model is denoted as  $\Sigma$ -PS, as we assume perfect synchronization of the portfolios in terms of their covariance dynamics. On the other hand, we assume that the expected excess return and the variance of each portfolio have identical regime dynamics, as in (3.9). In this case, the regime switches of the correlation are still determined by the regime indicator of the large-cap portfolio.<sup>3</sup> This covariance matrix specification is indicated by  $\Sigma$ -IS-APS, as it allows for imperfect synchronization with asymmetric phase shifts (IS-APS) in the volatility dynamics of the portfolios.

For the models with imperfect synchronization, we specify the set of admissible phase shift parameters  $\mathcal{C}$  such that all  $\kappa_j$ 's,  $j = 1, \dots, J$ , are in the interval  $[-2, 2]$ , i.e.  $c_j = 2$ , as it is unlikely that the delay exceeds two months. We also restrict the differences between the phase shift parameters for different regimes to be one month at most, i.e.  $d = 1$ .

Finally, we do not include autoregressive dynamics in any of the models. This may seem restrictive, especially in light of Boudoukh *et al.* (1994) who argue that lead-lag effects among size-based portfolios may be spurious as they are in fact, to a large extent, due to (contemporaneous cross-correlation and) own-autocorrelation of small-cap returns. This suggests a large stock portfolio does not necessarily lead a small stock portfolio once lagged small-cap portfolio returns are included to account for these autocorrelation effects. With this in mind, we also estimate all models with the first lag of the small-cap returns included, as well as models with unrestricted first-order autoregressive dynamics. The results, especially those related to lead-lag parameters, are very similar to those reported here. Hence, we do not display these results for the sake of brevity.<sup>4</sup>

In order to achieve a robust model selection we compute both the marginal likelihood using the whole sample period as well as the predictive likelihood. For evaluating the predictive likelihood we use the sample from July 1963 until December 2000 for estimation of the model parameters and the remaining period from January 2001 until December 2007 to compute the predictive likelihood.<sup>5</sup> The posterior distributions are based on 50,000 iterations of the Gibbs sampling algorithm. The first 20,000 draws are discarded as burn-in sample and the final 30,000 draws are used for inference and for predictive likelihood

---

<sup>3</sup>We also estimate models with the correlation regime being determined by the regime of the small-cap portfolio's expected return and volatility. Overall these models perform worse in terms of marginal likelihood and predictive likelihood. Detailed results are available upon request.

<sup>4</sup>Detailed results are available upon request.

<sup>5</sup>For the complete out-of-sample period the predictive likelihood is computed using the posterior distribution based on the initial estimation sample, as recursive updating is not feasible in terms of computation time.

computation. The convergence of the MCMC sampler is checked using statistical and visual inspection and in all the models convergence is assured.

Table 3.1: Comparison of the models estimated using LC and SC Data

|                    | <u>Log marg. likelihood</u> |                  | <u>Log pred. likelihood</u> |                  |
|--------------------|-----------------------------|------------------|-----------------------------|------------------|
|                    | $\Sigma$ -PS                | $\Sigma$ -IS-APS | $\Sigma$ -PS                | $\Sigma$ -IS-APS |
| Panel A: 2 regimes |                             |                  |                             |                  |
| IND                | -3356.3                     | -3306.5          | -465.1                      | -461.7           |
| PS                 | -3220.7                     | -3218.8          | -465.0                      | -464.8           |
| IS-SPS             | -3206.5                     | -3218.8          | -462.8                      | -464.8           |
| IS-APS             | -3211.3                     | -3210.4          | -461.2                      | -463.9           |
| Panel B: 3 regimes |                             |                  |                             |                  |
| IND                | -3508.4                     | -3443.6          | -464.7                      | -470.5           |
| PS                 | -3283.2                     | -3257.1          | -465.9                      | -463.2           |
| IS-SPS             | -3213.5                     | -3257.1          | -464.9                      | -463.2           |
| IS-APS             | -3225.9                     | -3206.1          | -463.7                      | -459.6           |
| Panel C: 4 regimes |                             |                  |                             |                  |
| IND                | -3984.6                     | -3770.8          | -469.8                      | -475.2           |
| PS                 | -3322.4                     | -3276.0          | -463.1                      | -463.3           |
| IS-SPS             | -3322.4                     | -3276.0          | -463.1                      | -462.6           |
| IS-APS             | -3326.9                     | -3230.5          | -463.5                      | -461.6           |

*Note:* The table presents log marginal likelihoods and log predictive likelihoods of the competing models with (i) different number of regimes; (ii) different types of synchronization in expected returns, and (iii) different types of regime switching of the variances and correlations. The models are applied to monthly excess returns on large-cap and small-cap portfolios for the sample period July 1963-December 2000. Marginal likelihoods are those obtained when the models are estimated for the complete sample period. Predictive likelihood values are for the period January 2001-December 2007, which are computed conditional on the posterior distributions of the model parameters obtained with an estimation sample ending in December 2000. The regime dynamics in the expected excess returns correspond with (i) independent cycles (IND); (ii) perfect synchronization (PS); (iii) imperfect synchronization with a single phase shift of the common cycle (IS-SPS); and (iv) imperfect synchronization with regime-specific phase shifts (IS-APS). The specifications of the covariance matrix dynamics correspond with (i) perfect synchronization ( $\Sigma$ -PS), according to the regime indicator for the expected excess returns of the large-cap portfolio; and (ii) imperfect synchronization with regime-specific phase shifts ( $\Sigma$ -IS-APS), where the correlation switches regimes according to the regime indicator for the expected excess returns of the large-cap portfolio. Posterior results for computing log marginal likelihoods and predictive likelihoods are based on 50,000 simulations of which the first 20,000 are discarded as burn-in sample.

Table 3.1 displays the log marginal likelihood and log predictive likelihood values of the competing models, which offer five clear-cut conclusions. First, independent of the number of regimes and the specification of the covariance matrix, the models with a (possibly shifted) common cycle convincingly outperform the model with independent cycles in terms of marginal likelihood. The differences become larger when the number of regimes increases or when perfect synchronization of the covariance dynamics is imposed. The model with independent cycles also performs worse in terms of predictive likelihood, with the exception of the two regime model with imperfect synchronization in the volatilities.

Second, independent of the number of regimes and the specification of the relation between the cycles in the two portfolio returns, the models with imperfect synchronization of regime shifts in volatilities ( $\Sigma$ -IS-APS) achieve higher marginal likelihood and predictive likelihood values than the corresponding models with perfect synchronization ( $\Sigma$ -PS). The only exception is the two-regime model with symmetric phase shifts (and the two-regime model with asymmetric regime shifts, but only in terms of predictive likelihood). Hence, the assumption of perfect synchronization in the variance dynamics does not seem appropriate.

Third, independent of the number of regimes, for the  $\Sigma$ -IS-APS models we find that the marginal likelihood and predictive likelihood values for the specifications with perfect synchronization and with imperfect synchronization due to symmetric phase shifts are (almost) identical. This occurs because the phase shift parameters in the IS-SPS models are estimated to be equal to zero, which reduces this specification to the PS model.

Fourth, independent of the number of regimes, for the  $\Sigma$ -IS-APS models we find that the marginal likelihood and predictive likelihood values for the specifications with imperfect synchronization due to asymmetric phase shifts are (substantially) higher than those for the PS and IS-SPS specifications. This suggests that allowing for regime-specific lead-lag times is appropriate for the large-cap and small-cap portfolios. While the differences in likelihood values are relatively small for the two-regime models, they become substantial when multiple regimes are considered.

Fifth, focusing on the IS-APS specification with  $\Sigma$ -IS-APS, the model with three regimes offers the best performance in terms of predictive likelihood.<sup>6</sup>

---

<sup>6</sup>Note that the model with three regimes not only achieves higher marginal likelihood and predictive likelihood values than the model with two regimes, but also than the model with four regimes. For the PS and IS-SPS specifications this is even more pronounced, with the marginal likelihood value declining monotonically as the number of regimes increases. While this may seem surprising at first, it can be understood by recalling that marginal likelihood (or Bayes factors) tends to select the more parsimonious model when improper priors are used for the regime dependent parameters, as discussed in Section 3.3.4.

In sum, the results in Table 3.1 lead us to select the model with three regimes and imperfect synchronization with asymmetric phase shifts in both the expected excess returns and the volatilities as the preferred specification. In the remainder of this section, we therefore focus on the estimation results for this model specification.

Table 3.2: Posterior results for three regime MS-VAR model with asymmetric synchronization and regime-dependent heteroskedasticity

|            |                           | Regime 1                          | Regime 2       | Regime 3       |
|------------|---------------------------|-----------------------------------|----------------|----------------|
| Large caps | Mean $\mu_{1,j}$          | 3.772 (0.626)                     | 0.114 (0.224)  | -1.337 (0.964) |
|            | Volatility $\sigma_{1,j}$ | 2.901 (0.395)                     | 3.353 (0.181)  | 7.804 (0.795)  |
| Small caps | Mean $\mu_{2,j}$          | 6.310 (0.767)                     | -0.519 (0.368) | -2.077 (1.480) |
|            | Volatility $\sigma_{2,j}$ | 4.159 (0.650)                     | 3.909 (0.241)  | 11.192 (1.228) |
|            | Correlation $\rho_j$      | 0.305 (0.198)                     | 0.724 (0.032)  | 0.709 (0.065)  |
|            |                           | Transition probabilities $p_{ij}$ |                |                |
|            |                           | Regime 1                          | 0.516 (0.010)  | 0.440 (0.098)  |
|            |                           | Regime 2                          | 0.064 (0.024)  | 0.875 (0.028)  |
|            |                           | Regime 3                          | 0.215 (0.081)  | 0.066 (0.057)  |
|            |                           | Phase shifts $\kappa_j$           |                |                |
|            |                           | 0.000 (0.000)                     | -1.000 (0.000) | -0.049 (0.217) |

*Note:* The table presents posterior means and standard deviations (in parentheses) of parameters in the MS-VAR model with three regimes, imperfect synchronization with asymmetric phase shifts, and regime-dependent heteroskedasticity, estimated for monthly excess returns of large cap and small-cap portfolios over the period July 1963 - December 2007. The volatilities  $\sigma_{l,j}$  have the same regime dynamics as the corresponding expected excess returns  $\mu_{l,j}$ ,  $l = 1, 2$ . The correlation  $\rho_j$  has the same regime dynamics as the mean and volatility of the large-cap portfolio. Posterior results are based on 50,000 simulations of which the first 20,000 are discarded as burn-in sample.

Table 3.2 summarizes the posterior results of the model parameters. These indicate that the three regimes in the model have distinct features. A general pattern is that in each regime the volatility of the small-cap portfolio is much higher than the corresponding volatility of the large-cap portfolio, as expected. However, this higher volatility of the small-cap portfolio comes with a higher mean return (indicating a risk-return trade-off) only for the first regime. Regime 1 is characterized by a large positive mean return for both portfolios together with low volatility (especially for large caps), capturing a ‘boom’ market. For both large-cap and small-cap portfolios regime 2 has mean return close to zero and relatively low volatility. An interesting result is that the volatility in the boom market is lower than the moderate regime for the large-cap portfolios. Regime 3 can be characterized as a ‘bear’ market, as it shows a substantially negative mean return accompanied by high volatility, indicating the large uncertainty during these time periods.

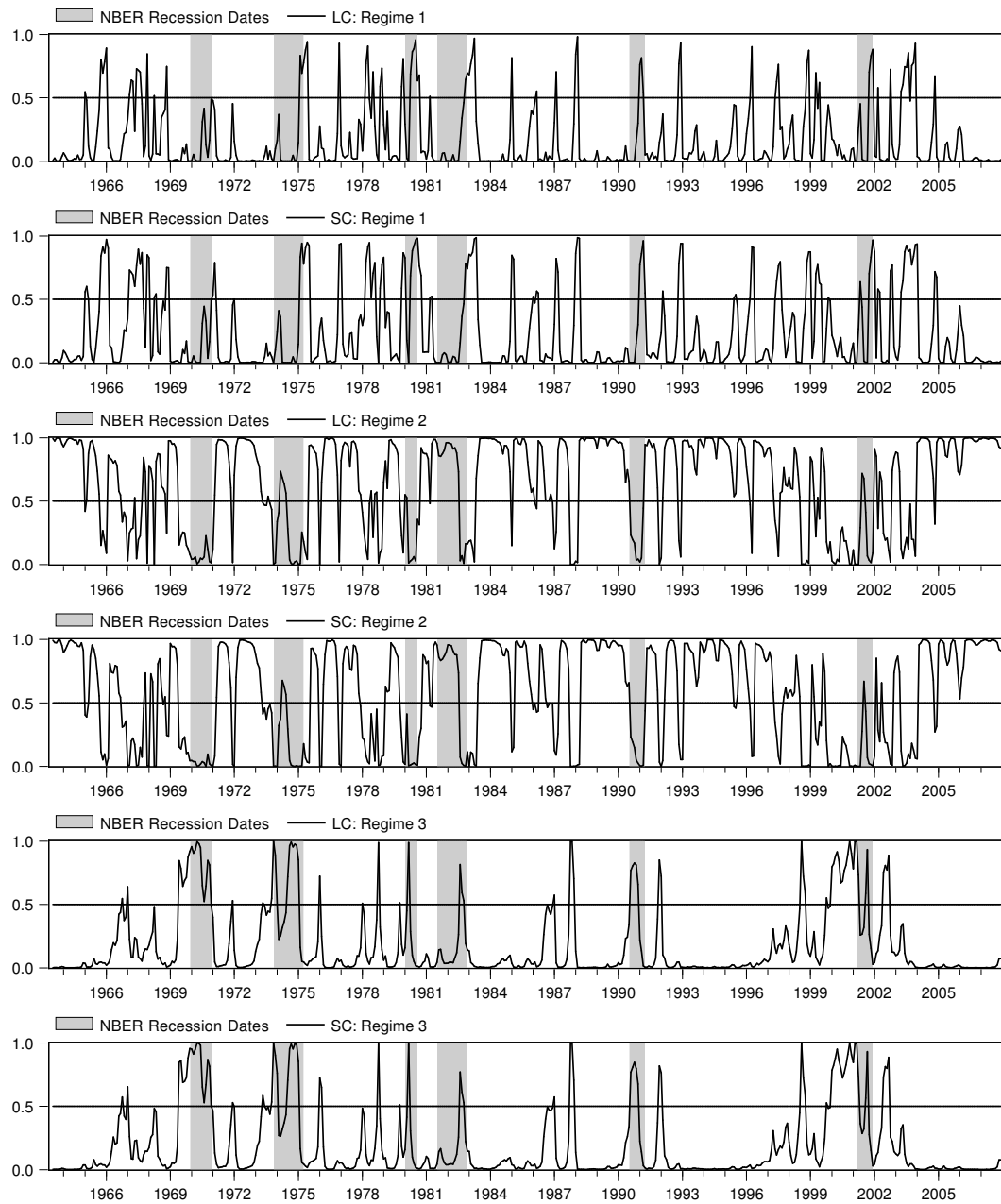
Summarizing, the first and third regimes capture turbulent times with abrupt fluctuations in the stock market, while the second regime captures normal times with either small gains or losses together with limited volatility.

The transition probability estimates also reveal an interesting pattern in the relative duration of the different regimes and the cyclical dynamics in the portfolio returns. The probability of staying in the boom regime is 0.52 and hence the duration of this regime is short, on average only two months. As the transition probability of 0.88 indicates, the average duration of the moderate regime 2 is longer than the other regimes with almost eight months. The probability of staying in the bear regime is 0.73 implying an average duration of about four months. Transition from the boom regime is ten times more likely to occur to the moderate regime than to the bear regime, with transition probabilities of 0.44 and 0.043, respectively. Hence, the boom regime typically does not end with a crash, but is followed by a period with moderate returns. By contrast, when the bear regime is left, this is much more likely to happen to the boom market regime than to the moderate regime, with transition probabilities of 0.215 and 0.066, respectively. This suggests that bear markets are most often followed by a (short) period with high returns, providing evidence of a ‘recovery’ or ‘bounce-back effect’ in these stock returns, similar to findings for the US business cycle in, for example, Kim *et al.* (2005).

We display the posterior regime probabilities in Figure 3.1. We also include recessions as dated by the NBER in shaded areas to elaborate the links between the regime timings with the business cycle. From the graphs we can infer that the regimes are estimated quite precisely in the sense that generally the regime probabilities are close to either zero or one, although there is some uncertainty about the regime assignment in a few periods. The conclusion drawn above about the regimes capturing turbulent and normal times is evident in Figure 3.1. The boom regime, for example, mostly occurs at the end of recessions (among others). In addition, except for the recession in 1981-2, the moderate regime appears considerably less frequently during recessions than during expansions. The bear market regime mostly occurs before and during the start of the recession. In addition, this regime captures all stock market crashes that occurred during the sample period, such as ‘Black Monday’ (October 1987), the Russia crises (August-September 1998), the collapse of the dot-com bubble (2000-1), and 9/11 (September 2001).

As an alternative way of assessing the regime dynamics, Figure 3.2 shows the implied model volatilities, obtained as the weighted average of the regime-specific volatilities with the posterior regime probabilities as weights, together with the monthly realized volatility constructed from the daily portfolio returns. It is seen that the patterns in realized

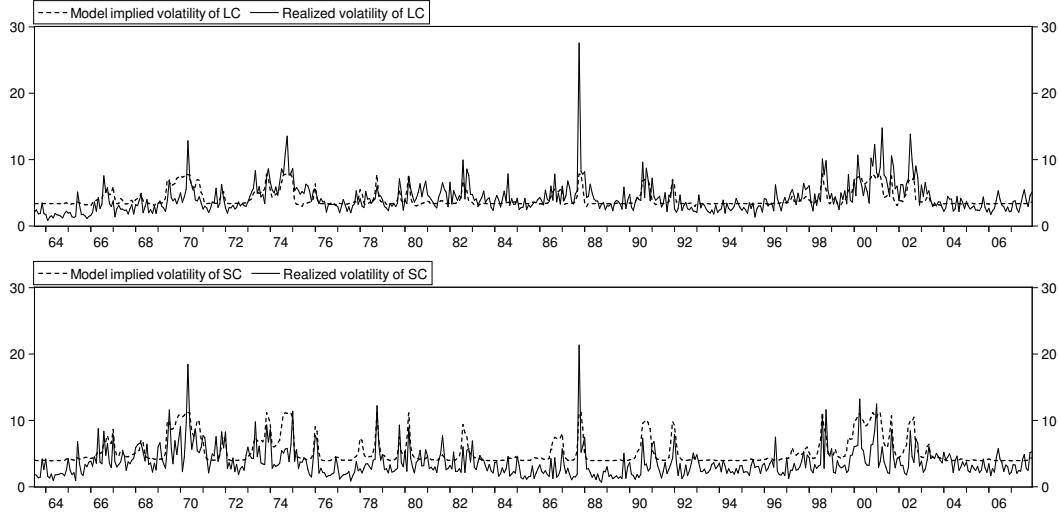
Figure 3.1: Posterior probabilities of regimes of LC and SC



*Note:* The solid line is the posterior regime probabilities for large-caps (LC) and small-caps (SC) regimes summarized in Table 3.2. The shaded areas indicate the recession dates of US recessions as determined by the the NBER's Business Cycle Dating Committee.

volatility are captured fairly well by the model, in the sense that periods with a higher level of realized volatility generally are matched by increased ‘model-implied’ volatility.

Figure 3.2: Comparison of volatilities implied by the model and realized volatilities



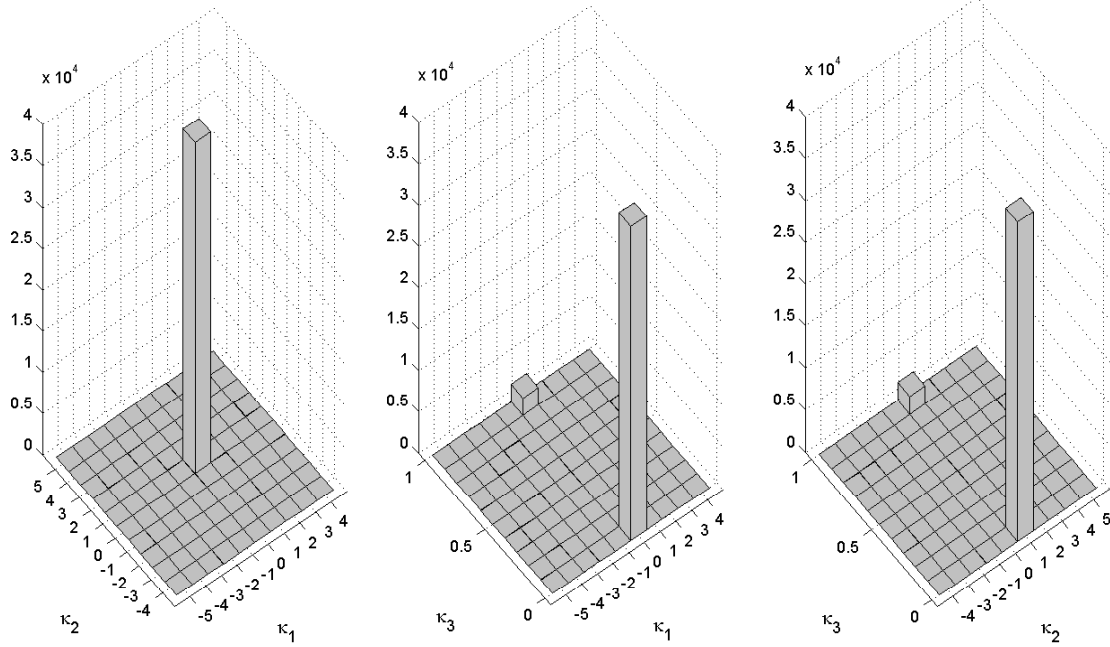
*Note:* The solid lines are the monthly realized volatilities of the portfolios computed as the sum of squared daily returns plus twice the sum of the cross-product of returns on consecutive trading days of the corresponding month. The latter provides a correction for the presence of autocorrelation in the daily returns, see Zhou (1996). The dashed lines are the monthly volatilities implied by the MS-VAR model with three regimes, imperfect synchronization with asymmetric phase shifts, and regime-dependent heteroskedasticity.

Finally, we turn to the estimates of the phase shift parameters to assess the implications of the model concerning the lead-lag effects of the large-cap and small-cap portfolios. The posterior results of the phase shift parameters  $\kappa_j$  are shown in Table 3.2 and in Figure 3.3.

The estimates provide convincing evidence for the presence of asymmetric phase shifts. Specifically, we find large and small stocks enter into the boom and bear regimes contemporaneously. For the boom regime we find that all posterior probability mass is located at  $\kappa_1 = 0$ . For regime 3, there is only a modest probability that  $\kappa_3 = 1$ . In contrast, for the moderate regime the large-cap portfolio leads the small-cap portfolio by one month. The posterior uncertainty for this parameter is very small.<sup>7</sup>

<sup>7</sup>Clearly, the posterior standard deviations of the  $\kappa$  parameters are very small. Note that we impose strong priors to deal with the uncertainty embodied in the data but still this prior assigns equal probability to all types of synchronization. Hence, the result is not affected by this reasonably strong prior, though the amount of variation may be affected by its strength. Nevertheless, the results with weaker priors are very similar to the results shown here. Detailed results are available upon request.

Figure 3.3: Distribution of lead/lag times for LC and SC regimes



*Note:* The graphs show the posterior distribution of the lead time parameters  $\kappa_j$ ,  $j = 1, 2, 3$ , in the three regime MS-VAR model with imperfect synchronization with asymmetric phase shifts, and regime-dependent heteroskedasticity for large-caps (LC) and small-caps (SC), summarized in Table 3.2.

We interpret our findings concerning the phase shift parameters as follows. The contemporaneous switches of the two portfolios to the boom and bear regimes suggest that during turmoil periods, with either good or bad news, information diffusion occurs rapidly and lead-lag effects are not present. The positive phase shift for the second regime indicates that news ‘travels slowly’ during more moderate times, in the sense that it is incorporated only gradually into small stocks, such that their prices adjust with a delay compared to large stocks. We note that our findings of asymmetric phase shifts are quite different to the conclusions reached by McQueen *et al.* (1996) and Hou (2007) concerning the asymmetric nature of lead-lag effects. McQueen *et al.* (1996) find that large caps lead small caps only in response to good news, but not bad news. Hou (2007) documents the opposite effect, namely that small stocks adjust slowly to negative news but not to positive news. Incorporating lead-lag effects in a regime-switching framework with multiple regimes, we obtain intuitively plausible results that possibly reconcile the findings of McQueen *et al.* (1996) and Hou (2007). Our posterior results suggest that information

diffusion in fact occurs rapidly for both good and bad news items that have considerable impact on stock prices, in the sense that large-caps and small-caps switch to the extreme boom and bear regimes at the same time. Small-caps respond with a delay during moderate times, possibly due to the fact that less ‘exciting’ news during those periods is incorporated in their prices less rapidly.

### 3.5 Conclusion

In this paper we have developed a Markov-Switching vector autoregressive model that allows for imperfect synchronization of cyclical regimes in multiple variables, due to phase shifts of a single common cycle. The model has three key features. First, the amount of phase shift can be different across regimes (as well as across variables). This feature is empirically relevant given the evidence for asymmetries in lead-lag effects in different types of assets, as well as in leading and coincident economic indicators. Second, the cycle can consist of any number of regimes  $J \geq 2$ . Allowing for multiple regimes can be necessary for accurate modeling of cyclical dynamics in financial asset returns, as well as business cycle dynamics in macroeconomic variables. Third, the model allows for regime-dependent volatilities and correlations. This is particularly important for identification of regimes in asset prices, where differences in volatilities (and correlations) often are more pronounced than differences in mean returns.

In an empirical application to monthly returns on size-based stock portfolios, we found that a three-regime model with asymmetric phase shifts and regime-dependent heteroscedasticity characterizes the joint distribution of returns most adequately. While large- and small-cap portfolios switch contemporaneously into boom and crash regimes, the large-cap portfolio leads the small-cap portfolio for switches to a moderate regime by a month. This suggests that during turmoil periods, with either good or bad news, information diffusion occurs rapidly and lead-lag effects are not present. News ‘travels slowly’ during more moderate times, in the sense that it is incorporated only gradually into small stocks, such that their prices adjust with a delay compared to large stocks.

## 3.A Appendices

### 3.A.1 Conditional Posterior Distributions

In this appendix, we derive the posterior distributions that we use in the sampling scheme described in Section 3.3.3. For notational convenience, let  $\mathcal{S}_{1,t}$  and  $\mathcal{S}_{2,t}$  be the  $J \times 1$  vectors of indicator functions for the first and second variable, respectively. The elements in these vectors take the values 0 or 1, indicating which of the  $J$  regimes occurs at time  $t$ .

We write the model (3.5) in matrix notation

$$Y_t - \mathcal{S}'_t \mu = \Phi(Y_{t-1} - \mathcal{S}'_{t-1} \mu) + \mathcal{E}_t, \quad \text{with } \mathcal{E}_t \sim N(0, \Sigma_t), \quad (\text{A-1})$$

where  $Y_t$  and  $\mathcal{E}_t$  are defined below (3.10) and

$$\mathcal{S}_t = \begin{pmatrix} \mathcal{S}_{1,t} & \mathbf{0}_J \\ \mathbf{0}_J & \mathcal{S}_{2,t} \end{pmatrix}, \quad (\text{A-2})$$

with  $\mathbf{0}_J$  a  $J \times 1$  vector of zeros.

#### Sampling of $\mu$

To sample  $\mu$  we first rewrite (A-1) as

$$\Sigma_t^{-\frac{1}{2}} (Y_t - \Phi Y_{t-1}) = \Sigma_t^{-\frac{1}{2}} (\mathcal{S}'_t - \Phi \mathcal{S}'_{t-1}) \mu + \Sigma_t^{-\frac{1}{2}} \mathcal{E}_t. \quad (\text{A-3})$$

This is a regression in the form of

$$Z_t = X_t \mu + \epsilon_t \quad \text{with } \epsilon_t \sim N(0, I_2), \quad (\text{A-4})$$

and hence the conditional distribution of  $\mu$  is multivariate normal with mean  $(X'X)^{-1} X'Z$  and variance  $(X'X)^{-1}$ , where  $X = (X'_2, \dots, X'_T)'$  and  $Z = (Z'_2, \dots, Z'_T)'$ , see Zellner (1971). The restrictions on the elements of  $\mu$  that are imposed for identification of the regimes, as discussed in Section 3.3.2, can be applied by sampling from the corresponding truncated distribution or by using acceptance rejection sampling.

#### Sampling of $\Phi$

Conditional on the remaining model parameters, the model in (A-1) is a multivariate regression of the form

$$Z = \Phi X + U, \quad (\text{A-5})$$

where  $Z = (Z_2, \dots, Z_T)$  with  $Z_t = Y_t - S'_t \mu$ , and  $X = (X_2, \dots, X_T)$  with  $X_t = (Y_{t-1} - S'_{t-1} \mu)$ . Using the fact that  $\text{vec}(\Phi X) = (X' \otimes I_2) \text{vec}(\Phi)$ , we can write this as a univariate regression

$$z = (X' \otimes I_2) \text{vec}(\Phi) + u, \quad (\text{A-6})$$

where  $z = \text{vec}(Z)$  and  $u = \text{vec}(U)$ . It then follows that the conditional posterior distribution of  $\text{vec}(\Phi)$  is multivariate normal with mean  $((X' \otimes I_2)' \Omega^{-1} (X' \otimes I_2))^{-1} ((X' \otimes I_2)' \Omega^{-1} z)$  and covariance matrix  $((X' \otimes I_2)' \Omega^{-1} (X' \otimes I_2))^{-1}$ , where

$$\Omega^{-1} = \text{Cov}(u)^{-1} = \text{diag}(\Sigma_2^{-1}, \dots, \Sigma_T^{-1}).$$

### Sampling of Variances and Correlations

To sample the variances we decompose the multivariate normal distribution of  $\mathcal{E}_t$  into the conditional distribution of  $\varepsilon_{2,t}$  given  $\varepsilon_{1,t}$  and the marginal distribution of  $\varepsilon_{1,t}$ . This results in

$$\prod_{t=2}^T f(\mathcal{E}_t) = \prod_{t=2}^T \frac{1}{\sigma_{1,t}} \phi\left(\frac{\varepsilon_{1,t}}{\sigma_{1,t}}\right) \frac{1}{\sigma_{2,t} \sqrt{(1 - \rho_t^2)}} \phi\left(\frac{\varepsilon_{2,t} - \rho_t \sigma_{2,t} / \sigma_{1,t} \varepsilon_{1,t}}{\sigma_{2,t} \sqrt{1 - \rho_t^2}}\right), \quad (\text{A-7})$$

where  $\sigma_{l,t}^2$  denotes the variance of  $\varepsilon_{l,t}$ ,  $l = 1, 2$ , which follows from (3.8) or (3.9). To sample  $\sigma_{1,j}^2$  we use an Metropolis-Hastings sampler (Metropolis *et al.*, 1953; Hastings, 1970), where we use as candidate an inverted  $\chi^2$  distribution with scale parameter  $\sum_{t=2}^T I[S_{1,t} = j] \varepsilon_{1,t}^2$  and degrees of freedom  $\sum_{t=2}^T I[S_{1,t} = j]$  for  $j = 1, \dots, J$  which follows from the first part of the decomposition in (A-7) in combination with the prior. To sample  $\sigma_{2,j}^2$  we use a similar approach where we consider the marginal of  $\varepsilon_{2,t}$  in the decomposition.

To sample  $\rho_j$  from its conditional posterior distribution we can again use (A-7). It is easy to see that only the time periods for which  $S_{1,t} = j$  are relevant

$$(1 - \rho_j^2)^{-\frac{3}{2}} \prod_{t=2}^T \left( \frac{1}{\sqrt{(1 - \rho_t^2)}} \phi\left(\frac{\varepsilon_{2,t} - \rho_t \sigma_{2,t} / \sigma_{1,t} \varepsilon_{1,t}}{\sigma_{2,t} \sqrt{1 - \rho_t^2}}\right) \right)^{I[S_{1,t}=j]}. \quad (\text{A-8})$$

We can easily implement the griddy Gibbs sampler approach of Ritter and Tanner (1992). Given that  $\rho_j \in (-1, 1)$  for all  $j = 1, \dots, J$ , we can setup a grid in this interval based on the precision we desire about the value of  $\rho_j$ .

### Sampling of $P$

From the conditional likelihood function (3.13), it follows that the transition probabilities from state  $i$  can be written as

$$f(p_{i,1}, \dots, p_{i,J} | S_1^T) \propto \prod_{j=1}^J p_{i,j}^{T_{i,j}} \quad \text{for } i = 1, \dots, J, \quad (\text{A-9})$$

where  $T_{ij}$  denotes the number of transitions from state  $i$  to state  $j$ . This corresponds to the kernel of a Dirichlet distribution and hence the transition probabilities can be sampled from a Dirichlet distribution with parameters  $T_{ij}$  for  $j = 1, \dots, J-1$ . The transition probability to state  $J$ ,  $p_{iJ}$ , follows from the restriction that  $\sum_{j=1}^J p_{ij} = 1$ .

### Sampling of Lead/Lag Parameters $\kappa$

As the  $\kappa_j$  parameters can only take discrete values we can compute the posterior probabilities for all  $\kappa \in \mathcal{C}$  and sample from a multinomial distribution. We can sample all  $\kappa_j$  parameters at once or conditional on each other, one at a time. When the number of admissible  $\kappa$ 's is large, the latter approach may be more attractive as the number of combinations increases only linearly in the number of states, whereas it increases exponentially in the first case.

### Sampling of Regimes

The conditional posterior density of  $S_{1,t}$  denoted by  $f(S_{1,t} | S_1^{-t}, \theta, Y^T)$  for  $t = 1, \dots, T$  and  $S_1^{-t} = S_1^T \setminus \{S_{1,t}\}$ , is proportional to the transition probabilities due to the Markov structure and to the density of  $Y$  conditional on the regimes. Hence, we can write the posterior density of  $S_{1,t}$

$$f(S_{1,t} | S_1^{-t}, \theta, Y^T) \propto f(S_{1,t} | S_{1,t-1}, \theta) f(S_{1,t+1} | S_{1,t}, \theta) \prod_{i=t-\kappa_{\min}}^{t+1+\kappa_{\max}} f(Y_i | Y^{i-1}, S^i, \theta), \quad (\text{A-10})$$

where  $f(Y_t | Y^{t-1}, S^t, \theta)$  is given in (3.11),  $\kappa_{\max} = \max(\kappa_1, \dots, \kappa_J)$  and  $\kappa_{\min} = \min(\kappa_1, \dots, \kappa_J)$ . At time  $t = T$  the term  $f(S_{1,t+1} | S_{1,t}, \theta)$  drops out. The regime at  $t = 1$  can be sampled from the conditional distribution

$$f(S_{1,1} | S_1^{-1}, \theta, Y^T) \propto f(S_1 | \theta) f(S_{1,2} | S_{1,1}, \theta) \prod_{i=2}^{t+1+\kappa_{\max}} f(Y_i | Y^{i-1}, S^i, \theta), \quad (\text{A-11})$$

where the unconditional density  $f(S_{1,1}|\theta)$  follows a multinomial distribution with ergodic probabilities of the Markov chain.

Sampling of the state variables can be implemented by starting from the most recent value of  $S_1^T$  and sampling the states backward in time, one after another. After each step, the  $t^{th}$  element of  $S_1^T$  is replaced by its most recent draw.



# Chapter 4

## Measuring and Predicting Heterogeneous Recessions

### 4.1 Introduction

A century of empirical evidence about the US business cycle has revealed that economic recessions come in a variety of shapes and sizes. While some downturns are relatively mild, others are rather severe. While some recessions are short, lasting only about half a year, the duration of others is substantially longer. While some contractions are followed by rapid and strong recoveries of economic activity, the effects of other recessions are felt much longer after they have ended. Interestingly, the recent recession, often labeled as the ‘Great Recession’, seems to have combined the worst of these different dimensions, in the sense that it was long lasting (from December 2007 until June 2009, according to the NBER business cycle dating committee), severe (with an average annualized quarterly GDP growth rate of  $-3.5\%$  between 2008Q1 and 2009Q2), and followed by a sluggish recovery.

Not surprisingly, a substantial body of research has been devoted to understanding the characteristics and dynamics of recessions. One of the key points of interest in this literature has been the determinants of the severity of recessions, dating back at least to Fisher (1933). Motivated mostly by the ‘Great Depression’ during 1929-1933, several theories on the dynamics of recessions have been put forward specifically addressing the question why some recessions turn out to be more severe than others. Mishkin (1978), Bernanke (1981), Romer (1993), Bernanke *et al.* (1996, 1999), and Gertler and Kiyotaki (2010), among others, discuss several mechanisms and channels having adverse effects on output and aggregate demand during recessions. These include the lack of financial in-

mediation, the uncertainty brought about by stock market crashes changing consumer preferences, and the deteriorated household balance-sheets caused by increasing real burdens due to deflationary periods during recessions, among others. A common theme in these theories is that recessions start with mild negative shocks but different mechanisms (such as the credit channel, the uncertainty channel, the balance sheet channel, and so forth) can amplify their effects or trigger more severe shocks and, hence, affect the severity of the recession.

Motivated by both empirical evidence and theory, in this paper we statistically analyze the relevance of distinguishing different types of economic downturns. Applying Markov-Switching (Vector) AutoRegressive (MS-(V)AR) models, we examine whether it is useful to discriminate between mild and severe recessions in monthly coincident and leading indicator variables. Our analysis consists of three parts. We first conduct a univariate analysis on the Conference Board Coincident Economic Index (CEI) and its four constituents (the number of employees on non-agricultural payrolls (ENP), personal income less transfer payments (PI), industrial production (IP) and manufacturing and trade sales (MTS)) as well as the Conference Board Leading Economic Index (LEI). By analyzing both individual coincident variables as well as the CEI, we aim to uncover the characteristics of regimes each individual variable embeds and to explore the effects of aggregating individual variables to a single index on the regime switching dynamics. Second, we employ a multivariate analysis using a coincident variable that represents the business cycle most accurately together with the LEI. Specifically, we use IP for that purpose, as it appears from our univariate analysis that this variable captures the NBER recessions and expansions most closely. Third and finally, we attempt to link our empirical findings to the theories concerning the determinants of the severity of recessions. In particular, we document empirical evidence in favor of the ‘financial accelerator’ theory (Bernanke *et al.*, 1996, 1999), by showing a link between the behavior of credit spreads and the severity of recessions.

Our univariate analysis shows that a three regime MS-AR model is most appropriate for most of the individual business cycle indicators as well as for the CEI and the LEI. In most of the cases the regimes may be characterized as expansions, mild recessions and severe recessions. This finding contrasts most of the previous studies that apply Markov (and other types of) regime-switching models in this context, which typically find that the third regime (beyond expansions and recessions) captures a recovery or ‘bounce back’ phase, see Sichel (1994), Boldin (1996), and Clements and Krolzig (2003), among others. A notable exception is Hamilton (2005), who reports a three regime Markov-switching model for the postwar unemployment rate also with separate regimes for mild and severe

recessions. For the CEI, we find that the two types of recessions actually correspond with recessions before and after the mid 1980s. This distinction, however, is mostly due to the occurrence of slow recoveries following the more recent recessions, leading to rather prolonged recession signals in the early 1990s and 2000s (as these mild recoveries are considered as part of the recessions), which is at odds with the NBER business cycle chronology. This effect mostly arises because of the sluggish improvement in labor market conditions (hence the name ‘jobless recoveries’, see Gordon (1993); Groshen and Potter (2003)), which is confirmed by the fact that we find the same phenomenon for employment. On the other hand, the three regime model for industrial production captures both mild and severe recessions successfully and produces signals very close to the NBER recession dates. The posterior probabilities reveal three severe recession periods, occurring during the recessions of 1974-5, 1980 and 2007-9. A three regime model of LEI results in similar findings.

In our multivariate analysis, we use the MS-VAR model developed in Paap *et al.* (2009) and Cakmakli *et al.* (2011). This model has two attractive features. First, by employing a multivariate model of similar variables it provides more precise regime signals and better predictions. Second, as the model identifies the degree of synchronization of the cycles in the IP and the LEI it allows us to assess the LEI’s ability to predict the severity of recessions. Here, by synchronization we mean that the different variables in fact share a single common cyclical component but subject to different phase shifts, where the lead-lag time can differ across regimes. We show that this type of MS-VAR model with three regimes describes the US business cycle better than the models with two regimes. Results indicate that LEI is most timely in predicting moderate recessions with a lead time of 12 months, while the lead time of severe recessions is (only) 6 months.

It is useful to note that in the context of two-regime Markov-switching models several specifications have been proposed to capture different recession shapes. Hamilton (1989)’s original formulation with regime-switching mean growth rates and homoskedastic errors implies that recessions are so-called ‘L’-shaped, as there is no subsequent fast-growth recovery phase. Kim *et al.* (2005) augment Hamilton’s original model with a so-called ‘bounce-back’ term, such faster growth in the quarters immediately following a recession can be generated. The strength of the recovery can be linked to the length of the preceding recession or its depth, or both. Depending on the exact specification, the model can capture ‘V’-shaped recessions, characterized by a strong recovery after a sharp contraction, or ‘U’-shaped recessions with a smoother transition from contractions to expansions. Morley and Piger (forthcoming) provide an excellent recent comparison of these bounce-back specifications. While the bounce-back approach has intuitive appeal, we do not adopt it

in this paper. Note that it focuses on the properties of the post-recession period, and its relation with the preceding recessions. The recessions themselves, however, are assumed to be all identical, in the sense that the recession regime is characterized by a constant mean growth rate, independent of the length or severity of the ongoing recession (or of anything else). Given that the main purpose of our analysis is to explore whether it is useful to distinguish different types of recessions, we consider Markov-switching models with multiple, unrestricted regimes.

Besides the stimulating effects of lower price levels and interest rates on consumption (wealth effect) as well as on investment (Keynes effect) during recessions, several mechanisms have been put forward that may lead to the deepening of a contraction (and hence the occurrence of a severe recession). Mishkin (1978) emphasizes the effect of deflation on aggregate demand during the Great Depression, related to the increasing real debt burden of households. This caused consumers to cut their spending on illiquid assets (durable goods, residential housing assets, and so forth) and/or to liquidation, thereby shifting household's balance-sheet toward liquidity. Considering the possible relation between the Great Crash in 1929 and the subsequent Great Depression, Romer (1990) points out the link between uncertainty and consumption decisions. The stock market crash in 1929 and the resulting long-lasting extreme variation in stock prices increased uncertainty for consumers about their future income stream. This, in general, caused people to defer their irreversible, durable goods consumption and to increase consumption of reversible (nondurable) goods. Bernanke *et al.* (1996, 1999) focus on the 'small shocks, large cycles' puzzle in their 'financial accelerator' theory. Referring to the amplification of initial shocks brought about by changes in credit-market conditions (the 'financial accelerator'), this theory suggests that first, borrowers facing relatively high agency costs in credit markets will bear the brunt of economic downturns (due to the fact that many investors attempt to move their money into relatively safe investments, i.e. the flight to quality). Second, reduced spending, production, and investment by high-agency-cost borrowers will exacerbate the effects of recessionary shocks. We document empirical evidence in favor of the financial accelerator theory by showing a link between the behavior of credit spreads and the severity of recessions.

The remainder of the paper is organized as follows. We first provide a description of the data and provide a preliminary analysis in Section 4.2. We describe the univariate Markov-switching AR model and discuss our univariate empirical findings in Section 4.3. In Section 4.4, we discuss the Markov-switching VAR model with different phase shifts for multiple regimes. We also provide the multivariate empirical results in this section. We document the link between the behavior of credit spreads and the severity of recessions

in Section 4.5. Finally, we conclude in Section 4.6. Technical details are deferred to a set of appendices.

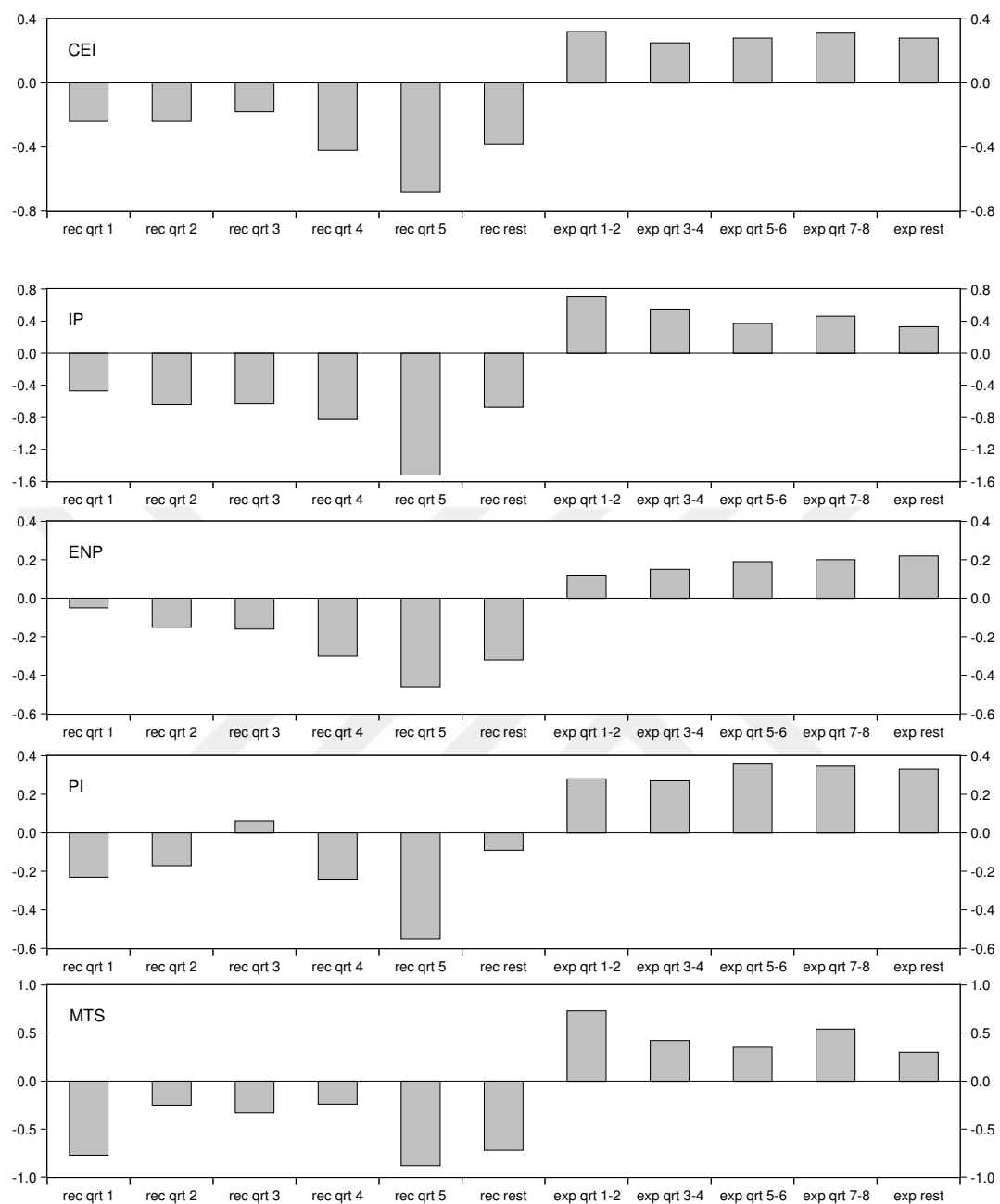
## 4.2 Stylized facts of US recessions

We analyze the Conference Board Coincident Economic Index (CEI), its four components, and the Conference Board Leading Economic Index (LEI). The CEI is composed of the number of employees on non-agricultural payrolls (ENP), personal income less transfer payments (PI), industrial production (IP) and manufacturing and trade sales (MTS). Monthly observations for all variables are available for the period from January 1960 until October 2010.

Figure 4.1 shows the average growth rates of the CEI and its four components during the course of the recessions that occurred during our sample period and their aftermath, see for example Sichel (1994) for a similar analysis. The graphs show the average growth rate in different quarters of recessions as defined by the NBER turning points, and during six-month periods of expansions. The graphs indicate that for the CEI, ENP and IP recessions progressively become more severe, in the sense that their average growth rates monotonically decline during the different quarters of the contraction periods. Note that this goes against the conventional wisdom that recessions typically start with a sharp downturn. While this sort of pattern in the first recession quarter does seem to be present for PI and MTS, we also observe large negative growth rates during the fourth and fifth quarters of recessions for those variables. Hence, contraction periods that last relatively long (i.e. more than three quarters) have a severe latter part also for PI and MTS.

For expansions, we do not observe a uniform pattern across variables. Only MTS follows a pattern consistent with high growth recoveries following recessions. The average growth rate during the first six months of expansions is equal to 0.75 percent, which gradually declines as the expansion continues. IP displays similar behavior, albeit much less pronounced. For ENP the pattern is reversed in the sense that average growth rate gradually increases during expansions periods. This indicates the sluggish adjustment process of the employment, which was particularly evident during the periods following the 1990-1 and 2001 recessions. When constructing the CEI it seems that these opposite patterns in the individual coincident variables are averaged out to a large extent, such that the average growth rate is approximately the same for all six-month subperiods of expansions.

Figure 4.1: Average growth rates of coincident economic indicators over the course of recessions and expansions during the period January 1960 - October 2010



*Note:* 'rec qrt  $j$ ',  $j = 1, \dots, 5$ , denotes the  $j$ -th quarter of recessions, while 'rec rest' denotes the remaining periods of recessions. 'exp qrt  $2j - 1$  to  $2j$ ',  $j = 1, \dots, 4$ , denotes quarters  $2j - 1$  and  $2j$  of expansions, while 'exp rest' denotes the remaining periods of expansions.

## 4.3 Univariate analysis

We analyze the business cycle dynamics in the individual variables by means of univariate Markov-Switching AutoRegressive (MS-AR) models. Our aim is to determine the number of regimes we should distinguish in order to adequately describe the cyclical features in these series, as well as the corresponding regime characteristics.

Let  $y_t$  denotes the growth rate of a given coincident or leading indicator in month  $t$ . We assume that the business cycle can be divided into  $J$  phases or regimes, which are characterized by different means of  $y_t$ . Unexpected growth, denoted by  $\varepsilon_t$ , is assumed to be normally distributed with time-varying volatility  $\sigma_t$ . The exact specification of the volatility dynamics is described in detail below. We assume that autoregressive coefficients are constant across regimes.<sup>1</sup> In case of first-order autoregressive dynamics, our assumptions imply the model specification

$$y_t - \mu_{S_t} = \phi(y_{t-1} - \mu_{S_{t-1}}) + \varepsilon_t \quad \text{with } \varepsilon_t \sim NID(0, \sigma_t^2), \quad (4.1)$$

where  $S_t$  is latent multinomial variable taking the value  $j$  if  $y_t$  is in regime  $j$  in month  $t$ , and  $\mu_{S_t} = E[y_t|S_t]$  denotes the unconditional mean of  $y_t$  in the regime indicated by  $S_t$ . The regime-indicator variable  $S_t$  is assumed to be a first-order homogenous Markov process with transition probabilities

$$\Pr(S_t = j|S_{t-1} = i) = p_{ij} \quad \text{for } i, j = 1, \dots, J. \quad (4.2)$$

As one of our aims is to assess the usefulness of additional regimes (on top of expansion and recession phases) to characterize the business cycle dynamics, the specification of the error variance  $\sigma_t^2$  in (4.1) is not innocuous. Specifically, ignoring regime-switching behavior in the variance may spuriously suggest that multiple regimes are present. For example, assume the business cycle is characterized by two regimes that not only have different mean growth rates but also different variances. Imposing homoskedasticity, we may then find that it is necessary to allow for a third regime, essentially approximating the high volatility regime with two regimes with different mean growth rates. To avoid this issue we consider two possibilities for the time-varying volatility  $\sigma_t$  in (4.1). In the first specification, we assume the volatility to be constant across regimes, while in the second we allow for regime-dependent heteroskedasticity. In both cases we incorporate a single structural break in the volatilities, to accommodate the Great Moderation, that is, the

---

<sup>1</sup>Extending the model to allow for regime-dependent autoregressive dynamics is straightforward. This is not pursued here to keep the complexity of the model at a feasible level. Notice that regime dependence in the mean growth rate already provides a great deal of flexibility.

large and persistent decline in macroeconomic volatility in the mid-1980s, see McConnell and Perez-Quiros (2000), among others. In case of regime-dependent heteroskedasticity we keep the volatility specification parsimonious by imposing that the proportional change in volatility at the time of the structural break is the same in all regimes. This corresponds with the specification

$$\sigma_t = (\delta I[t < \tau] + I[t \geq \tau])\sigma_{S_t}, \quad (4.3)$$

where  $I[A]$  is an indicator function for the event  $A$ ,  $\tau$  is the period when the structural break in volatility occurs, and  $\delta$  gives the ratio of the volatilities before and after the break. The homoskedastic specification boils down to imposing  $\sigma_{S_t} = \sigma$  in (4.3).

### 4.3.1 Estimation and model selection

We use a Bayesian approach for estimation and inference in the MS-AR model using Markov Chain Monte Carlo (MCMC) techniques. Specifically, we use Gibbs sampling together with data augmentation (see Geman and Geman, 1984; Tanner and Wong, 1987) to obtain posterior results. The estimation of MS-AR models with Bayesian techniques has become common practice in business cycle analysis. Therefore we do not provide a detailed exposition of the estimation procedure here, but we refer to Krolzig (1997) and Kim and Nelson (1999), among others. A description of the estimation of the variance process with a structural break, as specified in (4.3), is provided in the Appendix 4.A.4 for the multivariate setting considered in the next section.

A few details concerning our Bayesian estimation procedure are useful to note. First, we use noninformative priors for all model parameters, in order to let the data decide on the most appropriate specification. Second, we use the prior for the regime-dependent mean parameters to identify the regimes. Specifically, we set the prior specification for  $\mu = (\mu_1, \mu_2, \dots, \mu_J)'$  as

$$f(\mu) \propto \begin{cases} 1 & \text{if } \mu \in \{\mu \in \mathbb{R}^J | \mu_1 < \dots < \mu_j < \dots < \mu_J\} \\ 0 & \text{elsewhere.} \end{cases} \quad (4.4)$$

Note that this prior only involves inequality restrictions and does not impose a particular sign on any of the mean growth rates. Hence, we only rank order the regimes in terms of their mean growth rates but do not enforce any specific type of regimes such as recovery or severe recession to be present *a priori*.

For each of the variables we estimate models with two, three and four regimes with the two different variance specifications. Each model is estimated without and with first-order

autoregressive dynamics to account for possible autocorrelation that cannot be captured by the Markov-switching mean (and volatility).

For selecting the most appropriate model specification, we use several criteria. First, we consider the predictive likelihoods of the models. To compute these we first estimate the model parameters using only part of the full sample (denoted as ‘estimation sample’) and then based on these estimates we evaluate the likelihood of the remaining part. This approach has the advantage that it is not affected by the choice of prior distributions and over-fitting, while it is directly related to the posterior model probabilities, see Geweke and Amisano (2010, 2011). We use the sample from January 1960 until December 2000 for estimation of the model and the remaining period from January 2001 until October 2010 to compute the predictive likelihood. Computational details are provided in Appendix 4.A.5.

Second, we examine the robustness and stability of the models by comparing the estimation results based upon the initial estimation sample and the full sample.

Third, we check the ability of the models to produce a reasonable description of the business cycle. Using the NBER turning points to define recessions and expansions, we check the compatibility of the posterior regime probabilities with this chronology.

### 4.3.2 Empirical Results

The estimation results reveal very similar conclusions for three variables, namely CEI, IP and LEI. The results for ENP are also comparable except for the dating of turning points, in particular troughs. Due to the presence of several post-recession periods with sluggish recovery in labor market conditions, the MS-AR models for employment contain rather long recessions, ending considerably later than the NBER trough dates. The models for MTS and PI show much less correspondence with the NBER recession dates such that these variables seem to have their own regime-switching dynamics. Therefore, we provide only detailed results of CEI, IP and LEI for the sake of brevity.<sup>2</sup>

Table 4.1 displays the log predictive likelihoods of the univariate models for CEI, IP and LEI, giving rise to several interesting observations. First, for the models with constant variances the predictive likelihood values tend to increase dramatically when increasing the number of regimes from two to three. This is most pronounced for CEI and IP. This trend continues when we move to a four-regime model for the CEI, albeit only slightly. For the LEI we also find an increase in predictive likelihood with the number of regimes, but this is much more modest than for the CEI and IP.

<sup>2</sup>All remaining results are available upon request.

Table 4.1: Comparison of univariate MS-AR models: Log predictive likelihood values

|                        | AR(0)     |           |           | AR(1)     |           |           |
|------------------------|-----------|-----------|-----------|-----------|-----------|-----------|
|                        | 2 Regimes | 3 Regimes | 4 Regimes | 2 Regimes | 3 Regimes | 4 Regimes |
| <u>Results for CEI</u> |           |           |           |           |           |           |
| constant $\sigma$      | -28.6     | -12.3     | -9.3      | -22.4     | -11.7     | -12.8     |
| reg.-dep. $\sigma$     | -25.4     | -58.8     | -13.2     | -19.8     | -20.0     | -8.0      |
| <u>Results for IP</u>  |           |           |           |           |           |           |
| constant $\sigma$      | -130.1    | -112.1    | -123.0    | -123.5    | -113.7    | -123.7    |
| reg.-dep. $\sigma$     | -117.1    | -122.1    | -127.2    | -117.9    | -132.8    | -125.2    |
| <u>Results for LEI</u> |           |           |           |           |           |           |
| constant $\sigma$      | -88.0     | -86.8     | -85.8     | -88.5     | -83.3     | -84.7     |
| reg.-dep. $\sigma$     | -89.9     | -109.8    | -89.2     | -89.2     | -76.7     | -86.0     |

*Note:* The table presents log predictive likelihood values for MS-AR models estimated for monthly growth rates of the Conference Board coincident economic index (CEI), industrial production index (IP), and the Conference Board leading economic index (LEI). The estimation sample used for obtaining a posterior sample of the model parameters covers the period January 1960 - December 2000. The remaining period January 2001 - October 2010 is used to compute log predictive likelihood values. Constant  $\sigma$  stands for the model where in each subperiod (separated by a structural break in the variance) homoskedasticity is assumed. Regime-dependent (reg.-dep.)  $\sigma$  stands for the model with regime-dependent heteroskedasticity together with a structural break in variances. AR(0) and AR(1) indicate models with no and first-order autoregressive dynamics, respectively. Posterior results are based on 40,000 simulations of which the first 20,000 are discarded as burn-in sample. The convergence of the MCMC sampler is checked using statistical and visual inspection and in all model specifications convergence is assured.

When the aim is to capture recessions and expansions in the most general sense with no further refinements, Markov-switching models with two regimes are the obvious candidates. In that case, it seems that allowing for regime-dependent heteroskedasticity is beneficial, in the sense that for the CEI and IP the predictive likelihood is higher than for the models with constant variance. This does not hold, however, for the LEI. Interestingly, Table 4.1 shows that increasing the number of regimes in the model with regime-specific volatilities leads to a decline in predictive likelihood, which in many cases is quite substantial. In fact, for all three variables we find that three-regime models with constant variances provide superior predictive results than the two-regime models with regime-dependent heteroskedasticity.

When we compare the models with and without first-order autoregressive dynamics (denoted by AR(0) and AR(1)), we observe mixed results. Especially when the number of regimes is restricted to only two, adding AR(1) dynamics provides an improvement in model predictions. The same, however, does not apply uniformly when more regimes are considered. It seems that regime-switching means (and variances) only are not sufficient to describe the cyclical behavior in these series completely in case only two regimes are allowed. In that case, the first-order autoregressive dynamics is useful to capture (part of) the remaining structure. With the increase in the number of regimes this limited effect vanishes completely, however. It is also important to note that unreported results indicate that the models with AR(1) dynamics lack robustness, in the sense that we find rather different parameter estimates for these models when using the initial estimation

sample until December 2000 compared to using the complete sample period until October 2010.<sup>3</sup>

The results show that for all variables the models with three and four regimes and with constant variances with a single structural break and no autoregressive dynamics provide relatively higher predictive likelihood values.<sup>4</sup> Moreover, these models are robust against the sample chosen. Therefore, we display the results of these models in Tables 4.2, 4.3 and 4.4 for CEI, IP and LEI, respectively. We also include the results of the two regime model with regime-dependent heteroscedasticity and a single structural break for comparison. The posterior regime probabilities for these models are displayed in Figures 4.2, 4.3 and 4.4.

---

<sup>3</sup>Using the estimation sample until December 2000, the models with AR(1) dynamics give similar parameter estimates as the corresponding models without autoregressive dynamics. This is also the main source of the high predictive likelihood values for some of these models.

<sup>4</sup>A possible explanation for this finding may be that the regime-switching dynamics already provides sufficient flexibility to the model. Note that, a model with no AR terms but with Markov switching regime dynamics can be written as an ARMA specification, see Hamilton (1989) for details.

Table 4.2: Posterior results of selected univariate Markov-switching models for the CEI

| MS(2)AR(1) with regime-dependent $\sigma$         |          |             |             |              |              |
|---------------------------------------------------|----------|-------------|-------------|--------------|--------------|
|                                                   |          | Regime 1    | Regime 2    |              |              |
| Mean, $\mu_j$                                     |          | 0.22 (0.02) | 0.01 (0.13) |              |              |
| Variance after the structural break, $\sigma_j^2$ |          | 0.04 (0.01) | 0.16 (0.05) |              |              |
| Most likely break date, $\tau$                    |          | 1984-1      |             |              |              |
| Ratio between st.dev's, $\delta$                  |          | 1.51 (0.15) |             |              |              |
| Autoregressive coefficient, $\phi$                |          | 0.37 (0.08) |             |              |              |
|                                                   |          | Regime 1    | Regime 2    |              |              |
| Transition probabilities, $p_{ij}$                | Regime 1 | 0.93 (0.04) | 0.07 (0.04) |              |              |
|                                                   | Regime 2 | 0.25 (0.09) | 0.75 (0.09) |              |              |
| MS(3)AR(0) with constant $\sigma$                 |          |             |             |              |              |
|                                                   |          | Regime 1    | Regime 2    | Regime 3     |              |
| Mean, $\mu_j$                                     |          | 0.35 (0.03) | 0.10 (0.04) | -0.45 (0.06) |              |
| Variance after the structural break, $\sigma^2$   |          | 0.04 (0.00) |             |              |              |
| Most likely break date, $\tau$                    |          | 1984-2      |             |              |              |
| Ratio between st.dev's, $\delta$                  |          | 1.54 (0.10) |             |              |              |
|                                                   |          | Regime 1    | Regime 2    | Regime 3     |              |
| Transition probabilities, $p_{ij}$                | Regime 1 | 0.96 (0.02) | 0.04 (0.01) | 0.01 (0.01)  |              |
|                                                   | Regime 2 | 0.03 (0.02) | 0.92 (0.03) | 0.05 (0.02)  |              |
|                                                   | Regime 3 | 0.14 (0.06) | 0.11 (0.06) | 0.75 (0.07)  |              |
| MS(4)AR(0) with constant $\sigma$                 |          |             |             |              |              |
|                                                   |          | Regime 1    | Regime 2    | Regime 3     | Regime 4     |
| Mean, $\mu_j$                                     |          | 0.67 (0.20) | 0.25 (0.07) | -0.00 (0.02) | -0.52 (0.06) |
| Variance after the structural break, $\sigma^2$   |          | 0.04 (0.00) |             |              |              |
| Most likely break date, $\tau$                    |          | 1984-2      |             |              |              |
| Ratio between st.dev's, $\delta$                  |          | 1.48 (0.14) |             |              |              |
|                                                   |          | Regime 1    | Regime 2    | Regime 3     | Regime 4     |
| Transition probabilities, $p_{ij}$                | Regime 1 | 0.53 (0.28) | 0.31 (0.20) | 0.07 (0.09)  | 0.09 (0.09)  |
|                                                   | Regime 2 | 0.03 (0.03) | 0.92 (0.05) | 0.04 (0.03)  | 0.01 (0.02)  |
|                                                   | Regime 3 | 0.03 (0.07) | 0.08 (0.10) | 0.81 (0.17)  | 0.08 (0.06)  |
|                                                   | Regime 4 | 0.10 (0.06) | 0.11 (0.07) | 0.13 (0.09)  | 0.66 (0.10)  |

*Note:* The table presents posterior means and standard deviations (in parentheses) of parameters in the selected MS-AR models estimated for monthly growth rates of the Conference Board coincident economic index (CEI) over the period January 1960 - October 2010. MS( $J$ )AR( $K$ ) is the abbreviation for the Markov-switching model with  $J$  regimes and with  $K^{th}$  order autoregressive dynamics. Constant  $\sigma$  stands for the model where in each subperiod (separated by a structural break in the variance) homoskedasticity is assumed. Regime-dependent  $\sigma$  stands for the model with regime-dependent heteroskedasticity together with a structural break in variances. The most likely break date is defined as the mode of the posterior distribution of  $\tau$ . Posterior results are based on 40,000 simulations of which the first 20,000 are discarded as burn-in sample. The convergence of the MCMC sampler is checked using statistical and visual inspection and in all model specifications convergence is assured.

Table 4.3: Posterior results of selected univariate Markov-switching models for the IP

| MS(2)AR(1) with regime-dependent $\sigma$        |          |             |              |
|--------------------------------------------------|----------|-------------|--------------|
|                                                  |          | Regime 1    | Regime 2     |
| Mean, $\mu_j$                                    |          | 0.28 (0.04) | -0.28 (0.29) |
| Variance after the structural break $\sigma_j^2$ |          | 0.24 (0.03) | 1.54 (0.47)  |
| Most likely break date, $\tau$                   |          | 1984-2      |              |
| Ratio between st.dev's, $\delta$                 |          | 1.45 (0.23) |              |
| Autoregressive coefficient, $\phi$               |          | 0.27 (0.05) |              |
|                                                  |          | Regime 1    | Regime 2     |
| Transition probabilities, $p_{ij}$               | Regime 1 | 0.96 (0.02) | 0.05 (0.02)  |
|                                                  | Regime 2 | 0.29 (0.10) | 0.71 (0.10)  |

| MS(3)AR(0) with constant $\sigma$               |          |             |              |              |
|-------------------------------------------------|----------|-------------|--------------|--------------|
|                                                 |          | Regime 1    | Regime 2     | Regime 3     |
| Mean, $\mu_j$                                   |          | 0.39 (0.03) | -0.47 (0.10) | -2.45 (0.25) |
| Variance after the structural break, $\sigma^2$ |          |             |              |              |
| Most likely break date, $\tau$                  |          | 1984-2      |              |              |
| Ratio between st.dev's, $\delta$                |          | 1.53(0.10)  |              |              |
|                                                 |          | Regime 1    | Regime 2     | Regime 3     |
| Transition probabilities, $p_{ij}$              | Regime 1 | 0.96 (0.01) | 0.03 (0.01)  | 0.00 (0.00)  |
|                                                 | Regime 2 | 0.13 (0.05) | 0.80 (0.05)  | 0.07 (0.03)  |
|                                                 | Regime 3 | 0.28 (0.15) | 0.31 (0.16)  | 0.41 (0.15)  |

| MS(4)AR(0) with constant $\sigma$               |          |             |             |             |              |
|-------------------------------------------------|----------|-------------|-------------|-------------|--------------|
|                                                 |          | Regime 1    | Regime 2    | Regime 3    | Regime 4     |
| Mean, $\mu_j$                                   |          | 1.72 (0.36) | 0.50 (0.05) | 0.00 (0.04) | -1.77 (0.21) |
| Variance after the structural break, $\sigma^2$ |          | 0.24 (0.03) |             |             |              |
| Most likely break date, $\tau$                  |          | 1984-2      |             |             |              |
| Ratio between st.dev's, $\delta$                |          | 1.44 (0.30) |             |             |              |
|                                                 |          | Regime 1    | Regime 2    | Regime 3    | Regime 4     |
| Transition probabilities, $p_{ij}$              | Regime 1 | 0.16 (0.13) | 0.49 (0.19) | 0.24 (0.16) | 0.11 (0.10)  |
|                                                 | Regime 2 | 0.01 (0.01) | 0.93 (0.02) | 0.06 (0.02) | 0.01 (0.01)  |
|                                                 | Regime 3 | 0.03 (0.02) | 0.05 (0.03) | 0.89 (0.03) | 0.04 (0.02)  |
|                                                 | Regime 4 | 0.16 (0.08) | 0.14 (0.09) | 0.16 (0.12) | 0.54 (0.11)  |

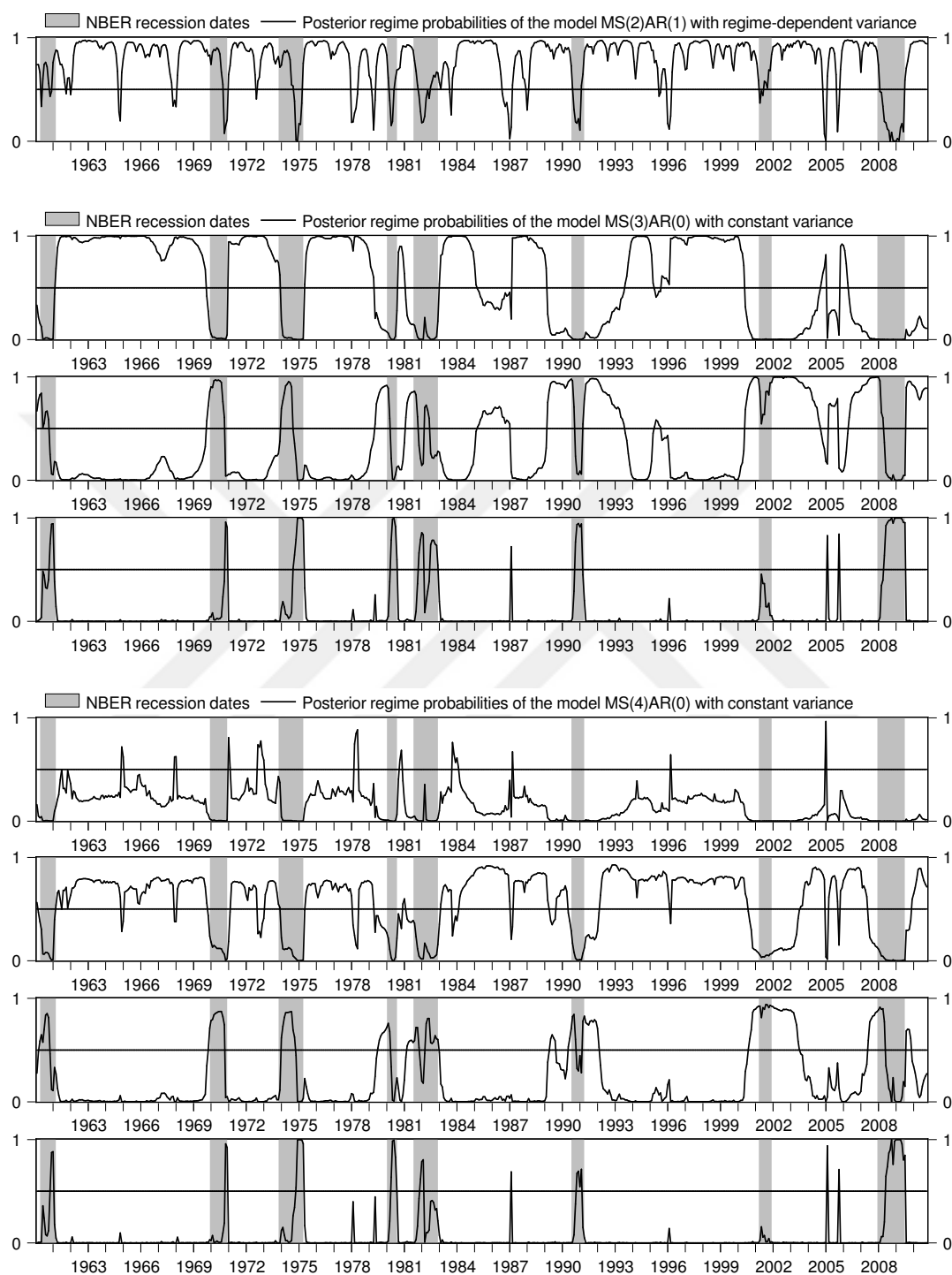
*Note:* The table presents posterior means and standard deviations (in parentheses) of parameters in the selected MS-AR models estimated for monthly growth rates of industrial production (IP) over the period January 1960 - October 2010. See Table 4.2 for further details.

Table 4.4: Posterior results of selected univariate Markov-switching models for the LEI

| MS(2)AR(1) with regime-dependent $\sigma$        |          |             |              |              |              |
|--------------------------------------------------|----------|-------------|--------------|--------------|--------------|
|                                                  |          | Regime 1    | Regime 2     |              |              |
| Mean, $\mu_j$                                    |          | 0.41 (0.06) | -0.33 (0.17) |              |              |
| Variance after the structural break $\sigma_j^2$ |          | 0.20 (0.03) | 0.40 (0.17)  |              |              |
| Most likely break date, $\tau$                   |          | 1984-2      |              |              |              |
| Ratio of st.dev's, $\delta$                      |          | 1.31 (0.12) |              |              |              |
| Autoregressive coefficient, $\phi$               |          | 0.26 (0.07) |              |              |              |
|                                                  |          | Regime 1    | Regime 2     |              |              |
| Transition probabilities, $p_{ij}$               | Regime 1 | 0.97 (0.01) | 0.03 (0.01)  |              |              |
|                                                  | Regime 2 | 0.08 (0.03) | 0.92 (0.03)  |              |              |
| MS(3)AR(0) with constant $\sigma$                |          |             |              |              |              |
|                                                  |          | Regime 1    | Regime 2     | Regime 3     |              |
| Mean, $\mu_j$                                    |          | 0.46 (0.03) | -0.27 (0.07) | -1.64 (0.30) |              |
| Variance after the structural break, $\sigma^2$  |          | 0.20 (0.02) |              |              |              |
| Most likely break date, $\tau$                   |          | 1983-7      |              |              |              |
| Ratio between st.dev's, $\delta$                 |          | 1.37 (0.09) |              |              |              |
|                                                  |          | Regime 1    | Regime 2     | Regime 3     |              |
| Transition probabilities, $p_{ij}$               | Regime 1 | 0.96 (0.01) | 0.03 (0.01)  | 0.00 (0.00)  |              |
|                                                  | Regime 2 | 0.07 (0.03) | 0.90 (0.03)  | 0.03 (0.02)  |              |
|                                                  | Regime 3 | 0.23 (0.12) | 0.11 (0.10)  | 0.66 (0.13)  |              |
| MS(4)AR(0) with constant $\sigma$                |          |             |              |              |              |
|                                                  |          | Regime 1    | Regime 2     | Regime 3     | Regime 4     |
| Mean, $\mu_j$                                    |          | 1.06 (0.11) | 0.40 (0.04)  | 0.00 (0.04)  | -0.98 (0.14) |
| Variance after the structural break, $\sigma^2$  |          | 0.17 (0.02) |              |              |              |
| Most likely break date, $\tau$                   |          | 1984-2      |              |              |              |
| Ratio between st.dev's, $\delta$                 |          | 1.31 (0.12) |              |              |              |
|                                                  |          | Regime 1    | Regime 2     | Regime 3     | Regime 4     |
| Transition probabilities, $p_{ij}$               | Regime 1 | 0.71 (0.08) | 0.18 (0.07)  | 0.04 (0.03)  | 0.03 (0.02)  |
|                                                  | Regime 2 | 0.01 (0.01) | 0.95 (0.02)  | 0.03 (0.02)  | 0.01 (0.01)  |
|                                                  | Regime 3 | 0.05 (0.02) | 0.02 (0.02)  | 0.83 (0.05)  | 0.10 (0.05)  |
|                                                  | Regime 4 | 0.07 (0.04) | 0.03 (0.03)  | 0.22 (0.10)  | 0.68 (0.10)  |

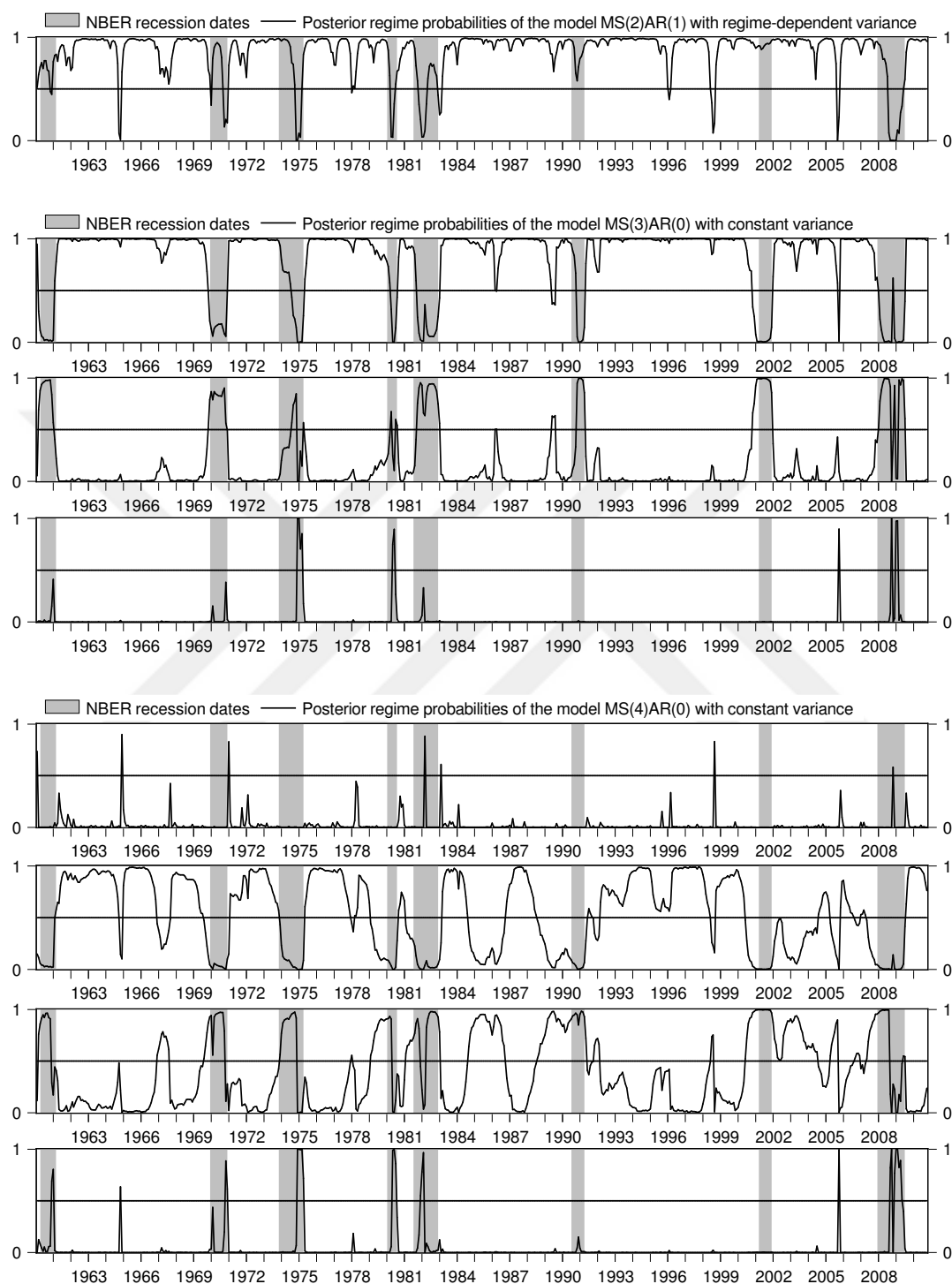
*Note:* The table presents posterior means and standard deviations (in parentheses) of parameters in the selected MS-AR models estimated for monthly growth rates of the Conference Board leading economic index (LEI) over the period January 1960 - October 2010. See Table 4.2 for further details.

Figure 4.2: Posterior regime probabilities for selected models of the CEI



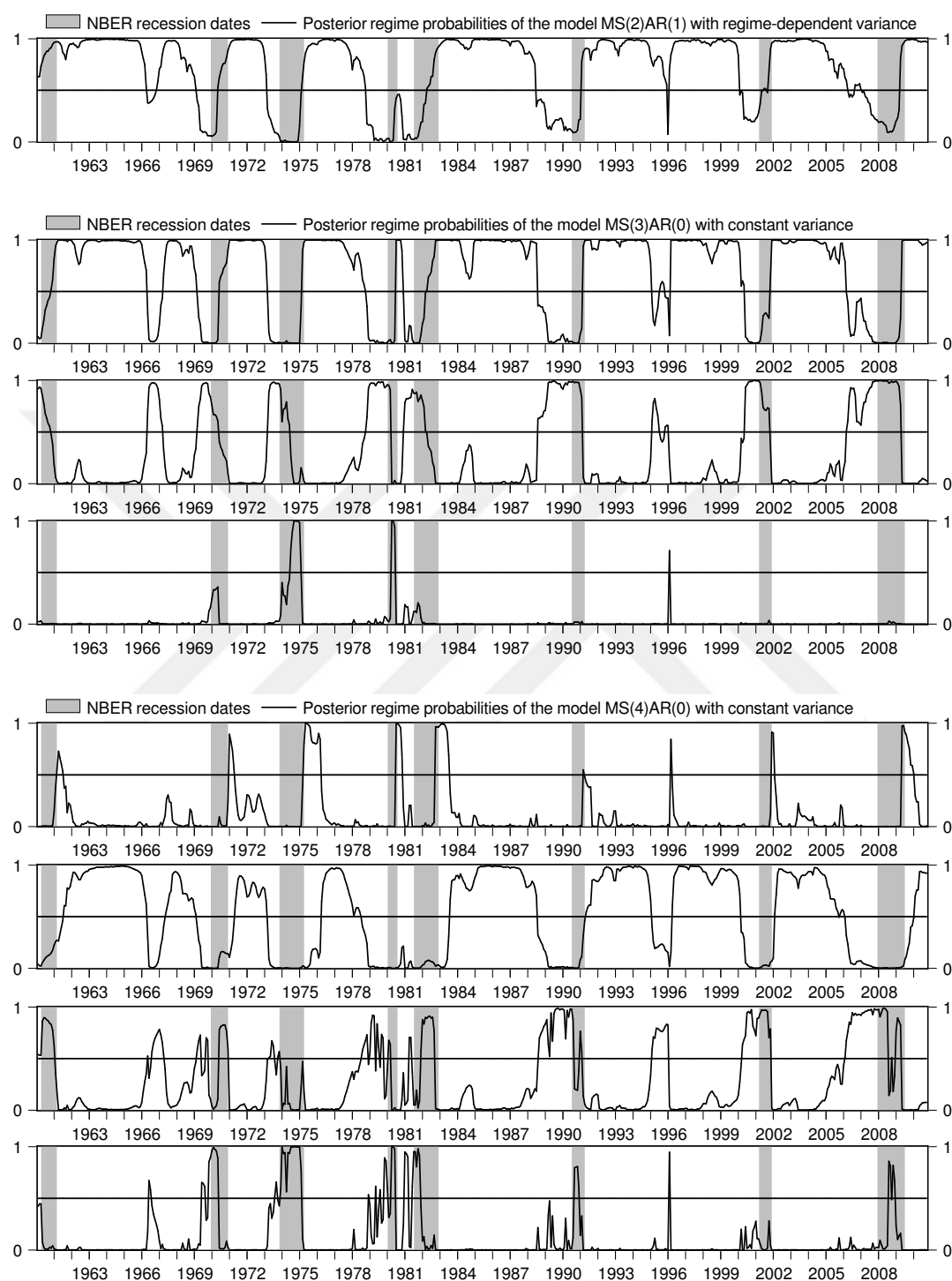
*Note:* The solid lines are the posterior regime probabilities for the selected models summarized in Table 4.2 for the Conference Board coincident economic index (CEI). The shaded areas indicate the US recessions as determined by the NBER Business Cycle Dating Committee.

Figure 4.3: Posterior regime probabilities for selected models of IP



*Note:* The solid lines are the posterior regime probabilities for the selected models summarized in Table 4.3 for industrial production (IP). The shaded areas indicate the US recessions as determined by the NBER Business Cycle Dating Committee.

Figure 4.4: Posterior regime probabilities for selected models of LEI



*Note:* The solid lines are the posterior regime probabilities for the selected models summarized in Table 4.4 for the Conference Board leading economic index (LEI). The shaded areas indicate the US recessions as determined by the NBER Business Cycle Dating Committee.

For the models with two regimes for CEI and IP, it seems that the variance dynamics is the driving force of the regimes rather than the mean dynamics. While the estimates for the mean growth rates are different in the two regimes, note that only for IP we find a negative growth rate for one of the regimes (but with a large posterior standard deviation), while for the CEI we find one regime with positive growth and one with essentially zero growth, on average. The differences in variances across regimes are much larger, in the sense that for CEI and IP the variances in the regimes with the highest mean growth are, respectively, four and six times smaller than the variances during the other regime. From Figures 4.2 and 4.3 we observe that the correspondence of the “low growth, high volatility” regime with the NBER recessions is far from perfect. The mild recessions of 1990-1 and 2001 are missed almost completely, while at other times temporary high volatility wrongfully signals a ‘recession’; for example in 1987, 1996 and 2005 for the CEI. From the first panel of Figure 4.4, we see that the two regime model for the LEI can provide more accurate signals for recessions. Obviously, the mismatch between the NBER recession dates is not because of the low quality of signals but because of the leading property of this indicator. Still, it seems that at the onset of LEI’s recessions there are mixed signals that are far from 0 or 1 especially for the period around 2006.

The most interesting finding for the three regime models is that, in contrast to the existing evidence, the additional third regime is a recession type of regime rather than a recovery phase as in Sichel (1994) and Boldin (1996), among others. For all three variables the estimation results indicate the presence of separate phases of mild and severe recessions. For the CEI, the mean growth rate during mild recessions is positive but rather close to zero, while severe recessions are characterized by a large negative mean growth rate of  $-0.45$  percent per month. One drawback of this specification, however, is that it produces false signals following the recessions of 1990-1 and 2001, in the sense that the probability of the second regime is fairly high for a considerable period of time after the end of these recessions. Unreported results indicate very similar results for ENP; hence, the second regime also captures those periods with basically no growth or limited growth in employment conditions, in particular the jobless recovery periods that are not consistent with the NBER recession dating chronology.

The three regime model of IP on the other hand signals very accurate predictions of the recessions consistent with NBER peak and trough dates. As shown by the graphs in the middle panel of Figure 4.3, the third regime captures the latter part of the recessions in 1974, 1980 and 2008. At  $-2.45$  percent the mean growth rate in this regime is much lower than that of the second regime, representing a mild recession phase with mean growth of  $-0.47$  percent. Clearly, the additional third regime does not capture isolated outlying

observations as the predictive likelihoods dramatically increase with the inclusion of this regime and also the probability of staying in this regime is far from 0.

The same conclusion applies to the LEI. The third regime has similar characteristics as that for IP with a very large mean contraction rate, but with phase shifts in terms of the timing of the start of this regime (as expected). It is also important to note that inclusion of the third regime actually improves the recession signals of the LEI compared to the two regime model, especially around 2006.

Unreported results for the other variables lead to similar conclusions. The main exception is that for MTS and PI the additional third regime indicates a recovery phase following recessions. Besides the fact that these models have very low predictive likelihoods, they also seriously lack robustness.<sup>5</sup>

Given the distinction between mild and severe recessions in the three regime models and previous evidence for the existence of a recovery phase, one might expect superior performance of four regime models, as these may be able to capture both phenomena simultaneously. However, this turns out not to be the case. The estimation results in Tables 4.2-4.4 and Figures 4.2-4.4 indicate that the additional fourth regime does not represent a recovery regime, except for the LEI to some extent. In particular for IP, the regime with the highest mean growth rate (1.72) captures only a few outliers, as shown by the occasional spikes in regime probability in Figure 4.3 and the low probability (0.16) of staying in the first regime. Albeit less dramatically, the same applies to the CEI. These results are in line with the predictive likelihoods in Table 4.1 where we find slight improvements for the four regime model (compared to the three regime model) for the LEI and the CEI, but a worsening for IP.

When the break dates of the variances are considered, there is a consensus across all the variables and across the various model specifications. The estimates indicate the first quarter of 1984 as the most likely break date, in line with existing evidence, see McConnell and Perez-Quiros (2000) and Sensier and van Dijk (2004), among others. The ratios of the standard deviations before and after the break are approximately 1.5 for the CEI and IP and 1.35 for the LEI, confirming the substantial reduction in volatility in macroeconomic fluctuations due to the Great Moderation.

The results from the univariate analysis indicate that the Markov-switching models with three regimes provide a much improved description of the business cycle dynamics compared to the models with two regimes only. While the models for the CEI show a dramatic increase in predictive likelihoods as the number of regimes is increased, this does not translate into an improved recession signal. We attribute this to the aggregation of

---

<sup>5</sup>Detailed results are available upon request.

different variables with distinct business cycle features to construct the CEI, in particular ENP showing slow ('jobless') recovery, IP showing clear signs of the occurrence of 'severe recessions', and PI and MTS showing quick recoveries following contractions. On the other hand, IP and LEI provide first, more accurate predictions brought about by the additional third regime (as indicated by the predictive likelihoods); second, robust models with estimates that are not sensitive to the sample chosen; and third, clear signals of the recession dates and recession types (see Figures 4.3 and 4.4). We therefore proceed in the next section with a bivariate analysis of IP and LEI using an MS-VAR model.

## 4.4 Multivariate analysis

The posterior regime probabilities of the univariate MS models for IP and LEI indicate that both variables display similar characteristics and dynamics in their business cycle regimes, see Figures 4.3 and 4.4. We therefore consider the joint modeling of these two series, to obtain a clearer picture of the US business cycle. Obviously, there is a phase shift in the timing of the regime changes for the two variables, as IP is a coincident variable while the LEI leads the business cycle. Hence, when constructing a bivariate MS-VAR model for IP and LEI this phase shift has to be taken into account. Moreover, the time shift seems to be asymmetric across the business cycle. Indeed, using a two-regime MS-VAR model Paap *et al.* (2009) show that the LEI leads business cycle peaks by twelve months and troughs by three months, on average. In this paper we extend the analysis of Paap *et al.* (2009) to more than two regimes using an adjusted version of the model proposed by Cakmakli *et al.* (2011).

The model of Cakmakli *et al.* (2011) is a bivariate MS-VAR model with 'imperfect synchronization' of the cycles in the two variables due to asymmetric phase shifts of a single underlying Markov regime-switching process. The key feature of this model is that it allows for any number of regimes  $J \geq 2$  and that the amount of the phase shifts can be different across regimes. Furthermore, regime-dependent heteroskedasticity can easily be implemented.

Let  $y_{1,t}$  and  $y_{2,t}$  denote the observations of the IP and LEI growth rates in month  $t$  for  $t = 1, \dots, T$ . As in the univariate case, we assume that the  $J$  phases of the business cycle are characterized by different means of  $y_{1,t}$  and  $y_{2,t}$ . We assume that autoregressive coefficients are constant across regimes. In case of first-order autoregressive dynamics this

leads to the model specification

$$\begin{aligned} y_{1,t} - \mu_{1,S_{1,t}} &= \phi_{1,1}(y_{1,t-1} - \mu_{1,S_{1,t-1}}) + \phi_{1,2}(y_{2,t-1} - \mu_{2,S_{2,t-1}}) + \varepsilon_{1,t}, \\ y_{2,t} - \mu_{2,S_{2,t}} &= \phi_{2,1}(y_{1,t-1} - \mu_{1,S_{1,t-1}}) + \phi_{2,2}(y_{2,t-1} - \mu_{2,S_{2,t-1}}) + \varepsilon_{2,t}, \end{aligned} \quad (4.5)$$

where  $S_{l,t}$  are latent multinomial variables taking the value  $j$  if  $y_{l,t}$  is in regime  $j$  at time  $t$ , and where  $\mu_{l,S_{l,t}} = \mathbb{E}[y_{l,t}|S_{l,t}]$  denotes the unconditional mean of  $y_{l,t}$  in regime  $S_{l,t}$  for  $l = 1, 2$ . The disturbances are assumed to be normally distributed with mean zero and time-varying covariance matrix  $\Sigma_t$ , that is,

$$(\varepsilon_{1,t}, \varepsilon_{2,t})' \sim NID(0, \Sigma_t). \quad (4.6)$$

Before we discuss the specification of the covariance matrix  $\Sigma_t$ , we first consider the properties of the regime indicators  $S_{l,t}$ .

To model the dynamics in the regime indicators, we use again a first-order homogenous Markov process. A natural and elegant approach to model the systematic phase shifts between the cycles of LEI and IP is to assume that the cycle in  $y_{2,t}$  leads the cycle in  $y_{1,t}$  by  $\kappa$  periods (Hamilton and Perez-Quiros, 1996)

$$S_{2,t-\kappa} = S_{1,t}. \quad (4.7)$$

In other words, there is a common cycle but it affects the different variables with a certain phase shift. We refer to this case as imperfect synchronization with ‘symmetric’ phase shifts (SPS), where symmetry refers to the fact that all possible regime transitions in the two variables differ by the same number of time periods  $\kappa$ .

The specification in (4.7) may still be too restrictive, in the sense that the phase shift of the cycle may well be different for different regimes transitions. For example, leading indicator variables typically have a considerably longer lead time at business cycle peaks than at troughs, see The Conference Board (2001). For this purpose, Paap *et al.* (2009) consider a two-regime model with possibly different phase shifts  $\kappa_1$  and  $\kappa_2$  for these two types of turning points. Cakmakli *et al.* (2011) generalize the idea of imperfect synchronization with ‘asymmetric’ phase shifts (APS) to multiple regimes ( $J > 2$ ) by specifying

$$S_{2,t-\kappa_{S_{1,t}}} = S_{1,t}, \quad (4.8)$$

where the subscript  $S_{1,t}$  to  $\kappa$  indicates that the regime indicator is shifted by a possibly different number of time periods  $\kappa_j$  for each regime  $j = 1, \dots, J$ . Put differently, we assume that the lead time is different per regime, such that each regime in the LEI starts

earlier by  $\kappa_j$  periods than the corresponding regimes in the IP, see Cakmakli *et al.* (2011) for model details. Note that the specification in (4.8) embeds imperfect synchronization with symmetric phase shifts (all  $\kappa_j \neq 0$  but equal) as a special case, as well as what we might call perfect synchronization (all  $\kappa_j = 0$ ).<sup>6</sup>

To finalize the model specification, we return to the specification of the covariance matrix  $\Sigma_t$  in (4.6). As in the univariate case, we consider constant and regime-dependent variances, while in both cases we allow for a single structural break in volatilities to accommodate the Great Moderation. To facilitate the specification, the covariance matrix is decomposed into variances and correlations, as

$$\Sigma_t = D_t R D_t, \quad (4.9)$$

where  $D_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t})$  is a diagonal matrix with the standard deviations of the error terms as diagonal elements and  $R$  is a matrix with ones on the diagonal and the correlation  $\rho$  as the off-diagonal element (see, for example, Barnard *et al.*, 2000). In case of regime-dependent variances, the elements of  $D_t$  are modeled as

$$\sigma_{l,t} = \begin{cases} \delta \sigma_{l,S_{l,t}} & \text{if } t < \tau \\ \sigma_{l,S_{l,t}} & \text{if } t \geq \tau \end{cases} \quad \text{for } l = 1, 2, \quad (4.10)$$

where we allow for a single break in the variances at time  $\tau$  captured by the single scaling factor  $\delta$ . To keep the complexity of the model at a feasible level, the correlation parameter  $\rho$  is assumed to be regime-independent and constant over time unlike the variances. Assuming constant variances across regimes boils down to imposing  $\sigma_{l,S_{l,t}} = \sigma_l$ . In this case, we also allow the correlation to change at the time of the structural break in the volatilities.

#### 4.4.1 Bayesian Inference

As in the univariate case, we opt for a Bayesian approach to estimate the parameters of the bivariate MS-VAR model with imperfect synchronization. Note that this greatly facilitates inference on the value of the discrete lead time parameters  $\kappa_j$ ,  $j = 1, \dots, J$  (and the timing of the structural break in variances  $\tau$ ), which is problematic in a frequentist analysis.

---

<sup>6</sup>The specification in (4.8) is not complete, in the sense that it may lead to situations where for some time periods  $S_{2,t}$  is assigned to multiple regimes (or to no regime at all). We impose the rule that the regime with the larger phase shift parameter determines the regime of  $S_{2,t}$  in such conflicting (or ‘empty’) periods, see Cakmakli *et al.* (2011) for discussion.

As we want our posterior results to be driven by the data rather than the prior distributions, we impose rather diffuse prior specifications for the model parameters. Details of the adopted priors are provided in Appendix 4.A.2. The posterior distribution of the model parameters and unobserved states is proportional to the product of the complete likelihood function and the joint prior, which is the product of the marginal prior distributions. The complete likelihood function is derived in Appendix 4.A.3. As in the univariate case, we use Gibbs sampling together with data augmentation to obtain posterior results. Details on the derivation of the conditional posterior distributions are given in Appendix 4.A.4. Model selection is carried out in a similar manner as in the univariate analysis. We compute marginal likelihood values to compare models with different types of synchronization.

## 4.4.2 Empirical Results

Following the evidence from our univariate analysis we estimate MS-VAR models with two or three regimes. For the relation between the regime dynamics in IP and LEI, we consider both possibilities discussed in the previous section, namely (i) imperfect synchronization with a single phase shift of the common cycle (SPS); and (ii) imperfect synchronization with asymmetric (or regime-specific) phase shifts (APS).

For the models with imperfect synchronization, we specify the set of admissible phase shift parameters to be in the interval  $[-15, 15]$ . Hence any regime in the cycle of the LEI can lead/lag the corresponding regime in the IP by at most 15 months. We also restrict the differences between the phase shift parameters for different regimes to be one year at most. For the autocorrelation dynamics, we impose  $\phi_{1,1} = \phi_{2,1} = 0$  as marginal likelihoods computed using mildly informative priors for these parameters are hardly affected by this restriction, see Paap *et al.* (2009) and Hamilton and Perez-Quiros (1996) for a similar result. The prior for the break date parameter  $\tau$  is set in such a way that the break in the variances can occur everywhere apart from the first and last 6 months of the sample.

We use the sample from January 1960 until December 2000 for estimation of the model and the remaining period from January 2001 until October 2010 to compute the predictive likelihood as in the univariate analysis. We focus on the predictive ability of the models for the IP series, as compared to LEI prediction of IP is of first-order importance.

The marginal and predictive likelihood results in Table 4.5 show three interesting features. First, both the marginal and the predictive likelihood increase when moving from the SPS to the APS specification, with the only exception of the three regime model with regime-dependent variances. This clearly shows the importance of allowing

for different phase shifts for the different regimes. Second, in line with our previous results inclusion of the third regime increases the predictive likelihood values in most cases. Third, imposing equal variances across regimes improves the predictive likelihood for all specifications.<sup>7</sup> Combining these three findings, we conclude that the three regime model with asymmetric phase shifts and equal variances across regimes performs best in terms of predictive ability. The parameter estimates (posterior means and standard deviations) for this model specification are displayed in Table 4.6 and posterior probabilities of regimes are shown in Figure 4.5.

Table 4.5: Comparison of multivariate MS-VAR models estimated for IP and LEI

|                  | Log marginal likelihood |                    | Log predictive likelihood |                    |
|------------------|-------------------------|--------------------|---------------------------|--------------------|
|                  | constant $\Sigma$       | reg.-dep. $\Sigma$ | constant $\Sigma$         | reg.-dep. $\Sigma$ |
| <u>2 regimes</u> |                         |                    |                           |                    |
| SPS              | -1196.4                 | -1144.8            | -132.2                    | -152.0             |
| APS              | -1164.7                 | -1120.5            | -124.8                    | -148.2             |
| <u>3 regimes</u> |                         |                    |                           |                    |
| SPS              | -1228.8                 | -1150.0            | -116.3                    | -144.1             |
| APS              | -1136.6                 | -1117.6            | -107.9                    | -144.4             |

*Note:* The table presents log marginal likelihoods and log predictive likelihoods of the competing MS-VAR models with (i) two or three regimes; (ii) imperfect synchronization with a single phase shift (SPS) or regime-specific (or asymmetric) phase shifts (APS) of the common cycle and (iii) homoskedastic error terms (constant  $\Sigma$ ) or with regime-dependent variances (regime-dependent  $\Sigma$ ), where in both cases the variances are subject to a single structural break and the correlation is assumed to be the same across regimes and constant over time. The models are applied to monthly growth rates of industrial production (IP) and the Conference Board leading economic index (LEI) for the sample period January 1960 - October 2010. Marginal likelihoods are those obtained when the models are estimated for the complete sample period. Predictive likelihood values are for the period January 2001 - October 2010, which are computed conditional on the posterior distributions of the model parameters obtained with an estimation sample ending in December 2000. Posterior results for computing log marginal likelihoods and predictive likelihoods are based on 40,000 simulations of which the first 20,000 are discarded as burn-in sample. The convergence of the MCMC sampler is checked using statistical and visual inspection and in all model specifications convergence is assured.

The estimation results in Table 4.6 indicate once again the necessity of distinguishing between two different types of economic declines.<sup>8</sup> On the one hand, the regime with mean

<sup>7</sup>This observation is not valid for marginal likelihoods, however, marginal likelihoods of the models with constant and regime-dependent variances are not comparable as we specify improper prior distributions for the variance parameters.

<sup>8</sup>Note that the estimation results also show that it is worthwhile to combine the information in IP and the LEI by jointly modeling these two series. The precision of the parameter estimates is improved compared to the univariate models, as well as the accuracy of the posterior regime probabilities. Finally, the predictive likelihood values obtained from the bivariate model are also higher than those from the corresponding univariate three regime model for IP.

growth rates of  $-0.5$  and  $-0.25$  for IP and the LEI can be labeled as ‘mild recessions’. On the other hand, the regime with much larger declines of  $-2.4$  and  $-1.5$  percent could be called ‘severe recessions’.

Table 4.6: Posterior results for three regime MS-VAR model for IP and LEI with asymmetric synchronization, homoskedastic shocks and a structural break in the variances

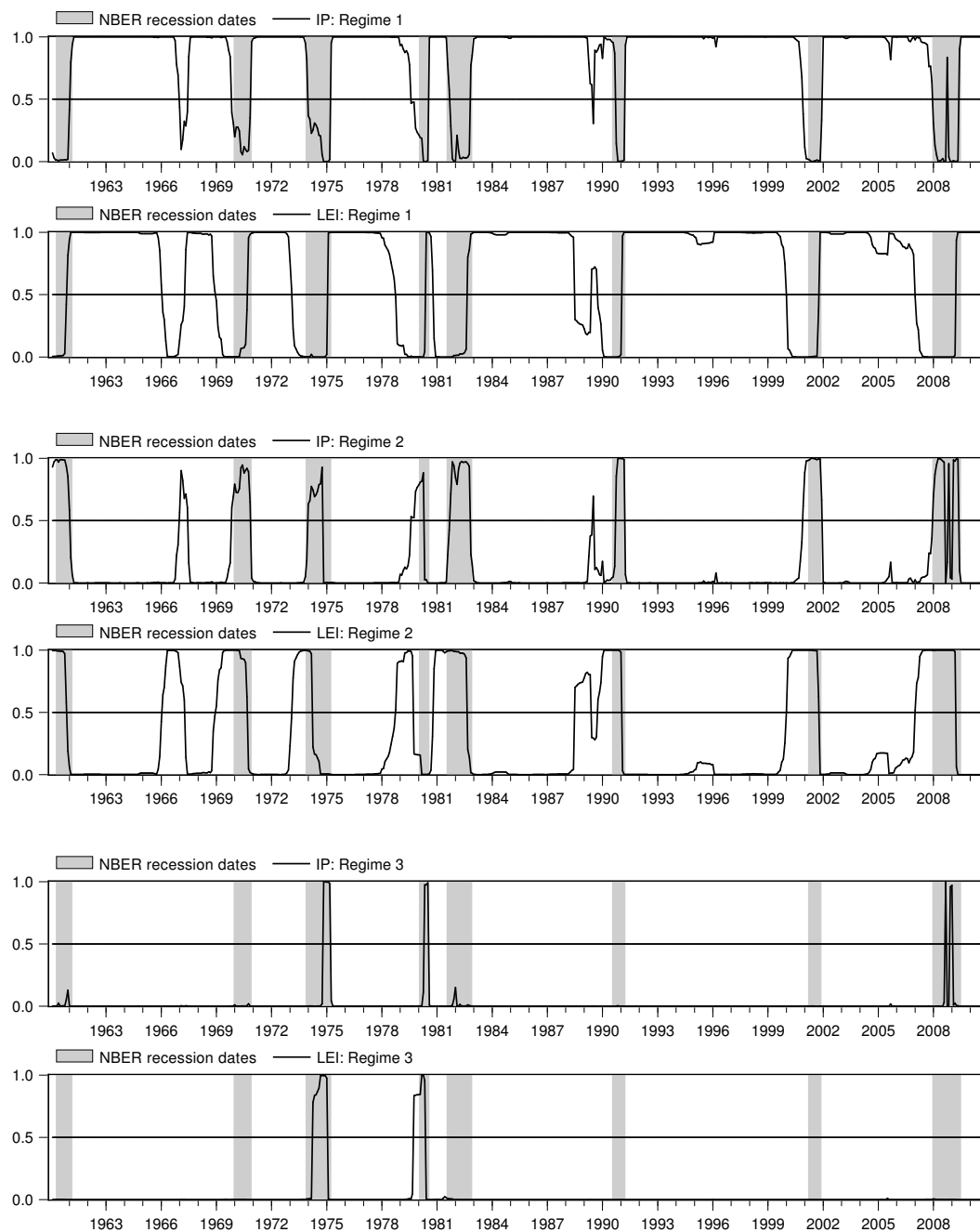
|                                                   |              | IP             | LEI            |               |
|---------------------------------------------------|--------------|----------------|----------------|---------------|
| Growth Rates, $\mu_{i,j}$                         | Regime 1     | 0.371 (0.028)  | 0.407 (0.033)  |               |
|                                                   | Regime 2     | -0.535 (0.087) | -0.242 (0.060) |               |
|                                                   | Regime 3     | -2.395 (0.200) | -1.508 (0.219) |               |
|                                                   |              |                |                |               |
| Most likely break date, $\tau$                    | 1984-2       |                |                |               |
| Variance before break                             |              | 0.526 (0.049)  | 0.388 (0.035)  |               |
| Variance after break                              |              | 0.245 (0.022)  | 0.221 (0.019)  |               |
| Correlation between the error terms of IP and LEI | before break |                | 0.457 (0.052)  |               |
|                                                   | after break  |                | 0.263 (0.057)  |               |
| Autoregressive dynamics                           |              | $\phi_{1,2}$   | $\phi_{2,2}$   |               |
|                                                   |              | 0.059 (0.050)  | 0.162 (0.043)  |               |
| Transition Probabilities                          |              | Regime 1       | Regime 2       | Regime 3      |
|                                                   | Regime 1     | 0.973 (0.008)  | 0.025 (0.008)  | 0.002 (0.002) |
|                                                   | Regime 2     | 0.112 (0.041)  | 0.831 (0.050)  | 0.057 (0.027) |
|                                                   | Regime 3     | 0.272 (0.119)  | 0.169 (0.105)  | 0.559 (0.130) |
| Lead/Lag times                                    |              | Regime 1       | Regime 2       | Regime 3      |
|                                                   |              | 2.042 (0.201)  | 10.838 (1.326) | 6.231 (1.722) |

*Note:* The table presents posterior means and standard deviations (in parentheses) of parameters of the MS-VAR model with three regimes and a single structural change in the covariance matrix, estimated using monthly growth rates of IP and LEI over the period January 1960 - October 2010. The most likely break date is defined as the mode of the posterior distribution of  $\tau$ . Posterior results are based on 40,000 draws from the parameter distribution and convergence is verified using visual and statistical tools. Number of burn-in simulations is 20,000. The convergence of the MCMC sampler is checked using statistical and visual inspection and in all model specifications convergence is assured.

The posterior regime probabilities in Figure 4.5 indicate that severe recessions occurred in 1974-5, 1980, and 2008-9, with all three being preceded by a period of mild recession. We return to this issue in detail in the next section. Interestingly, the first two severe recessions were followed immediately by expansion periods, while the most recent one was followed by another mild recession period before entering the expansion regime in June 2009. Note that the LEI did signal the severe recessions in 1974-5 and 1980 in advance, but it failed to provide an early warning for the most recent one in 2008-9. This is due to the fact that this phase was followed by a mild recession period again, unlike the previous crises in 1974-5 and 1980. This indicates the necessity for more sophisticated leading indicators or at least multiple indicators each of which pre-signaling specific phases of the state of the economy as earlier as possible.

The regime probabilities in 2008-9 show a peculiar pattern anyway. As seen in Figure 4.5, the posterior probabilities convincingly indicate a severe recession during the

Figure 4.5: Posterior probabilities of regimes of IP and LEI



*Note:* The solid lines are the posterior regime probabilities for industrial production (IP) and the Conference Board leading economic index (LEI) for the three regime MS-VAR model with imperfect synchronization with asymmetric (or regime-specific) phase shifts and with a single structural break in the covariance matrix, summarized in Table 4.6. The shaded areas indicate the US recessions as determined by the NBER Business Cycle Dating Committee.

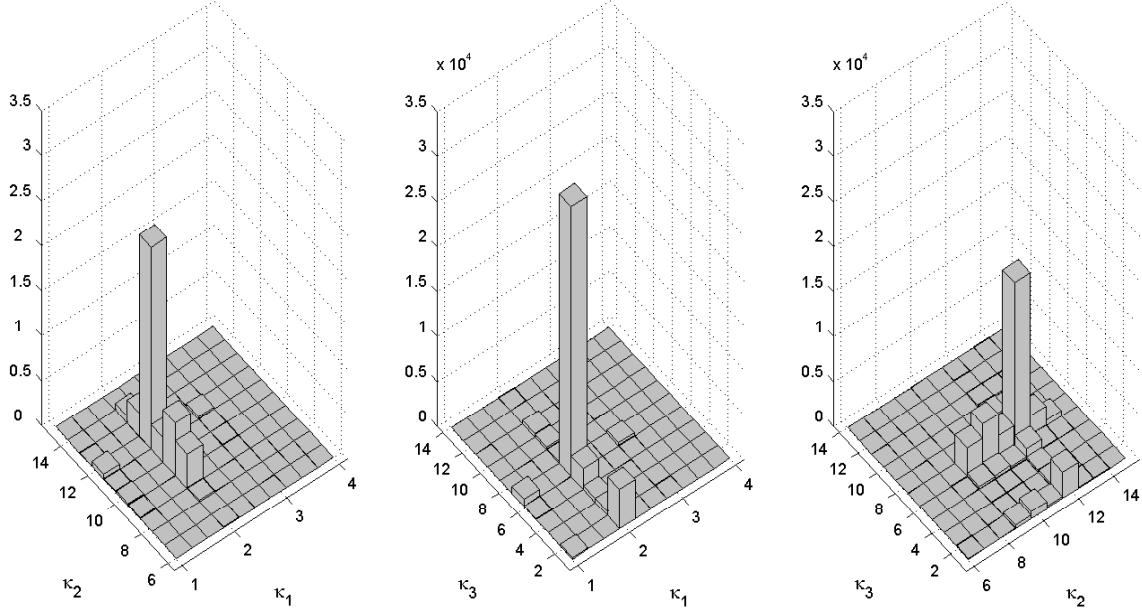
period September 2008 - January 2009, but interrupted by a spike in the probability of the expansion regime in October 2008 (and in the probability of the mild recession regime in November and December). This corresponds with a 1.26% increase in IP during that month, following the large decline of 4.04% in September. One possible explanation for this event might be the very low base level for the October growth rate. A more theoretical explanation is based on Romer (1990)'s uncertainty hypothesis, that is, increasing stock market volatility gives rise to consumers deferring their consumption on durable goods and increasing their consumption on nondurables. Indeed, calculating monthly realized stock market volatility based on daily S&P 500 returns we find that volatility was at its highest level in October and November 2008 (the second highest level during the entire postwar period after the October 1987 crash). Interestingly, durable goods IP data in these months showed a 2.3% decline, whereas nondurable goods increased by 2.7%, the only positive realization between May 2008 and March 2009.

While showing much stronger declines in output, severe recessions last considerably shorter than mild contractions. According to the transition probability estimates the average duration of a typical mild recession is  $1/(1 - 0.83) \approx 6$  months, whereas severe recessions only last on average  $1/(1 - 0.56) \approx 2.2$  months.

The posterior distribution of the lead/lag parameters  $\kappa_j$ ,  $j = 1, \dots, 3$  is shown in Figure 4.6. For the expansion, mild recession and severe recession phases, the mean lead times are 2, 10.8 and 6.2 months, respectively. Posterior mode values point out a very similar picture with values of 2, 12 and 7 months for the three regimes. As the posterior standard deviations are quite small and all the marginal distributions are unimodal, mode and mean values are very close to each other. These results demonstrate that the LEI is more successful in signaling upcoming (mild) recessions than expansions, in agreement with results in Paap *et al.* (2009) based on a two-regime model for the CEI and LEI. In addition, our results suggest that the LEI is somewhat less timely for severe recessions, although the lead time of just over half a year is still substantial.

Based on the estimates of the  $\kappa_j$  parameters we can also draw inference about the relative duration of each regime. Recall that the estimation results show that all recession periods start as a mild downturn. Hence, based on the modes of the posterior distributions a typical expansion regime in the LEI is  $(2-12=-10)$  10 months shorter than that of IP. Recessions can either remain mild for their entire duration, or worsen into a severe downturn. In the former case, a typical (mild) recession in the LEI is  $(12-2=10)$  10 months longer than that of IP while in the latter case it is only  $(12-7=5)$  5 months longer.

Figure 4.6: Posterior distribution of lead time parameters



*Note:* The graphs show the posterior distribution of the lead time parameters  $\kappa_j$ ,  $j = 1, 2, 3$ , in the three regime MS-VAR model with imperfect synchronization with asymmetric (or regime-specific) phase shifts for industrial production (IP) and the Conference Board leading economic index (LEI), summarized in Table 4.6.

As before, the structural break in volatility is dated in February 1984. The complete posterior distribution for  $\tau$  shows that this estimate is quite precise. Hence, the uncertainty about the timing of the Great Moderation is small.

Before closing this section, some remarks about a possible ‘recovery’ phase (instead of or in addition to a severe recession regime) in the US business cycle are in order. First of all, it is important to note that our model specification does not enforce the nature of the third (or any other) regime. The restrictions on the mean parameters in (4.4) pin down the regimes only in terms of rank order but not in terms of the sign of the mean growth rates. Hence our three regime model could also have resulted in identifying a recovery regime. In fact, we estimated three regime models with certain restrictions on the mean growth rates to force the model to describe a recovery phase (in addition to expansions and recessions). The estimation results (available upon request) show that, when imposed, a recovery regime occurs mostly following the recessions before 1990. This indicates that recoveries follow more severe recessions, suggesting that an extension of the model in that direction may be useful, see Morley and Piger (forthcoming) for various

possibilities. Interestingly, the lead time of the LEI for the recovery regime is almost the same as for the expansion regime. Hence, it conveys limited information about the prediction of the end of recessions beyond the information embodied in the expansion regime. Moreover, this model resulted in poor predictive likelihood value, below that of the two and three regime models. Hence, we conclude that the model we present here explains the business cycle characteristics of IP and the LEI better than that model with a third recovery regime.

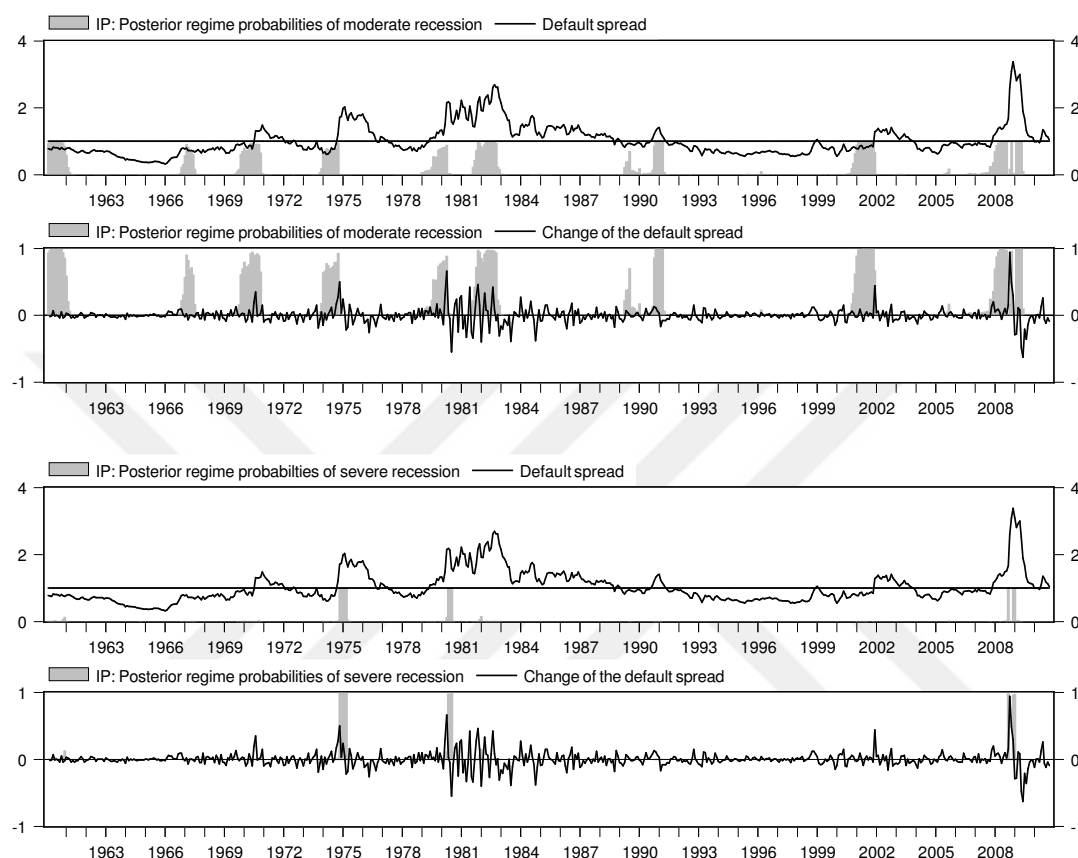
## 4.5 Severe recessions and the theory of the financial accelerator

The economic downturns identified in our previous analysis have in common that they all start as a mild recession. Differences occur as the downturn proceeds, in the sense that some recessions remain relatively mild while others turn into severe contractions. This empirical result establishes a link to economic theories about the severity of recessions. In particular, the finding that all recessions start with relatively small shocks but then differ in terms of the amplification of these adverse shocks and/or the magnitude of the following shocks is consistent with the theory of the ‘financial accelerator’. This states that borrowers facing relatively high agency costs in credit markets will bear the brunt of economic downturns (due to the flight to quality mechanism). This suggests that credit spreads, that is the difference in yields on corporate bonds issued by firms with low and high credit quality, should widen during recessions. Furthermore, the increase in these spreads may be expected to be larger for deeper contractions. To examine this possibility, Figure 4.7 displays the posterior probabilities of the mild and severe recession regimes together with the (level and first difference of the) spread between Moody’s Baa and Aaa corporate bond yields. Note that this spread should indicate the severity of the credit squeeze in extreme market conditions, as it measures the yield difference between two groups of relatively safe investment-grade bonds.

Figure 4.7 reveals an interesting pattern. The highest levels of the credit spread are clearly associated with the severe recessions identified in 1974-5, 1980 and 2007-9. Especially the most pronounced increases of the spread correspond to the start of the severe recession phases. The relation between the mild recession regime and the credit spread is much less clear. If an increase in the spread can be observed at all, it is very minor compared to the level shifts around severe recessions. The behavior of the credit

spread clearly provides evidence of the relation between the financial accelerator and the third regime captured in our analysis.

Figure 4.7: Posterior recession probabilities and Moody's Baa-Aaa credit spread



*Note:* The graphs show the posterior regime probabilities for mild and severe recessions in the three regime MS-VAR model with imperfect synchronization with asymmetric (or regime-specific) phase shifts for industrial production (IP) and the Conference Board leading economic index (LEI), summarized in Table 4.6, together with the level and first difference of the spread between Moody's Baa and Aaa corporate bond yields.

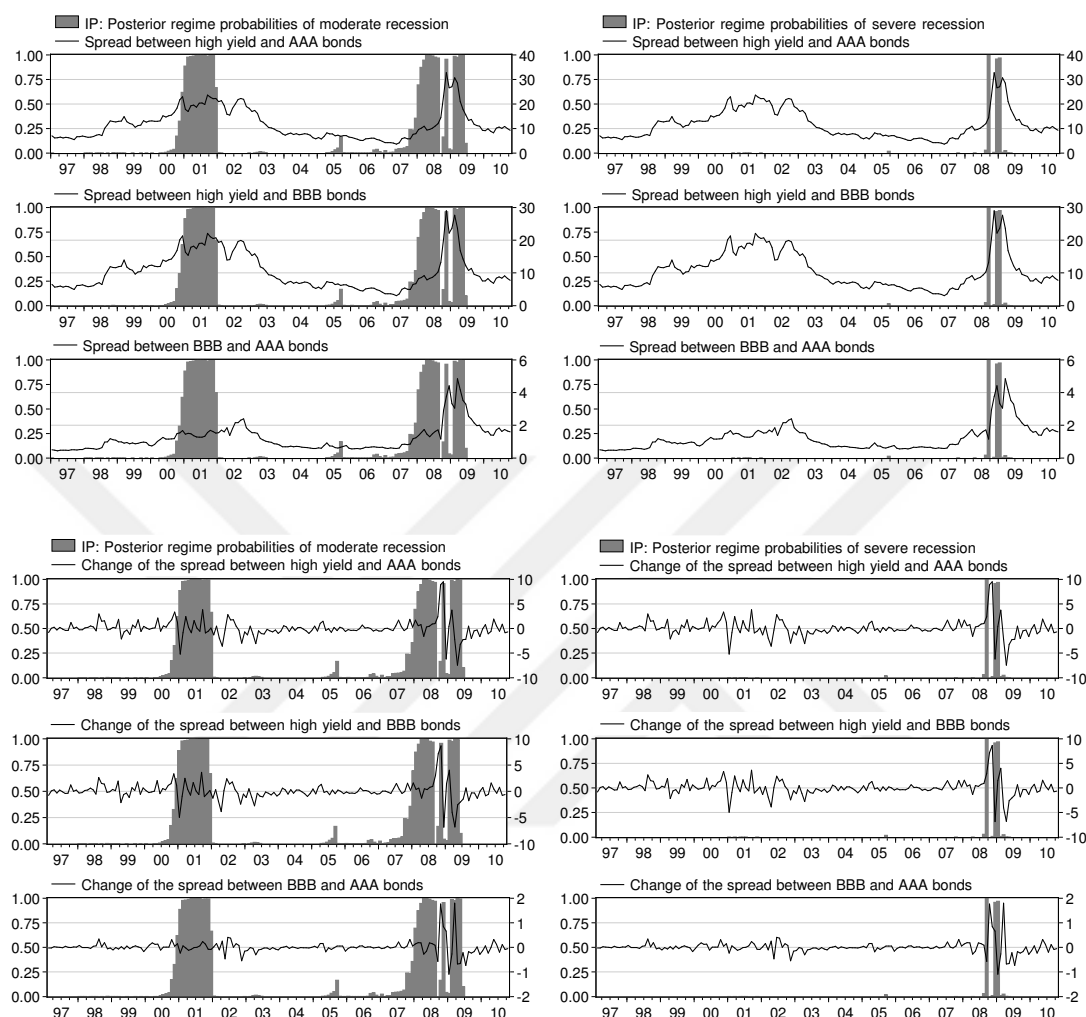
We further analyze this effect using more specific data on corporate bonds, comprising option-adjusted-spreads (OAS) of investment-grade corporate bonds with ratings AAA and BBB and high yield corporate bonds with ratings CCC and below.<sup>9</sup> As the benchmark

<sup>9</sup>This data represents the OASs of subsets of the Bank of America Merrill Lynch US Corporate Master Index, tracking the performance of US dollar denominated investment grade rated corporate debt publicly issued in the US domestic market. The BofA Merrill Lynch OASs are the calculated spreads between a computed OAS index of all bonds in a given rating category and a spot Treasury curve. The data is obtained from the FRED database of Federal Reserve Bank of St. Louis.

yields for these spreads are mostly identical (namely treasuries with similar maturities), the differences between these OASs of corporate bonds with different credit ratings provide detailed measures of the credit risk premiums in the corporate bond market, which may provide more insight in the relevance of the financial accelerator effect for the severity of recessions. We compute three spreads, namely the spread between the high yield (CCC) corporate bonds and the investment-grade bonds with either AAA or BBB rating, as well as the spread between yields on BBB- and AAA-rated bonds. A limitation of this dataset is that it only dates back until January 1997. Still the period from January 1997 until October 2010 covers the two most recent recessions with distinct characteristics, hence, it may provide sufficient information for our analysis. We display the posterior regime probabilities of the mild and severe recession regimes together with these three spreads and their first differences in Figure 4.8.

The results confirm our previous analysis. The spread between the two groups of investment grade bond yields increases substantially around the severe recession of 2007-9, while it does not exhibit a clear departure from its path during the recession in 2001. Additionally, the spreads between the high-yield and investment-grade bonds (both BBB and AAA) increase in both of the recessions, albeit more pronounced during the severe recession. Put differently, the spreads between relatively risky and safe bonds always increase during recessions, as expected, but the magnitude (as well as the speed) of the changes is positively related to the severity of the recession. These two different degrees of disfunctioning of the credit channel and its reflection on the economy is nicely captured by our two different types of recession regimes. Our findings corroborate the conclusions of Claessens *et al.* (2009), who focus on the role of interactions between macroeconomic and financial variables in determining the severity and duration of a recession. Analyzing recessions world-wide their evidence indicates that severe and deep recessions mostly are triggered by (or at least coincide with) credit crunches and house price busts.

Figure 4.8: Posterior recession probabilities and credit spreads between corporate bonds of different quality



*Note:* The graphs show the posterior regime probabilities for mild and severe recessions in the three regime MS-VAR model with imperfect synchronization with asymmetric (or regime-specific) phase shifts for industrial production (IP) and the Conference Board leading economic index (LEI), summarized in Table 4.6, together with the level and first difference of credit spreads between high yield and investment grade corporate bonds.

## 4.6 Conclusion

The macroeconomic experience during the post-World War II period has provided distinct evidence on US recessions. While recessions before the 1980s were short but relatively severe, subsequent recessions in 1990-1 and 2001 were shorter in duration but followed by sluggish recoveries. Interestingly, the most recent recession of 2007-9 appears to be unique

in its characteristics in the sense that it comprises the worst parts of both previous types. Motivated by these observations, we provide empirical evidence for two distinct types of US recessions using Markov-switching models. First, using univariate analysis of monthly coincident and leading economic indicators we show that a three regime model renders a more accurate characterization of their business cycle regimes and dynamics than a model with two regimes only. Interestingly, the third regime captures ‘severe recessions’, contrasting the conventional view of a ‘recovery’ phase as the additional third regime.

This empirical finding is confirmed in a bivariate analysis of industrial production and the Conference Board’s Leading Economic Index, using Markov-switching vector autoregressive models that allow for phase shifts between the cyclical regimes of these two variables. Results indicate that a three regime model with a severe recession phase (next to regimes for expansions and mild recessions) describes the data better than the two regime models with only recession and expansion regimes and a three regime model with a recovery phase. The occurrence of the severe recession regime mostly corresponds with periods of substantial credit squeezes and dramatic increases in credit spreads, providing evidence of the financial accelerator effect.

## 4.A Appendices

In this appendix we describe the different aspects involved in the Bayesian analysis of Markov-switching models. We discuss the model specification details and the specification of prior distributions in Section 4.A.1 and 4.A.2, respectively. This is followed by a derivation of the complete data likelihood function in Section 4.A.3 and a description of the posterior simulator in Section 4.A.4. Finally, we consider model selection based on the predictive likelihood in Section 4.A.5.

### 4.A.1 Model specification details

We follow the common practice in the literature to treat the unobserved regimes in Markov-switching models as parameters to be estimated. For this purpose we derive the complete data likelihood function. Here we provide details for the bivariate MS-VAR model as given in (4.5)–(4.6), with imperfect synchronization due to asymmetric phase shifts as specified in (4.8) and covariance matrix as in (4.9)–(4.10). The likelihood for the model with symmetric phase shifts follows straightforwardly as it is nested in this general specification. For notational convenience, let  $\mathcal{S}_{1,t}$  and  $\mathcal{S}_{2,t}$  be the  $J \times 1$  vectors of regime indicators for the first and second variable respectively. The elements in these vectors take the values 0 or 1, indicating which of the  $J$  regimes occurs at time  $t$ . We write the model (4.5) in vector notation

$$Y_t - \mathcal{S}_t' \mu = \Phi(Y_{t-1} - \mathcal{S}_{t-1}' \mu) + \mathcal{E}_t \quad \mathcal{E}_t \sim N(0, \Sigma_{\mathcal{S}_t}), \quad (\text{A-1})$$

where  $Y_t = (y_{1,t}, y_{2,t})'$ ,  $\mathcal{E}_t = (\varepsilon_{1,t}, \varepsilon_{2,t})'$ ,  $\Phi$  is a  $(2 \times 2)$  matrix containing the autoregressive parameters and where

$$\mathcal{S}_t = \begin{pmatrix} \mathcal{S}_{1,t} & \mathbf{0}_J \\ \mathbf{0}_J & \mathcal{S}_{2,t} \end{pmatrix}, \quad (\text{A-2})$$

with  $\mathbf{0}_J$  a  $J \times 1$  vector of zeros.

### 4.A.2 Prior distributions

As we want the data to drive our estimation results we specify diffuse prior distributions for the model parameters. The parameters of key interest are the lead/lag times or phase shifts  $\kappa = (\kappa_1, \dots, \kappa_J)$ . We use a discrete uniform prior where we assign equal probability

to each possible value of  $\kappa$  in a predefined set  $\mathcal{C}$ , that is

$$f(\kappa) \propto \begin{cases} 1 & \text{for all } \kappa \in \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A-3})$$

The specification of the set  $\mathcal{C}$  determines the type of synchronization that is assumed. For example, when  $\mathcal{C} = \{\kappa \in \mathbb{Z}^J \mid -c_j \leq \kappa_j \leq c_j, j = 1, \dots, J\}$  for certain positive valued  $c_j$ ,  $j = 1, \dots, J$ , we allow for imperfect synchronization with the phase shifts being restricted only by the bounds  $c_j$ . Additional restrictions may be imposed such that, for example, the difference between the phase shifts of the distinct regimes cannot exceed a certain threshold  $d$ , that is,  $\mathcal{C} = \{\kappa \in \mathbb{Z}^J \mid -c_j \leq \kappa_j \leq c_j, |\kappa_i - \kappa_j| \leq d, i, j = 1, \dots, J\}$ . Note that setting  $d$  equal to zero (while  $c_j > 0$  for  $j = 1, \dots, J$ ) results in a model where phase shift parameters are identical across regimes as given in (4.7). Setting  $c_j = 0$  implies an MS-VAR model with perfect synchronization.

For the transition probabilities we use an uninformative Dirichlet prior

$$p_{1j}, \dots, p_{Jj} \sim \text{Dir}(1, \dots, 1). \quad (\text{A-4})$$

When this non-informative prior for the transition probabilities is used special attention must be paid to the prior specifications of the regime-dependent parameters. This follows from the fact that the value of the complete data likelihood function is the same if we switch all regime dependent parameters together with the corresponding transition probabilities. This ‘label switching problem’ complicates proper posterior analysis as the posterior distributions of the regime dependent parameters become multimodal, see Frühwirth-Schnatter (2001) and Geweke (2007) for discussion. To circumvent this problem, we define the prior for the regime dependent mean parameters  $\mu = (\mu'_1, \mu'_2)'$  in such a way that it identifies the regimes. Specifically, we set the prior specification for  $\mu$  as

$$f(\mu) \propto \begin{cases} 1 & \text{if } \mu_1 \in \{\mu_1 \in \mathbb{R}^J \mid \mu_{1,1} < \dots < \mu_{1,j} < \dots < \mu_{1,J}\} \\ 0 & \text{elsewhere.} \end{cases} \quad (\text{A-5})$$

For the regime-dependent variance parameters  $\sigma^2 = \{\sigma_{1,j}^2, \sigma_{2,j}^2\}_{j=1}^J$  as well as for the correlation parameter  $\rho$  we take uninformative priors

$$f(\sigma_{l,j}^2) \propto \sigma_{l,j}^{-2} \quad \text{for } l = 1, 2 \text{ and } j = 1, \dots, J, \quad (\text{A-6})$$

$$f(\rho) \propto (1 - \rho^2)^{-\frac{3}{2}}, \quad (\text{A-7})$$

which correspond to the uninformative prior  $|\Sigma|^{-3/2}$  for  $\Sigma$  in case of constant variances. In case we assume a single structural break in the covariance structure, then the prior distribution of the structural break parameter  $\tau$  has a discrete uniform distribution

$$f(\tau) = \begin{cases} \frac{1}{T-1-2b} & \text{if } \tau \in \{k+b+1, \dots, T-b\} \\ 0 & \text{elsewhere,} \end{cases} \quad (\text{A-8})$$

and hence we do not allow for a break in the first and last  $b$  observations of the sample period. For the size of the break  $\delta$  we use an uninformative prior

$$f(\delta^2) \propto \delta^{-2}. \quad (\text{A-9})$$

Finally, the prior for the autoregressive coefficients is flat on the stationary region, that is,

$$f(\Phi) \propto 1, \quad (\text{A-10})$$

when the characteristic roots of  $\Phi$  lie outside the unit circle and 0 otherwise.

### 4.A.3 Complete data likelihood function

Given the assumption of multivariate normality for the shocks  $\mathcal{E}_t$ , the conditional density of  $Y_t$  given the past observations  $Y^{t-1} = \{Y_1, \dots, Y_{t-1}\}$  and given the past and current states  $S^t = \{S_1, \dots, S_t\}$  is given by

$$f(Y_t|Y^{t-1}, S^t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \tau, \delta, \Phi, \kappa) = \left( \frac{1}{\sqrt{2\pi}} \right)^2 |\Sigma_t|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \mathcal{E}_t' \Sigma_t^{-1} \mathcal{E}_t \right), \quad (\text{A-11})$$

where  $\mu_l = (\mu_{l,1}, \dots, \mu_{l,J})'$  and  $\sigma_l^2 = (\sigma_{l,1}^2, \dots, \sigma_{l,J}^2)'$  for  $l = 1, 2$ ,  $\kappa = (\kappa_1, \dots, \kappa_J)'$ . The complete data likelihood function for model (A-1) conditional on the first observation equals

$$f(Y^T, S^T|Y_1, \theta) = \left( \prod_{i=1}^J \prod_{j=1}^J (p_{ij}^{T_{ij}}) \right) \prod_{t=2}^T f(Y_t|Y^{t-1}, S^t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \tau, \delta, \Phi, \kappa), \quad (\text{A-12})$$

where  $T$  denotes the sample size,  $T_{ij}$  is the number of transitions from regime  $i$  to regime  $j$ , and  $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho, \delta, \text{vec}(\Phi), \kappa, \text{vec}(P))$  represents all the model parameters, with  $P = \{p_{ij}\}_{i,j=1}^J$  the matrix with transition probabilities. The likelihood function conditional only on the model parameters can be obtained by summing (A-12) over all the possible

states as

$$f(Y^T|Y_1, \theta) = \sum_{S_{1,1}=1}^J \sum_{S_{1,2}=1}^J \cdots \sum_{S_{1,T}=1}^J f(Y^T, S^T|Y_1, \theta). \quad (\text{A-13})$$

#### 4.A.4 The Gibbs sampler and full conditional posterior distributions

##### Gibbs sampler

The simulation scheme of the Gibbs sampler is as follows

1. Sample  $\mu$  from  $f(\mu|\Phi, \sigma^2, \rho, \tau, \delta, S^T, Y^T)$ .
2. Sample  $\Phi$  from  $f(\Phi|\mu, \sigma^2, \rho, \tau, \delta, S^T, Y^T)$ .
3. Sample  $\sigma_l^2$  from  $f(\sigma_l^2|\mu, \Phi, \rho, \tau, \delta, S^T, Y^T)$  for  $l = 1, 2$ .
4. Sample  $\rho$  from  $f(\rho|\mu, \Phi, \sigma^2, \tau, \delta, S^T, Y^T)$ .
5. Sample  $\tau$  from  $f(\tau|\mu, \Phi, \sigma^2, \rho, \delta, S^T, Y^T)$ .
6. Sample  $\delta$  from  $f(\delta|\mu, \Phi, \sigma^2, \rho, \tau, S^T, Y^T)$ .
7. Sample  $P$  from  $f(p_{ij}|S_1^T)$ .
8. Sample  $\kappa_j$  from  $f(\kappa_j|\mu, \Phi, \sigma^2, \rho, \tau, \delta, S_1^T, Y^T)$  for  $j = 1, \dots, J$ .
9. Sample  $S_1^T$  from  $f(S_1^T|P, \mu, \Phi, \sigma^2, \rho, \tau, \delta, \kappa, Y^T)$ .

##### Sampling of $\mu$

To sample  $\mu$  we first rewrite (A-1) as

$$\Sigma_t^{-\frac{1}{2}} (Y_t - \Phi Y_{t-1}) = \Sigma_t^{-\frac{1}{2}} \{S'_t - \Phi S'_{t-1}\} \mu + \Sigma_t^{-\frac{1}{2}} \mathcal{E}_t. \quad (\text{A-14})$$

This is a multivariate regression in the form of

$$Z_t = X_t \mu + \epsilon_t \quad \text{with } \epsilon_t \sim N(0, I_2), \quad (\text{A-15})$$

and hence the conditional distribution of  $\mu$  is multivariate normal with mean  $(X'X)^{-1} X'Y$  and variance  $(X'X)^{-1}$ , where  $X = (X'_2, \dots, X'_T)'$  and  $Z = (Z'_2, \dots, Z'_T)'$ , see Zellner (1971). The restrictions on the elements of  $\mu$  that are imposed for identification of the regimes can be applied by sampling from the corresponding truncated distribution or using acceptance rejection sampling.

### Sampling of $\Phi$

Conditional on remaining model parameters the model in (A-1) is a multivariate regression of the form

$$Z = \Phi X + U, \quad (\text{A-16})$$

where  $Z = (Z_2, \dots, Z_T)$  with  $Z_t = Y_t - \mathcal{S}'_t \mu$ , and  $X = (X_2, \dots, X_T)$  with  $X_t = (Y_{t-1} - \mathcal{S}'_{t-1} \mu)$  and

$$\Omega^{-1} = \text{Cov}(\text{vec}(U))^{-1} = \text{diag}(\Sigma_2^{-1}, \dots, \Sigma_T^{-1}). \quad (\text{A-17})$$

Using the fact that

$$\text{vec}(Z) = (X' \otimes I_2) \text{vec}(\Phi) + \text{vec}(U), \quad (\text{A-18})$$

we can simply write the regression model as a univariate regression

$$\mathbf{z} = (X' \otimes I_2) \text{vec}(\Phi) + \mathbf{u}. \quad (\text{A-19})$$

It then follows that the conditional posterior distribution of  $\text{vec}(\Phi)$  is multivariate normal with mean  $((X' \otimes I_2)' \Omega^{-1} (X' \otimes I_2))^{-1} ((X' \otimes I_2)' \Omega^{-1} \mathbf{z})$  and covariance matrix  $((X' \otimes I_2)' \Omega^{-1} (X' \otimes I_2))^{-1}$ , where

$$\Omega^{-1} = \text{Cov}(u)^{-1} = \text{diag}(\Sigma_2^{-1}, \dots, \Sigma_T^{-1}).$$

In case of zero restrictions on the elements of  $\Phi$ , one can easily modify  $(X' \otimes I_2)$  by dropping the corresponding columns of  $(X' \otimes I_2)$ .

### Sampling of $P$

From the conditional likelihood function (A-12), it follows that the transition probabilities from state  $i$  can be written as

$$f(p_{i,1}, \dots, p_{i,J}) \propto \prod_{j=1}^J p_{i,j}^{T_{i,j}} \quad \text{for } i = 1, \dots, J, \quad (\text{A-20})$$

where  $T^{ij}$  denotes the number of transitions from state  $i$  to state  $j$ . This corresponds to the kernel of a Dirichlet distribution and hence the transition probabilities can be sampled from a Dirichlet distribution with parameters  $T_{ij}$  for  $j = 1, \dots, J-1$ . The transition probability to state  $J$ ,  $p_{iJ}$ , follows from the restriction that  $\sum_{j=1}^J p_{ij} = 1$ .

### Sampling of lead/lag parameters $\kappa$

As the  $\kappa_j$  parameters can only take discrete values we can compute the posterior probabilities for all  $\kappa \in \mathcal{C}$  and sample from a multinomial distribution. We can sample all  $\kappa_j$  parameters at once or conditional on each other, one at a time. When the number of admissible  $\kappa$ 's is large, the latter approach may be more attractive as the number of combinations increases only linearly in the number of states, whereas it increases exponentially in the first case.

### Sampling of regimes

The conditional posterior density of  $S_{1,t}$  denoted by  $f(S_{1,t}|S_1^{-t}, \theta, Y^T)$  for  $t = 1, \dots, T$  and  $S_1^{-t} = S_1^T \setminus \{S_{1,t}\}$ , is proportional to the transition probabilities due to the Markov structure and to the density of  $Y$  conditional on the regimes. Hence, we can write the posterior density of  $S_{1,t}$

$$f(S_{1,t}|S_1^{-t}, \theta, Y^T) \propto f(S_{1,t}|S_{1,t-1}, \theta) f(S_{1,t+1}|S_{1,t}, \theta) \prod_{i=t-\kappa_{\min}}^{t+1+\kappa_{\max}} f(Y_i|Y^{i-1}, S^i, \theta), \quad (\text{A-21})$$

where  $f(Y_i|Y^{i-1}, S^i, \theta)$  is given in (A-11),  $\kappa_{\max} = \max(\kappa_1, \dots, \kappa_J)$  and  $\kappa_{\min} = \min(\kappa_1, \dots, \kappa_J)$ . At time  $t = T$  the term  $f(S_{1,t+1}|S_{1,t}, \theta)$  drops out. The regime at  $t = 1$  can be sampled from the conditional distribution

$$f(S_{1,1}|S_1^{-1}, \theta, Y^T) \propto f(S_1|\theta) f(S_{1,2}|S_{1,1}, \theta) \prod_{i=2}^{t+1+\kappa_{\max}} f(Y_i|Y^{i-1}, S^i, \theta), \quad (\text{A-22})$$

where the unconditional density  $f(S_{1,1}|\theta)$  follows a multinomial distribution with ergodic probabilities of the Markov chain.

Sampling of the state variables can be implemented by starting from the most recent value of  $S_1^T$  and sampling the states backward in time, one after another. After each step, the  $t^{\text{th}}$  element of  $S_1^T$  is replaced by its most recent draw.

### Sampling of $\tau$

The conditional posterior density of  $\tau$  is given by

$$f(\tau|s^T, \theta \setminus \{\tau\}, Y^T) \propto \mathbb{I}[b + k + 1 < \tau \leq T - b] \times |\Sigma_t|^{-\frac{1}{2}(\tau+2)} \exp\left(-\frac{1}{2}\text{tr}\left(\Sigma_t^{-1} \left(\sum_{k+1}^{\tau-1} \mathcal{E}_t' \mathcal{E}_t\right)\right)\right), \quad (\text{A-23})$$

where  $\Sigma_t$  is defined by (4.9)-(4.10). As  $\tau$  can take discrete values on the range  $[b + k + 1, T - b]$ , we can sample from its full posterior distribution using the posterior probabilities of each of these discrete values.

### Sampling of variances and correlations

To sample the variances we decompose the multivariate normal distribution of vector of error terms in a conditional distribution of  $\varepsilon_{2t}$  given  $\varepsilon_{1t}$  and the marginal distribution of  $\varepsilon_{1t}$ . This results in

$$\begin{aligned} \prod_{t=2}^T f(\mathcal{E}_t) &= \prod_{t=2}^{\tau-1} \frac{1}{\delta\sigma_{1,t}} \phi\left(\frac{\varepsilon_{1,t}}{\delta\sigma_{1,t}}\right) \frac{1}{\delta\sigma_{2,t}\sqrt{(1-\rho^2)}} \phi\left(\frac{\varepsilon_{2,t}-\rho\sigma_{2,t}/\sigma_{1,t}\varepsilon_{1,t}}{\delta\sigma_{2,t}(1-\rho^2)}\right) \\ &\quad \prod_{t=\tau}^T \frac{1}{\sigma_{1,t}} \phi\left(\frac{\varepsilon_{1,t}}{\sigma_{1,t}}\right) \frac{1}{\sigma_{2,t}\sqrt{(1-\rho^2)}} \phi\left(\frac{\varepsilon_{2,t}-\rho\sigma_{2,t}/\sigma_{1,t}\varepsilon_{1,t}}{\sigma_{2,t}(1-\rho^2)}\right), \end{aligned} \quad (\text{A-24})$$

where  $\sigma_{l,t}^2$  denotes the variance of  $\varepsilon_{l,t}$ ,  $l = 1, 2$ , which follows from (4.10) or (4.10) with  $\sigma_{l,S_{l,t}}^2$  replaced by  $\sigma_{l,t}^2$ . To sample  $\sigma_{1,j}^2$  we use a Metropolis-Hastings sampler (Metropolis *et al.*, 1953; Hastings, 1970), where we use as candidate an inverted  $\chi^2$  distribution with scale parameter  $\sum_{t=2}^{\tau-1} I[S_{1,t} = j] \delta^{-2} \varepsilon_{1,t}^2 + \sum_{t=\tau}^T I[S_{1,t} = j] \varepsilon_{1,t}^2$  and degrees of freedom  $\sum_{t=2}^T I[S_{1,t} = j]$  for  $j = 1, \dots, J$  which follows from the first part of the decomposition in (A-24) in combination with the prior. To sample  $\sigma_{2,j}^2$  we use a similar approach where we consider the marginal of  $\varepsilon_{2,t}$  in the decomposition.

To sample  $\rho$  from its conditional posterior distribution we can again use (A-24), which leads to the following as

$$(1-\rho^2)^{-3/2} \prod_{t=2}^{\tau-1} \frac{1}{\sqrt{(1-\rho^2)}} \phi\left(\frac{\varepsilon_{2,t}-\rho\sigma_{2,t}/\sigma_{1,t}\varepsilon_{1,t}}{\delta\sigma_{2,t}(1-\rho^2)}\right) \prod_{\tau}^T \frac{1}{\sqrt{(1-\rho^2)}} \phi\left(\frac{\varepsilon_{2,t}-\rho\sigma_{2,t}/\sigma_{1,t}\varepsilon_{1,t}}{\sigma_{2,t}(1-\rho^2)}\right). \quad (\text{A-25})$$

We can easily implement the griddy Gibbs sampler approach of Ritter and Tanner (1992). Given that,  $\rho \in (-1, 1)$  we can setup a grid in this interval based on the precision we desire about the value of  $\rho$ .

(A-24) shows that the full conditional posterior distribution of  $\delta^2$  is an inverted  $\chi^2$  distribution with scale parameter  $\sum_{t=2}^{\tau-1} \sum_{j=1}^J I[S_{1,t} = j] \sigma_{1,t}^{-2} \varepsilon_{1,t}^2 + \sigma_{2,t}^{-2} (1-\rho^2)^{-2} (\varepsilon_{2,t} - \rho\sigma_{2,t}/\sigma_{1,t}\varepsilon_{1,t})^2$  and  $2(\tau-1)$  degrees of freedom. Hence, we sample from this distribution for drawing  $\delta$ .

### 4.A.5 Evaluation of predictive likelihood

Selecting the best model is not an easy task when the competing alternatives embody regime-switching dynamics. In most cases standard testing procedures apply only when the number of regimes is the same in the models being compared. In this case marginal likelihood based comparisons, e.g. Bayes factors, can be implemented. When the task is determination of the number of the regimes, however, a complication arises when improper priors such as ours are used for the regime dependent parameters of interest. In this case, Bayes factors are not properly defined and tend to select the more parsimonious model (see Gelfand and Dey, 1994, for details). An alternative way of model comparison is by means of predictive Bayes factors computed using predictive likelihoods. This approach has the advantage that it is not affected by the choice of prior distributions and over-fitting, while it is directly related to the posterior model probabilities. Moving away from a two regime model to a model with higher number of regimes may increase the model fit, however, the resulting increase may simply because of model over-fitting brought by the additional regimes. If this is the case this would signal itself in the predictions of the model. In this sense, predictive likelihood evaluation provides also an immune methodology of model selection against over-fitting.

For a given model, the predictive likelihood of the observation at  $t_0 + 1$ ,  $Y_{t_0+1}$ , conditional on the previous observations  $Y^{t_0}$ , is given by

$$f(Y_{t_0+1}|Y^{t_0}) = \int f(Y_{t_0+1}|\theta)f(\theta|Y^{t_0})d\theta, \quad (\text{A-26})$$

where  $p(\theta|Y^{t_0})$  is the posterior distribution of the model parameters  $\theta$  given the observations until  $t_0$ , and  $p(Y_{t_0+1}|\theta)$  is the density of the observation  $y_{t_0+1}$ , which can be written as

$$f(Y_{t_0+1}|\theta) = \sum_{j=1}^J f(Y_{t_0+1}|S_{1,t_0+1} = j, \theta)f(S_{1,t_0+1} = j|\theta, Y^{t_0}). \quad (\text{A-27})$$

We can use the posterior simulator to obtain the distribution of the model parameters and estimate the predictive likelihood by  $G^{-1} \sum_{g=1}^G f(Y_{t_0+1}|Y^{t_0}, \theta^{(g)})$  where  $G$  is a large number of draws from the posterior distribution. This can be extended to compute the predictive likelihood of a sequence  $Y^{t_0+1, t_0+h} = (Y'_{t_0+1}, \dots, Y'_{t_0+h})'$  where

$$f(Y^{t_0+1, t_0+h}|Y^{t_0}) = \int \prod_{l=1}^h f(Y_{t_0+l}|Y^{t_0+l-1}, \theta)f(\theta|Y^{t_0})d\theta, \quad (\text{A-28})$$

see Frühwirth-Schnatter (2006), for details.



# Chapter 5

## Bayesian Semiparametric Dynamic Nelson-Siegel Model

### 5.1 Introduction

Modeling the evolution of large cross sections of treasury bond yields has been an intensive area of econometric research for many decades. Examining the patterns of bond yields is of key importance for many reasons including understanding the expectations of economic agents, managing fiscal debt, pricing bonds and related derivatives, pricing sovereign risk, and portfolio allocation. Typically, the common patterns observed in bond yields can be summarized using a few key factors. Indeed, existing statistical evidence shows that only a few factors can explain the bulk of the variation of bond yields, see for example the analysis in de Pooter (2007). Besides statistical evidence, financial theory often postulates a factor structure on the dynamic evaluation of the yield curve, i.e. the cross section of the bond yields with varying maturities. Not surprisingly, models that exploit this factor structure have been the major workhorse in modeling the common variation of treasury bond yields.

Among these models, dynamic Nelson-Siegel model (DNS) is a popular class of factor models that has been heavily used among practitioners and academics to analyze the dynamic evolution of the yield curve, see Diebold and Li (2006). Based on the parsimonious nature of the Nelson-Siegel framework, see Nelson and Siegel (1987), the DNS model is proved to be very useful in practice (see Svensson, 1994; Gurkaynak *et al.*, 2007, for example). Consequently, the DNS model constitutes an attractive class of yield curve models with desirable statistical properties, while still preserving a parsimonious model structure.

While the cross section of bond yields or in other words, the dependence structure of the yields on the underlying factors is well studied in the DNS model, the precise density of bond yields is not sufficiently explored yet. One typical assumption on the evolution of the bond yields factors, for example, is the Gaussianity of the error distributions. This assumption often serves as a means for constructing a feasible estimation procedure for the optimal point predictions (in the least squares sense). When the density of the yield series, however, is not Gaussian or unknown, assumption of Gaussianity for the error distributions does not apply. This is especially the case for the returns of many asset classes including bond returns where this assumption seems hard to be fulfilled.<sup>1</sup> Although bond markets bear relatively limited risk compared to equity markets, asymmetric shocks, changing behavior in different economic conditions and expectations, and changes in monetary policy are likely to be the potential causes of this non-Gaussianity, among others. On the one hand, evidence from the recent global financial crises 2007-9 and the aftermath of it indicates that the ‘zero bound’ is likely to be a binding constraint for nominal yields as treasury bond yields in many countries fell essentially to their zero bound. From a statistical perspective, existence of the zero lower bound as a binding constraint leads to a skewed distribution for bond yields. On the other hand, extremely high bond yields observed in the beginning of 1980s (and never observed again) indicate that an upper bound for nominal yields does not exist and extreme shocks are more likely to occur in the upper tail of the yields’ distribution. Hence, the assumption of Gaussianity may not be adequate to capture these observed phenomena.

Unfortunately, it is not a straightforward task to model the non-Gaussianity or to estimate the densities of the yields because of the high dimensionality of the DNS model. When the non-Gaussianity is modeled for each bond yield separately, overparametrization is a potential problem that may complicate econometric inference.<sup>2</sup> One elegant and natural way to model this non-Gaussianity and to estimate the density of the series is to exploit the factor structure, i.e. estimating the density of the factors instead of the density of the individual variables. This facilitates the econometric inference substantially

---

<sup>1</sup>Potential underlying reasons for the observed non-Gaussianity of the financial asset returns range from distributions being mixture of underlying distributions, see for example Guidolin and Timmermann (2006), Dai *et al.* (2007) and Ang *et al.* (2008), time varying parameters, see Bianchi *et al.* (2009), excess skewness and/or kurtosis (fat-tails and potential outliers), see Engle (1982), Bollerslev (1986), Laurent *et al.* (2006) and Koopman *et al.* (2010), to nonlinear dependencies between variables, see Longin and Solnik (1995), Patton (2006) and Junker *et al.* (2006).

<sup>2</sup>For some specific cases the model may still remain parsimonious. For example, if the true data generating process follows a multivariate  $t$  distribution then only a single parameter of degrees of freedom is required to estimate additionally. Unfortunately, we do not observe the true data generating process in reality and such a modeling approach may impose unrealistic assumptions such as assuming identical degrees of freedom for each cross-section. It also excludes many types of behavior observed in many asset classes such as underlying mixture distributions and nonlinear dependence structure.

by-passing the problem of the curse of dimensionality without losing much explanatory power.<sup>3</sup> This is especially the case for the DNS model where a few key factors explain the variation in the entire cross section of the yields quite successfully.

In this paper, we propose a Bayesian semiparametric alternative to the assumption of Gaussianity for yield curve factors in the dynamic Nelson-Siegel model. Using the semiparametric structure we estimate the factor density along with the factors. Specifically, we specify the distribution of the factors using a Dirichlet process mixture, that is proved to be very useful in many applications of econometrics, see for example Chib and Hamilton (2002); Hirano (2002); Griffin and Steel (2004); Jensen (2004); Conley *et al.* (2008); Jensen and Maheu (2010), among others. The resulting model does not rely on the assumption of Gaussianity of the factors nor we presume a specific distribution for the factors. We propose an efficient and computationally tractable algorithm to obtain Bayesian inference. For analyzing the potential of the Bayesian semiparametric dynamic Nelson-Siegel model, we focus on the panel time-series of US zero-coupon bond yields.

Traditional approaches of modeling the dynamics of the yield curve mostly focus either on the no-arbitrage relations in equilibrium to model the yield curve<sup>4</sup> or focus on functions that can approximate the observed yield curve well (see McCulloch, 1975; Nelson and Siegel, 1987, for example). For the latter, Nelson and Siegel (1987) provide a parsimonious three-component exponential approximation to the yield curve. However, the framework of Nelson and Siegel (1987) concerns only about fitting the yield curve of bonds for a given point in time. Diebold and Li (2006) and Diebold *et al.* (2006) extend this framework by imposing a first order autocorrelation on each component and thereby adding a time dimension. They show that by doing that the model can be cast in a state-space framework with three factors where these factors can be interpreted as the level, slope and curvature of the yield curve.

While the main focus is on modeling the cross section of yields with point predictions of the factors, the distribution of the factors is usually neglected. In most cases, a vector autoregressive process of the first order with Gaussian innovations is assumed in the state equation in which case standard Kalman filter techniques can easily be applied.

---

<sup>3</sup>Indeed, using a similar intuition Patton and Oh (2011) model the nonlinear dependence in high dimensional models using factor copulas where factors are modeled using some non-Gaussian distributions.

<sup>4</sup>See Vasicek (1977); Cox *et al.* (1985); Hull and White (1990) for example. Duffie and Kan (1996) provide a multifactor affine (yields are affine functions of underlying factors) term structure model in this vein. Dai and Singleton (2000) classify the canonical representation of the affine no-arbitrage term-structure models and focus on three factor models. However, many issues plague the estimation of these models including the existence of local maxima each having a different economic interpretation, see Orphanides and Kim (2005), and overparametrization. Indeed, Dai and Singleton (2002) and Duffie (2002), among others, impose some further restrictions on model parameters to alleviate these problems.

However, the Gaussianity assumption may not be appropriate for many reasons. First, for US bond yields the volatility before the Great Moderation<sup>5</sup> was much higher than the succeeding period. Indeed, Hautsch and Ou (2009) and Koopman *et al.* (2010) for the U.S. and Bianchi *et al.* (2009) for the U.K., among others, document the time variation in the volatility of the factors. These studies mostly employ stochastic volatility (Jacquier *et al.*, 2004; Omori *et al.*, 2007) or GARCH (Bollerslev, 1986) type approaches to model fat tails and the time variation in the volatility, which may not be efficient for the typical monthly data of yields as these approaches are mostly suitable for higher frequency data. Second, these approaches usually do not take nonlinear dependencies into account; hence, correlations are typically assumed to be constant. Junker *et al.* (2006), for example, use asymmetric copulas to model the nonlinear dependencies between the factors extracted from an arbitrage-free affine term structure model. Third, distributions of yields are likely to be skewed as interest rates cannot be negative. Hence, extreme observations and potential outliers are likely to be in the upper tail of the distribution. Therefore, we provide a unified approach that can take all of these concerns into account when modeling the term structure of interest rates.

In the empirical application using US yield data, we show that the Bayesian semiparametric dynamic Nelson-Siegel model captures two different periods of high and low volatility. While the period with low volatility coincides with the times of Great Moderation, the period with high volatility captures the times of severe recessions of 1970s and the period of Federal Reserve's monetary experiments in the beginning of 1980s (see Friedman, 1984, for example) together with the recent recession of 2007-9. An interesting finding is that the model identifies the two historical peak points of the yields in the beginning of 1980s as outliers in a separate cluster leading to more precise estimates of the parameters. Moreover, the estimate of the correlation reveals an interesting pattern in the sense that the estimates of the models with the DPM structure generally reveal a higher degree of correlation compared to the benchmark with Gaussian innovations. However, during turbulent times correlation coefficients decrease substantially or switch signs. Finally, we perform several sensitivity analyses regarding to prior specifications and the model structure that confirm the robustness of our results.

The remainder of the paper is organized as follows. Section 5.2 describes the Bayesian semiparametric dynamic Nelson-Siegel model. We discuss the Bayesian estimation methodology together with the prior and posterior specifications in Section 5.3, with full details being provided in Appendix 5.A.2. Section 5.4 displays and discusses the results of our

---

<sup>5</sup>Persistent decline in the volatility of many US macroeconomic and financial indicators after mid 1984 is referred as Great Moderation.

model using US data, with the results of the model using simulated datasets being provided in Appendix 5.A.3. Finally, Section 5.5 concludes.

## 5.2 Model specification

### 5.2.1 Dynamic Nelson-Siegel model

Fitting a parsimonious parametric curve on the large set of yields that can approximate the observed yield curve successfully facilitates the analysis of the entire space of the yields. Accordingly, such an approach enables to capture the space of the yields using only a few key parameters. Following this idea, Nelson and Siegel (1987) provide a parsimonious yet flexible specification that can accommodate various forms of observed yield curves. The method of Nelson and Siegel (1987), extended by Siegel and Nelson (1988), is to specify the forward curve using an exponential approximation.

More explicitly, let  $y_t(\tau)$  be the yield of a zero-coupon bond at time  $t$  with the maturity date  $t + \tau$ . The approximation of the forward curve translates into the following specification of the yield curve with only three components<sup>6</sup> as

$$y_t(\tau) = f_{1,t} + f_{2,t} \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + f_{3,t} \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right). \quad (5.1)$$

Although (5.1) is governed by only four parameters;  $(\lambda, f_{1,t}, f_{2,t}, f_{3,t})'$ , the main advantage of this parsimonious specification is its flexibility to accommodate different forms of the yield curve successfully. The parameter  $\lambda$  in (5.1) governs the exponential decay rate, i.e. while small values produce slow decay and provide a better fit to the curve at long end, large values produce fast decay and provide a better fit to the curve at the short end. Moreover, it is the key factor in determining the maximum loading of the third coefficient. Once the decay rate is determined, the key parameters  $(f_{1,t}, f_{2,t}, f_{3,t})'$  act as the main determinants of the cross section of yields.<sup>7</sup>

The framework of Nelson and Siegel (1987), however, concerns only with fitting the yield curve of bonds for a given point in time. As a result, while in the classical specifi-

<sup>6</sup>There is a one-to-one mapping between the forward curve and the yield curve as the yield is an equally weighted average of the forward rates.

<sup>7</sup>Indeed, in many applications of the Nelson-Siegel model the value of  $\lambda$  is fixed to 0.0609, see for example Diebold and Li (2006). This value maximizes the loading on the medium-term factor, i.e. curvature, at exactly 30 months. Fixing the value of  $\lambda$  prior to estimation facilitates the estimation of the model without losing any explanatory power. Indeed, the estimates of  $\lambda$ , for example, in Hautsch and Yang (2010) are very similar to this value with a limited uncertainty. Throughout this paper we follow this common practice by fixing  $\lambda$  at 0.0609.

cation of Nelson and Siegel (1987),  $(f_{1,t}, f_{2,t}, f_{3,t})'$  are interpreted as static coefficients to be estimated at a given point in time, Diebold and Li (2006) interpret these parameters as factors governing the evolution of the cross section of maturities over time. The first factor loads on each yield equally and it can be interpreted as a level factor. Additionally, as  $\tau \rightarrow \infty$  the loadings of the second and third factors approach to zero leaving only the first factor to determine the long end. Hence, this factor can also be interpreted as long-term factor. The loading on the second factor starts with 1 at the short end. As  $\tau \rightarrow \infty$  it decays quickly to 0, thus, the effect of the second factor is mostly on the short end leading to the interpretation of this factor as short-term factor. Furthermore, this factor is closely related to the yield curve slope as the change of it affects the short end more than the long end changing the slope of the yield curve. The loading on the third factor starts at 0 as  $\tau \rightarrow 0$  and then it increases as the maturity increases and after some maximum level it again decreases to 0 as  $\tau \rightarrow \infty$ . This implies that the third factor affects mostly the medium term maturities leading to the interpretation as a medium-term factor. Moreover, this effect in the medium term determines the curvature of the yield curve. In summary, the factors can be interpreted as long-term, short-term and medium term factors or alternatively as level, slope and curvature factors.

Given this interpretation of the Nelson-Siegel factors, the evolution of these factors shapes the evolution of the entire yield curve over time. Indeed, to explore this time variation Diebold and Li (2006) and Diebold *et al.* (2006) extend this framework by imposing a first order autocorrelation to these factors. They show that by doing this the model can be cast into a state-space (or into a dynamic factor model as a more specific case) framework with three factors where these factors can be interpreted as dynamic level, slope and curvature factors. The complete state-space specification for time period  $t$  is given by

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{\lambda\tau_1}}{\lambda\tau_1} - e^{\lambda\tau_1} \\ 1 & \frac{1-e^{\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{\lambda\tau_2}}{\lambda\tau_2} - e^{\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{\lambda\tau_N}}{\lambda\tau_N} - e^{\lambda\tau_N} \end{pmatrix} \begin{pmatrix} f_{1,t} \\ f_{2,t} \\ f_{3,t} \end{pmatrix} + \begin{pmatrix} \eta_t(\tau_1) \\ \eta_t(\tau_2) \\ \vdots \\ \eta_t(\tau_N) \end{pmatrix}, \quad (5.2)$$

and the evolution of the factors (state equation) takes the following form

$$\begin{pmatrix} f_{1,t} \\ f_{2,t} \\ f_{3,t} \end{pmatrix} = \begin{pmatrix} \mu_{1,t} \\ \mu_{2,t} \\ \mu_{3,t} \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix} \begin{pmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{3,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix}. \quad (5.3)$$

In the standard dynamic Nelson-Siegel model the error terms in both equations are distributed according to a multivariate normal distribution. Given the Gaussianity of error terms, standard Kalman filter\smoother techniques can easily be implemented to estimate the unobserved states and model parameters. When the error distributions are unknown, however, assumption of Gaussian distribution for the error term in the state equation may not be an appropriate assumption. For example, when the actual distribution has in fact fatter tails than those implied by the Gaussian distribution or when it is a skewed distribution, then the model with the Gaussianity assumption would be misspecified and the inference obtained using this assumption would be misleading. Instead, our approach is to keep the error distribution of the state equation as unknown and to estimate this density using Bayesian density estimation techniques, specifically, Dirichlet process mixtures. To simplify notation let

$$y_t = \begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix}, H = \begin{pmatrix} 1 & \frac{1-e^{\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{\lambda\tau_1}}{\lambda\tau_1} - e^{\lambda\tau_1} \\ 1 & \frac{1-e^{\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{\lambda\tau_2}}{\lambda\tau_2} - e^{\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{\lambda\tau_N}}{\lambda\tau_N} - e^{\lambda\tau_N} \end{pmatrix}, f_t = \begin{pmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{3,t-1} \end{pmatrix},$$

$$\eta_t = \begin{pmatrix} \eta_t(\tau_1) \\ \eta_t(\tau_2) \\ \vdots \\ \eta_t(\tau_N) \end{pmatrix}, \mu_t = \begin{pmatrix} \mu_{1,t} \\ \mu_{2,t} \\ \mu_{3,t} \end{pmatrix}, F = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix}, \varepsilon_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix}.$$

Then the model can be written in the compact notation as

$$\begin{aligned} y_t &= Hf_t + \eta_t & \eta_t &\sim N(0, R) \\ f_t &= Ff_{t-1} + \varepsilon_t & \varepsilon_t &\sim N(\theta_t) \text{ where } \theta_t = (\mu_t, Q_t) \sim G(.), \end{aligned} \tag{5.4}$$

where  $\eta_t$  follows a multivariate normal distribution with mean zero and the covariance matrix  $R$ . For the error distribution of the factors the above specification indicates that conditional on the model parameters  $\theta_t$  the system is Gaussian, where  $\theta_t$  includes the mean and the variance of the factor error distribution in period  $t$ . However,  $\theta_t$  is distributed according to some unknown prior distribution  $G$ . The constant term in the state equation is suppressed in the error distribution as the mean of the error distribution and it will be explained in detail in the next section.

### 5.2.2 Dirichlet process prior and Dirichlet process mixture

In this section, we introduce the semiparametric specification of the distribution of  $\theta_t$  using the Dirichlet process mixture, which leads us to the estimation of the error distribution of factors. We start our discussion of Dirichlet process by pointing the analogy between the Bayesian inference of a parameter and a density. The prior distribution specification of a parameter is of key importance in obtaining Bayesian inference. The specification of the prior distribution has to be general enough to cover all potential regions of the parameter space without any constraint. When the estimation concerns a density instead of a parameter, the parameter space is replaced by a space of probability distribution functions. Analogous to the parameter estimation, when the aim is density estimation, a prior distribution that spans a space of probability distribution functions that is large enough to capture all possible candidates is required. Following this aim, Ferguson (1973) and Antoniak (1974) proposed Dirichlet process (prior) and Dirichlet process mixture model that is flexible enough to provide a prior specification satisfying this condition successfully.

Let  $G_0$  be a probability distribution function over some parameter space  $\Theta$  such that  $\theta \in \Theta$  and let  $\alpha$  be a positive scalar. A distribution  $G$  is distributed with Dirichlet process  $DP(G_0, \alpha)$  with base distribution  $G_0$  and precision parameter  $\alpha$  if for any measurable  $J$  partition of the parameter space,  $\theta_{1:J} = (\theta_1, \theta_2, \dots, \theta_J)$ , the distribution  $G = G(\theta_{1:J})$  follows a Dirichlet distribution  $Dir(\alpha G_0(\theta_1), \alpha G_0(\theta_2), \dots, \alpha G_0(\theta_J))$ . To clarify the underlying intuition and to demonstrate the functional form of the DP we resort to the so-called Polya's urn representation of the DP, see Blackwell and MacQueen (1973). Suppose we have a continuous sample of colored balls of size  $\alpha$  and the distribution of the colors of these balls is denoted as  $G_0(\Lambda)$  that depends on the hyperparameters  $\Lambda$ . Suppose a ball is sampled from the urn and upon observing the color of the sampled ball another ball of exactly the same color is added to the urn along with the sampled ball. Putting back the drawn ball in the urn keeps the initial sample size  $\alpha$  and the color distribution  $G_0(\Lambda)$  unchanged. However, as one more ball with identical color is added to the urn the sample size increases by one. This implies that for the next round, either a ball with a new color from the initial sample of  $\alpha$  will be drawn (given that the base distribution  $G_0(\Lambda)$  is continuous) or the ball that is identical to the previously drawn ball will be drawn. Repeating this  $T$  times results in a  $T$  sample of balls (ignoring the initial sample size) with  $J$  different colors, where the distribution of this combination of  $J$  colors are distributed according to a Dirichlet distribution, see Ferguson (1973) for a formal proof.

When the state equation in dynamic Nelson-Siegel model is considered, the parameters of the factor error distribution,  $\theta_t$ , are analogous to the balls of size  $T$  in the urn and the

$J$  partition of the parameters space,  $\theta_{1:J} = (\theta_1, \theta_2, \dots, \theta_J)$ , is analogous to the resulting  $J$  colors. This Polya's urn scheme suggests that the conditional DP prior  $f(\theta_t|\theta_{-t})|G = G \sim DP(G_0(\Lambda), \alpha)$  has the following form

$$f(\theta_t|\theta_{-t}) = \frac{\alpha}{\alpha + T - 1} G_0(\Lambda) + \sum_{i=1, i \neq t}^T \frac{\delta_j(\theta_i)}{\alpha + T - 1}, \quad (5.5)$$

where  $\delta(\cdot)$  is the Dirac function indicating the point mass such that  $\delta(\theta) = 1$  if  $\theta \in \theta_j$ . (5.5) indicates that at each step either a new  $\theta_t$  will be drawn from  $G_0(\Lambda)$  with probability proportional to  $\alpha$  or  $\theta_t$  will be a member of an existing cluster  $j$  with probability proportional to  $\sum_{i=1, i \neq t}^T \delta_j(\theta_i)$ . Notice that (5.5) can be rewritten as

$$f(\theta_t|\theta_{-t}) = \frac{\alpha}{\alpha + T - 1} G_0(\Lambda) + \frac{T - 1}{\alpha + T - 1} \sum_{i=1, i \neq t}^T \frac{\delta_j(\theta_i)}{T - 1}. \quad (5.6)$$

This functional form indicates a weighted average of the base prior and the empirical distribution, where the weights are the functions of the precision parameter and the sample size. As  $T \rightarrow \infty$  or  $\alpha \rightarrow 0$ ,  $f(\theta_t|\theta_{-t}) \rightarrow \sum_{i=1, i \neq t}^T \frac{\delta_{\theta_i}^j}{T-1}$ ; hence, the distribution of  $\theta_t$  becomes observed empirical distribution. As  $\alpha \rightarrow \infty$ ,  $f(\theta_t|\theta_{-t}) \rightarrow G_0(\Lambda)$ ; hence, any realization of DP is  $G_0(\Lambda)$ .

As a flexible specification, Dirichlet process spans the entire distribution of the discrete probability functions with probability 1. This can be seen from the stick breaking representation of the DP, see Sethuraman (1994). The stick breaking representation corresponds to the following scheme. Suppose we break a part of  $W_1\%$  of a stick of length 1 where  $W_1 \sim \text{Beta}(1, \alpha)$ . Repeating this  $j$  times, at step  $j$ , we break  $W_j\%$  of the remaining stick with  $W_j \sim \text{Beta}(1, \alpha)$  and the length of this part becomes  $W_j \prod_{i=1}^{j-1} (1 - W_i)(1 - W_2), \dots, (1 - W_{j-1})$ . Let denote the broken parts as  $\pi_1 = W_1$ ,  $\pi_2 = W_2(1 - W_1)$  and  $\pi_j = W_j \prod_{i=1}^{j-1} (1 - W_i)(1 - W_2), \dots, (1 - W_{j-1})$ . As  $j \rightarrow \infty$  the unit length of the stick can be written as  $1 = \sum_{j=1}^{\infty} \pi_j$ . Since any partition can be the result of this stick breaking process, Dirichlet process spans the entire distribution of the discrete probability functions with probability 1. As the functional form also suggests, the length of the broken parts is analogous to the probabilities of each partition. Suppose we denote each part of the stick, i.e. each partition, as  $\theta_j$ , which is a draw from the base distribution  $G_0$ . Then, we can write a draw of distribution  $G$  from the DP as

$$G = \sum_{j=1}^{\infty} \pi_j \delta(\theta_j). \quad (5.7)$$

(5.7) points out the DP as an *infinite mixture* of point mass functions. Unfortunately, as the DP spans the entire distribution of the discrete probability functions with probability 1, the space of continuous probability functions is excluded. Although any distribution over  $\theta$  can be approximated arbitrarily accurately in the pointwise sense by a sequence of draws from the DP, the smoothness of the distribution is still of concern. Dirichlet process mixture model remedies this drawback by convolving the DP with a continuous kernel. This is accomplished by replacing the point mass function  $\delta(\theta_j)$  in (5.7) with a continuous function  $f(\cdot|\theta_j)$  which is a continuous kernel function conditional on the parameters  $\theta$  with the DP prior. By doing that we obtain a specification that does span the entire set of continuous probability distributions with probability 1, which is denoted as the Dirichlet process mixture (DPM) model. The DPM takes the form of

$$\sum_{j=1}^{\infty} \pi_j f(\cdot|\theta_j). \quad (5.8)$$

This implies that (5.8) corresponds to an (*countably*) *infinite number of mixtures* of underlying distributions and it provides a very intuitive and tractable way of estimation of the unknown distributions. More explicitly, the density of a random variable, say  $x$ , can be estimated by the DPM quite accurately using the mixture components  $f(x|\theta_j)$  with parameters  $\theta$  where  $\theta$  has a DP prior.

Notice that, in practice the sample size is finite. This means that the number of mixtures will be limited with the sample size in practice. Moreover, the discrete nature of the DP prior often results in a low number of mixture components that can successfully approximate the target density. This suggests the DPM as a very flexible yet parsimonious modeling approach to estimate an unknown density and to model the non-Gaussianities observed in the bond yields data.

### 5.2.3 Bayesian semiparametric dynamic Nelson-Siegel model

If the distribution of the factors in (5.4) is of an unknown form, a specification that is flexible enough to estimate the distribution of the factors along with other model parameters is crucial. DPM provides a convenient and tractable way of estimation of this distribution. By assigning a DP prior as the prior of the parameters  $\theta_t$  in (5.4), the error terms of the state equation are distributed according to a DPM and we achieve a model where the error distribution in state equation (factor evolution) is unknown and estimated along with other parameters by estimating the number of mixtures and mixture components. Specifically, in our Bayesian semiparametric dynamic Nelson-Siegel model

(DNS-DPM), we assume  $f(\cdot|\theta_j)$  to be a Gaussian density, where  $\theta_t = (\mu_t, Q_t)$  includes the mean and the variance of the error distribution, respectively. As a result the full model specification takes the following hierarchical form

$$\begin{aligned} y_t &= Hf_t + \eta_t & \eta_t &\sim N(0, R) \\ f_t &= Ff_{t-1} + \varepsilon_t & \varepsilon_t &\sim N(\theta_t) \\ & & \theta_t|G &\sim G \\ & & G &\sim DP(G_0(\Lambda), \alpha). \end{aligned} \tag{5.9}$$

The hierarchical representation indicates that the prior distribution  $G$  is distributed according to  $DP(G_0(\Lambda), \alpha)$  and together with the multivariate normal density of the data conditional on  $\theta_t$  it forms the posterior distribution of  $\theta_t$  as a DPM. Notice that when the resulting number of mixtures is exactly one, we end up in the standard DNS model. Hence, the above specification also involves the standard model with Gaussian innovations as a special case.

Without specifying the autoregressive coefficient matrix  $F$  and the covariance matrix of the measurement error  $R$  the model is not complete. Existing evidence on the factor evolution suggests that persistence of the factors is one of the most prominent stylized facts of the yield curve models. In many applications of the dynamic Nelson-Siegel model estimates of the autoregressive coefficient matrix  $F$  indicate uniformly this persistence with diagonal values above 0.9 or around it regardless of the model specification or estimation technique.<sup>8</sup> Following this evidence, we keep  $F$  in (5.9) out of the set of the parameters with DP prior.

Another parameter that we do not assign a DP prior is the covariance matrix of the measurement error,  $R$ . In fact, the extension of the model to a DPM structure in the measurement error distribution is a straightforward extension of our modeling approach. However, as the dimension of the yields is much higher than the dimension of the underlying factors such an approach would result in estimation of many parameters and identification of the densities could be difficult. On the other hand, one could also impose the semiparametric structure only in the measurement equation. Our strategy, however, is to estimate the density of the yields and to model the observed non-Gaussianities in a parsimonious yet flexible specification using the underlying factors as discussed in Section 5.1. Given that factors can explain the variation in a large set of yields quite successfully, we follow the approach of estimation the factor density in this paper.

---

<sup>8</sup>See for example Diebold and Li (2006); Diebold *et al.* (2006); Christensen *et al.* (2011); Koopman *et al.* (2010); Hautsch and Yang (2010), among others, for a comparison of the estimates of the autoregressive coefficient matrix.

### 5.3 Bayesian inference

In this section we provide details on econometric inference of the Bayesian semiparametric dynamic Nelson-Siegel model. The joint posterior distribution of the parameters with DPM distribution is constructed by the product of the DP prior together with the likelihood. Unfortunately, the resulting joint posterior distribution of these parameters,  $p(\theta_{1:T}|y_{1:T})$  does not have a closed-form solution except for some trivial cases or in case of very low sample sizes. Instead, Escobar and West (1995) propose to use MCMC samplers to sample from the conditional distributions by exploiting the Polya's urn representation of the DP as in (5.5). In Section 5.3.1, we discuss the prior distributions. In Section 5.3.2 we discuss the full conditional posterior distributions related to the DPM and the resulting simulation scheme together with the DPM sampler. We discuss the posterior conditional distributions which do not follow a DPM distribution in Appendix 5.A.2 as these distributions are identical with the parameter distributions in the standard DNS model with Gaussian innovations. The Bayesian model selection procedures for model comparison are discussed in Section 5.3.3.

#### 5.3.1 Prior distributions

We start to the specifications of the prior distributions with the specifications that are related to the DP. This includes the specification of the base prior and the specification of the hyperprior for the precision parameter  $\alpha$ . For the base prior  $G_0$  we use a normal inverse Wishart distribution as follows

$$G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, a, \nu, S) \mapsto \mu|\Sigma \sim N(\mu_0, \frac{\Sigma}{a}) \text{ and } \Sigma \sim IW(\nu, S), \quad (5.10)$$

where  $E(\mu) = \mu_0$  and  $E(\Sigma^{-1}) = \nu S^{-1}$ . Moreover,  $S$  can be decomposed further into  $S = \nu \nu I_3$  where  $I_n$  is an identity matrix of the dimension of  $n$ . This implies that a priori we assume uncorrelated errors. With very limited information on the correlation structure among the mixture parameters this is a natural assumption, see for example Conley *et al.* (2008) for a similar specification.

For the precision parameter of the DP,  $\alpha$ , we use the hyperprior distribution of the form

$$f(\alpha) \propto \left(1 - \frac{\alpha - \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}}\right)^p. \quad (5.11)$$

We take the power parameter  $p = 0.8$  and  $\alpha \in [0.1, 3]$ , see Conley *et al.* (2008) for details about this hyperprior. The power value implies that the prior has more probability mass

on more parsimonious specifications. We prefer such a hyperprior specification since we want to keep a balance between the estimation of the density and model overfitting. As a high value of  $\alpha$  implies a high number of clusters, a more flexible specification could easily result in model overfitting.

The remaining model parameters do not involve a DP prior. For the autoregressive parameters in transition equation 5.3 we assume the standard multivariate normal distribution as

$$f(\text{vec}(F)) = N(\mathbf{0}_9, I_9), \quad (5.12)$$

where  $\text{vec}(F)$  is the vec operator stacking each columns of  $F$  into a single column vector.  $\mathbf{0}_9$  is a vector of zeros of dimension nine. Unreported results of prior sensitivity analysis suggests that the likelihood information dominates prior information for  $F$  and thus, the prior specification is of secondary importance for this parameter.

For  $R$ , i.e. the covariance matrix of the measurement error in (5.2), we assume a noninformative diagonal structure as all the cross sectional dependence is cast to the factor structure. For each individual elements a noninformative diagonal structure on the prior distribution of  $R$  implies

$$f(\sigma_n^2) \propto \frac{1}{\sigma_n^2} \quad \text{for all } n = 1, 2, \dots, N, \quad (5.13)$$

where  $\sigma_n^2$ 's are the diagonal elements of  $R$  and  $N$  is the number of the cross sections of yields, i.e number of the maturities.

### 5.3.2 Posterior simulation

The fact that the number of mixtures as well as the mixture components of the distribution of the factors are unknown leads to an intractable likelihood function. Instead, we set up an MCMC algorithm to sample from the full conditional posterior distributions. Specifically, we use Gibbs sampling together with data augmentation (see Geman and Geman, 1984; Tanner and Wong, 1987) to obtain posterior results. The resulting simulation scheme is as follows

1. Initialize the parameters by drawing  $\theta_t$  from the base prior  $G_0$  and  $f_t$  from standard normal distribution for  $t = 0, 1, \dots, T$ . At step  $(m)$  of the iteration
2. Sample  $F^{(m)}$  from  $p(F|f_{0:T}^{(m-1)}, \theta_{1:T}^{(m-1)})$ .
3. Sample  $R^{(m)}$  from  $p(\sigma_n^2|y_{1:T}, f_{1:T}^{(m-1)}, \theta_{1:T}^{(m-1)})$  for  $n = 1, 2, \dots, N$ .

4. Sample  $\theta_{1:T}^{(m)}$  from  $p(\theta_t|f_{0:T}^{(m-1)}, \theta_{-t}^{(m-1)}, F^{(m)})$  for  $t = 1, 2, \dots, T$ .
5. Sample  $f_{0:T}^{(m)}$  from  $p(f_{0:T}|y_{1:T}, \theta_{1:T}^{(m)}, F^{(m)})$ .
6. Sample  $\alpha^{(m)}$  from  $p(\alpha|\theta_{1:T}^{(m)})$ .
7. Repeat (2)-(6)  $M$  times.

Alternatively, steps (3) and (4) can be replaced by sampling  $\theta_t^{(m)}$  from  $p(\theta_t|y_{1:T}, \theta_{-t}^{(m-1)}, F^{(m)})$  which will be discussed in detail in the next section and in Appendix 5.A.1. Steps (2), (3), and (5) are common to many models in the Bayesian state space framework, see for example Kim and Nelson (1999). Details of the full conditional distributions for these parameters are given in Appendix 5.A.2. Here we focus on the DPM sampler in step (4) and on the posterior simulation of the precision parameter  $\alpha$  in step (6).

### Sampling of error distribution parameters in the transition equation, $\theta_t$

Sampling the parameters with DPM distribution is not an easy task as both the factors and these parameters are unobserved, and  $\theta_{1:T}$  relates to the observed data through the unobserved factors. First, we start with the general idea of the DPM sampler as proposed in Escobar and West (1995) and then we show how we can integrate this into state space framework to obtain inference in the Bayesian semiparametric dynamic Nelson-Siegel (DNS-DPM) model.

The posterior distribution of the parameters with the DPM distribution,  $p(\theta_t|y_{1:T})$  does not have a closed form solution. This is due to the fact that obtaining the posterior distribution requires the integration out of the all remaining model parameters as  $p(\theta_t|y_{1:T}) = \int p(\theta_t, \theta_{-t}|y_{1:T})d\theta_{-t}$ . Unfortunately this is not possible except for some trivial cases or for cases of very low dimensions. Instead, one can resort to MCMC sampler using the conditional distributions, see Escobar and West (1995). Notice that, the joint posterior distribution can be defined in terms of the conditionals as

$$p(\theta_{1:T}|y_{1:T}) = p(\theta_1|y_{1:T})p(\theta_2|\theta_1, y_{1:T}) \dots p(\theta_T|\theta_{1:T-1}, y_{1:T}). \quad (5.14)$$

Equation (5.14) implies that each of the parameters at time  $t$  can be sampled from a distribution conditional on the parameters sampled so far as implied in (5.5) replacing  $T$  with  $t$ . Moreover, as the sequence of these parameters are exchangeable, each of the parameters can be treated as if it is the last observation and an efficient Gibbs sampler can be used to obtain a sample from the joint posterior distribution  $p(\theta_{1:T}|y_{1:T})$  by replacing the conditionals with  $p(\theta_t|\theta_{-t}, y_{1:T})$ .

In the resulting simulation scheme, iteration ( $m$ ) of the sampler to obtain  $p(\theta_{1:T}^{(m)}|y_{1:T})$  can be implied by sequentially sampling from  $p(\theta_t|\theta_{-t}, y_{1:T})$  for  $t = 1, 2, \dots, T$  with each posterior being

$$\begin{aligned} p(\theta_t|\theta_{-t}, y_{1:T}) &\propto p(y_{1:T}|\theta_t, \theta_{-t}^{(m-1)})p(\theta_t|\theta_{-t}^{(m-1)}) \\ &= p(y_{1:T}|\theta_{1:T})p(\theta_t|\theta_{-t}^{(m-1)}). \end{aligned} \quad (5.15)$$

The estimation of  $\theta_t$  implies to draw from (5.15) conditional on the  $\theta_{-t}$  and the data. Specifically (5.15) corresponds to the following specification

$$p(y_{1:T}|\theta_{1:T}) \propto \left( \frac{\alpha}{\alpha + T - 1} G_0(\Lambda) + \sum_{i=1, i \neq t}^T \frac{\delta_{\theta_i}^j}{\alpha + T - 1} \right). \quad (5.16)$$

A generic way to sample from the distribution in (5.16) is to setup a Metropolis-Hastings (MH) algorithm with the conditional prior as the candidate distribution, see Neal (2000). This would mean to evaluate the following acceptance probability  $\rho^{(m)}$  at iteration ( $m$ ) of the MCMC sampler

$$\rho^{(m)} = \min \left( 1, \frac{p(y_{1:T}|\theta_{1:T})p(\theta_t^*|\theta_{-t})q(\theta_t^*|\theta_{-t})}{p(y_{1:T}|\theta_{1:T})p(\theta_t^{(m-1)}|\theta_{-t})q(\theta_t^{(m-1)}|\theta_{-t})} \right). \quad (5.17)$$

Note that both  $p(\theta_t^*|\theta_{-t})$  and  $p(\theta_t^{(m-1)}|\theta_{-t})$  are identical as  $\frac{\alpha}{\alpha + T - 1} G_0(\Lambda) + \sum_{i=1, i \neq t}^T \frac{\delta_{\theta_i}^j}{\alpha + T - 1}$  and independent across draws. Hence (5.17) corresponds to an independence chain MH algorithm with the following acceptance probability

$$\rho^{(m)} = \min \left( 1, \frac{p(y_{1:T}|\theta_t^*, \theta_{-t})}{p(y_{1:T}|\theta_t^{(m-1)}, \theta_{-t})} \right), \quad (5.18)$$

see Caron *et al.* (2008). The acceptance probability indicates that for sampling  $\theta_t$  a full Kalman filter step is required to evaluate the likelihood. This implies that drawing a single  $\theta_t$ , requires  $O(T)$  operations. As we sample  $\theta_t$  for  $t = 1, 2, \dots, T$ , a full sweep requires  $O(T^2)$  operations. To decrease the number of operations Caron *et al.* (2008) suggest a decomposition based on Doucet and Andrieu (2001). This decomposition and the resulting algorithm is described in detail in Appendix 5.A.1.

A drawback of this procedure, however, is that the behavior of the MH algorithm may differ substantially from the Gibbs sampling algorithms with conjugate priors, see Neal (2000) for details. One reason is that while other methods consider all mixture components when deciding for a new draw, for MH method it is more likely to consider changing the draw to a component with many observations rather than with few observations

because the candidate density is the DP prior. Changing a draw to a new component is proportional to  $\alpha$  hence with small values of  $\alpha$  the probability of a new cluster may be inefficiently low. Notice that even if a new cluster is sampled from the candidate density whether it will be accepted depends on the acceptance probability. This may dramatically slow down the mixing capability of the chain further unless a high value for  $\alpha$  is chosen.

Closely related, the choice of the base prior plays a crucial role in the sense that the efficiency of the independence chain algorithm depends on the appropriate choice of the DP prior (candidate density) for posterior distribution in (5.16). The reason is that convolving the DP prior with the data likelihood may result in a posterior that is quite different from the prior. For example, in (5.18) while a base prior that has fat tails compared to the likelihood function would result in frequent rejections, a base prior that covers the parts of the likelihood with high probability mass may result in clusters that are very close to each other even if the acceptance rate is high.

In this paper we depart from this algorithm based on MH step and propose another algorithm that is computationally more tractable and depend on the Gibbs sampling using the posterior distribution directly by exploiting the state-space structure of the model. The intuition is that the dependence of the  $\theta_t$  to the data is through the unobserved states. Hence, the likelihood function can be decomposed into the distribution of the yields conditional on the factors and the distribution of the factors conditional on the parameters with the DP prior. During the steps of the Gibbs sampler the states are estimated for each round of the sampler, and thus, for estimation of the  $\theta_t$ , these estimated states can be treated as observed data. That would facilitate the estimation of the  $\theta_t$  greatly as in this case posterior distribution has a closed form solution and MH algorithm is not required.

Formally, the conditional likelihood of the observation at period  $t$  can be decomposed as

$$\begin{aligned} p(y_t|\theta_t, \theta_{-t}) &= \int p(y_t|f_t, \theta_t, \theta_{-t})p(f_t|\theta_t, \theta_{-t})df_t \\ &= \int p(y_t|f_t)p(f_t|\theta_t)df_t. \end{aligned} \quad (5.19)$$

and using (5.19) the posterior distribution conditional on yields in (5.15) takes the following form

$$p(\theta_t|\theta_{-t}, y_{1:T}) = \int p(y_t|f_t)p(f_t|\theta_t)p(\theta_t|\theta_{-t})df_t. \quad (5.20)$$

Equation (5.20) suggests that at each step of the Gibbs sampler the posterior distribution of the parameters with the DP prior can be evaluated using the part  $p(f_t|\theta_t)p(\theta_t|\theta_{-t})$  by treating the factors as observed data. The resulting posterior distribution conditional on

the factors becomes

$$\begin{aligned} p(\theta_t | \theta_{-t}, f_{1:T}) &\propto p(f_t | \theta_t) p(\theta_t | \theta_{-t}) \\ &\propto \frac{\alpha q_{0,f}}{\alpha + T - 1} p(f_t | \theta_t) G_0(\theta_t) + \frac{\sum_{i=1, i \neq t}^T \delta_{\theta_i}^j}{\alpha + T - 1} p(f_t | \theta_j) \end{aligned} \quad (5.21)$$

where  $q_{0,f} = \int p(f_t | \theta_t) G_0(\theta_t) d\theta_t$ ,  $p(f_t | \theta_t) G_0(\theta_t)$  is the posterior density using the factor likelihood and the base prior, and  $p(f_t | \theta_j)$  is the factor density conditional on the parameters in the  $j^{th}$  cluster. Fortunately, as we use factors instead of the observed yields the integral in  $q_{0,f}$  has a closed-form solution. As  $\theta_t = (\mu_t, Q_t)$ ,  $q_{0,f}$  takes the following form

$$\begin{aligned} q_{0,f} &= \int p(f_t | \mu_t, Q_t) G_0(\mu_t, Q_t | \Lambda) d\mu_t dQ_t \\ &= \frac{p(\varepsilon_t | \mu_t, Q_t) p(\mu_t | Q_t) p(Q_t)}{p(\mu_t | \varepsilon_t, Q_t) p(Q_t | \varepsilon_t)} \\ &= 2^{\frac{3}{2}} (2\pi)^{-\frac{3}{2}} \left(\frac{a}{a+1}\right)^{\frac{3}{2}} \frac{\Gamma_3\left(\frac{\nu+1}{2}\right)}{\Gamma_3\left(\frac{\nu}{2}\right)} \frac{|V|^{\frac{\nu}{2}}}{\left|V + \frac{a}{a+1}(\varepsilon_t - \mu_0)(\varepsilon_t - \mu_0)'\right|^{\frac{\nu+1}{2}}}. \end{aligned} \quad (5.22)$$

where  $\Gamma_3(\cdot)$  is the multivariate Gamma function of dimension 3 (the dimension of the factors) and  $\varepsilon_t = f_t - Ff_{t-1}$ , see, for example, Escobar and West (1995); Conley *et al.* (2008). Since in the first equation we implicitly condition on the estimates of  $F$  and  $f_{t-1}$ , we switch from  $f_t$  to  $\varepsilon_t$  when we proceed from the first equation to the second.

Plugging (5.21) together with (5.22) into (5.20) results in the following posterior distribution conditional on the observed yields

$$\begin{aligned} p(\theta_t | \theta_{-t}, y_{1:T}) &\propto \int p(y_t | f_t) \left( \frac{\alpha}{\alpha + T - 1} q_{0,f} p(f_t | \theta_t) G_0(\theta_t) \right) df_t \\ &\quad + \int p(y_t | f_t) \frac{\sum_{i=1, i \neq t}^T \delta_{\theta_i}^j}{\alpha + T - 1} p(f_t | \theta_j) df_t \\ &= \int p(y_t | f_t) \left( \frac{\alpha}{\alpha + T - 1} q_{0,f} p(f_t | \theta_t) G_0(\theta_t) + \frac{\sum_{i=1, i \neq t}^T \delta_{\theta_i}^j}{\alpha + T - 1} p(f_t | \theta_j) \right) df_t. \end{aligned} \quad (5.23)$$

This implies that we can add one more step to the MCMC sampler by treating the estimated factors from the previous step as the observed data and carry on the sampling of the DPM distributed parameters using these factors. To obtain a sample from the posterior distribution (5.21), let  $c_t$  be the indicator function taking the value  $j$  if observation at period  $t$  belongs to the  $j^{th}$  cluster. Following this, at step ( $m$ ) of the iteration, conditional on the factors,  $f_{0:T}^{m-1}$ ,  $\theta_t^m$  for  $t = 1, 2, \dots, T$  can be sampled using the following steps

1. For  $t = 1, 2, \dots, T$  sample  $c = (c_1, c_2, \dots, c_T)'$  from  $p(c_t | c_{-t}^{(m-1)}, f_{0:T}^{(m-1)})$  for  $t = 1, 2, \dots, T$  using the Dirichlet process mixture posterior using (5.21),
2. For all  $c = c_j$  sample  $\theta_j$  from  $p(\theta_j | c^{(m)}, f_{0:T}^{(m-1)}) \propto \prod_{i:c_i=j} [f(f_t | \theta_j)] G_0(\theta_j)$ .

The second step of the algorithm is often called as the ‘remix’ step and it is to ensure that the values of  $\theta_j$  can change using all the observations in that cluster rather than a single observation. This is accomplished by drawing it from an equivalent distribution to  $p(\theta_1, \dots, \theta_T | f_{0:T})$  as in the second step, see West *et al.* (1994); MacEachern and Müller (1998).

### Sampling DP precision parameter, $\alpha$

The posterior for the  $\alpha$ ,  $p(\alpha | J)$ , is the product of the hyperprior in (5.11) and the likelihood function  $p(J | \alpha)$  where  $J$  is the number of clusters. The resulting posterior is

$$p(\alpha | J) \propto p(\alpha) T! \alpha^J \frac{\Gamma(\alpha)}{\Gamma(\alpha + T)} \quad (5.24)$$

see Antoniak (1974). Given (5.24) we can easily implement the griddy Gibbs sampler approach of Ritter and Tanner (1992), see also Conley *et al.* (2008). Given that  $\alpha \in [0.1, 3]$  we can setup a grid in this interval based on the precision we desire about the value of  $\alpha$ .

Escobar (1994) shows that the expected number of clusters for a given  $\alpha$ ,  $E(J | \alpha) = \sum_{i=1}^T \frac{\alpha}{\alpha + i - 1}$  approaches to  $\alpha \ln\left(\frac{\alpha + T}{\alpha}\right)$  as  $T \rightarrow \infty$ . This implies that the interval for  $\alpha$  corresponds to the number of clusters to be between 1 and 16 with more support for lower values.<sup>9</sup>

### 5.3.3 Model selection

The Bayesian semiparametric dynamic Nelson-Siegel model encompasses the standard DNS model with Gaussian innovations as a specific case. Hence, comparison of the standard model as a restricted version of the semiparametric model with the unrestricted semiparametric model would indicate whether there is a need for the semiparametric structure. Following this aim, in this section we discuss several Bayesian model comparison techniques.

Since the key task in our methodology is density estimation we use predictive density as a first metric for model comparison. We, therefore, use the in-sample predictive density to measure the fit of the predicted densities by the two competing models to the data.

---

<sup>9</sup>West (1992) shows that the distribution of  $\alpha$  conditional on the number of clusters can be cast as a mixture of the two Gamma distributions with the mixing proportion distributed according to Beta distribution. Hence, a Gamma prior for  $\alpha$  also leads to a tractable posterior distribution of mixture of two Gamma distributions and a sample from  $\alpha$  can be obtained by sequentially sampling first the mixing proportion from the Beta distribution and then  $\alpha$  from the Gamma distribution.

The in-sample predictive posterior density can be computed as

$$\begin{aligned} f(y_t|y_{1:T}) &= \int p(y_t|f_t, \theta_{1:T}, \alpha) p(f_t|\theta_{1:T}, \alpha, y_{1:T}) df_t d\theta d\alpha \\ f(y_t|y_{1:T}) &= \int p(y_t|f_t) p(f_t|\theta_{1:T}, \alpha, y_{1:T}) df_t d\theta d\alpha \\ f(y_t|y_{1:T}) &\approx \frac{1}{M} \sum_{m=1}^M p(y_t|f_t^m), \end{aligned} \quad (5.25)$$

see Gelfand and Mukhopadhyay (1995); Jensen and Maheu (2010). The density of the data conditional on the factors is a multivariate normal distribution as stated in (5.9).

As a second metric for evaluation of the models we use the posterior predictive p-value (*ppp*). The *ppp* provides a framework to understand whether the data simulated from the model (using the sample from the posterior distributions of the model parameters) can replicate various features of the actual data. These features of the data can usually be formulated as functions of data,  $g(y)$ , such as the skewness or the kurtosis measures. If this is the case, it implies that indeed the model specification can approximate the underlying data generating process of the actual data well. As we want to capture the non-Gaussian features of the bond yields data, this metric is very suitable for comparison of the dynamic Nelson-Siegel models with a semiparametric factor structure and with only Gaussian innovations. The posterior predictive p-value can be computed as

$$\begin{aligned} p(g(y^{sim})|y_t) &= \int p(g(y^{sim})|y_t, \Gamma) p(\Gamma|y_t) \\ &= \int p(g(y^{sim})|\Gamma) p(\Gamma|y_t), \end{aligned} \quad (5.26)$$

where  $\Gamma$  denotes the parameter space. Computing (5.26) requires to generate simulated data,  $y^{sim}$ , using the posterior distribution of model parameters. For that we use

$$\begin{aligned} f_t &\sim N(\mu_t + F f_{t-1}, Q_t) \\ y_t &\sim N(H f_t, R), \end{aligned} \quad (5.27)$$

that is, first, we sample the states using the state equation and then we use these simulated states to sample simulated data. The second equation in (5.26) stems from the fact that conditional on the model parameters, observations do not provide any additional information for the simulated data. The *ppp* can then be computed as

$$\begin{aligned} ppp &= p(g(y_t^{sim}) \geq g(y_t)) \\ &\approx \frac{1}{S} \sum_i^S \mathbf{1}[g(y_t^{sim}) \geq g(y_t)]. \end{aligned} \quad (5.28)$$

A  $ppp$  value close to 0.5 indicates that the model can generate the features of the actual data well while a  $ppp$  value close to 0 or 1 indicates that the model is not able to generate these features observed in the actual data.

## 5.4 Empirical results using US yield data

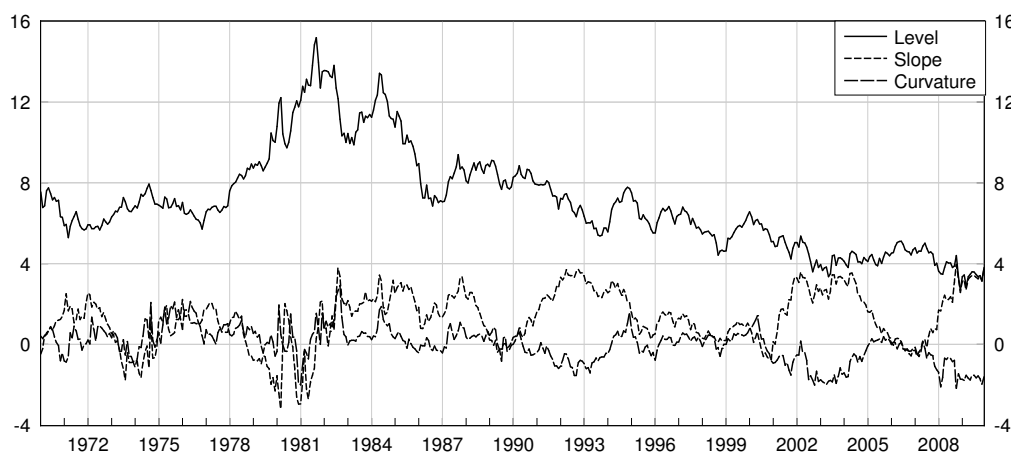
In this section we examine the potential of our semiparametric framework using a real bond yields dataset. Following this aim, in our empirical investigation we use the US treasury bond data and accordingly, we investigate the US yield curve. This dataset is analyzed extensively in many applications of the dynamic Nelson-Siegel model, and thus, it provides us ample opportunities to compare our results with the standard applications of dynamic Nelson-Siegel model. The data consists of all Treasury bills, noncallable notes and bonds over the period January 1970 through December 2009 from the CRSP Monthly Treasury Cross-Sectional File. The instruments with special liquidity problems, which are the bills with less than one month to maturity and all notes and bonds with less than one year to maturity are eliminated, see Diebold and Li (2006) for details. Forward rates are computed using the unsmoothed Fama-Bliss method (see Fama and Bliss (1987)). These forward rates are converted into unsmoothed Fama-Bliss zero yields.<sup>10</sup> The resulting dataset consists of 22 yields with maturities 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 60, 72, 84, 96, 108, 120 months and each of the yields has 480 monthly observations, i.e.  $N = 22$  and  $T = 480$ .

We start our analysis with a preliminary investigation of the distribution of the US yield curve factors. We first construct crude estimates of factors using linear combinations of the yields and we use these factors as if they were observed. Such a preliminary analysis would provide us some clues on the factor densities before the comprehensive analysis of the yield curve factors using the dynamic Nelson-Siegel model where factors are latent variables to be estimated along with other model parameters. We follow the common practice in the literature by computing the level factor using the 10-year yield, the slope factor by taking the difference between the 10-year and 3-month yields, and the curvature by taking twice the 2-year yield minus the sum of the 3-month and 10-year yields, see for example Diebold and Li (2006); Koopman *et al.* (2010). We take these combinations as observed data and analyze the distribution of the factors. Figure 5.1 displays the resulting factors.

Following the structure of the state equation in the dynamic Nelson-Siegel model we first estimate a Vector AutoRegression of the first order (VAR(1)) using these observed

<sup>10</sup>We thank Michel van der Wel for making the data available to us.

Figure 5.1: Estimates of factors using linear combinations of the yields



*Note:* The solid line is the level factor computed using the 10-year yield. The dotted line is the slope factor computed by taking the difference between the 10-year and 3-month yields and the dashed line is the curvature factor computed by taking twice the 2-year yield minus the sum of the 3-month and 10-year yields. We use the zero-coupon yields over the period January 1970 - December 2009 for computations.

factors. We restrict the off-diagonal elements of the autoregressive coefficients to be zero but results with the unrestricted model are very similar. For the estimation we fix the parameters of the base distribution at  $\mu_0 = (0.05, -0.04, -0.003)'$ ,  $a = 5$ ,  $\nu = 5$  and  $v = 1$  and estimate  $\alpha$  where  $\alpha \in [0.1, 3]$ . For the specific values of  $\mu_0$ , we first filter the crude estimates of the factors using an AR(1) filter with autoregressive coefficients close to 1 and then we take the average of the filtered values. We could also set all the values to zero but such a setting is unrealistic as it implies that the unconditional mean of the factors (especially the level factor) is zero. The results are qualitatively very similar with different setups of hyperparameters. Standard VAR(1) model with a multivariate normal error distribution is denoted as VAR(1)-N model and the model where the error distributions follow a DPM is denoted as VAR(1)-DPM. The results of the VAR(1)-N and VAR(1)-DPM models are presented in Table 5.1.

The results in Table 5.1 confirm the findings about the persistence of the factors, see Diebold and Li (2006), Diebold *et al.* (2006) and Christensen *et al.* (2009) among others. All of the autoregressive coefficient estimates (posterior means) are above or around 0.90 with the estimates of VAR(1)-DPM model higher than that of VAR(1)-N models. An interesting result is that the precision of the VAR(1)-DPM parameters is much higher compared to VAR(1)-N model. This may also be due to the fact that we use crude

estimates of the factors and VAR(1)-DPM model can account for the measurement errors much better than VAR(1)-N model. This is also reflected to the much higher marginal likelihood value of the VAR(1)-DPM model.

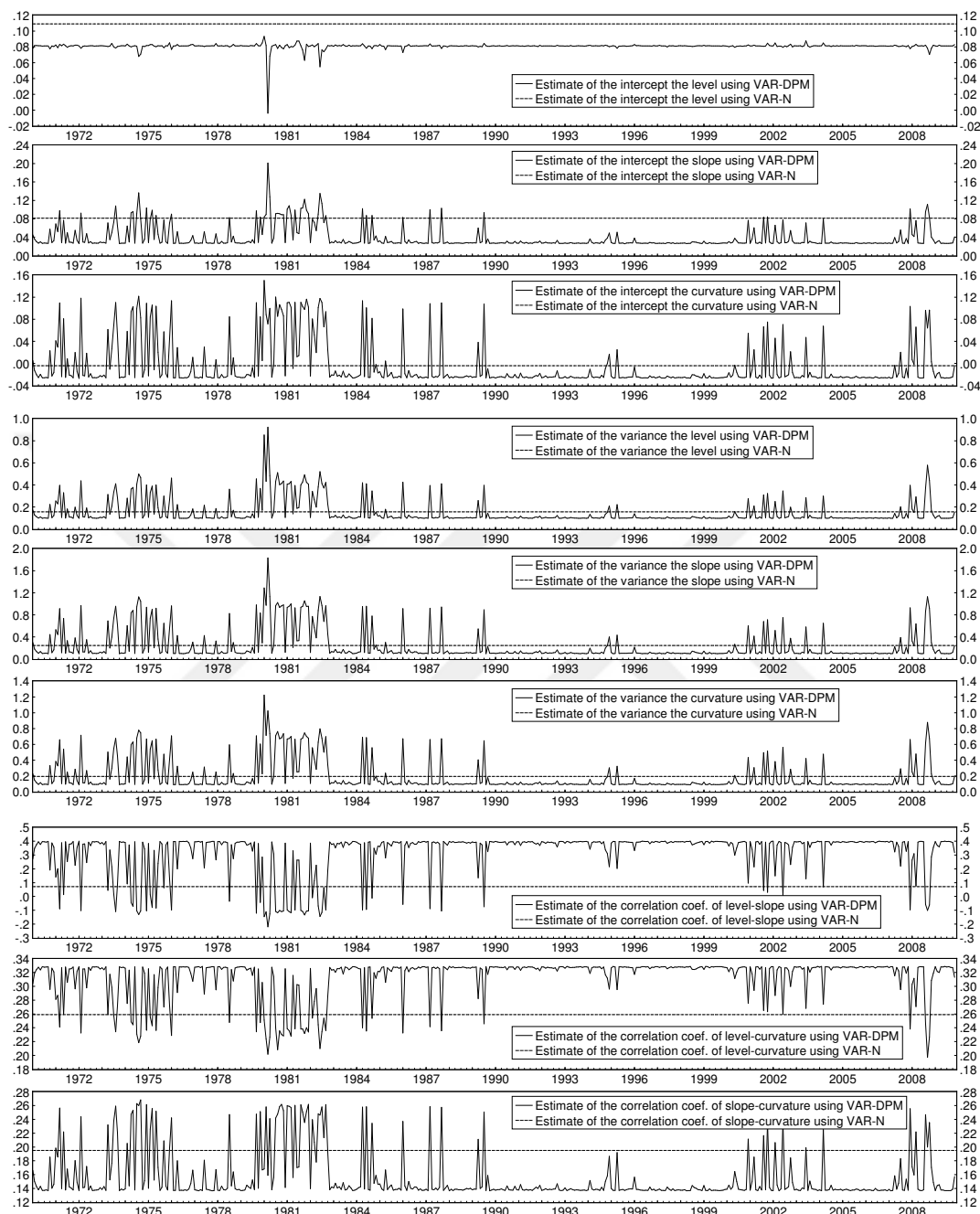
Table 5.1: Posterior results for VAR(1)-DPM and VAR(1)-N models

|                     | Level            | VAR(1)-N         |                   | Level            | VAR(1)-DPM       |                  |
|---------------------|------------------|------------------|-------------------|------------------|------------------|------------------|
|                     |                  | Slope            | Curvature         |                  | Slope            | Curvature        |
| $\mu_f$             | 0.105<br>(0.034) | 0.078<br>(0.030) | -0.004<br>(0.019) |                  |                  |                  |
| $\phi_f$            | 0.984<br>(0.004) | 0.945<br>(0.016) | 0.903<br>(0.022)  | 0.987<br>(0.003) | 0.975<br>(0.009) | 0.921<br>(0.019) |
| $\sigma_f^2$        | 0.138<br>(0.009) | 0.229<br>(0.015) | 0.178<br>(0.012)  |                  |                  |                  |
| $\rho_{1j}$         | 1                | 0.079<br>(0.046) | 0.295<br>(0.041)  |                  |                  |                  |
| $\rho_{2j}$         | 0.079<br>(0.046) | 1                | 0.216<br>(0.044)  |                  |                  |                  |
| $\rho_{3j}$         | 0.295<br>(0.041) | 0.216<br>(0.044) | 1                 |                  |                  |                  |
| $\alpha$            |                  |                  |                   |                  | 0.45<br>(0.32)   |                  |
| # of clusters       |                  |                  |                   |                  | 2.88<br>(1.03)   |                  |
| Marginal Likelihood |                  | -729.58          |                   |                  | -508.29          |                  |

*Note:* VAR(1) is the Vector AutoRegressive (VAR) model with first-order autoregressive dynamics. We restrict the autoregressive coefficient matrix  $F$  to be a diagonal matrix and thus, the off-diagonal elements are set to zero. VAR(1)-N stands for the VAR(1) model with Gaussian innovations, while VAR(1)-DPM stands for the VAR(1) model with DPM innovations.  $\mu_f$  for  $f = 1, 2, 3$  is the unconditional mean of the factor errors (intercept of the factors) and  $\phi_f$  is the first-order autoregressive coefficients of the factors, i.e diagonal elements of the autoregressive coefficient matrix.  $\sigma_f^2$  is the variance of the factor error terms, whereas  $\rho_{i,j}$  denotes the correlation of the error term of the  $i^{th}$  factor for  $i = 1, 2, 3$  with the  $j^{th}$  factor error term  $j \in \{1, 2, 3, \setminus i\}$ .  $\alpha$  is the precision parameter of the DP. The values indicate the posterior means of the parameters with the posterior standard deviations in parenthesis. For the estimation of the models we use the sample period from January 1970 through December 2009. 60,000 simulations from the posterior distribution are used for inference. The first 10,000 draws are discarded as burn-in sample and the remaining 50,000 draws, where we kept every  $10^{th}$  draw, are used for posterior inference in the models. The convergence of the sampler is checked using statistical and visual inspection and in all the model specifications convergence is assured.

The estimates of the DPM parameters and comparison with that of the VAR(1)-N model are displayed in Figure 5.2.

Figure 5.2: Posterior means of the parameters for VAR-N and VAR-DPM model



*Note:* The solid lines correspond to the estimates of the time varying parameters due to the DPM structure which are not displayed in Table 5.1. The dotted lines represent the corresponding estimates using VAR(1)-N model displayed also in Table 5.1. See Table 5.1 for further details.

The results indicate two different types of periods. The first period captures the years between mid 1980s and mid 2000s which is the Great Moderation period with a

persistence decline in the volatility of many US macroeconomic and financial variables, see McConnell and Perez-Quiros (2000), and Barnett and Chauvet (2011). This period is reflected as low values of the variance estimates in VAR(1)-DPM model opposite to the second period excluding the great moderation. In the other period the volatility of the factors is relatively higher which is also reflected to the estimates of the intercepts, most notably for the slope and curvature factors, and to the estimates of the correlation coefficients.

An interesting finding from the preliminary analysis is that the intercept of the level factor is quite stable trough the sample period. This leads us to consider a DNS-DPM model, where only the variance parameter is assumed to follow a DPM but the intercept parameters are assumed to be constant, in our analysis with the DNS model. Following Jensen and Maheu (2010) we denote this model as DNS-DPMP. The estimation of this model is identical to the DNS-DPM model explained in detail in Section 5.3 with only (5.22) replaced by

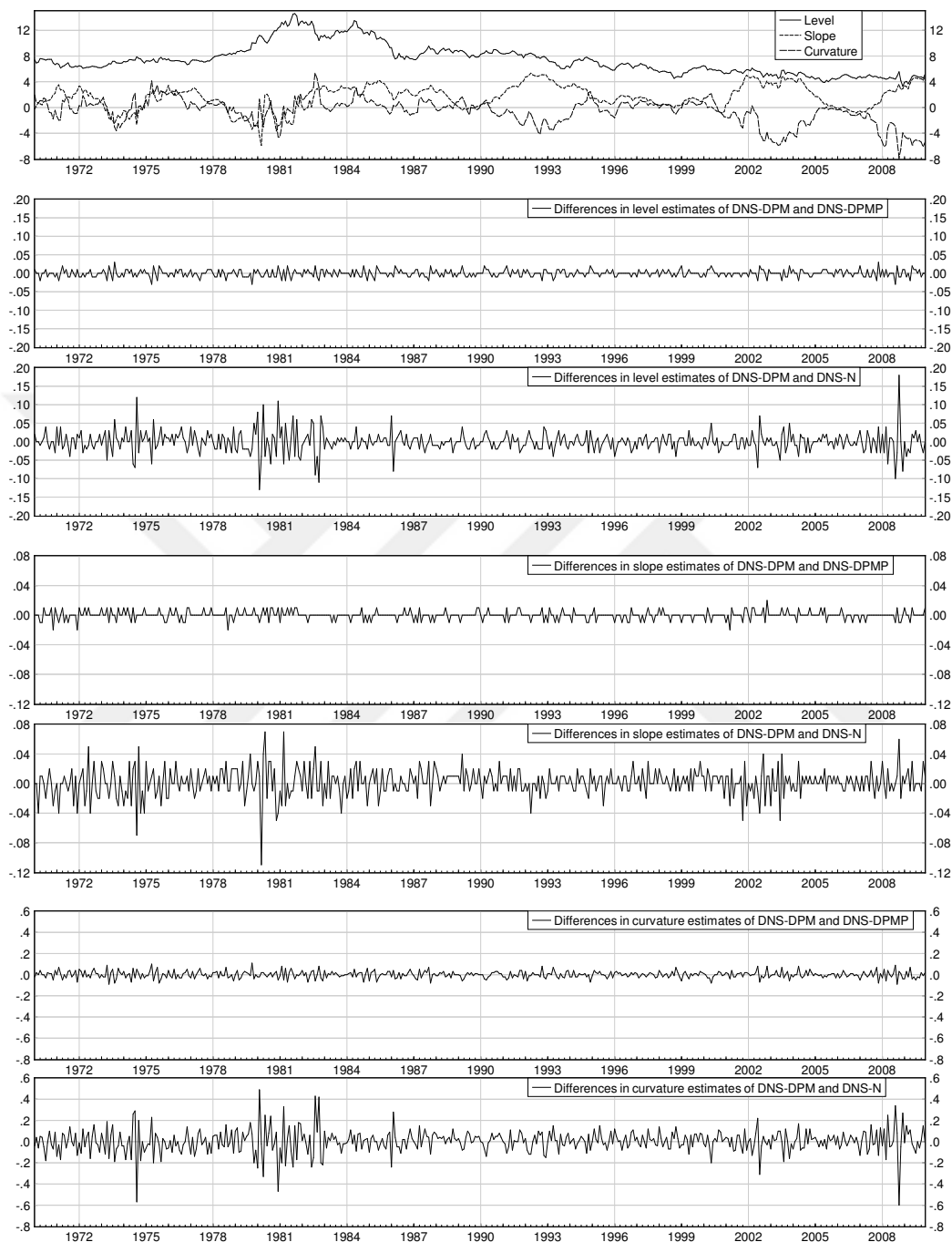
$$q_{0,f} = 2^{\frac{3}{2}} (2\pi)^{-\frac{3}{2}} \frac{\Gamma_3\left(\frac{\nu+1}{2}\right)}{\Gamma_3\left(\frac{\nu}{2}\right)} \frac{|V|^{\frac{\nu}{2}}}{|V + (\varepsilon_t - \mu)(\varepsilon_t - \mu)'|^{\frac{\nu+1}{2}}}, \quad (5.29)$$

and the state equation in (5.3) replaced by

$$\begin{pmatrix} f_{1,t} \\ f_{2,t} \\ f_{3,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix} \begin{pmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{3,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix}. \quad (5.30)$$

The estimates of the factors using the DNS-DPM model and the differences of these estimates from those of the DNS-DPMP and DNS-N models are displayed in Figure 5.3. The first graph in the figure shows the factor estimates using DNS-DPM model. The factor estimates follow a similar path with the factors constructed using the linear combinations of the yields in Figure 5.1. However, the factors estimated using DNS-DPM model are much smoother than the linear combinations as expected. The differences of the estimates from the competing models reveal an interesting pattern. While the differences between the factor estimates of the DNS-DPM model and the DNS-DPMP model are very limited the same does not hold for the DNS-N model, most notably, for the periods with the high volatility in the preliminary analysis. Nevertheless, except these periods the point estimates of the factors are close across different models.

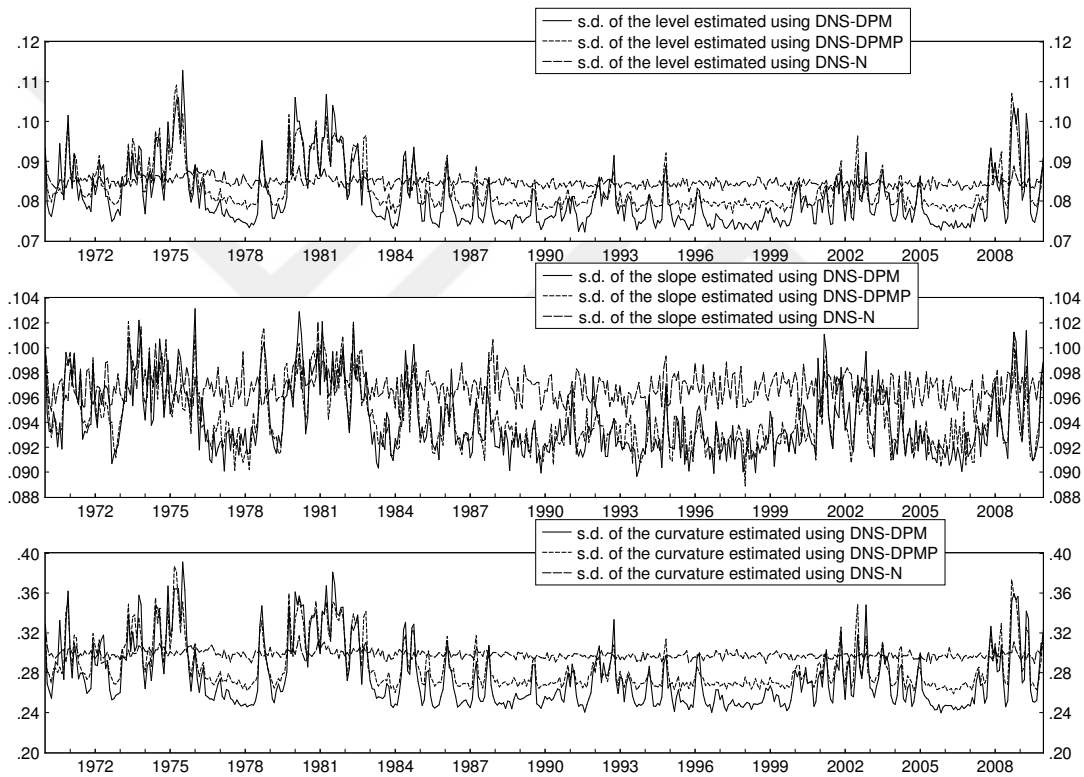
Figure 5.3: Posterior means of the factors for DNS-DPM model and differences with those for DNS-DPMP and DNS-N models



*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM(P) is the model with DPM innovations where both the mean and the variance of the innovations (only the variance of the innovations) have a DP prior. In the first graph, the solid, the dotted, and the dashed lines represent posterior means of the level, slope, and curvature factors, respectively, estimated using the DNS-DPM model. In the remaining graphs the solid lines correspond to the differences of factor estimates (posterior means) of the competing models.

In Figure 5.4 we display the course of the posterior standard deviations of the factors through the sample period. Interestingly, there is a unique pattern for all three factors in the sense that posterior standard deviations of the factors are much lower for the models that use the DPM structure compared to the standard model. Moreover, for the level and curvature factor, the precision is even higher for the DNS-DPM model compared to the DNS-DPMP model. This shows the increasing precision of the factor estimates when we model the distribution of the factors semiparametrically using the DPM.

Figure 5.4: Posterior standard deviations of the factors for DNS-DPM, DNS-DPMP and DNS-N models



*Note:* The solid, the dotted, and the dashed lines represent posterior standard deviations of the factors estimated using DNS-DPM, DNS-DPMP, and DNS-N models, respectively. The first graph displays the results for the level factors, whereas the second and the third graph display the results for the slope and curvature factors, respectively. See Figure 5.3 for further details.

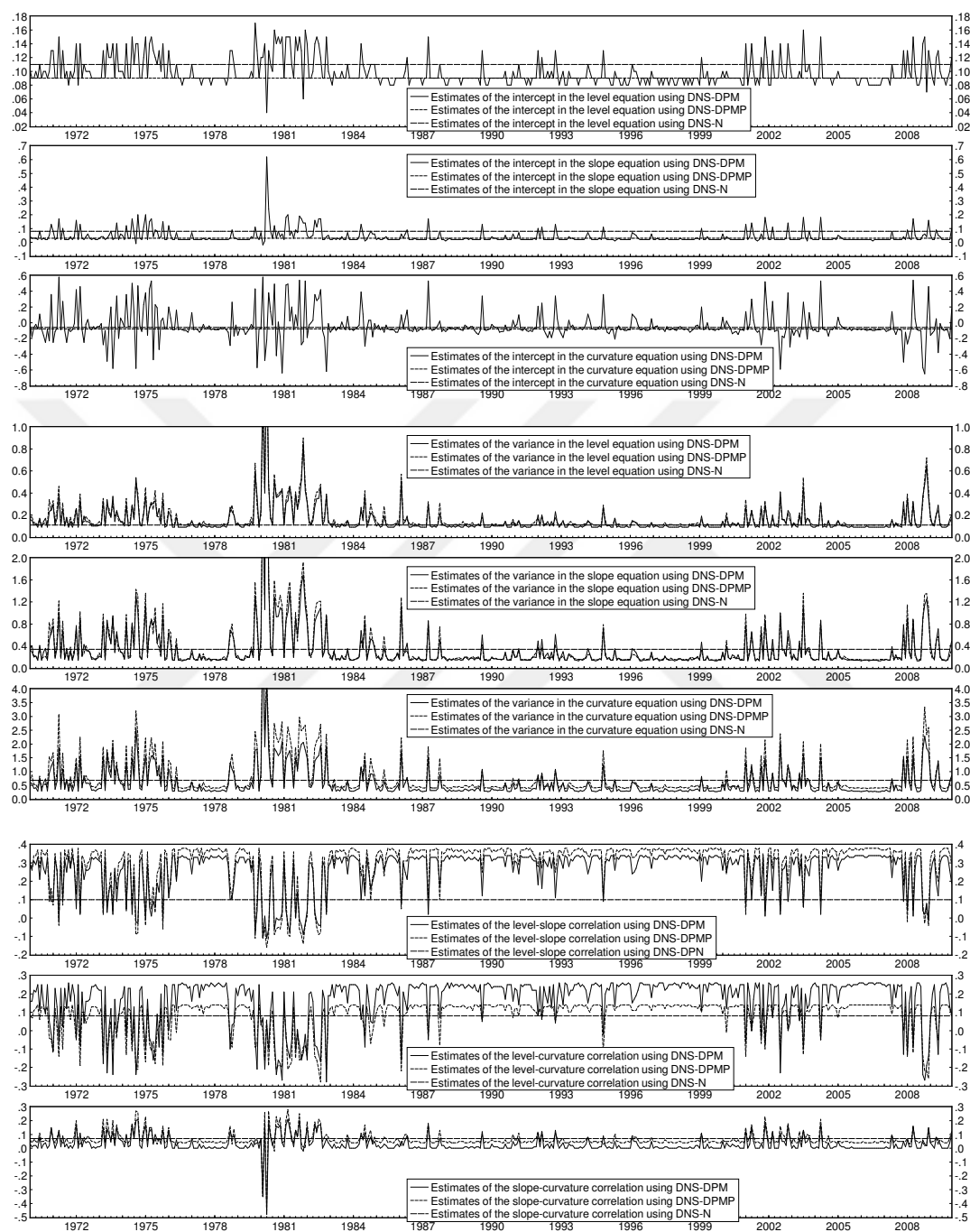
The parameter estimates of the competing models are shown in Table 5.2 and the evolution of the parameters that follow the DPM and comparison of those with the estimates of the competing models are displayed in Figure 5.5.

Table 5.2: Posterior results for the DNS-N, DNS-DPMP and DNS-DPM models

|                   | DNS-N            |                   |                   | DNS-DPMP         |                   |                   | DNS-DPM          |                  |                  |
|-------------------|------------------|-------------------|-------------------|------------------|-------------------|-------------------|------------------|------------------|------------------|
|                   | Level            | Slope             | Curvature         | Level            | Slope             | Curvature         | Level            | Slope            | Curvature        |
| $\mu_f$           | 0.106<br>(0.031) | -0.075<br>(0.034) | -0.062<br>(0.039) | 0.093<br>(0.026) | -0.026<br>(0.025) | -0.084<br>(0.035) |                  |                  |                  |
| $\phi_f$          | 0.985<br>(0.004) | 0.959<br>(0.013)  | 0.919<br>(0.020)  | 0.987<br>(0.003) | 0.979<br>(0.008)  | 0.939<br>(0.019)  | 0.986<br>(0.003) | 0.978<br>(0.008) | 0.947<br>(0.018) |
| $\sigma_f^2$      | 0.114<br>(0.008) | 0.337<br>(0.023)  | 0.684<br>(0.057)  |                  |                   |                   |                  |                  |                  |
| $\rho_{1j}$       | 1                | 0.097<br>(0.049)  | 0.079<br>(0.056)  |                  |                   |                   |                  |                  |                  |
| $\rho_{2j}$       | 0.097<br>(0.049) | 1                 | 0.069<br>(0.052)  |                  |                   |                   |                  |                  |                  |
| $\rho_{3j}$       | 0.079<br>(0.056) | 0.069<br>(0.052)  | 1                 |                  |                   |                   |                  |                  |                  |
| $\alpha$          |                  |                   |                   |                  | 0.92<br>(0.50)    |                   |                  | 0.94<br>(0.47)   |                  |
| # of clusters     |                  |                   |                   |                  | 5.77<br>(1.99)    |                   |                  | 5.84<br>(1.65)   |                  |
| I-S pred. density |                  | 9189              |                   |                  | 9515              |                   |                  | 9538             |                  |

*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM(P) is the model with DPM innovations where both the mean and the variance of the innovations (only the variance of the innovations) have a DP prior.  $\mu_f$  for  $f = 1, 2, 3$  is the unconditional mean of the factor errors (intercept of the factors) and  $\phi_f$  is the first-order autoregressive coefficients of the factors, i.e diagonal elements of the autoregressive coefficient matrix.  $\sigma_f^2$  is the variance of the factor error terms, whereas  $\rho_{i,j}$  denotes the correlation of the error term of the  $i^{th}$  factor for  $i = 1, 2, 3$  with the  $j^{th}$  factor error term  $j \in \{1, 2, 3, \setminus i\}$ .  $\alpha$  is the precision parameter of the DP. The values indicate the posterior means of the parameters with the posterior standard deviations in parenthesis. I-S pred. density is the abbreviation for the In-Sample predictive density, which is computed using (5.25). For the estimation of the models we use the sample period from January 1970 through December 2009. 60,000 simulations from the posterior distribution are used for inference. The first 10,000 draws are discarded as burn-in sample and the remaining 50,000 draws, where we kept every 10<sup>th</sup> draw, are used for posterior inference in the models. The convergence of the sampler is checked using statistical and visual inspection and in all the model specifications convergence is assured.

Figure 5.5: Posterior means of the parameters for DNS-DPM, DNS-DPMP and DNS-N models



*Note:* The solid lines correspond to the estimates of the time varying parameters due to the DPM structure which are not displayed in Table 5.2. The dotted lines represent the corresponding estimates using VAR(1)-N model displayed also in Table 5.2. See Table 5.2 for further details.

First, in line with the findings from the preliminary analysis, the estimates of the autoregressive parameters are increasing when the DPM structure is used. Despite the increasing persistence, or in other words, despite the higher values of the estimates of the autoregressive parameters, the posterior standard deviations of these parameters are much lower than those in the standard model. This leads us to the conclusion that the autoregressive parameters can be estimated much precisely using the model with DPM structure.

Second, the same also applies to the estimates of the intercept parameters when these are considered for the DNS-N and DNS-DPMP model in the sense that posterior standard deviations of the intercept parameters are much lower when DNS-DPMP model is considered. The increase in the persistence of the level and slope factor is accommodated with lower values (in absolute terms) in the intercept estimates leading to a similar mean level for these factors when these two models are considered. For the curvature factor, although the persistence increases, the intercept also increases in absolute value leading to a mean that is higher than (in absolute terms) its counterpart using DNS-N model.

When the parameters that have a semiparametric structure are analyzed, we observe that the estimates of the parameters related to the covariance matrix of the factor errors are very similar for DNS-DPM and DNS-DPMP model. Interestingly, the estimates of the intercept parameters for DNS-DPM model are fluctuating around the value estimated by the DNS-DPMP model, while the estimate of the DNS-N model is greater than that of the DNS-DPMP accommodating these fluctuations. Nevertheless, these fluctuations are limited in magnitude with the only exceptions around 1980 which are the two extreme peaks of the historical interest rates observed during the change of the Federal Reserve's monetary target. It seems that the DPM model categorizes these observations as a separate cluster leading to the interpretation of these observations as outliers. The volatility estimates are in line with the preliminary analysis in the sense that the effect of the great moderation, the change in the Federal Reserve's monetary policy and the effect of the recessions of 1970s, 1980s, and the recent 2007-9 crisis are nicely captured by the changing volatility with both DNS-DPM and DNS-DPMP. The volatility of the curvature factor indicated by the DNS-DPMP model is slightly higher than the volatility indicated by the DNS-DPM model. This is due to the fact that part of the volatility is captured by the variation in the intercept parameter. As in the estimates of the intercept the estimate of the volatility around 1980 is extremely high (the estimate in the figure is truncated to increase the visual quality of the graph) again leading to the interpretation of this observation as an outlier. As a result of this changing volatility, the volatility estimate of the DNS-N model accommodates this with a volatility that is higher than the estimates

of the volatility of the DPM models during the periods of Great Moderation and with a lower volatility during volatile periods.

The estimates of the correlation coefficients show that while the estimates of the DNS-N model are close to zero, the estimates of the DNS-DPMP (DNS-DPMP) model are around 0.3 (0.3) for the correlation between the level and slope factors' error terms and around 0.2 (0.1) for the correlation between the level and curvature factors' error terms. An interesting finding is that these correlations decrease significantly during the volatile periods. This implies that during the turbulent times the correlation between the long term yields (level factor) and short and medium term yields (slope and curvature factors) are decreasing. As the turbulent times mostly correspond to the recessions, the dynamics of the yields with different maturities changes as the short term yields and corresponding spreads are more affected by the changes in the monetary policy to stimulate the economy while longer yields are affected mostly by the inflation expectation in the longer term, see Diebold *et al.* (2006), Ang *et al.* (2008) and Christensen *et al.* (2010) among others. The correlation coefficients of the slope and curvature factors' error terms are closer to zero with limited fluctuations. Again the extreme observations of 1980s enter here as a negative correlation during these periods unlike the other periods. These results show a nonlinear dependence structure between the factors. As in the volatility case, the estimate of the DNS-N model balances between the changing correlations leading to moderate estimates of the correlations.

We display the estimates of the precision parameter  $\alpha$  and the average number of clusters together with the in-sample posterior predictive density (I-S predictive density) in the bottom panel of Table 5.2. The estimates of the DNS-DPMP and DNS-DPM models are very close to each other indicating that indeed the second moment is the main driving force of the DPM structure. In both cases the estimate of  $\alpha$  is close to one resulting in the average number of clusters around six. Unreported results show that the mode of the distribution of the number of clusters with more than 30 observations is two indicating that on average there are two major periods with different characteristics and the remaining clusters capture aberrant observations. The I-S predictive density points a dramatic increase when switching from DNS-N model to DNS-DPMP and a further increase when switching from DNS-DPMP to DNS-DPM, albeit limited. This indicates the necessity of a flexible modeling approach that can capture various different characteristics of the US sovereign bond yields.

As a final measure of model fit we display the posterior predictive p (*ppp*) value of the models for each maturity in Table 5.3. The objective functions are the sample

kurtosis and skewness as we want to measure how the model can capture the non-Gaussian characteristics of the data.

Table 5.3: Posterior predictive p-values ( $ppp$ ) of the models with skewness and kurtosis of the yields as criterion functions

|      | <u>DNS-N</u> |          | <u>DNS-DPMP</u> |          | <u>DNS-DPM</u> |          |
|------|--------------|----------|-----------------|----------|----------------|----------|
|      | Skewness     | Kurtosis | Skewness        | Kurtosis | Skewness       | Kurtosis |
|      | $ppp$        | $ppp$    | $ppp$           | $ppp$    | $ppp$          | $ppp$    |
| M3   | 0.04         | 0.45     | 0.02            | 0.23     | 0.03           | 0.32     |
| M6   | 0.24         | 0.42     | 0.41            | 0.31     | 0.45           | 0.40     |
| M9   | 0.05         | 0.23     | 0.33            | 0.46     | 0.31           | 0.43     |
| M12  | 0.06         | 0.42     | 0.36            | 0.36     | 0.34           | 0.42     |
| M15  | 0.04         | 0.30     | 0.29            | 0.21     | 0.30           | 0.26     |
| M18  | 0.02         | 0.47     | 0.20            | 0.32     | 0.22           | 0.36     |
| M21  | 0.00         | 0.27     | 0.14            | 0.49     | 0.15           | 0.42     |
| M24  | 0.04         | 0.40     | 0.26            | 0.41     | 0.29           | 0.44     |
| M27  | 0.10         | 0.27     | 0.41            | 0.23     | 0.43           | 0.23     |
| M30  | 0.09         | 0.28     | 0.38            | 0.23     | 0.41           | 0.23     |
| M33  | 0.10         | 0.32     | 0.39            | 0.26     | 0.43           | 0.26     |
| M36  | 0.06         | 0.47     | 0.33            | 0.36     | 0.37           | 0.35     |
| M39  | 0.04         | 0.48     | 0.25            | 0.40     | 0.29           | 0.38     |
| M42  | 0.02         | 0.41     | 0.19            | 0.45     | 0.23           | 0.42     |
| M45  | 0.01         | 0.49     | 0.17            | 0.40     | 0.20           | 0.37     |
| M48  | 0.02         | 0.35     | 0.21            | 0.29     | 0.25           | 0.27     |
| M60  | 0.09         | 0.31     | 0.36            | 0.27     | 0.41           | 0.24     |
| M72  | 0.01         | 0.14     | 0.13            | 0.17     | 0.16           | 0.13     |
| M84  | 0.10         | 0.36     | 0.38            | 0.50     | 0.43           | 0.44     |
| M96  | 0.13         | 0.20     | 0.43            | 0.36     | 0.47           | 0.42     |
| M108 | 0.08         | 0.03     | 0.34            | 0.13     | 0.37           | 0.17     |
| M120 | 0.29         | 0.00     | 0.13            | 0.02     | 0.13           | 0.03     |

*Note:*  $M\tau$  indicates the yields with maturity  $\tau$ . The  $ppp$  values are computed using (5.26) with the objective functions of sample skewness and kurtosis. The  $ppp$  values correspond to the significance of the null hypothesis of the criterion functions' equality to that of the data. A  $ppp$  value closer to 0.5 indicates the better conformity of the model to the observed data.  $ppp$  values greater than 0.5 are subtracted from 1 to facilitate comparison. See Table 5.2 for further details.

The results for DNS-N model show that while the state-space structure can still capture the kurtosis of the data for various maturities to some extent it cannot do so for the skewness of the data as many of the  $ppp$  values are close to 0. On the contrary, the semiparametric models perform quite well in capturing both the kurtosis and skewness. Additionally, the DNS-DPM model can capture the skewness of the data better than the DNS-DPMP although the increase in  $ppp$  values are limited. This is due to the fact the DNS-DPM specification models the intercepts semiparametrically allowing for skewed distributions.

### 5.4.1 Sensitivity check of DP hyperparameters and of the model structure

The estimation of our Bayesian semiparametric dynamic Nelson-Siegel model involves the estimation of the precision parameter  $\alpha$  but does not involve the estimation of the hyperparameters,  $\Lambda$ , of the base prior  $G_0$ . Estimation of the hyperparameters is a straightforward extension of the model given that we can assign conjugate hyperpriors on these parameters. However, as the likelihood function related to the hyperparameters involves the number of clusters as observations the informative content of the likelihood is limited and a careful selection of the hyperprior distribution is crucial. Instead, we prefer to fix the hyperparameters and perform a prior sensitivity analysis, see Jensen and Maheu (2010) for a similar analysis. We, therefore, re-estimate the model using the following six different hyperparameter settings

- Prior 2:  $G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, 10, 5, 5 * 1 * I_3)$
- Prior 3:  $G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, 15, 5, 5 * 1 * I_3)$
- Prior 4:  $G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, 5, 10, 10 * 1 * I_3)$
- Prior 5:  $G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, 5, 15, 15 * 1 * I_3)$
- Prior 6:  $G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, 10, 10, 10 * 1 * I_3)$
- Prior 7:  $G_0(\mu_0, a, \nu, S) = N - IW(\mu_0, 15, 15, 15 * 1 * I_3)$ .

We display posterior results in Table 5.4. Results show that the autoregressive parameters are not so sensitive to the different settings of hyperparameters. The autoregressive parameter of the curvature is marginally sensitive to the choice of the degrees of freedom  $\nu$  of the variance parameter. With the increasing  $\nu$  the autoregressive parameter of the curvature is decreasing, albeit only marginally. This also affects the I-S predictive density to some extent causing a limited decrease. It seems that the choice of  $a$  does not influence the results as from Table 5.4 changes in the value of  $a$  are not as effective as changes in the value of  $\nu$ . A similar pattern can also be observed for the estimates of the precision parameter  $\alpha$  and the number of clusters. While changes in  $a$  do only have a minor effect on these estimates, the effect of the changes in  $\nu$  is relatively higher. Note that an increase in  $\nu$  implies a more informative prior on the variance structure causing to less variation in the clusters, i.e. to fewer number of clusters. Still this effect is limited in the sense that in most extreme cases the number of clusters only decreases from six to four. As a

result the estimates of  $\alpha$  also decrease to a limited extent. In general, we conclude that the results are quite robust to the different parameter settings.

Table 5.4: Posterior results for the DNS-DPM models with different hyperparameter settings

|                   | Prior 2          | Prior 3          | Prior 4          | Prior 5          | Prior 6          | Prior 7          |
|-------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| $\phi_{11}$       | 0.987<br>(0.003) | 0.987<br>(0.003) | 0.985<br>(0.004) | 0.985<br>(0.004) | 0.986<br>(0.003) | 0.986<br>(0.003) |
| $\phi_{22}$       | 0.978<br>(0.008) | 0.977<br>(0.008) | 0.978<br>(0.008) | 0.978<br>(0.008) | 0.978<br>(0.008) | 0.978<br>(0.008) |
| $\phi_{33}$       | 0.947<br>(0.018) | 0.946<br>(0.018) | 0.937<br>(0.018) | 0.932<br>(0.018) | 0.937<br>(0.018) | 0.932<br>(0.019) |
| $\alpha$          | 0.97<br>(0.50)   | 0.98<br>(0.50)   | 0.77<br>(0.46)   | 0.64<br>(0.43)   | 0.75<br>(0.46)   | 0.57<br>(0.41)   |
| # of clusters     | 6.01<br>(1.84)   | 6.05<br>(1.92)   | 4.79<br>(1.75)   | 4.04<br>(1.69)   | 4.69<br>(1.82)   | 3.64<br>(1.61)   |
| I-S pred. density | 9540             | 9532             | 9506             | 9444             | 9497             | 9445             |

*Note:* The values indicate the posterior means of the parameters with the posterior standard deviations in parenthesis. See Table 5.2 for further details.

One of the assumptions we impose on the model is that the autoregressive coefficient matrix is diagonal hence we do not allow for interactions between the factors in the autoregressive coefficient matrix. Christensen *et al.* (2011) notes that a full autoregressive coefficient matrix increase the in-sample fit of the frequentist DNS model. To examine whether our results are mainly driven by the diagonal autoregressive coefficient matrix in our model sofar, we also estimate the DNS-DPM model with a full autoregressive coefficient matrix. The results are displayed in Table 5.5. The results show that indeed the cross autoregressive coefficients are very close to zero. In fact, only for  $\phi_{23}$ , zero is outside the 95% highest posterior density interval with a posterior mean of 0.033. Moreover, inclusion of the cross autoregressive coefficients deteriorates the precision of the diagonal estimates in the sense that standard deviations of the diagonal elements are higher compared to the DNS-DPM model with the off-diagonal elements set to zero. The estimates of the precision parameter  $\alpha$  and the average number of clusters remain similar indicating the DPM dynamics of the model remains unchanged. Still inclusion of the cross autoregressive coefficients increases the I-S predictive density, albeit marginally, in line with Christensen *et al.* (2011).

Table 5.5: Posterior results for the DNS-DPM model with full autoregressive coefficient matrix,  $F$

|                   | $\phi_{i1}$       | $\phi_{i2}$      | $\phi_{i3}$       |
|-------------------|-------------------|------------------|-------------------|
| $\phi_{1j}$       | 0.992<br>(0.006)  | 0.020<br>(0.009) | -0.008<br>(0.009) |
| $\phi_{2j}$       | -0.012<br>(0.008) | 0.960<br>(0.011) | 0.033<br>(0.013)  |
| $\phi_{3j}$       | 0.014<br>(0.014)  | 0.025<br>(0.020) | 0.922<br>(0.022)  |
| $\alpha$          |                   | 0.92<br>(0.47)   |                   |
| # of clusters     |                   | 5.69<br>(1.61)   |                   |
| I-S pred. density |                   | 9556             |                   |

*Note:* The values indicate the posterior means of the parameters with the posterior standard deviations in parenthesis. See Table 5.2 for further details.

## 5.5 Conclusion

In this paper, we put forward a Bayesian semiparametric dynamic Nelson-Siegel model to model various characteristics of the US sovereign bond yields such as skewness, fat tails and nonlinear dependence between the underlying factors. Specifically, we model the distribution of the underlying factors semiparametrically using Dirichlet process mixtures. We provide an algorithm that is easy to implement and that makes use of the full conditional posterior distribution. Consequently, the sensitivity of the algorithm to the choice of the prior distributions is very limited.

In the empirical application we show that our unified methodology of the Bayesian semiparametric dynamic Nelson-Siegel model captures several aspects of the US treasury bond yields data. We locate two different periods with different characteristics of volatility dynamics. Moreover, our results indicate that the correlation between the errors of the factors are usually higher than the correlations indicated by a standard dynamic Nelson-Siegel model. However, this dependence decreases or switches signs during turbulent periods pointing out nonlinear dependence between the factors.

The model and the resulting algorithm proposed in this paper can be used for other applications as well. In fact, the model falls into the class of the dynamic factor models which in turn is a special case of the general state space models. As the Bayesian inference as well as the accompanying algorithm are not dependent on the specific Nelson-Siegel

factor loadings structure, several applications of the state space models can be extended to allow for semiparametric structure in the state equation. Therefore, our methodology provides a unified approach to model different types of non-Gaussian behavior in a wide variety of applications of state space models.



## 5.A Appendices

### 5.A.1 Backward-forward recursion

In this section we describe the algorithm to sample the DPM parameters conditional on the data (bond yields) rather than factors. The algorithm is based on Caron *et al.* (2008) and Doucet and Andrieu (2001). First, we decompose the likelihood as

$$\begin{aligned} p(y_{1:T}|\theta_{1:T}) &= p(y_{t+1:T}|y_{1:t}, \theta_{1:T})p(y_{1:t}|\theta_{1:T}) \\ p(y_{1:T}|\theta_{1:T}) &= p(y_{t+1:T}|y_{1:t}, \theta_{1:T})p(y_t|y_{1:t-1}, \theta_{1:T})p(y_{1:t-1}|\theta_{1:T}) \\ p(y_{1:T}|\theta_{1:T}) &= p(y_{t+1:T}|y_{1:t}, \theta_{1:T})p(y_t|y_{1:t-1}, \theta_{1:t})p(y_{1:t-1}|\theta_{1:t-1}). \end{aligned} \quad (\text{A-1})$$

The last two terms in equation in (A-1) can easily be evaluated using the Kalman Filter. Doucet and Andrieu (2001) decompose the first term further as

$$\begin{aligned} p(y_{t+1:T}|y_{1:t}, \theta_{1:T}) &= \int p(y_{t+1:T}, f_t|y_{1:t}, \theta_{1:T})df_t \\ p(y_{t+1:T}|y_{1:t}, \theta_{1:T}) &= \int p(y_{t+1:T}|f_t, y_{1:t}, \theta_{1:T})p(f_t|y_{1:t}, \theta_{1:T})df_t \\ p(y_{t+1:T}|y_{1:t}, \theta_{1:T}) &= \int p(y_{t+1:T}|f_t, \theta_{1:T})p(f_t|y_{1:t}, \theta_{1:t})df_t \\ p(y_{t+1:T}|y_{1:t}, \theta_{1:T}) &= \int p(y_{t+1:T}|f_t, \theta_{t+1:T})p(f_t|y_{1:t}, \theta_{1:t})df_t. \end{aligned} \quad (\text{A-2})$$

Note that the term  $p(f_t|y_{1:t}, \theta_{1:t})$  can be evaluated using the Kalman filter as well. Doucet and Andrieu (2001) note that

$$\begin{aligned} p(y_{t+1:T}|f_t, \theta_{t+1:T}) &= \int p(y_{t+2:T}|f_t, \theta_{t+1:T})p(y_{t+1}, f_{t+1}|f_t, \theta_{t+1:T})df_{t+1} \\ p(y_{t+1:T}|f_t, \theta_{t+1:T}) &= \int p(y_{t+2:T}|f_t, \theta_{t+1:T})p(y_{t+1}, f_{t+1}|f_t, \theta_{t+1})df_{t+1}. \end{aligned} \quad (\text{A-3})$$

leading to a backward recursion. By plugging (A-2) into (A-1) results in

$$p(y_{1:T}|\theta_{1:T}) = \left( \int p(y_{t+1:T}|f_t, \theta_{t+1:T})p(f_t|y_{1:t}, \theta_{1:t})df_t \right) p(y_t|y_{1:t-1}, \theta_{1:t})p(y_{1:t-1}|\theta_{1:t-1}). \quad (\text{A-4})$$

The first term of (A-4) comes from a backward recursion using (A-3) and it is identical for both the numerator and denominator of the (5.18) as for both likelihood evaluation we condition on the previous draws of  $\theta_{t+1:T}^{(m-1)}$ . The second and third component come from the Kalman filter (forward recursion) steps for both the previous draw and for the candidate draw. The resulting sampler for sampling from  $p(y_{1:T}|\theta_{1:T})$  at iteration  $(m)$  becomes

- **Backward recursion:** Compute  $\int p(y_{t+1:T}|f_t, \theta_{t+1:T})p(f_t|y_{1:t}, \theta_{1:t})df_t$  for  $t = T, \dots, 1$

$$\begin{aligned} \text{Let } B_{t+1}^{(m-1)} * B_{t+1}^{(m-1)} &= Q_{t+1}^{(m-1)}, \text{ i.e. } B \text{ is the upper triangular root of } Q_{t+1} \\ \Delta_{t+1}(\theta_{t+1:T}^{(m-1)}) &= \left[ I_3 - B_{t+1}^{(m-1)} G_{t+1|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) B_{t+1}^{(m-1)} \right]^{-1} \end{aligned} \quad (\text{A-5})$$

then compute and store  $G_{t|t+1}^{-1}(\theta_{t+1:T}^{(m-1)})$  and  $G_{t|t+1}^{-1} a_{t|t+1}(\theta_{t+1:T}^{(m-1)})$  as follows

$$\begin{aligned} - \text{ For } t = T \\ G_{T|T}^{-1}(\theta_{T:T}^{(m-1)}) &= H' R^{-1} H \\ G_{T|T}^{-1} a_{T|T}(\theta_{T:T}^{(m-1)}) &= H' R^{-1} y_T \end{aligned} \quad (\text{A-6})$$

- For  $t = T - 1, \dots, 1$

$$\begin{aligned} G_{t|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) &= F' G_{t+1|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) \\ &\times \left[ I_3 - B_{t+1}^{(m-1)} \Delta_{t+1}(\theta_{t+1:T}^{(m-1)}) B_{t+1}'^{(m-1)} G_{t+1|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) \right] F \\ G_{t|t+1}^{-1} a_{t|t+1}(\theta_{t+1:T}^{(m-1)}) &= F' \left[ I_3 - G_{t+1|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) B_{t+1}^{(m-1)} \Delta_{t+1}(\theta_{t+1:T}^{(m-1)}) B_{t+1}'^{(m-1)} \right] \\ &\times \left[ G_{t+1|t+1}^{-1} a_{t+1|t+1}(\theta_{t+1:T}^{(m-1)}) - G_{t+1|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) \mu_{t+1}^{(m-1)} \right] \\ G_{t|t}^{-1}(\theta_{t:T}^{(m-1)}) &= G_{t|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) + H' R^{-1} H \\ G_{t|t}^{-1} a_{t|t}(\theta_{t:T}^{(m-1)}) &= G_{t|t+1}^{-1} a_{t|t+1}(\theta_{t+1:T}^{(m-1)}) + H' R^{-1} y_t \end{aligned} \quad (\text{A-7})$$

- **Forward recursion:** Compute  $p(y_t|y_{1:t-1}, \theta_{1:t}^{(m)})p(y_{1:t-1}|\theta_{1:t-1}^{(m)})$ , for  $t = 1, \dots, T$

Note that for the observation at  $t$   $p(y_{1:t-1}|\theta_{1:t-1}^{(m)})$  is same for both numerator and denominator of the (5.18), hence we only need to compute  $p(y_t|y_{1:t-1}, \theta_{1:t}^{(m)}) = p(y_t|y_{1:t-1}, \theta_t^{(m)}, \theta_{1:t-1}^{(m)})$  for  $t = 1, \dots, T$ . The remaining part of the algorithm is as follows

- Draw  $\theta_t^{(*)}$  from  $\theta_t^{(*)} \sim p(\theta_t|\theta_{1:t-1}^{(m)}, \theta_{t+1:T}^{(m-1)})$
- Perform Kalman filter and store  $\hat{f}_{t|t}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})$ ,  $\hat{P}_{t|t}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})$ ,  $\hat{y}_{t|t-1}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})$  and  $\hat{\Sigma}_{y,t|t-1}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})$
- Compute the likelihood  $p(y_{1:T}|\theta_{1:t-1}^{(m)}, \theta_t^{(*)}, \theta_{t+1:T}^{(m-1)})$  as follows

Define  $C_{t|t}(\theta_{1:t})$ ,  $\Pi_{t|t}(\theta_{1:t})$  and  $A_{t|t}(\theta_{1:t})$  such that

$$\begin{aligned} P_{t|t}(\theta_{1:t}) &= C_{t|t}(\theta_{1:t}) \Pi_{t|t}(\theta_{1:t}) C_{t|t}'(\theta_{1:t}) \\ A_{t|t}(\theta_{1:t}) &= C_{t|t}(\theta_{1:t}) \left[ \Pi_{t|t}^{-1}(\theta_{1:t}) + C_{t|t}'(\theta_{1:t}) G_{t|t+1}^{-1}(\theta_{t+1:T}^{(m-1)}) C_{t|t}(\theta_{1:t}) \right]^{-1} C_{t|t}'(\theta_{1:t}) \end{aligned}$$

then

$$\ln(p(y_{1:T}|\theta_{1:t-1}^{(m)}, \theta_t^{(*)}, \theta_{t+1:T}^{(m-1)})) \propto$$

$$\begin{aligned}
& -\frac{1}{2}n_y \log(2\pi) - \frac{1}{2} \log(|\hat{\Sigma}_{y,t|t-1}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})|) \\
& - \frac{1}{2}(y_t - \hat{y}_{t|t-1}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)}))' \hat{\Sigma}_{y,t|t-1}^{-1}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})(y_t - \hat{y}_{t|t-1}(\theta_{1:t-1}^{(m)}, \theta_t^{(*)})) \\
& - \frac{1}{2} \log\left(\left|\Pi_{t|t}(\theta_{1:t})C'_{t|t}(\theta_{1:t})G_{t|t+1}^{-1}(\theta_{t+1:T}^{(m-1)})C_{t|t}(\theta_{1:t}) + I_3\right|\right) \\
& - \frac{1}{2}\hat{f}'_{t|t}(\theta_{1:t}^{(m)})G_{t|t+1}^{-1}(\theta_{t+1:T}^{(m-1)})\hat{f}_{t|t}(\theta_{1:t}^{(m)}) \\
& - 2\hat{f}'_{t|t}(\theta_{1:t}^{(m)})G_{t|t+1}^{-1}a_{t|t+1}(\theta_{t+1:T}^{(m-1)}) \\
& - \left(G_{t|t+1}G_{t|t+1}^{-1}a_{t|t+1}(\theta_{t+1:T}^{(m-1)}) - \hat{f}_{t|t}(\theta_{1:t}^{(m)})\right)' G_{t|t+1}^{-1}A_{t|t}(\theta_{1:t})G_{t|t+1}^{-1} \\
& \left(G_{t|t+1}G_{t|t+1}^{-1}a_{t|t+1}(\theta_{t+1:T}^{(m-1)}) - \hat{f}_{t|t}(\theta_{1:t}^{(m)})\right)
\end{aligned} \tag{A-8}$$

– With probability  $\rho$  accept  $\theta_t^{(m)} = \theta_t^*$  where

$$\rho^{(m)} = \min\left(1, \frac{p(y_{1:T}|\theta_t^*, \theta_{-t})}{p(y_{1:T}|\theta_t^{(m-1)}, \theta_{-t})}\right) \tag{A-9}$$

otherwise set  $\theta_t^{(m)} = \theta_t^{(m-1)}$

### 5.A.2 Full conditional posterior distributions

In this section we describe the full conditional posterior distributions of the parameters that are common to both the model with DPM structure and the model with Gaussian innovations. Conditional on the remaining parameters these distributions can be obtained by the product of the prior distributions and the likelihood function described in Section 5.3.1 and 5.3.2.

#### Sampling of autoregressive parameters, $F$

Conditional on the states  $f_{0:T}$  and the parameters  $\theta_{1:T}$  the state equation in (5.3) can be rewritten as Vector AutoRegression (VAR) of first order as follows

$$f_t = \mu_t + Ff_{t-1} + \epsilon_t \text{ and } \epsilon_t \sim N(0, Q_t) \quad (\text{A-1})$$

If we redefine the variables such that

$$\begin{aligned} z_t &= f_t - \mu_t & \text{for } t = 1, \dots, T \\ x_t &= f_{t-1} & \text{for } t = 0, \dots, T-1 \end{aligned} \quad (\text{A-2})$$

we can rewrite (A-1) as

$$z_t = Fx_t + \epsilon_t \text{ and } \epsilon_t \sim N(0, Q_t) \quad (\text{A-3})$$

In compact form the model becomes

$$Z = F * X + U \quad (\text{A-4})$$

where  $Z = (z_1, z_2, \dots, z_T)$ ,  $X = (x_1, x_2, \dots, x_T)$  and  $U = (\epsilon_1, \epsilon_2, \dots, \epsilon_T)$ . Using the fact that  $\text{vec}(AB) = (B' \otimes I)\text{vec}(A)$ , where  $\otimes$  stands for the Kronecker product, we can write the following univariate model

$$\begin{aligned} \text{vec}(Z) &= (X' \otimes I_3) * \text{vec}(F) + \text{vec}(U) \\ z &= (X' \otimes I_3) * \text{vec}(F) + u \\ \text{where Cov}(u) &= \Omega = \text{diag}(Q_1, Q_2, \dots, Q_T) \end{aligned} \quad (\text{A-5})$$

Using the multivariate normal prior  $N(\underline{\mu}_{\text{vec}(F)}, \underline{\Sigma}_{\text{vec}(F)})$  in (5.12) the posterior distribution is Gaussian,  $\text{vec}(F)|z \sim N(\bar{\mu}_{\text{vec}(F)}, \bar{\Sigma}_{\text{vec}(F)})$ , with the following parameters

$$\begin{aligned}\bar{\Sigma}_{\text{vec}(F)} &= \left( \underline{\Sigma}_{\text{vec}(F)}^{-1} + (X' \otimes I_3)' \Omega^{-1} (X' \otimes I_3) \right)^{-1} \\ \bar{\mu}_{\text{vec}(F)} &= \bar{\Sigma}_{\text{vec}(F)} \left( \underline{\Sigma}_{\text{vec}(F)}^{-1} \underline{\mu}_{\text{vec}(F)} + (X' \otimes I_3)' \Omega^{-1} (X' \otimes I_3) \text{vec}(F)_{OLS} \right) \\ \text{where } \text{vec}(F)_{OLS} &= ((X' \otimes I_3)' \Omega^{-1} (X' \otimes I_3))^{-1} (X' \otimes I_3)' \Omega^{-1} z\end{aligned}\tag{A-6}$$

### Sampling of covariance matrix in measurement equation, $R$

First, we rewrite the measurement equation

$$y_t = H f_t + \eta_t \tag{A-7}$$

Since  $R$  is the covariance matrix of  $\eta_t$ , we can sample  $R$  from the inverse-Wishart distribution  $IW(\bar{\nu}_R, \bar{S}_R)$  with the degrees of freedom parameter  $\bar{\nu}_R = \nu_R + T$  and with the shape parameter  $\bar{S}_R = \underline{S}_R + \sum_{t=1}^T (y_t - H f_t)(y_t - H f_t)'$ .

### Sampling factors, $f_t$

Conditional on the remaining model parameters, drawing  $f_{1:T}$  can be implemented using standard Bayesian inference. This constitutes running the Kalman filter first and running a simulation smoother using the filtered values for drawing smoothed states as in Carter and Kohn (1994) and Frühwirth-Schnatter (1994). In the first step, start the recursion for  $t = 1, \dots, T$

$$\begin{aligned}f_{t|t-1} &= \mu_t + F f_{t-1|t-1} \\ P_{t|t-1} &= F P_{t-1|t-1} F' + Q_t \\ \eta_{t|t-1} &= y_t - H f_{t|t-1} \\ \zeta_{t|t-1} &= H P_{t|t-1} H' + R \\ K_t &= P_{t|t-1} H' \zeta_{t|t-1}' \\ f_{t|t} &= f_{t|t-1} + K_t \eta_{t|t-1} \\ P_{t|t} &= P_{t|t-1} - K_t H' \zeta_{t|t-1}',\end{aligned}\tag{A-8}$$

and store  $f_{t|t}$  and  $P_{t|t}$ . For  $f_{0|0}$  and  $P_{0|0}$ , notice that, in the steady state  $E(f) = E(\mu) + F E(f)$  and  $E(P) = F E(P) F' + E(Q)$  leading to  $E(f) = (I_3 - F)^{-1} E(\mu)$  and  $\text{vec}(E(P)) = (I_3 - F \otimes F)^{-1} \text{vec}(Q)$ . Moreover, since the initial state by definition is the state before observing data we can sample  $\mu$  and  $Q$  from the base prior  $G_0$ .

The last filtered state  $f_{T|T}$  and its covariance matrix  $P_{T|T}$  correspond to the smoothed estimates of the mean and the covariance matrix for period  $T$ . Having stored all the filtered values, simulation smoother involves the following backward recursions for  $t = T - 1, \dots, 1$

$$\begin{aligned}
 \eta_{t+1|t}^* &= f_{t+1} - \mu_{t+1} - Ff_{t|t} \\
 \zeta_{t+1|t}^* &= FP_{t|t}F' + Q_{t+1} \\
 f_{t|t, f_{t+1}} &= f_{t|t} + P_{t|t}F'\zeta_{t+1|t}^{*-1}\eta_{t+1|t}^* \\
 P_{t|t, P_{t+1}} &= P_{t|t} - P_{t|t}F'\zeta_{t+1|t}^{*-1}F P_{t|t}
 \end{aligned} \tag{A-9}$$

Intuitively, the simulation smoother updates the states using the same principle as in the Kalman filter with state equation replaced by the measurement equation in the Kalman filter and updates are with respect to all available information using backward recursion. For updating the initial states, using the state equation for the first observation we can write  $f_{0|t, f_1} = F^{-1}(f_1 - \mu_1)$  and  $P_{0|t, P_1} = F^{-1}(P_1 + Q_1)F'^{-1}$ . Given the mean  $f_{t|t, f_{t+1}}$  and the covariance matrix  $P_{t|t, P_{t+1}}$ , the states can be sampled from  $f_t \sim N(f_{t|t, f_{t+1}}, P_{t|t, P_{t+1}})$  for  $t = 0, \dots, T$ .

### 5.A.3 Examples using simulated datasets

In this section, we examine the potential of the Bayesian semiparametric dynamic Nelson-Siegel model using two sets of simulated yield data. Using these simulated yields, we want to analyze, first, whether the proposed estimation algorithm indeed can capture the specific properties of the data we simulate, and second, whether the DPM framework is able to estimate the factor densities and to capture several possible causes of non-Gaussianities such as excess skewness, excess kurtosis and nonlinear dependence structure between the factors. In case of any problems relating to the model or algorithm, displaying the posterior results of the parameters with DPM distribution as well as with other distributions would easily indicate whether the proposed algorithm converges to the true density of the factors and whether the estimated parameters are related to the parameter set we use in data generating processes (DGP).

Following this aim, we generate two samples of datasets and estimate these data, first, using a standard Bayesian dynamic Nelson-Siegel model with Gaussian innovations (DNS-N) and second using a Bayesian semiparametric dynamic Nelson-Siegel model with factors distributed according to the DPM (DNS-DPM). For the latter model, we use  $\mu_0 = (0, 0, 0)'$ ,  $a = 5$ ,  $\nu = 5$  and  $v = 1$  as the hyperparameters of the base prior  $G_0(\theta|\mu_0, a, \nu, v)$ . By setting a low value of  $a$ , the variance of the mean parameter is kept relatively high to capture different values of  $\mu_t$  and to mitigate the effect of our ignorance about the exact value of the mean parameters. We also apply a similar approach for the variance parameter by setting the degrees of freedom parameter to be relatively low. This setting corresponds to a standard normal distribution with relatively high variances of parameters. For both of the simulated data, we use  $N = 22$  and  $T = 480$  to match the real data, i.e. we simulate two datasets of 22 cross sections with each having 480 observations.

#### Example 1

In the first example, we focus on the distributions with excess skewness and on the nonlinear dependence between factors. Mixture distributions can easily create such phenomena and indeed they are widely used in practice, see for example Guidolin and Timmermann (2006), Dai *et al.* (2007) and Ang *et al.* (2008). Therefore, we simulate a dataset with two underlying mixture distributions. For constructing this dataset, we first generate the states using a mixture of two multivariate normal distributions as follows

$$f_t \sim \begin{cases} N(\mu_1 + Ff_{t-1}, \Sigma_1) & \text{with probability } p \\ N(\mu_2 + Ff_{t-1}, \Sigma_2) & \text{with probability } 1 - p. \end{cases} \quad (\text{A-1})$$

The parameter values are  $\mu_1 = (2.00, 1.50, 1.25)'$  and  $\mu_2 = (1.00, -1.50, -1.25)'$ , the covariance matrices are constructed such that the variances  $\sigma_1^2 = (3.00, 3.00, 3.00)'$  and  $\sigma_2^2 = (1.50, 1.50, 1.50)'$  and the correlations  $\rho_1 = (-0.10, 0.00, 0.01)'$  and  $\rho_2 = (0.10, 0.10, -0.01)'$ . The probability of being in the first cluster is set to 0.2. We construct the data according to (5.2) with all variances set to 0.25. The diagonal of  $F$  is set to  $(0.50, 0.50, 0.50)'$  and the off-diagonal elements of  $F$  are set to 0, and  $\lambda = 0.0609$ . The parameter settings are chosen to closely match the properties of the real data. In the first cluster where the frequency of occurrence is lower, the variance parameter is set higher to accommodate highly volatile periods which occur less frequently than calm periods. The correlations between the factors are set to low values as observed in reality but they switch signs between the clusters to create nonlinear dependence structure. The value of  $\lambda$  is fixed to the same value as in Diebold and Li (2006). This value maximizes the curvature loading of the yield with 30 months maturity. However, we set the diagonal of  $F$  deliberately to lower values than actually observed in the real data. In the next example we will set this to higher values to investigate whether increasing persistence of the factors has an influence on parameter estimation.

The resulting simulated data has a sample skewness of the level factor (error) as 0.30 (0.46) and the sample kurtosis of this factor is 0.65 (1.01). The sample skewness of the slope factor (error) is 0.75 (1.06) and the sample kurtosis is 1.77 (1.55). The sample skewness of the curvature factor (error) is 0.91 (1.05) and the sample kurtosis is 2.73 (2.99). The degree of skewness is far from being zero for all of the factors with the slope and curvature factors having highest skewness. The degree of kurtosis is high across all of the factors as the variances in one cluster is twice high compared to the variance in the other cluster.

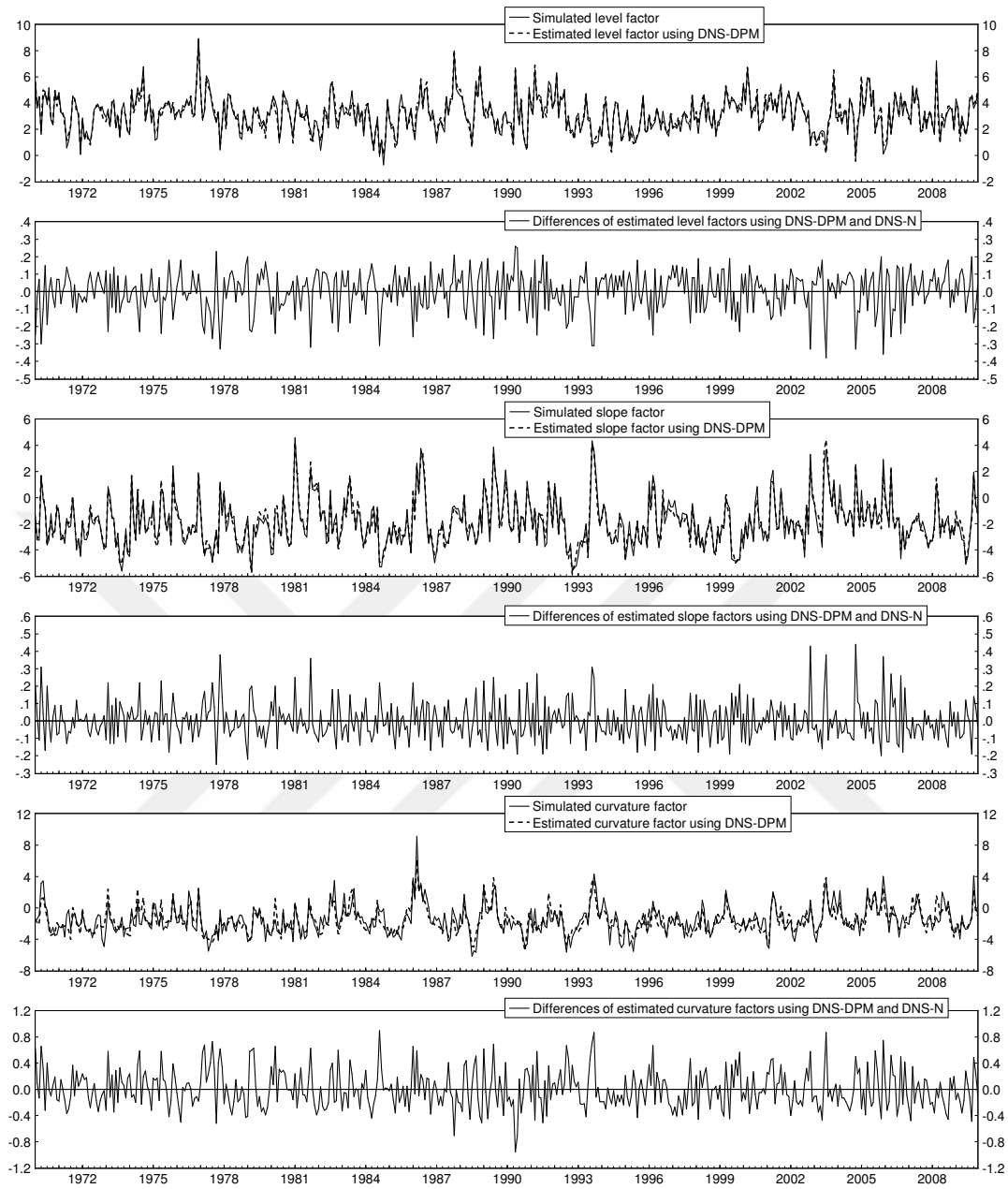
The posterior means of the factors and the differences between these means for both models are displayed in Figure A.1. The estimates of the factors using DNS-DPM model closely match with their actual values. The differences between the factor estimates of the DNS-DPM model and the DNS-N model fluctuates around zero but the distributions of these differences are skewed indicating a specific pattern. For example, the differences of the level factor estimates are negatively skewed with the degree of skewness of -0.65. This indicates that usually the factor estimates using DNS-DPM model are lower than their counterpart. This is due to the fact that while the standard model with Gaussian innovations balances between the two cluster variances and the mean (of the error term), the DNS-DPM model can capture both of the variance and mean in both clusters and thus produces lower estimates (the probability of being in the state with low variance is 0.8). On the other hand, the distribution of the differences for both slope and curvature factors are

positively skewed with skewness values of 0.88 and 0.48 implying greater factor estimates for the DNS-DPM model compared to the estimates using DNS-N model. This is due to the fact that both models can capture the cluster with negative mean error but DNS-DPM can also capture the cluster with positive mean error resulting in greater estimates for these states. In general, the DNS-DPM model can accommodate the skewness in the state distributions due to the alternating signs of mean error and changes in the variances quite well.

The parameter estimates for both models are displayed in Table A.1. The estimates of the DPM parameters for each observation are displayed in Figure A.2 together with actual values that are used in simulation and with the estimates for the DNS-N model. The results in Table A.1 indicate that the parameter estimates of the DNS-N model are usually in between the actual parameter values of the two clusters. The error mean estimates are closer to the second cluster's value especially for the slope and curvature factor. For the variance parameters the results are opposite in the sense that the variance of the slope and curvature factors are closer to 3, the actual value of the variance of the first cluster. Another reason for this high variance is due to the flexible structure of the state space models. As mentioned before, the model accommodates to the data with a higher variance in order to mitigate the effect of alternating signs of error means and the effect of changing variances.

In line with the previous findings, Figure A.2 shows that the DNS-DPM model can capture the observations with different clusters quite accurately for all the parameters. The graph of the estimated parameters closely resembles to the graph of actual parameter values, hence the DNS-DPM model can accommodate the changes in the mean error as well as variances and correlations quite accurately and timely. Note that, state space models are already flexible in the sense that point estimates of the factors are not extremely different from each other, however, as Figure A.2 shows DNS-DPM model provides a very accurate estimate of the distribution of the factors by taking the changes in all of the parameters into account. This fact translates into higher values of the model fit as the average in-sample predictive density is much higher than the DNS-N model with more than 300 points difference indicating almost 15 points difference per maturity class.

Figure A.1: Simulated factors and posterior means of the factors estimated using DNS-DPM and differences with those estimated using DNS-N in Example 1



*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM is the model with DPM innovations where both the mean and the variance of the innovations have a DP prior. The data is generated using (A-1) with the parameter values provided in the text.

Table A.1: Posterior results of the DNS models in Example 1

|                   | Level            | DNS-N             |                   | Level            | DNS-DPM          |                  |
|-------------------|------------------|-------------------|-------------------|------------------|------------------|------------------|
|                   |                  | Slope             | Curvature         |                  | Slope            | Curvature        |
| $\mu_j$           | 1.640<br>(0.148) | -0.933<br>(0.105) | -0.811<br>(0.126) |                  |                  |                  |
| $\phi_j$          | 0.473<br>(0.044) | 0.492<br>(0.041)  | 0.485<br>(0.059)  | 0.494<br>(0.040) | 0.523<br>(0.036) | 0.514<br>(0.055) |
| $\sigma_j^2$      | 1.552<br>(0.116) | 2.471<br>(0.179)  | 2.841<br>(0.391)  |                  |                  |                  |
| $\rho_{1j}$       | 1                | 0.064<br>(0.052)  | 0.070<br>(0.071)  |                  |                  |                  |
| $\rho_{2j}$       | 0.064<br>(0.052) | 1                 | 0.562<br>(0.053)  |                  |                  |                  |
| $\rho_{3j}$       | 0.070<br>(0.073) | 0.562<br>(0.053)  | 1                 |                  |                  |                  |
| $\alpha$          |                  |                   |                   |                  | 0.27<br>(0.34)   |                  |
| # of clusters     |                  |                   |                   |                  | 2.96<br>(1.21)   |                  |
| I-S pred. density |                  | -9571.4           |                   |                  | -9332.4          |                  |

*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM is the model with DPM innovations where both the mean and the variance of the innovations have a DP prior.  $\mu_f$  for  $f = 1, 2, 3$  is the unconditional mean of the factor errors (intercept of the factors) and  $\phi_f$  is the first-order autoregressive coefficients of the factors, i.e diagonal elements of the autoregressive coefficient matrix.  $\sigma_f^2$  is the variance of the factor error terms, whereas  $\rho_{i,j}$  denotes the correlation of the error term of the  $i^{th}$  factor for  $i = 1, 2, 3$  with the  $j^{th}$  factor error term  $j \in \{1, 2, 3, \setminus i\}$ .  $\alpha$  is the precision parameter of the DP. The values indicate the posterior means of the parameters with the posterior standard deviations in parenthesis. I-S pred. density is the abbreviation for the In-Sample predictive density, which is computed using (5.25). The data is generated using (A-1) with the parameter values provided in the text. 60,000 simulations from the posterior distribution are used for inference. The first 10,000 draws are discarded as burn-in sample and the remaining 50,000 draws, where we kept every  $10^{th}$  draw, are used for posterior inference in the models. The convergence of the sampler is checked using statistical and visual inspection and in all the model specifications convergence is assured.

The estimation of the model parameters with low posterior standard deviations results also in more accurate estimation of the parameters with distributions other than the DPM. This can be seen for the autoregressive parameters in Table A.1. The estimates of these parameters are closer to the actual values and posterior standard deviations of the estimates are around 10% lower for the DNS-DPM model. One interesting result is the estimates of the correlation between the error terms of the slope and curvature factors. While the actual values are very close to 0 with 0.01 and -0.01, the estimate of this parameter is 0.56 for the DNS-N model and for the DNS-DPM model on average

0.12 in one cluster and between 0.25 and 0.35 in the other cluster. Indeed, the sample correlation of the simulated errors has the correlation -0.01 for one cluster and 0.31 for the other. Although the estimate for the first cluster is not so accurate, its difference with the actual value is much lower compared to the DNS-N model.

### Example 2

In the second example we focus on the effects of fat-tailed distributions and potential outliers. Following this aim, we simulate a yield dataset where the distributions of the factors have fat tails. For constructing this dataset, we first generate the factors using the following multivariate  $t$  distribution

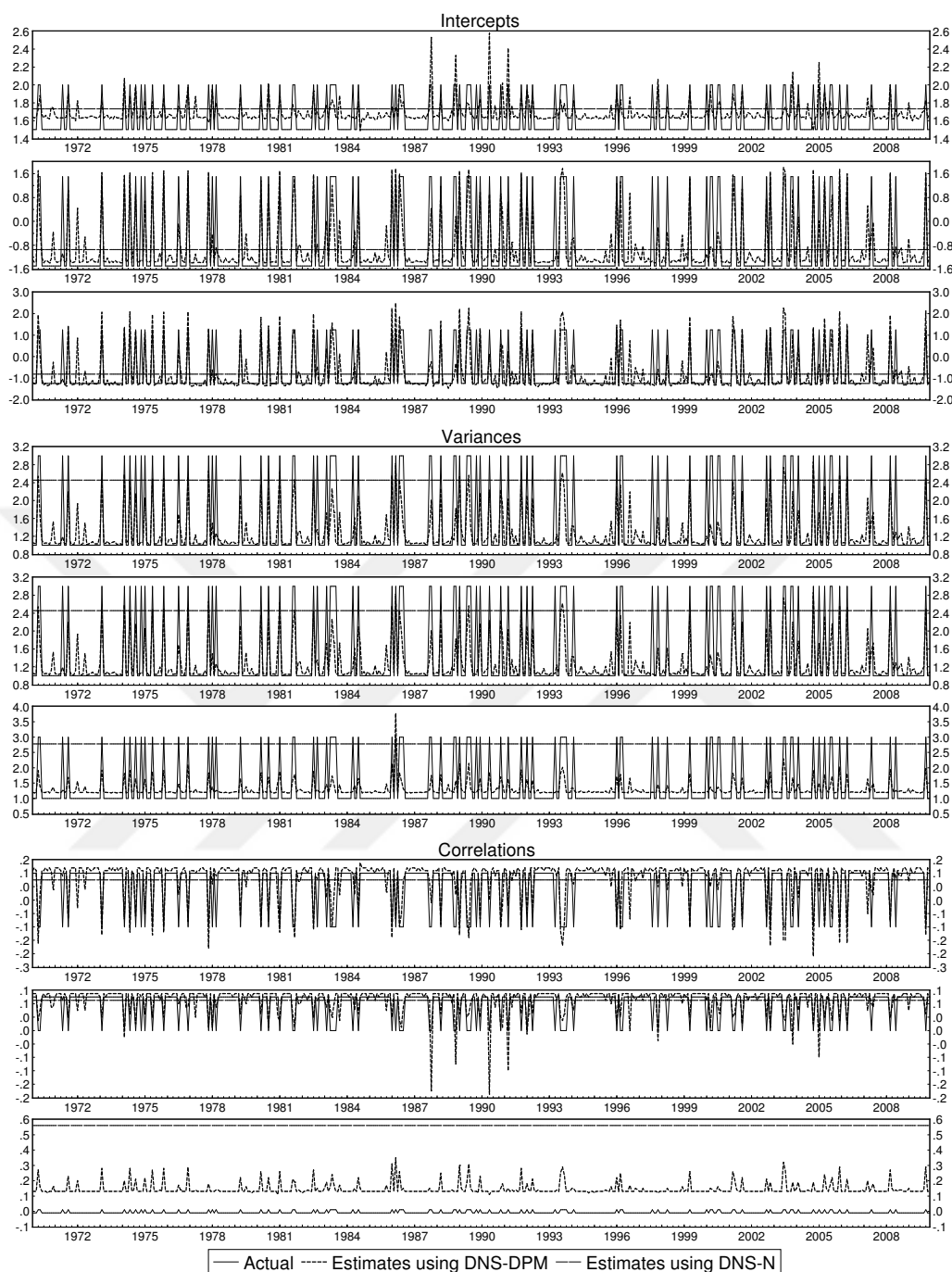
$$f_t \sim t(\mu + F f_{t-1}, \Sigma, \psi), \quad (\text{A-2})$$

where  $t(\mu + F f_{t-1}, \Sigma, \psi)$  indicates a multivariate  $t$  distribution with  $\psi$  degrees of freedom, with mean  $\mu + F f_{t-1}$  and with variance  $\frac{\psi}{\psi-2} \Sigma$ .

The parameter values for the intercepts are  $\mu = (0.50, -0.50, 0.05)'$ . The covariance matrices are constructed such that the variances are  $\sigma^2 = (2.00, 2.00, 2.00)'$  and the correlations are  $\rho = (-0.20, 0.20, 0.10)'$ . The values are similar to the expectation of their counterparts in the first example except that the correlations are slightly higher than those in the first example. We construct the data according to (5.2) with all variances set to 0.2 except for the variances of the first and the last maturities' yield which are set to 0.6 and the variance of the second and the last second maturities' yield which are set to 0.4. This is to mimic the features of the real data as usually the variation in the short and the long end of the yield curve is higher than the variation in medium term yields. The diagonal of  $F$  is set to  $(0.95, 0.95, 0.90)'$  and  $\lambda = 0.0609$ . The high values of autocorrelation coefficients of the factors enable us to investigate the effects of the persistence in the factors on the parameter estimates. We set the degrees of freedom parameter  $\psi$  equal to 6. The resulting sample skewness of the level factor (error) is -0.16 (0.04) and the kurtosis is -0.07 (5.00). The sample skewness of the slope factor (error) is 0.04 (0.37) and the sample kurtosis is -0.58 (1.95). The sample skewness of the curvature factor (error) is -0.18 (-0.15) and the sample kurtosis is -0.24 (4.10). The error distributions are not so skewed but the kurtosis values are quite high as a result of the  $t$  distribution with a low degrees of freedom of 6.

The posterior means of the factors and the differences between the factor estimates (posterior means) for both models are displayed in Figure A.3. The factors estimated using the DNS-DPM model are closely tracking the actual factors. This is especially

Figure A.2: Simulated parameter values and posterior means of these parameters for DNS-DPM and DNS-N in Example 1

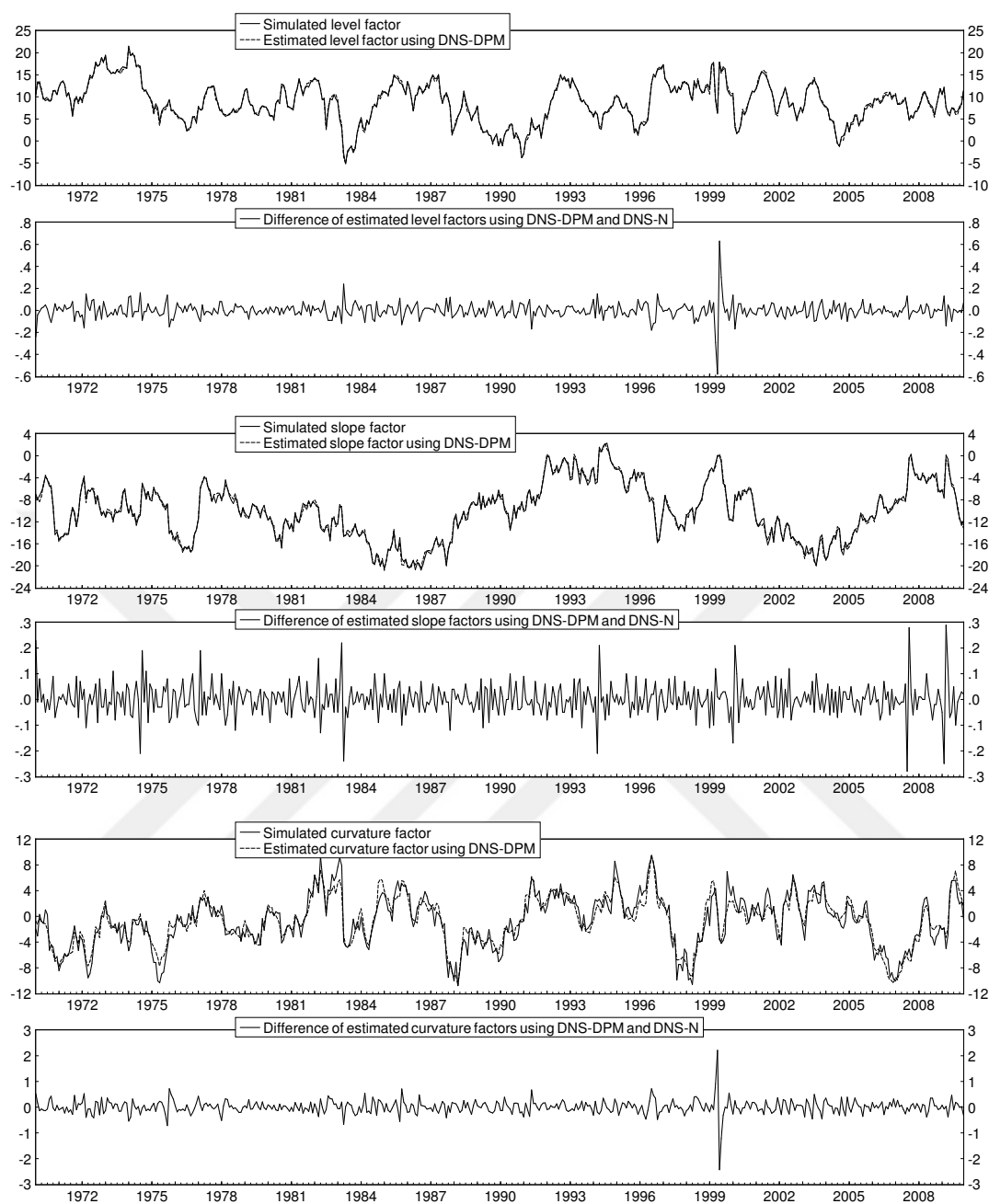


*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM is the model with DPM innovations where both the mean and the variance of the innovations have a DP prior. The solid lines correspond to the estimates of the time varying parameters due to the DPM structure which are not displayed in Table A.1. The dotted lines represent the corresponding estimates using VAR(1)-N model displayed also in Table A.1. The data is generated using (A-1) with the parameter values provided in the text. 60,000 simulations from the posterior distribution are used for inference. The first 10,000 draws are discarded as burn-in sample and the remaining 50,000 draws, where we kept every 10<sup>th</sup> draw, are used for posterior inference in the models. The convergence of the sampler is checked using statistical and visual inspection and in all the model specifications convergence is assured.

the case for the level and slope factors and to a more limited extent for the curvature factor. The differences between the factors estimated using DNS-DPM and DNS-N model reveal the effect of the  $t$  distribution used for simulating the factors. The differences are fluctuating around zero without any specific pattern contrasting the first example. On the contrary, the effect of this fat-tailed distribution shows itself in some occasional large differences. Results show that the DNS-DPM model can capture the extreme factor realizations caused by the  $t$  distribution better than the DNS-N model, which assumes a Gaussian distribution for the factors, and thus, cannot capture these extremes.

The parameter estimates with distributions other than the DPM for both models are displayed in Table A.2. The parameter estimates of the DPM parameters for each observation are displayed in Figure A.4 together with the actual values that are used in simulation and together with the estimates for the DNS-N model. Results of the intercept estimates in Table A.2 and in Figure A.4 show that while the estimates of the slope factor for the DNS-N model is quite close to the actual value it overestimates the intercept of the level factor and underestimates the intercept of the curvature factor. This is also similar for the DNS-DPM model while there is some variation of the estimates around the actual value. A specific pattern for the variance parameters can be observed when the corresponding estimates are analyzed. Except for the slope factor, the variance of the level and curvature factors are overestimated by the DNS-N model as the extreme realizations of the  $t$  distribution is reflected as a higher variance of the Gaussian distribution. On the contrary, for the DNS-DPM model part of the variance is captured by the variation of the intercept estimates. This causes the variance parameter to be underestimated though the effect of the extreme observations are nicely reflected by the spikes in parameter estimates. A similar pattern can also be seen in the correlation estimates while the effects of the extreme observations are reflected to the correlation estimates of the DNS-DPM model as well.

Figure A.3: Simulated factors and posterior means of the factors estimated using DNS-DPMP and differences between DNS-N in Example 2



*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM is the model with DPM innovations where both the mean and the variance of the innovations have a DP prior. The data is generated using (A-2) with the parameter values provided in the text.

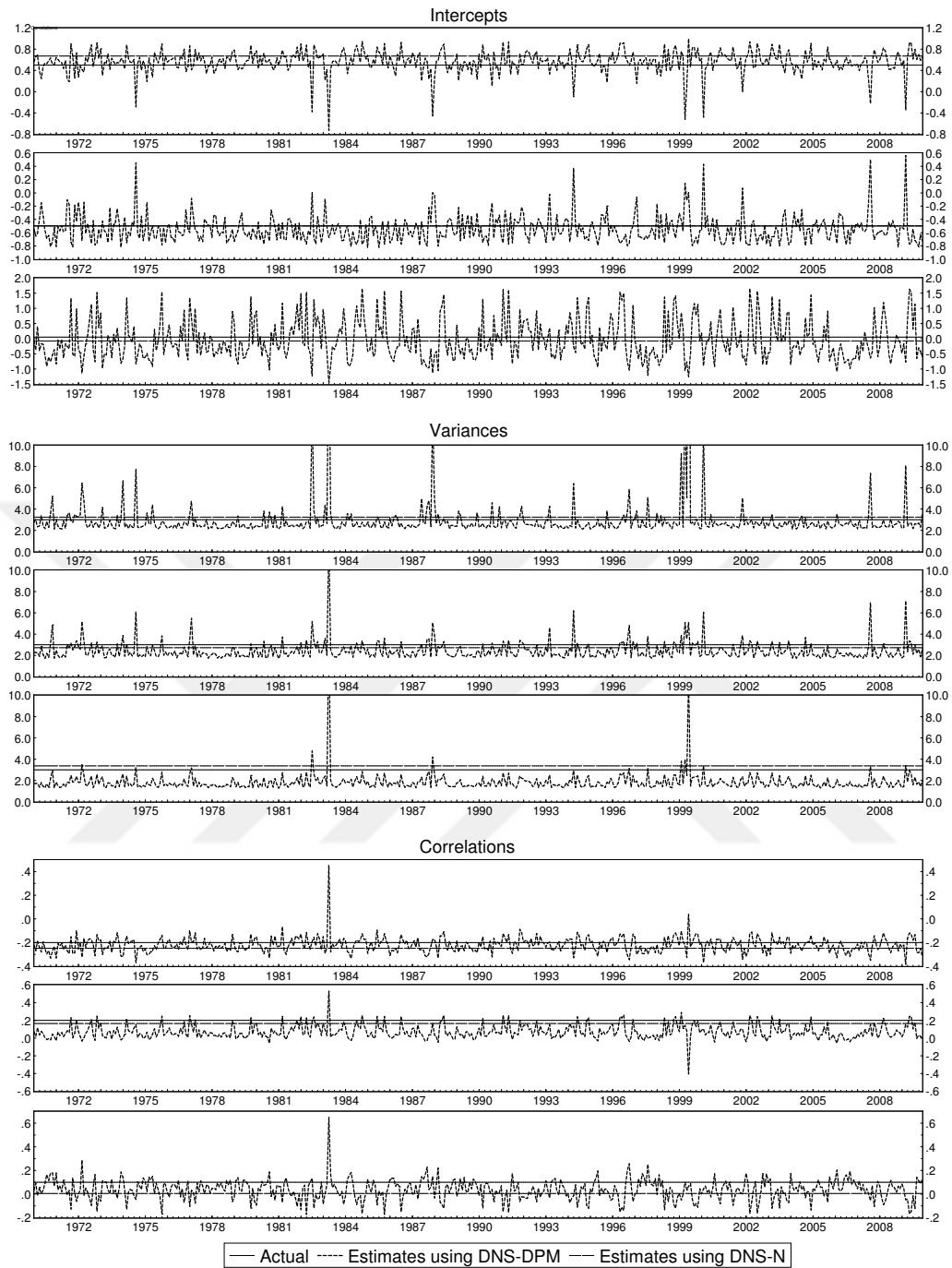
Table A.2: Posterior results of the DNS models in Example 2

|                   | Level             | DNS-N             |                   | Level            | DNS-DPM          |                  |
|-------------------|-------------------|-------------------|-------------------|------------------|------------------|------------------|
|                   |                   | Slope             | Curvature         |                  | Slope            | Curvature        |
| $\mu_j$           | 0.673<br>(0.171)  | -0.496<br>(0.159) | -0.078<br>(0.086) |                  |                  |                  |
| $\phi_j$          | 0.923<br>(0.017)  | 0.951<br>(0.014)  | 0.884<br>(0.025)  | 0.934<br>(0.015) | 0.945<br>(0.013) | 0.896<br>(0.023) |
| $\sigma_j^2$      | 3.239<br>(0.226)  | 2.723<br>(0.206)  | 3.339<br>(0.445)  |                  |                  |                  |
| $\rho_{1j}$       | 1                 | -0.249<br>(0.048) | 0.163<br>(0.070)  |                  |                  |                  |
| $\rho_{2j}$       | -0.249<br>(0.048) | 1                 | 0.005<br>(0.071)  |                  |                  |                  |
| $\rho_{3j}$       | 0.163<br>(0.070)  | 0.005<br>(0.071)  | 1                 |                  |                  |                  |
| $\alpha$          |                   |                   |                   |                  | 1.83<br>(0.62)   |                  |
| # of clusters     |                   |                   |                   |                  | 12.56<br>(4.13)  |                  |
| I-S pred. density |                   | -9504.5           |                   |                  | -9216.4          |                  |

*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM is the model with DPM innovations where both the mean and the variance of the innovations have a DP prior.  $\mu_f$  for  $f = 1, 2, 3$  is the unconditional mean of the factor errors (intercept of the factors) and  $\phi_f$  is the first-order autoregressive coefficients of the factors, i.e diagonal elements of the autoregressive coefficient matrix.  $\sigma_f^2$  is the variance of the factor error terms, whereas  $\rho_{i,j}$  denotes the correlation of the error term of the  $i^{th}$  factor for  $i = 1, 2, 3$  with the  $j^{th}$  factor error term  $j \in \{1, 2, 3, \setminus i\}$ .  $\alpha$  is the precision parameter of the DP. The values indicate the posterior means of the parameters with the posterior standard deviations in parenthesis. I-S pred. density is the abbreviation for the In-Sample predictive density, which is computed using (5.25). The data is generated using (A-2) with the parameter values provided in the text. 60,000 simulations from the posterior distribution are used for inference. The first 10,000 draws are discarded as burn-in sample and the remaining 50,000 draws, where we kept every  $10^{th}$  draw, are used for posterior inference in the models. The convergence of the sampler is checked using statistical and visual inspection and in all the model specifications convergence is assured.

The fact that the extreme realizations of the factors caused by the  $t$  distribution are nicely accommodated by the flexible DNS-DPM model is translated into more precise estimates of the autoregressive coefficients. While in most of the cases the estimates of the DNS-DPM model are much closer to the actual values compared to the DNS-N model, standard deviations of these are also on average 10% lower than that of the DNS-N model. This precision is also reflected to the higher in-sample predictive density of the DNS-DPM model. This is an important finding in the sense that modeling the distributions of the

Figure A.4: Simulated parameter values and posterior means of these parameters estimated using DNS-DPM and DNS-N in Example 2



*Note:* DNS is the dynamic Nelson-Siegel model proposed by Diebold and Li (2006). DNS-N is the DNS model with Gaussian innovations, while DNS-DPM is the model with DPM innovations where both the mean and the variance of the innovations have a DP prior. The solid lines correspond to the estimates of the time varying parameters due to the DPM structure which are not displayed in Table A.2. The dotted lines represent the corresponding estimates using VAR(1)-N model displayed also in Table A.2. The data is generated using (A-2) with the parameter values provided in the text. 60,000 simulations from the posterior distribution are used for inference. The first 10,000 draws are discarded as burn-in sample and the remaining 50,000 draws, where we kept every  $10^{th}$  draw, are used for posterior inference in the models. The convergence of the sampler is checked using statistical and visual inspection and in all the model specifications convergence is assured.

factors accurately results also in more accurate estimates of the evolution of the factors regardless the persistence of the factors.

One important observation from both examples concerns the estimates of the precision parameter  $\alpha$  and the number of clusters. In the first example the estimate of the  $\alpha$  is as low as 0.27 causing a low number of clusters of 2.96. This indicates that the DPM structure in the state equation can nicely capture the mixture structure of the state error distribution. Conversely, in the second example the estimate of the  $\alpha$  is as high as 1.83 and the resulting average number of clusters is 12.56. This can be explained by the fact that the distribution in the second example does not exhibit a mixture structure and the DPM structure in the state equation captures the fat tails and outliers using a high number of clusters. This shows the flexibility of the DPM structure in accommodating various different structures when estimating the dynamic Nelson-Siegel model.



# Nederlandse Samenvatting

## (Summary in Dutch)

Economische agenten hebben In het huidige informatietijdperk toegang tot een enorme hoeveelheid publieke informatie waarop zij hun beslissingen kunnen baseren. Centrale banken vereisen tijdige en accurate voorspellingen van de marktomstandigheden; beleggers diversifiëren hun portefeuilles op basis van hun risico/rendementsverwachtingen; en beleidsmakers hebben een precieze inschatting nodig van de huidige en verwachte toekomstige economische activiteit. Helaas is de informatie die agenten nodig hebben voor het maken van accurate beslissingen niet precies waarneembaar. Daarom moeten relevante statistische maatstaven en voorspellingen afgeleid worden uit de waargenomen macro-economische en financiële variabelen. Begrijpelijkwijz zullen betere schattingen en voorspellingen leiden tot meer accurate beslissingen.

Het centrale thema van deze dissertatie is het modelleren van macro-economische en financiële datasets om de vereiste informatie en niet-waarneembare kenmerken uit deze datasets te kunnen distilleren. Voor dit doel wordt gebruik gemaakt van zogenaamde 'modellen met niet-waarneembare componenten', in het bijzonder twee veelgebruikte varianten van deze modellen namelijk 'mixture' modellen en dynamische factor modellen. Het gemeenschappelijke kenmerk van deze modellen is dat aangenomen wordt dat een dynamisch proces ten grondslag ligt aan niet-waarneembare variabelen, ook wel aangeduid als de toestand. Dit wordt gebruikt om de ontwikkeling van de geobserveerde gegevens, die afhangen van deze verborgen toestand, te beschrijven.

Deze dissertatie ontwikkelt nieuwe methodieken en uitbreidingen op bestaande methodieken voor het extraheren van niet-waarneembare kenmerken uit macro-economische en financiële datasets. De resulterende modellen en schattingsmethodes zijn flexibel genoeg om verschillende processen in deze datasets te kunnen beschrijven.

In hoofdstuk 2 maken we gebruik van dynamische factor modellen om de gemeenschappelijke variatie in een grote macro-economische dataset te kunnen extraheren en om de macro-economische omstandigheden te kunnen beschrijven aan de hand van slechts

enkele belangrijke variabelen, aangeduid als factoren. Deze factoren worden vervolgens gebruikt om de rendementen en volatiliteit in de aandeelmarkt te voorspellen om hiermee winstgevende portefeuilles samen te stellen.

In hoofdstukken 3 en 4 maken we gebruik van 'mixture' modellen om het dynamisch gedrag te beschrijven van macro-economische en financiële variabelen die onderhevig zijn aan een gemeenschappelijke cyclus, waarbij de invloed van deze cyclus op sommige variabelen mogelijk vertraagd is. Met behulp van deze modellen, modelleren we de ontwikkeling van de rendementen van portefeuilles met verschillende karakteristieken. Vergelijkbare modellen worden ook gebruikt, met behulp van leidende indicatoren, bij het modelleren en voorspellen van heterogene economische recessies. Tenslotte, in het laatste hoofdstuk van deze dissertatie, modelleren we, gebruikmakend van de factor modellen en 'mixture' modellen, de ontwikkeling van een grote dataset van obligatie-yields door middel van het schatten van hun dichtheden.

# Bibliography

- Aiolfi, M. and C. A. Favero (2005), Model Uncertainty, Thick Modeling and The Predictability of Stock Returns, *Journal of Forecasting*, 24, 233–254.
- Akgiray, V. (1989), Conditional Heteroscedasticity in Time Series of Stock Returns: Evidence and Forecasts, *Journal of Business*, 62, 55–80.
- Andersen, T. G., T. Bollerslev, P. F. Christoffersen, and F. X. Diebold (2006), Volatility and Correlation Forecasting, in G. Elliot, C. W. J. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Amsterdam: North-Holland, pp. 777–878.
- Andrews, D. W. K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator, *Econometrica*, 60, 953–966.
- Ang, A. and G. Bekaert (2002), International Asset Allocation With Regime Shifts, *Review of Financial Studies*, 15, 1137–1187.
- Ang, A. and G. Bekaert (2007), Stock Return Predictability: Is it There?, *The Review of Financial Studies*, 20, 651–707.
- Ang, A., G. Bekaert, and M. Wei (2008), The Term Structure of Real Rates and Expected Inflation, *Journal of Finance*, 63, 797–849.
- Antoniak, C. E. (1974), Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems, *The Annals of Statistics*, 2, 1152–1174.
- Artis, M., H.-M. Krolzig, and J. Toro (2004), The European Business Cycle, *Oxford Economic Papers*, 56, 1–44.
- Artis, M. J., Z. G. Kontolemis, and D. R. Osborn (1997), Business Cycles for G7 and European Countries, *Journal of Business*, 70, 249–279.
- Avramov, D. (2002), Stock Return Predictability and Model Uncertainty, *Journal of Financial Economics*, 64, 423–458.

- Badrinath, S., J. Kale, and T. Noe (1995), Of Shepherds, Sheep, and the Cross-Autocorrelations in Equity Returns, *Review of Financial Studies*, 8, 401–430.
- Bai, J. (2003), Inferential Theory for Factor Models of Large Dimensions, *Econometrica*, 71, 135–171.
- Bai, J. (2010), Equity Premium Predictions with Adaptive Macro Indexes, Staff Reports 475, Federal Reserve Bank of New York.
- Bai, J. and S. Ng (2008), Forecasting Economic Time Series Using Targeted Predictors, *Journal of Econometrics*, 146, 304–317.
- Bair, E., T. Hastie, D. Paul, and R. Tibshirani (2006), Prediction by Supervised Principal Components, *Journal of the American Statistical Association*, 101, 119–137.
- Barnard, J., R. McCulloch, and X. Meng (2000), Modeling Covariance Matrices In Terms Of Standard Deviations And Correlations, With Application To Shrinkage, *Statistica Sinica*, 10, 1281–1311.
- Barnett, W. A. and M. Chauvet (2011), How Better Monetary Statistics Could Have Signaled the Financial Crisis, *Journal of Econometrics*, 161, 6–23.
- Bekaert, G., C. R. Harvey, and A. Ng (2005), Market Integration and Contagion, *Journal of Business*, 78, 39–70.
- Bernanke, B. (1981), Bankruptcy, Liquidity, and Recession, *American Economic Review*, 71, 155–59.
- Bernanke, B., M. Gertler, and S. Gilchrist (1996), The Financial Accelerator and the Flight to Quality, *The Review of Economics and Statistics*, 78, 1–15.
- Bernanke, B. S., M. Gertler, and S. Gilchrist (1999), The Financial Accelerator in A Quantitative Business Cycle Framework, in J. B. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, vol. 1 of *Handbook of Macroeconomics*, chap. 21, Elsevier, pp. 1341–1393.
- Bianchi, F., H. Mumtaz, and P. Surico (2009), Dynamics of the Term Structure of UK Interest Rates, Bank of England working papers 363, Bank of England.
- Blackwell, D. and J. MacQueen (1973), Ferguson Distributions via Polya Urn Schemes, *The Annals of Statistics*, 1, 353–355.

- Boivin, J. and S. Ng (2006), Are More Data Always Better for Factor Analysis?, *Journal of Econometrics*, 132, 169–194.
- Boldin, M. (1996), A Check on the Robustness of Hamilton’s Markov Switching Model Approach to the Economic Analysis of the Business Cycle, *Studies in Nonlinear Dynamics & Econometrics*, 1, 35–46.
- Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31, 307–327.
- Bordo, M. D. and T. Helbling (2003), Have National Business Cycles Become More Synchronized?, NBER Working Paper No. 10130.
- Bossaerts, P. and P. Hillion (1999), Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?, *Review of Financial Studies*, 12, 405–428.
- Boudoukh, J., M. P. Richardson, and R. F. Whitelaw (1994), A Tale of Three Schools: Insights on Autocorrelations of Short-Horizon Stock Returns, *Review of Financial Studies*, 7, 539–573.
- Breen, W., L. R. Glosten, and R. Jagannathan (1989), Economic Significance of Predictable Variations in Stock Returns, *Journal of Finance*, 44, 1177–1189.
- Brennan, M., N. Jegadeesh, and B. Swaminathan (1993), Investment Analysis and the Adjustment of Stock Prices to Common Information, *Review of Financial Studies*, 6, 799–824.
- Cakmakli, C., R. Paap, and D. J. van Dijk (2011), Modeling and Estimation of Synchronization in Multistate Markov-Switching Models, Tinbergen Institute Discussion Papers 11-002/4, Tinbergen Institute.
- Cakmakli, C. and D. van Dijk (2010), Getting the Most out of Macroeconomic Information for Predicting Stock Returns and Volatility, Tinbergen Institute Discussion Papers 10-115/4, Tinbergen Institute.
- Campbell, J. Y. and R. J. Shiller (1988), The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies*, 1, 195–228.
- Campbell, J. Y. and S. B. Thompson (2008), Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies*, 21, 1509–1531.

- Campbell, S. D. and F. X. Diebold (2009), Stock Returns and Expected Business Conditions: Half a Century of Direct Evidence, *Journal of Business & Economic Statistics*, 27, 266–278.
- Candelon, B., J. Piplack, and S. Straetmans (2008), On Measuring Synchronization of Bulls and Bears: The Case of East Asia, *Journal of Banking and Finance*, 32, 1022–1035.
- Candelon, B., J. Piplack, and S. Straetmans (2009), Multivariate Business Cycle Synchronization in Small Samples, *Oxford Bulletin of Economics and Statistics*, 71, 715–737.
- Canova, F., M. Ciccarelli, and E. Ortega (2007), Similarities and Convergence in G-7 Cycles, *Journal of Monetary Economics*, 54, 850–878.
- Caron, F., M. Davy, A. Doucet, E. Duflos, and P. Vanheeghe (2008), Bayesian Inference for Linear Dynamic Models With Dirichlet Process Mixtures, *IEEE Transactions on Signal Processing*, 56, 71–84.
- Carter, C. K. and R. Kohn (1994), On Gibbs Sampling for State Space Models, *Biometrika*, 81, 541–553.
- Chauvet, M. and S. Potter (2000), Coincident and Leading Indicators of the Stock Market, *Journal of Empirical Finance*, 7, 87–111.
- Chib, S. and B. H. Hamilton (2002), Semiparametric Bayes Analysis of Longitudinal Data Treatment Models, *Journal of Econometrics*, 110, 67–89.
- Chordia, T. and B. Swaminathan (2000), Trading Volume and Cross-Autocorrelations in Stock Returns, *Journal of Finance*, 55, 913–935.
- Christensen, J. H., F. X. Diebold, and G. D. Rudebusch (2011), The Affine Arbitrage-free Class of Nelson-Siegel Term Structure Models, *Journal of Econometrics*, 164, 4–20.
- Christensen, J. H. E., F. X. Diebold, and G. D. Rudebusch (2009), An Arbitrage-free Generalized Nelson–Siegel Term Structure Model, *Econometrics Journal*, 12, C33–C64.
- Christensen, J. H. E., J. A. Lopez, and G. D. Rudebusch (2010), Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields, *Journal of Money, Credit and Banking*, 42, 143–178.
- Claessens, S., M. A. Kose, and M. E. Terrones (2009), What Happens During Recessions, Crunches and Busts?, *Economic Policy*, 24, 653–700.

- Clements, M. P. and H.-M. Krolzig (2003), Business Cycle Asymmetries: Characterization and Testing Based on Markov-Switching Autoregressions, *Journal of Business & Economic Statistics*, 21, 196–211.
- Conley, T. G., C. B. Hansen, R. E. McCulloch, and P. E. Rossi (2008), A Semi-parametric Bayesian Approach to the Instrumental Variable Problem, *Journal of Econometrics*, 144, 276–305.
- Cox, J. C., J. Ingersoll, Jonathan E, and S. A. Ross (1985), A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385–407.
- Cremers, K. J. M. (2002), Stock Return Predictability: A Bayesian Model Selection Perspective, *Review of Financial Studies*, 15, 1223–1249.
- Dai, Q. and K. J. Singleton (2000), Specification Analysis of Affine Term Structure Models, *Journal of Finance*, 55, 1943–1978.
- Dai, Q. and K. J. Singleton (2002), Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics*, 63, 415–441.
- Dai, Q., K. J. Singleton, and W. Yang (2007), Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields, *Review of Financial Studies*, 20, 1669–1706.
- de Pooter, M. (2007), Examining the Nelson-Siegel Class of Term Structure Models, Tinbergen Institute Discussion Papers 07-043/4, Tinbergen Institute.
- de Pooter, M., M. Martens, and D. van Dijk (2008), Predicting the Daily Covariance Matrix for S&P 100 Stocks Using Intraday Data—But Which Frequency to Use?, *Econometric Reviews*, 27, 199–229.
- Del Negro, M. and C. Otrok (2008), Dynamic Factor Models with Time-Varying Parameters: Measuring Changes in International Business Cycles, Staff Reports 326, Federal Reserve Bank of New York.
- Della Corte, P., L. Sarno, and I. Tsiakas (2009), An Economic Evaluation of Empirical Exchange Rate Models, *Review of Financial Studies*, 22, 3491–3530.
- Della Corte, P., L. Sarno, and I. Tsiakas (2010), A Century of Equity Premium Predictability and the Consumption?Wealth Ratio: An International Perspective, *Journal of Empirical Finance*, 17, 313–331.

- Diebold, F. X. and C. Li (2006), Forecasting the Term Structure of Government Bond Yields, *Journal of Econometrics*, 130, 337–364.
- Diebold, F. X., G. D. Rudebusch, and S. Boragan Aruoba (2006), The Macroeconomy and The Yield Curve: a Dynamic Latent Factor Approach, *Journal of Econometrics*, 131, 309–338.
- Doucet, A. and C. Andrieu (2001), Iterative Algorithms for State Estimation of Jump Markov Linear Systems, *IEEE Transactions on Signal Processing*, 49, 1216–1227.
- Duffee, G. R. (2002), Term Premia and Interest Rate Forecasts in Affine Models, *Journal of Finance*, 57, 405–443.
- Duffie, D. and R. Kan (1996), A Yield-Factor Model Of Interest Rates, *Mathematical Finance*, 6, 379–406.
- Durbin, J. and S. J. Koopman (2001), *Time Series Analysis by State Space Methods*, Oxford University Press.
- Edwards, S., J. G. Biscarri, and F. Perez de Gracia (2003), Stock Market Cycles, Financial Liberalization and Volatility, *Journal of International Money and Finance*, 22, 925–955.
- Efron, B., T. Hastie, L. Johnstone, and R. Tibshirani (2004), Least Angle Regression, *Annals of Statistics*, 32, 407–499.
- Engle, R. F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987–1007.
- Escobar, M. D. (1994), Estimating Normal Means With a Dirichlet Process Prior, *Journal of the American Statistical Association*, 89, 268–277.
- Escobar, M. D. and M. West (1995), Bayesian Density Estimation and Inference Using Mixtures, *Journal of the American Statistical Association*, 90, 577–588.
- Fama, E. F. and R. R. Bliss (1987), The Information in Long-Maturity Forward Rates, *American Economic Review*, 77, 680–92.
- Fama, E. F. and K. R. French (1988), Dividend Yields and Expected Stock Returns, *Journal of Financial Economics*, 22, 3–25.
- Ferguson, T. S. (1973), A Bayesian Analysis of Some Nonparametric Problems, *The Annals of Statistics*, 1, 209–230.

- Ferson, W. and C. R. Harvey (1991), The Variation of Economic Risk Premiums, *Journal of Political Economy*, 99, 385–415.
- Fisher, I. (1933), The Debt-Deflation Theory of Great Depressions, *Econometrica*, 333–357.
- Fleming, J., C. Kirby, and B. Ostdiek (2001), The Economic Value of Volatility Timing, *Journal of Finance*, 56, 329–352.
- Fleming, J., C. Kirby, and B. Ostdiek (2003), The Economic Value of Volatility Timing Using ‘Realized’ Volatility, *Journal of Financial Economics*, 67, 473–509.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000), The Generalized Dynamic-Factor Model: Identification And Estimation, *The Review of Economics and Statistics*, 82, 540–554.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003), Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area?, *Journal of Monetary Economics*, 50, 1243–1255.
- French, K. R., G. W. Schwert, and R. F. Stambaugh (1987), Expected Stock Returns and Volatility, *Journal of Financial Economics*, 19, 3–29.
- Friedman, M. (1984), Lessons from the 1979-82 Monetary Policy Experiment, *American Economic Review*, 74, 397–400.
- Frühwirth-Schnatter, S. (1994), Data Augmentation and Dynamic Linear Models, *Journal of Time Series Analysis*, 15, 183–202.
- Frühwirth-Schnatter, S. (2001), Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models, *Journal of the American Statistical Association*, 96, 194–209.
- Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*, New York: Springer.
- Gelfand, A. E. and D. K. Dey (1994), Bayesian Model Choice: Asymptotics and Exact Calculations, *Journal of the Royal Statistical Society, Series B*, 56, 501–514.
- Gelfand, A. E. and S. Mukhopadhyay (1995), On Nonparametric Bayesian Inference for the Distribution of a Random Sample, *The Canadian Journal of Statistics*, 23, 411–420.

- Geman, S. and D. Geman (1984), Stochastic Relaxations, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6, 721–741.
- Gertler, M. and N. Kiyotaki (2010), Financial Intermediation and Credit Policy in Business Cycle Analysis, in B. M. Friedman and M. Woodford (eds.), *Handbook of Monetary Economics*, vol. 3 of *Handbook of Monetary Economics*, chap. 11, Elsevier, pp. 547–599.
- Geweke, J. (2007), Interpretation and Inference in Mixture Models: Simple MCMC works, *Computational Statistics & Data Analysis*, 51, 3529–3550.
- Geweke, J. and G. Amisano (2010), Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns, *International Journal of Forecasting*, 26, 216–230.
- Geweke, J. and G. Amisano (2011), Hierarchical Markov Normal Mixture Models with Applications to Financial Asset Returns, *Journal of Applied Econometrics*, 26, 1–29.
- Geweke, J. F. and K. J. . Singleton (1981), Maximum Likelihood ‘Confirmatory’ Factor Analysis of Economic Time Series, *International Economic Review*, 22, pp. 37–54.
- Gomes, J. F., L. Kogan, and M. Yogo (2009), Durability of Output and Expected Stock Returns, *Journal of Political Economy*, 117, 941–986.
- Gordon, R. J. (1993), The Jobless Recovery: Does It Signal a New Era of Productivity-led Growth?, *Brookings Papers on Economic Activity*, 24, 271–316.
- Griffin, J. E. and M. F. J. Steel (2004), Semiparametric Bayesian Inference for Stochastic Frontier Models, *Journal of Econometrics*, 123, 121–152.
- Groshen, E. L. and S. Potter (2003), Has Structural Change Contributed to a Jobless Recovery?, *Current Issues in Economics and Finance*.
- Guidolin, M. and A. Timmermann (2006), An Econometric Model of Nonlinear Dynamics in the Joint Distribution of Stock and Bond Returns, *Journal of Applied Econometrics*, 21, 1–22.
- Gurkaynak, R. S., B. Sack, and J. H. Wright (2007), The U.S. Treasury Yield Curve: 1961 to the Present, *Journal of Monetary Economics*, 54, 2291–2304.
- Hamilton, J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, 57, 357–84.

- Hamilton, J. D. (2005), What's Real About the Business Cycle?, NBER Working Papers 11161, National Bureau of Economic Research, Inc.
- Hamilton, J. D. and G. Lin (1996), Stock Market Volatility and the Business Cycle, *Journal of Applied Econometrics*, 11, 573–593.
- Hamilton, J. D. and G. Perez-Quiros (1996), What Do the Leading Indicators Lead?, *Journal of Business*, 69, 27–49.
- Harding, D. and A. Pagan (2006), Synchronization of Cycles, *Journal of Econometrics*, 132, 59–79.
- Harvey, A. C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, no. 9780521321969 in Cambridge Books, Cambridge University Press.
- Hastings, W. (1970), Monte Carlo Sampling using Markov Chains and their Applications, *Biometrika*, 57, 97–10.
- Hautsch, N. and Y. Ou (2009), Analyzing Interest Rate Risk: Stochastic Volatility in the Term Structure of Government Bond Yields, CFS Working Paper Series 2009/03, Center for Financial Studies.
- Hautsch, N. and F. Yang (2010), Bayesian inference in a Stochastic Volatility Nelson-Siegel model, *Computational Statistics & Data Analysis*, In Press, Corrected Proof, —.
- Hirano, K. (2002), Semiparametric Bayesian Inference in Autoregressive Panel Data Models, *Econometrica*, 70, 781–799.
- Hou, K. (2007), Industry Information Diffusion and the Lead-Lag Effect in Stock Returns, *Review of Financial Studies*, 20, 1113–1138.
- Huang, H. and T.-H. Lee (2010), To Combine Forecasts or to Combine Information?, *Econometric Reviews*, 29, 534–570.
- Hull, J. and A. White (1990), Pricing Interest-Rate-Derivative Securities, *Review of Financial Studies*, 3, 573–92.
- Jacquier, E., N. G. Polson, and P. E. Rossi (2004), Bayesian Analysis of Stochastic Volatility Models with Fat-tails and Correlated Errors, *Journal of Econometrics*, 122, 185–212.
- Jensen, M. J. (2004), Semiparametric Bayesian Inference of Long-Memory Stochastic Volatility Models, *Journal of Time Series Analysis*, 25, 895–922.

- Jensen, M. J. and J. M. Maheu (2010), Bayesian Semiparametric Stochastic Volatility Modeling, *Journal of Econometrics*, 157, 306–316.
- Jungbacker, B. and S. J. Koopman (2008), Likelihood-based Analysis for Dynamic Factor Models, Tinbergen Institute Discussion Papers 08-007/4, Tinbergen Institute.
- Junker, M., A. Szimayer, and N. Wagner (2006), Nonlinear Term Structure Dependence: Copula Functions, Empirics, and Risk Implications, *Journal of Banking & Finance*, 30, 1171–1199.
- Keim, D. B. and R. F. Stambaugh (1986), Predicting Returns in the Stock and Bond Markets, *Journal of Financial Economics*, 17, 357–390.
- Kim, C.-J., J. Morley, and J. Piger (2005), Nonlinearity and the Permanent Effects of Recessions, *Journal of Applied Econometrics*, 20, 291–309.
- Kim, C.-J. and C. R. Nelson (1999), *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, vol. 1, 1 edn., The MIT Press.
- Kole, E., K. Koedijk, and M. Verbeek (2006), Portfolio Implications of Systemic Crises, *Journal of Banking & Finance*, 30, 2347–2369.
- Koopman, S. J. and J. V. E. Azevedo (2008), Measuring Synchronization and Convergence of Business Cycles for the Euro Area, UK and US, *Oxford Bulletin of Economics and Statistics*, 70, 23–51.
- Koopman, S. J. and A. Harvey (1997), Multivariate Structural Time Series Models, in C. Heij, H. Schumacher, B. Hanzon, and C. Praagman (eds.), *Systematic Dynamics in Economic and Financial Models*, New York: John Wiley and Sons, pp. 269–298.
- Koopman, S. J., M. I. P. Mallee, and M. Van der Wel (2010), Analyzing the Term Structure of Interest Rates Using the Dynamic Nelson–Siegel Model With Time-Varying Parameters, *Journal of Business & Economic Statistics*, 28, 329–343.
- Kose, M. A., C. Otrok, and C. H. Whiteman (2003), International Business Cycles: World, Region, and Country-Specific Factors, *American Economic Review*, 93, 1216–1239.
- Kose, M. A., C. Otrok, and C. H. Whiteman (2008), Understanding the Evolution of World Business Cycles, *Journal of International Economics*, 75, 110–130.
- Krolzig, H.-M. (1997), *Markov-Switching Vector Autoregressions*, New York: Springer.

- Laurent, S., L. Bauwens, and J. V. K. Rombouts (2006), Multivariate GARCH Models: a Survey, *Journal of Applied Econometrics*, 21, 79–109.
- Leitch, G. and J. E. Tanner (1991), Economic Forecast Evaluation: Profits versus the Conventional Error Measures, *American Economic Review*, 81, 580–590.
- Lettau, M. and S. C. Ludvigson (2001), Consumption, Aggregate Wealth, and Expected Stock Returns, *Journal of Finance*, 56, 815–849.
- Lettau, M. and S. Van Nieuwerburgh (2008), Reconciling the Return Predictability Evidence, *Review of Financial Studies*, 21, 1607–1652.
- Lo, A. W. and A. C. MacKinlay (1990a), Contrarian Profits and Stock Market Overreaction, *Review of Financial Studies*, 3, 175–205.
- Lo, A. W. and A. C. MacKinlay (1990b), Data-Snooping Biases in Tests of Financial Asset Pricing Models, *Review of Financial Studies*, 3, 431–467.
- Longin, F. and B. Solnik (1995), Is the Correlation in International Equity Returns Constant: 1960–1990?, *Journal of International Money and Finance*, 14, 3–26.
- Ludvigson, S. C. and S. Ng (2007), The Empirical Risk-Return Relation: A Factor Analysis Approach, *Journal of Financial Economics*, 83, 171–222.
- Ludvigson, S. C. and S. Ng (2009), Macro Factors in Bond Risk Premia, *Review of Financial Studies*, 22, 5027–5067.
- MacEachern, S. N. and P. Müller (1998), Estimating Mixture of Dirichlet Process Models, *Journal of Computational and Graphical Statistics*, 7, 223–238.
- Maheu, J. M. and T. H. McCurdy (2000), Identifying Bull and Bear Markets in Stock Returns, *Journal of Business & Economic Statistics*, 18, 100–112.
- Maheu, J. M., T. H. McCurdy, and Y. Song (2010), Components of Bull and Bear Markets: Bull Corrections and Bear Rallies, Working Paper No. 402, University of Toronto.
- Marquering, W. and M. Verbeek (2004a), The Economic Value of Predicting Stock Index Returns and Volatility, *Journal of Financial and Quantitative Analysis*, 39, 407–429.
- Marquering, W. and M. Verbeek (2004b), A Multivariate Nonparametric Test for Return and Volatility Timing, *Finance Research Letters*, 1, 250–260.

- McConnell, M. and G. Perez-Quiros (2000), Output Fluctuations in the United States: What Has Changed since the Early 1980's?, *American Economic Review*, 90, 1464–1476.
- McCulloch, J. H. (1975), The Tax-Adjusted Yield Curve, *Journal of Finance*, 30, 811–30.
- McQueen, G., M. Pinegar, and S. Thorley (1996), Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns, *Journal of Finance*, 51, 889–919.
- Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953), Equations of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, 21, 1087–1092.
- Mishkin, F. (1978), The Household Balance Sheet and the Great Depression, *Journal of Economic History*, 38, 918–937.
- Mönch, E. (2008), Forecasting the Yield Curve in a Data-Rich Environment: A No-arbitrage Factor-augmented VAR Approach, *Journal of Econometrics*, 146, 26–43.
- Mönch, E., S. Ng, and S. Potter (2009), Dynamic Hierarchical Factor Models, Staff Reports 412, Federal Reserve Bank of New York.
- Morley, J. and J. Piger (forthcoming), The Asymmetric Business Cycle, *The Review of Economics and Statistics*.
- Neal, R. (2000), Markov Chain Sampling Methods for Dirichlet Process Mixture Models, *Journal of computational and graphical statistics*, 249–265.
- Nelson, C. R. and A. F. Siegel (1987), Parsimonious Modeling of Yield Curves, *Journal of Business*, 60, 473–89.
- Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007), Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference, *Journal of Econometrics*, 140, 425–449.
- Orphanides, A. and D. H. Kim (2005), Term Structure Estimation with Survey Data on Interest Rate Forecasts, *Computing in Economics and Finance 2005* 474, Society for Computational Economics.
- Otrok, C. and C. H. Whiteman (1998), Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa, *International Economic Review*, 39, 997–1014.

- Owyang, M., J. Piger, and H. Wall (2005), Business Cycle Phases in U.S. States, *Review of Economics and Statistics*, 87, 604–616.
- Paap, R., R. Segers, and D. van Dijk (2009), Do Leading Indicators Lead Peaks More Than Troughs?, *Journal of Business & Economic Statistics*, 27, 528–543.
- Patton, A. J. (2006), Modelling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47, 527–556.
- Patton, A. J. and D. H. Oh (2011), Modelling Dependence in High Dimensions with Factor Copulas, Manuscript.
- Paye, B. S. and A. Timmermann (2006), Instability of Return Prediction Models, *Journal of Empirical Finance*, 13, 274–315.
- Pelletier, D. (2006), Regime Switching for Dynamic Correlations, *Journal of Econometrics*, 131, 445–473.
- Perez-Quiros, G. and A. Timmermann (2000), Firm Size and Cyclical Variations in Stock Returns, *Journal of Finance*, 55, 1229–1262.
- Perez-Quiros, G. and A. Timmermann (2001), Business Cycle Asymmetries in Stock Returns: Evidence from Higher Order Moments and Conditional Densities, *Journal of Econometrics*, 103, 259–306.
- Pesaran, M. H. and A. Timmermann (1992), A Simple Nonparametric Test of Predictive Performance, *Journal of Business & Economic Statistics*, 10, 561–565.
- Pesaran, M. H. and A. Timmermann (1995), Predictability of Stock Returns: Robustness and Economic Significance, *Journal of Finance*, 50, 1201–1228.
- Pesaran, M. H. and A. Timmermann (2002), Market Timing and Return Prediction under Model Instability, *Journal of Empirical Finance*, 9, 495–510.
- Pesaran, M. H. and A. Timmermann (2004), How Costly is it to Ignore Breaks When Forecasting the Direction of a Time Series?, *International Journal of Forecasting*, 20, 411–425.
- Pesaran, M. H. and A. Timmermann (2007), Selection of Estimation Window in the Presence of Breaks, *Journal of Econometrics*, 137, 134–161.

- Pesaran, M. H. and A. Timmermann (2009), Testing Dependence Among Serially Correlated Multicategory Variables, *Journal of the American Statistical Association*, 104, 325–337.
- Pettenuzzo, D. and A. Timmermann (forthcoming), Predictability of Stock Returns and Asset Allocation under Structural Breaks, *Journal of Econometrics*.
- Piazzesi, M., M. Schneider, and S. Tuzel (2007), Housing, Consumption and Asset Pricing, *Journal of Financial Economics*, 83, 531–569.
- Rapach, D. E., J. K. Strauss, and G. Zhou (2010), Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy, *Review of Financial Studies*, 23, 821–862.
- Rapach, D. E. and M. E. Wohar (2006), Structural Breaks and Predictive Regression Models of Aggregate U.S. Stock Returns, *Journal of Financial Econometrics*, 4, 238–274.
- Ravazzolo, F., D. van Dijk, R. Paap, and P. H. Franses (2008), Bayesian Model Averaging in the Presence of Structural Breaks, in M. Wohar and D. E. Rapach (eds.), *Frontiers of Economics and Globalization Series Vol.3*, Emerald Group Publishing, pp. 561–594.
- Richardson, T. and D. R. Peterson (1999), The Cross-Autocorrelation of Size-Based Portfolios is not an Artifact of Portfolio Autocorrelation, *Journal of Financial Research*, 22, 1–13.
- Ritter, C. and M. A. Tanner (1992), Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler, *Journal of the American Statistical Association*, 87, 861–868.
- Romer, C. D. (1990), The Great Crash and the Onset of the Great Depression, *The Quarterly Journal of Economics*, 105, 597–624.
- Romer, C. D. (1993), The Nation in Depression, *Journal of Economic Perspectives*, 7, 19–39.
- Sargent, T. J. and C. A. Sims (1977), Business Cycle Modeling without Pretending to Have Too Much a Priori Economic Theory, Working Papers 55, Federal Reserve Bank of Minneapolis.
- Schwert, G. W. (1989), Why Does Stock Market Volatility Change over Time?, *Journal of Finance*, 44, 1115–1153.

- Sensier, M. and D. van Dijk (2004), Testing for Volatility Changes in U.S. Macroeconomic Time Series, *The Review of Economics and Statistics*, 86, 833–839.
- Sethuraman, J. (1994), A Constructive Definition of Dirichlet Priors, *Statistica Sinica*, 4, 639–650.
- Sichel, D. E. (1994), Inventories and the Three Phases of the Business Cycle, *Journal of Business & Economic Statistics*, 12, 269–77.
- Siegel, A. F. and C. R. Nelson (1988), Long-Term Behavior of Yield Curves, *Journal of Financial and Quantitative Analysis*, 23, 105–110.
- Stock, J. H. and M. W. Watson (2002a), Forecasting Using Principal Components From a Large Number of Predictors, *Journal of the American Statistical Association*, 97, 1167–1179.
- Stock, J. H. and M. W. Watson (2002b), Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business & Economic Statistics*, 20, 147–162.
- Stock, J. H. and M. W. Watson (2005), Implications of Dynamic Factor Models for VAR Analysis, NBER Working Papers 11467, National Bureau of Economic Research, Inc.
- Svensson, L. E. (1994), Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994, NBER Working Papers 4871, National Bureau of Economic Research, Inc.
- Tanner, M. A. and W. H. Wong (1987), The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association*, 82, 528–550.
- The Conference Board (2001), *Business Cycle Indicators Handbook*, available at <http://www.conference-board.org/publications/describebook.cfm?id=852>.
- Timmermann, A. (2000), Moments of Markov Switching Models, *Journal of Econometrics*, 96, 75–111.
- Timmermann, A. (2008), Elusive Return Predictability, *International Journal of Forecasting*, 24, 1–18.
- Turner, C. M., R. Startz, and C. R. Nelson (1989), A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market, *Journal of Financial Economics*, 25, 3–22.
- Vasicek, O. (1977), An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177–188.

- Wachter, J. A. and M. Warusawitharana (2009), Predictable Returns and Asset Allocation: Should a Skeptical Investor Time the Market?, *Journal of Econometrics*, 148, 162–178.
- Welch, I. and A. Goyal (2008), A Comprehensive Look at The Empirical Performance of Equity Premium Prediction, *Review of Financial Studies*, 21, 1455–1508.
- West, K. D., H. J. Edison, and D. Cho (1993), A Utility-Based Comparison of Some Models of Exchange Rate Volatility, *Journal of International Economics*, 35, 23–45.
- West, M. (1992), Hyperparameter Estimation in Dirichlet Process Mixture Models, ISDS Discussion Paper 92-A03, Duke University.
- West, M., P. Müller, and M. D. Escobar (1994), Hierarchical priors and mixture models with applications in regression and density estimation, *Aspects of Uncertainty*, 363–386.
- Zellner, A. (1971), *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.
- Zhou, B. (1996), High-Frequency Data and Volatility in Foreign-Exchange Rates, *Journal of Business & Economic Statistics*, 14, 45–52.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

- 476 A. PARAKHONYAK, *Essays on Consumer Search, Dynamic Competition and Regulation.*
- 477 S. GUPTA, *The Study of Impact of Early Life Conditions on Later Life Events: A Look Across the Individual's Life Course.*
- 478 J. LIU, *Breaking the Ice between Government and Business: From IT Enabled Control Procedure Redesign to Trusted Relationship Building.*
- 479 D. RUSINOVA, *Economic Development and Growth in Transition Countries.*
- 480 H. WU, *Essays on Top Management and Corporate Behavior.*
- 481 X. LUI, *Three Essays on Real Estate Finance.*
- 482 E. L. W. JONGEN, *Modelling the Impact of Labour Market Policies in the Netherlands.*
- 483 M. J. SMIT, *Agglomeration and Innovations: Evidence from Dutch Microdata*
- 484 S. VAN BEKKUM, *What is Wrong With Pricing Errors? Essays on Value Price Divergence.*
- 485 X. HU, *Essays on Auctions.*
- 486 A. A. DUBOVIK, *Economic Dances for Two (and Three).*
- 487 A. M. LIZYAYEV, *Stochastic Dominance in Portfolio Analysis and Asset Pricing.*
- 488 B. SCHWAAB, *Credit Risk and State Space Methods.*
- 489 N. BASTURK, *Essays on Parameter Heterogeneity and Model Uncertainty.*
- 490 E. GUTIÉRREZ PUIGARNAU, *Labour markets, commuting and company cars.*
- 491 M. W. VORAGE, *The Politics of Entry.*
- 492 A. N. HALSEMA, *Essays on Resource Management: Ownership, Market Structures and Exhaustibility.*
- 493 R. E. VLAHU, *Three Essays on Banking.*
- 494 N. E. VIKANDER, *Essays on Teams and the Social Side of Consumption.*
- 495 E. DEMIREL, *Economic Models for Inland Navigation in the Context of Climate Change.*
- 496 V. A. C. VAN DEN BERG, *Congestion pricing with Heterogeneous travelers.*
- 497 E. R. DE WIT, *Liquidity and Price Discovery in Real Estate Assets.*
- 498 C. LEE, *Psychological Aspects of the Disposition Effect: An Experimental Investigation.*
- 499 M. H. A. RIDHWAN, *Regional Dimensions of Monetary Policy in Indonesia.*
- 500 J. GARCÍA, *The moral herd: Groups and the Evolution of Altruism and Cooperation*
- 501 F. H. LAMP, *Essays in Corporate Finance and Accounting.*
- 502 J. SOL, *Incentives and Social Relations in the Workplace.*

- 503 A.I.W. HINDRAYANTO, *Periodic Seasonal Time Series Models with applications to U.S. macroeconomic data.*
- 504 J.J. DE HOOP, *Keeping Kids in School: Cash Transfers and Selective Education in Malawi.*
- 505 O. SOKOLINSKIY, *Essays on Financial Risk: Forecasts and Investor Perceptions.*
- 506 T. KISELEVA, *Structural Analysis of Complex Ecological Economic Optimal Management Problems.*
- 507 U. KILINC, *Essays on Firm Dynamics, Competition and Productivity.*
- 508 M.J.L. DE HEIDE, *R&D, Innovation and the Policy Mix.*
- 509 F. DE VOR, *The Impact and Performance of Industrial Sites: Evidence from the Netherlands.*
- 510 J.A. NON, *Do ut Des: Incentives, Reciprocity, and Organizational Performance.*
- 511 S.J.J. KONIJN, *Empirical Studies on Credit Risk.*
- 512 H. VRIJBURG, *Enhanced Cooperation in Corporate Taxation.*
- 513 P.ZEPPINI, *Behavioural Models of Technological Change.*
- 514 P.H.STEFFENS, *It's Communication, Stupid! Essays on Communication, Reputation and (Committee) Decision-Making.*
- 515 K.C. YU, *Essays on Executive Compensation - Managerial Incentives and Disincentives.*
- 516 P. EXTERKATE, *Of Needles and Haystacks: Novel Techniques for Data-Rich Economic Forecasting.*
- 517 M. TYSZLER, *Political Economics in the Laboratory.*
- 518 Z. WOLF, *Aggregate Productivity Growth under the Microscope.*
- 519 M.K. KIRCHNER, *Fiscal Policy and the Business Cycle - The Impact of Government Expenditures, Public Debt, and Sovereign Risk on Macroeconomic Fluctuations.*
- 520 P.R. KOSTER, *The cost of travel time variability for air and car travelers.*
- 521 Y.ZU, *Essays of nonparametric econometrics of stochastic volatility.*
- 522 B.KAYNAR, *Rare Event Simulation Techniques for Stochastic Design Problems in Markovian Setting.*
- 523 P. JANUS, *Developments in Measuring and Modeling Financial Volatility.*
- 524 F.P.W. SCHILDER, *Essays on the Economics of Housing Subsidies.*
- 525 S. M MOGHAYER, *Bifurcations of Indifference Points in Discrete Time Optimal Control Problems.*